THE UNIVERSITY *of* EDINBURGH

Unclonability and Quantum Cryptanalysis: From Foundations to Applications

Mina Doosti

Doctor of Philosophy
Laboratory for Foundations of Computer Science
School of Informatics
University of Edinburgh
2022

Abstract

The impossibility of creating perfect identical copies of unknown quantum systems is a fundamental concept in quantum theory and one of the main non-classical properties of quantum information. This limitation imposed by quantum mechanics, famously known as the no-cloning theorem, has played a central role in quantum cryptography as a key component in the security of quantum protocols. In this thesis, we look at *Unclonability* in a broader context in physics and computer science and more specifically through the lens of cryptography, learnability and hardware assumptions. We introduce new notions of unclonability in the quantum world, namely *quantum physical unclonability*, and study the relationship with cryptographic properties and assumptions such as unforgeability, randomness and pseudorandomness. The purpose of this study is to bring new insights into the field of quantum cryptanalysis and into the notion of unclonability itself. We also discuss applications of this new type of unclonability as a cryptographic resource for designing provably secure quantum protocols.

First, we study the unclonability of quantum processes and unitaries in relation to their learnability and unpredictability. The instinctive idea of unpredictability from a cryptographic perspective is formally captured by the notion of *unforgeability*. Intuitively, unforgeability means that an adversary should not be able to produce the output of an unknown function or process from a limited number of input-output samples of it. Even though this notion is almost easily formalized in classical cryptography, translating it to the quantum world against a quantum adversary has been proven challenging. One of our contributions is to define a new unified framework to analyse the unforgeability property for both classical and quantum schemes in the quantum setting. This new framework is designed in such a way that can be readily related to the novel notions of unclonability that we will define in the following chapters. Another question that we try to address here is "What is the fundamental property that leads to unclonability?" In attempting to answer this question, we dig into the relationship between unforgeability and learnability, which motivates us to repurpose some learning tools as a new cryptanalysis toolkit. We introduce a new class of quantum attacks based on the concept of 'emulation' and learning algorithms, breaking new ground for more sophisticated and complicated algorithms for quantum cryptanalysis.

Second, we formally represent, for the first time, the notion of physical unclonability in the quantum world by introducing *Quantum Physical Unclonable Functions (qPUF)* as the quantum analogue of Physical Unclonable Functions (PUF). PUF is a hardware assumption introduced previously in the literature of hardware security, as physical devices with unique behaviour, due to manufacturing imperfections and natural uncontrollable disturbances that make them essentially hard to reproduce. We deliver the mathematical model for qPUFs, and we formally study their main desired cryptographic property, namely unforgeability, using our previously defined unforgeability framework. In light of these new techniques, we show several possibility and impossibility results regarding the unforgeability of qPUFs. We will also discuss how the quantum version of physical unclonability

relates to randomness and unknownness in the quantum world, exploring further the extended notion of unclonability.

Third, we dive deeper into the connection between physical unclonability and related hardware assumptions with quantum pseudorandomness. Like unclonability in quantum information, pseudorandomness is also a fundamental concept in cryptography and complexity. We uncover a deep connection between Pseudorandom Unitaries (PRU) and quantum physical unclonable functions by proving that both qPUFs and the PRU can be constructed from each other. We also provide a novel route towards realising quantum pseudorandomness, distinct from computational assumptions.

Next, we propose new applications of unclonability in quantum communication, using the notion of physical unclonability as a new resource to achieve provably secure quantum protocols against quantum adversaries. We propose several protocols for mutual entity identification in a client-server or quantum network setting. Authentication and identification are building-block tasks for quantum networks, and our protocols can provide new resource-efficient applications for quantum communications. The proposed protocols use different quantum and hybrid (quantum-classical) PUF constructions and quantum resources, which we compare and attempt in reducing, as much as possible throughout the various works we present. Specifically, our hybrid construction can provide quantum security using limited quantum communication resources that cause our protocols to be implementable and practical in the near term.

Finally, we present a new practical cryptanalysis technique concerning the problem of approximate cloning of quantum states. We propose variational quantum cloning (VarQlone), a quantum machine learning-based cryptanalysis algorithm which allows an adversary to obtain optimal (approximate) cloning strategies with short depth quantum circuits, trained using the hybrid classical-quantum technique. This approach enables the end-to-end discovery of hardware efficient quantum circuits to clone specific families of quantum states, which has applications in the foundations and cryptography. In particular, we use a cloning-based attack on two quantum coin-flipping protocols and show that our algorithm can improve near term attacks on these protocols, using approximate quantum cloning as a resource. Throughout this work, we demonstrate how the power of quantum learning tools as attacks on one hand, and the power of quantum unclonability as a security resource, on the other hand, fight against each other to break and ensure security in the near term quantum era.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(Mina Doosti)

Lay summary

One of the most routine tasks we do almost every day on our computers is copying a file. A computer file contains information in the form of a string of zeros and ones. But, what if instead of a normal file, data was encoded inside a tiny physical system? In fact, a system from the subatomic world where different rules of physic would apply to it. The set of rules in that scale is known as the theory of quantum mechanics, and quantum mechanics says that if you have a 'quantum file', it is forbidden to copy it! This fundamental rule of physics called the 'no-cloning theorem', or unclonability, while seemingly very limiting, is a very convenient property of nature. It allows us to conceal information and share it securely. Controlling quantum systems for the secure transmission of information and similar cryptographic tasks is called quantum cryptography.

On the other hand, we can also control these subatomic systems to perform computation, leading to physical devices known as quantum computers that perform computational tasks in a fundamentally different way from any 'classical' computer. Despite applications in many areas, such as solving some mathematical problems, optimizing some operations, and simulating complex molecules, quantum computers do not bring good news for our cryptosystems. For the same reason that they are efficient in solving some mathematical problems, they can break many of today's cryptosystems as they are based on the assumption that solving those problems would take a too long time to be feasible.

Even given the long and challenging technological road ahead of building quantum computers, there has been incredible progress in recent years that has brought the idea of quantum computing to reality. Todays's quantum computers, although 'small' in scale and 'low' in quality, can perform interesting tasks even now. Plus, we believe that sooner or later, we will get to the regime where quantum computers will surpass the limit of computation for any classical computers. Thus we need to be prepared for the threats that they will bring on.

The study of cryptography, in the near future, where we can both exploit quantum systems in our favour and will be at risk due to their computational power, is the art and science called quantum cryptanalysis. To master this art, one needs to understand the strengths and limitations of quantum systems. As such, the unclonability will be at the heart of it.

In this thesis, we study quantum unclonability beyond its usual scope. We explore other forms of natural unclonability that not only, are fundamentally connected to no-cloning, but can also be exploited for cryptography. An example of unclonable objects is optical devices that are particularly unique since their formation or manufacturing processes involve factors that we cannot control. As a result, they become physical devices that are not reproducible. This uniqueness makes them also a physical key. These physically unclonable objects can be modelled in the regime of quantum mechanics. A major part of this thesis includes the comprehensive study of them in the quantum regime, their several interesting properties, and finally, their applications in cryptography.

We also explore the relationship between unclonability and learning, that is how efficiently one can learn a quantum or classical system. In this research area, we use different tools from other fields of physics and computer science, such as machine learning. Specifically, we show that we can make a quantum machine to learn how to efficiently create an 'almost' satisfactory copy of a quantum system. This machine-learning algorithm can be used to attack the security of protocols. These attack analyses give a better perspective on the security of cryptosystems with current and future quantum technology and help us design our systems more securely.

Publications and manuscripts

The author has contributed to the following papers and publications in the course of her doctoral program.

Fully or partially included in this thesis:

1. Quantum physical unclonable functions: Possibilities and impossibilities. Quantum 5 (2021)[ADDK21]
   Myrto Arapinis, Mahshid Delavar, **Mina Doosti**, and Elham Kashefi
   DOI:10.22331/q-2021-06-15-475

2. On the connection between quantum pseudorandomness and quantum hardware assumptions. Quantum Science and Technology 7.3 (2022)[DKKC22]
   Mina Doosti, Niraj Kumar, Elham Kashefi, and Kaushik Chakraborty
   DOI:10.1088/2058-9565/ac66fb

3. Client-server identification protocols with quantum PUF. ACM Transactions on Quantum Computing 2.3 (2021)[DKDK21]
   Mina Doosti, Niraj Kumar, Mahshid Delavar, and Elham Kashefi
   DOI:10.1145/3484197

4. Progress toward practical quantum cryptanalysis by variational quantum cloning. Physical Review A 105.4 (2022)[CDKK22]
   Brian Coyle, **Mina Doosti**, Elham Kashefi, and Niraj Kumar
   DOI:10.1103/PhysRevA.105.042604

5. A Unified Framework For Quantum Unforgeability. arXiv preprint (2021)[DDKA21]
   Mina Doosti, Mahshid Delavar, Elham Kashefi, and Myrto Arapinis
   Arxiv:2103.13994

6. Quantum Lock: A Provable Quantum Communication Advantage. arXiv preprint (2021)[CDM+21]
   Kaushik Chakraborty, **Mina Doosti**, Yao Ma, Myrto Arapinis, and Elham Kashefi
   Arxiv:2110.09469

Excluded from this thesis:

7. Differential Privacy Amplification in Quantum and Quantum-inspired Algorithms. arXiv preprint (2022) / workshop paper (SRML)[ADK22]
   Armando Angrisani, **Mina Doosti**, Elham Kashefi
   Arxiv:2203.03604

Acknowledgements

I do not believe that PhD is a one-man (or, in this case, one-woman) journey, or at least it has not been so for me. If I have reached the centre of this spiral, it has been the result of many fortunate 'environmental factors' and 'lucky interactions' with amazing people I met along the way (just like a lucky quantum state that has finally collapsed to the right state). So I do not intend to keep this acknowledgement short, by no means.

First and foremost, I have to thank my supervisors, Elham Kashefi and Myrto Arapinis, the two brilliant, enthusiastic, inspiring and, certainly lovely women I have had the honour to spend my PhD under their supervision. I can never be thankful enough for their non-stop support, their patience, their insightful pieces of advice, and for being much more than just advisors to me. Starting my PhD and coming from a physics background, I have been an outsider to computer science and cryptography. This thesis is the product of their tireless effort in transforming me into the hybrid creature that I am now. Elham, who never came short on thrilling ideas, whose energy always awed me, and who taught me how to be an independent researcher. Myrto, who patiently educated me to appreciate 'formal' mathematical frameworks and taught me how to think like a cryptographer, and was there for me in all the ups and downs. I also greatly thank my examiners Anne Broadbent, and Petros Wallden, my PhD examiners for their valuable comments.

Although I cannot only thank Petros for being my examiner, a special thank goes to him, as he is one of the most intelligent men I have met during my PhD, who never withheld me his time, someone whom I could always go to with my questions about almost any topic, and whom I learned a lot from, and with whom I have had countless hours of fruitful discussions, and pleasant conversations. My PhD would have not been the same without him. I should also thank other faculties of the quantum group at the University of Edinburgh: Raul Garcia-Patron Sanchez and Chris Heunen, as I greatly benefited from their wisdom and insight during my time as a PhD student. I appreciate having the opportunity to study and research in such a generous environment. While almost half of my PhD collided with the covid pandemic and was spent at home, in the other half, I gathered many good memories from my time at the Informatic Forum and all the people there. I also would like to thank Marc Kaplan, my mentor at VeriQloud, for his support and mentorship.

Then, of course, I should thank the comrades, the members of our excellent quantum group over these years, a group that not even a global pandemic could diminish its merit: First, Alex Cojocaru, who is like my little brother, who knows how much I am thankful for his kindness, his friendship and his help all the way to the end, and I thank him also for giving me feedback on this thesis. Mahshid Delavar and Meisam Tarabkhah, for being such all-in friends, and in a way, my family here. Niraj Kumar, who not only was a mentor to me but a fantastic colleague and friend. Brian Coyle, as he truly deserves his title of "the most efficient man in the group", as he is a brilliant researcher and caring friend whom I have learned a lot from. Kaushik Chakraborty, my smart, patient and wonderful col-

league, office-mate and friend, whom I greatly enjoyed working with. Atul Mantri, who despite his short time in the group, was one of the most memorable persons, as he is clever in his work and humour equally. Ellen Derbyshire, for the unique kindness and company she always offered me. Yao Ma, my awesome colleague and friend with all the exciting memories we made togetehr. Rawad Mehzer, an equally clever and deep man whom I always enjoyed having a conversation with. Daniel Mills, a character that I will always respect and never forget. James Mills, with whom I spent memorable times. Ieva Cepaite, who is a cheerful and free soul. And I also thank Ross Grassie, Sima Bahrani, Theodoros Kapourniotis, Debasis Sadhukhan, other amazing postdocs that I had the pleasure to know in the group and discuss with, as well as Pablo Andrés-Martínez and Nuiok Dicaire from the extended Edinburgh quantum group.

And that is not all, since our group is a highly non-local one, with half of the entangled pair in Paris. I should extend my thanks to Dominik Leichtle, Léo Colisson, Ulysse Chabaud, Shraddha Singh, Armando Angrisani, Jonas Landman, Pierre-Emmanuel Emariau, Luka Music, Raja Yehia, Damian Markham, Harold Ollivier, Fred Groshans, Eleni Diamanti with all of whom I have great memories and interactions.

I am perhaps one of the luckiest human beings in having friends whose friendship expands over space-time. I start from my friends in Iran: to Armita, my best friend, the ever-present being in my life despite the distance, and my companion through all the pains and joys, and she knows what else. To Alma and Aria, because thinking about them (and our memories of 225) has always been the most heart-warming thought. To Laaya, Nima, Babak, Alireza and Ghazal, as they all have their own special place for me. I also thank Azadeh, as spending time with her was ever refreshing. Then I thank my lovely people I have met here: Mohammad and Lucy, for all the pleasing times we have spent in these years, and to Ségo, for being such a beautiful person who she is and for all the music we played together and all the connection we have shared. And finally, among friends, I thank my very dear friend, who knows who he is, and although he might find it dull if he ever reads this page, I can't go on without thanking him as he has been such a good friend during the toughest parts of this.

I give my sincere thanks to my mom and dad, my biggest supporters, who never ceased to encourage me during this time, even from the other side of the world, and I thank them for everything they have done for me in my life and for enduring the best and worst of me.

And as I will keep the best for the last, I thank Ramin, my love, my partner in life and crime, and my closest, most reliable company. Without him and his endless support, his love, and his beautiful mind, I probably would not be where I am now, let alone finishing this thesis and this period. Although no gratitude would do justice to what he has been for me all these years, I am glad that I have this chance to properly thank him.

To Ramin, my love,
And to every being who seeks knowledge and beauty.

Table of Contents

List of Figures

1

Introduction

"See that the imagination of nature is far, far greater than the imagination of man."

– Richard Feynman

One of the most impactful scientific revolutions of the 20th century was the development of quantum mechanics. Revolutionary discoveries about the behaviour of light and matter in the late 19th and early 20th centuries, not explainable by existing knowledge of physics by that time, or what we call today *classical physics*, led to a need for a completely new theory. Arguably, the most important among these was the illumination of the nature of light by Planck in 1900 and the photoelectric effect between 1883 and 1900. These physical phenomena that classical physics has still to date failed to explain have created one of those mutually horrifying and exciting situations in science: the sparks of *maybe we got it all wrong!*

Fanning the flames of this idea led to the birth of the main concepts of quantum theory and eventually to a well-established formalisation of quantum mechanics as we know it today. Quantum theory has changed our mindset and understanding of nature to a great degree. Even though it has gracefully and even surprisingly explained those phenomena which classical physics fell short in describing, it did so at the cost of being counter-intuitive and odd to our *classical* minds. It comes with predictions about nature, not as it appears in our everyday life, but as Richard Feynman famously put it "nature, ... she is absurd" [Ric88]. Quantum mechanics left our human mind with questions and mysteries about *probabilistic nature of observation*, *non-locality*, *unclonability* (which is the core topic of this thesis), and many more to ponder about over the years.

Attempting to unravel some of the mysteries of quantum mechanics led to the appearance of quantum information theory [NC10], which takes an information theory approach to study quantum systems. This new field, together with the rise of quantum computation, has engaged physicists, mathematicians, and computer scientists with new fundamental questions about the concept of computation and the differences between classical and quantum versions of it [AC16]. Quantum information, in its simplest form, starts with the idea of considering discrete quantum systems as carriers of information and treating them as quantum

versions of the binary systems we use for the storage and manipulation of information. However, the field expands extensively beyond this humble foundation and incorporates the full framework of information theory and many other powerful mathematical tools such as probability theory, group theory, representation theory and so on. Using all this powerful machinery, quantum information sheds a light on complicated problems of dealing with quantum properties of nature. Thanks to quantum information, we have now developed a much better understanding of these problems to the point that most of them are no longer mysteries, even if still strange. The idea of quantum computing, on the other hand, emerged from the idea of using physical quantum mechanical systems to simulate themselves. A task that seemed to be too hard to simulate using classical digital computers. This idea was introduced by Feynman[1] in 1981, at a conference where he talks about the difficulties of such simulations and asks "Can you do it with a *new kind* of computer? A quantum computer?"[Pre21].

To realise Feynman's groundbreaking idea would require us to understand this *new kind* of computation and to eventually acquire the ability to control quantum systems for performing our desired computational or simulation task. Obtaining this ability, as Feynman has predicted "doesn't look easy", and almost 40 years of relentless research (from his talk) has proved to be truly the case. Notwithstanding the unresolved challenges of controlling quantum systems, there has been remarkable progress in this area, especially in recent years. One of the main challenges is to achieve quantum computers able to perform useful computational tasks, outside the reach of classical computers, which requires a considerably large scale and a high level of control over such systems. In 2019 the Google AI Quantum group announced their quantum computer, with 53 working qubits, has surpassed this limit and has achieved what is famously known as *quantum supremacy* or *quantum advantage* in the realm of computation [AAB+19]. Despite the scepticism and critics about this result [Kal21, HS21, RSK21], it is an undeniable indication of an important fact: quantum computers are no longer just an idea, and we have entered a new era. More specifically, this new era is named NISQ, standing for *Noisy Intermediate-Scale Quantum* devices [Pre18]. The NISQ devices provide on the orders of 10s to 100 noisy qubits, but they can exploit the quantum behaviours of light or matter to execute *some limited* quantum programs. Although limited, they can provide a *laboratory* for theoretical research in the field of quantum computation and quantum information that was not possible until very recently.

The future large-scale quantum computers, on the other hand, are believed to have significant quantum advantages over classical computers for some specific problems, which are not limited to the simulation of physical systems. The range of problems we hope to be able to solve more efficiently with quantum computers extends to decision problems, search problems and learning problems. The ability to solve these problems, not just faster, but perhaps *in a different way*, using quantum properties, has already and will continue to impact many areas of

[1] Although not all the credit should be given to Mr Feynman! The idea of quantum computing in other forms, was also mentioned by Yuri Manin [man80] and Paul Benioff [Ben80] in 1980.

physics, computer science, chemistry [LWG+10, CRO+19, MEAG+20], biology [EWA+21, CRAG18, OSS+21], and even linguistics [HSG13, CK18, MTdFC20, MGdF+21]. One of the fields that has been hugely affected by quantum computing and quantum information is *cryptography*. Quantum mechanics has a rather fascinating and somewhat contradictory relationship with cryptography. When quantum steps into the realm of cryptography, not only does it threaten security, but it may as well enhance it! But most definitely, it has changed the way one can look at cryptography, as it has done with computation. Let us first start with the negative side of the story.

It is well-known that Shor's quantum algorithm [Sho94] menaces many widely-used cryptographic schemes that are based on the mathematical hardness assumptions of factoring and the discrete logarithm problems (such as RSA). A sufficiently large, fault-tolerant and universal quantum computer will be able to run this algorithm and solve these problems efficiently [WK19]. Furthermore, Shor's algorithm is not the only quantum algorithm that can be used as an attack on classical cryptography schemes. Other famous algorithms, such as Grover or Simon, have also been used for this purpose [BHT98, KLLNP16, JL18, LM17, GLRS16, AL13]. Generally speaking, a quantum computer can be seen as a powerful computational resource in the hands of an attacker. Yet, this extra computational power is not the only aspect that can raise an issue. An adversary who has been given the possibility to exploit non-classical properties of quantum data may as well, have other advantages. For example, a quantum adversary can also use entanglement to extract crucial information from a system or use the power of quantum superposition while interacting with cryptosystems. Hence a very central question to ask is *What are the possible advantages of an attacker equipped with quantum capabilities?* Answering this question, in full generality, is brutally challenging and requires a profound understanding of the underlying assumptions (both computational and physical) of cryptographic schemes, as well as the new potential ways for these quantum capabilities to be exploited to break them. Nevertheless, partially addressing this question is one of the central ideas of this thesis.

On the bright side, quantum mechanics and its odd properties provide us with a new way of achieving cryptographic functionalities, or that is to say, a fundamentally distinct type of cryptography: one that is based on the laws of quantum mechanics and the limitations that it imposes on the adversary. The field of research that studies this direction is known as *quantum cryptography*. The most well-known problem studied in this field is *Quantum Key Distribution (QKD)*, a protocol that enables two remote parties to establish a secure key by relying on the characteristics of quantum mechanics [BB14] in the presence of the most possibly powerful quantum adversary. Perhaps one of the most intriguing aspects of QKD is that, under a carefully specified set of assumptions and requirements about the underlying physical systems, it provably achieves the strongest known level of security without any computational assumptions. QKD, however, is not the only example of what quantum cryptography can bring to the table. For example, the wide range of capabilities that quantum features equip us with has motivated the construction of networks where the nodes are armed with the ability to transmit,

store and manipulate small quantities of quantum information[2][WEH18]. These quantum networks (also called a quantum internet) can enable applications fundamentally impossible for purely classical networks and systems, such as quantum money, quantum multi-party computations, and delegated quantum computing. For a better overview of these functionalities and developed protocols, we refer the reader to the *quantum protocol zoo* [Ver19], an open repository for quantum protocols.

In quantum cryptanalysis[3], which is the other main topic of this thesis, we walk a thin line between these two sides, trying to win the everlasting battle between making and breaking cryptosystems in a world (probably not too far away in the future) where both sides can make the most of quantum mechanical systems and computers. As such, quantum cryptanalysis encompasses many subfields of quantum sciences from quantum computing and quantum information to the foundations of quantum mechanics, to better manipulate them for designing secure systems. Furthermore, it has even merged with relatively younger fields, such as quantum machine learning and quantum learning theory, since they provide a new ground for cryptanalysis. Also, given that quantum technology is usually more expensive and resource-intensive than the usual classical computers and existing systems, another balance to maintain in quantum cryptanalysis is between the security guarantees and the required quantum resources. Maintaining this balance, although challenging, is what makes the design of such quantum systems and protocols thought-provoking and theoretically satisfying.

As mentioned above, a key factor for building efficient and secure quantum cryptosystems in the presence of a quantum attacker is to deeply understand the fundamental and non-classical aspects of quantum systems in general. In this spirit, we can ask: *What are the key elements of quantum security?* or in other words, *What is it that leads to the security (of primitives or protocols) in the regime of quantum mechanics?* Depending on the required functionality or protocol, different quantum features have been used, such as entanglement and non-locality, the probabilistic nature of measurements, the indistinguishability of quantum states and conjugate coding, and most of all, *unclonability*.

The unclonability of the quantum state is one of the most exploited and most common non-classical features in any cryptographic functionality that uses quantum systems. This fundamental limitation of quantum mechanics that forbids creating perfect copies of unknown quantum systems is one of the most central properties of quantum mechanics. Maybe that is why it is inherent in almost all quantum protocols and functionalities, even if not consciously employed. It is perhaps safe to say that unclonability is a *resource* for achieving quantum security. However, many questions still remain regarding unclonability, such as *Is the no-cloning of quantum states the only existing form of unclonability?* If not, *what are*

[2]Often referred to as *quantum communication*.

[3]In cryptography, the term 'cryptanalysis' is commonly used for referring to the study of attacks on cryptosystems and often at a practical or implementation level. However, in this thesis, we use this term in a slightly different sense to address both cases of breaking cryptosystems and designing secure ones, or more generally for *analysing* the cryptographic properties of a system

the other notions of unclonability? and *how can we relate them to cryptographic properties?* Or maybe we can ask more fundamental questions such as *Is there any deeper level to the concept of unclonability?* These are some of the questions that we try to tackle in this thesis as we continue to uncover the relationship between the broader notion of unclonability and quantum cryptanalysis. Finally, we aim to use the tools and concepts that we develop and gather along the way for practical applications.

1.1 Thesis overview

We give a brief summary of our contributions and the structure of the thesis. We exclude Chapter 2, which includes the preliminaries and background materials for the various tools we used in this thesis.

- Chapter 3: We start from the foundations while focusing on three primary notions: unclonability, unforgeability and learnability. We first lay an argument about the relationship between unclonability and unknownness, a concept that we formally define for unitary transformations, and then we bring unclonability into a greater regime which encompasses both cryptography and learning theory. This chapter serves as a roadmap for the rest of the thesis, where we focus on different aspects of each of the concepts we will discuss. Moreover, we introduce two main contributions which we will widely employ in the rest of the thesis. The first one is a new class of quantum attacks based on the concept of *emulation*, and the next one is a new unified framework for quantum unforgeability. Within this framework, we enclose the notion of unforgeability for both quantum and classical primitives, and we also provide a hierarchy of definitions. Finally, as a case study of our framework, several impossibility results are given, and some quantum-secure constructions have been introduced. The content of this chapter is the combination of two papers, Quantum physical unclonable functions: possibilities and impossibilities. Quantum 5 (2021)[ADDK21] and A Unified Framework For Quantum Unforgeability." arXiv preprint arXiv:2103.13994 (2021)[DDKA21], and some unpublished results which were excluded from the mentioned papers.

- Chapter 4: This chapter focuses on defining the notion of *quantum Physical Unclonable Functions (qPUF)* and studying its cryptographic properties using the unforgeability framework that has been introduced in the previous chapter. PUFs are a concept borrowed from the hardware security literature. However, as we will see in this chapter, defining a quantum counterpart is not a straightforward translation to the quantum regime but a rather more fundamental generalisation of the notion of physical unclonability. Here, we answer one of the questions we have asked before, *i.e.* we confirm the existence of other forms of unclonability, not unrelated to the unclonability of quantum states in quantum mechanics, while this relationship is more lucid

in the context of quantum PUF compared to classical ones. We focus on the unitary subclass of quantum PUFs, and we prove general no-go results about the unforgeability property of any such primitives. Finally, we formally prove that a large class of them can satisfy a level of quantum unforgeability powerful enough to make them strong and useful hardware tokens for cryptography. This chapter is based on the paper Quantum physical unclonable functions: Possibilities and impossibilities. Quantum 5 (2021)[ADDK21] as a part of a collaboration with *Mahshid Delavar*, *Myrto Arapinis* and *Elham Kashefi*.

- Chapter 5: This chapter is concerned with quantum pseudorandomness and its relationship with physical unclonability and, more generally, quantum hardware assumptions. Pseudorandomness is also one of the most rudimentary building blocks of modern cryptography as it provides the randomness required for cryptographic schemes in an efficient manner. In the quantum world, quantum pseudorandomness has been recently introduced by [JLS18] via the notions of pseudorandom quantum states and pseudorandom unitaries. The pseudorandom quantum objects provide an efficient and computational form of perfect uniform randomness over the Hilbert spaces known as Haar-randomness. In this chapter, we first study the connection between pseudorandom quantum states and unforgeability. We prove that using quantum pseudorandomness will allow the same level of security guarantee for unforgeability while improving efficiency. Then we delve into the relationship between quantum physical unclonability and pseudorandom unitaries, and we show that they are closely connected to the point that they can be derived from each other in terms of functionality. We also show that, interestingly, considering some assumptions over the family of qPUFs, even without assuming the full extent of quantum physical unclonability, will lead to quantum pseudorandom objects. This chapter is the result of a collaboration between *Kaushik Chakraborty*, *Niraj Kumar* and *Elham Kashefi*, published in On the connection between quantum pseudorandomness and quantum hardware assumptions. Quantum Science and Technology 7.3 (2022)[DKKC22].

- Chapter 6: This chapter which includes three main results from three projects is dedicated to applications of quantum physical unclonability and quantum-enhanced physical unclonable functions. In the first part of the chapter, we introduce two new identification protocols based on qPUFs as we have defined and studied in earlier chapters. Our protocols include client-server scenarios: that is, in the first one, a quantum server intends to identify a low-resource client who only owns a qPUF device, and in the second protocol, the client identifies a quantum server with a qPUF device, while we manage to delegate the quantum verification to the server as well such that the client only needs to run a classical verification test. Amid the security proof of these two protocols lies one of our leading arguments earlier concerning unclonability, since we will see how quantum physical unclon-

ability serves as a resource for these protocols to achieve exponential security in only a polynomial number of rounds of quantum communication. We also thoroughly discuss the role of differed quantum testing algorithms as our verification subroutines and compare them, which can be of interest even outside the scope of the presented protocols and more generally for other quantum communication protocols. Furthermore, to provide sufficient theoretical ground and benchmark for the experimental realisation of these protocols in the future, we give a resource analysis in terms of quantum memory, quantum communication and quantum computation resources. In the second part of the chapter, we will show that the results we have demonstrated in Chapter 5 bring on efficiency and practicality to our qPUF-based protocols. Finally, in the last part of this chapter, we introduce a new quantum-enhanced PUF construction that combines classical physical unclonability with quantum communication and combines the best of both worlds. This construction, although weaker than a full quantum PUF, allows for an efficient identification protocol that is implementable with today's technology, for instance, the existing QKD infrastructure, while it achieves a high level of security against quantum adversaries. This application also shows a particular *provable* advantage of quantum communication vs classical ones, as we will discuss through different properties that the protocol achieves. To prove the security of our construction and protocol, we use many tools and previous results from quantum information theory, including entropic uncertainty relations. The first part of the chapter is based on the work done in collaboration with *Niraj Kumar*, *Mahshid Delavar* and *Elham Kashefi*, which resulted in this publication Client-server identification protocols with quantum puf." ACM Transactions on Quantum Computing 2.3 (2021)[DKDK21]. The second part is from a small section of the paper mentioned before [DKKC22], while it was more appropriate to be included in this chapter. Lastly, the third part of this chapter is from a collaboration with *Kaushik Chakraborty*, *Yao Ma*, *Myrto Arapinis*, and *Elham Kashefi*, resulted in the paper Quantum Lock: A Provable Quantum Communication Advantage. arXiv preprint arXiv:2110.09469 (2021)[CDM+21]. We note that this last paper is included partially for more coherence and brevity of the chapter.

- Chapter 7: Finally, we turn to another aspect of the relation between quantum unclonability and quantum cryptanalysis, this time from a machine learning perspective. This chapter introduces a new cryptanalysis toolkit and method based on approximate quantum cloning and variational algorithms. We introduce our machine learning algorithm, VarQlone, that can efficiently learn to (approximately) clone quantum states of a specified family optimally and in a hardware-friendly manner. This algorithm can have several applications in the context of quantum foundation and quantum computing, specifically since it can be run on NISQ devices, as we have done so. However, in this thesis, we are particularly interested in its application for

cryptanalysis. For this purpose, we take two classes of quantum protocols, QKD and quantum coin-flipping, for case studies, and we relate their security to cloning-based attacks based on VarQlone. We argue that cryptanalysis in this new fashion, even for protocols that have been information-theoretically proven secure like QKD, is beneficial since it allows for benchmarking the state of the art of the current technology with the state of the art of sophisticated attacks that are also implementable on current hardware. Besides the relevance in the application, this chapter allows us to come back to the foundations. In the course of this chapter, we connect the security of certain classes of quantum protocols to specific classes of approximate cloning. This type of cloning-based cryptanalysis brings us one step closer to understanding the role of unclonability as a source of security in quantum cryptography. We even offer several theoretical guarantees and results on the specifications of this algorithm that could be of interest to the quantum machine learning community. The content of this chapter is from a collaboration with *Brian Coyle*, *Niraj Kumar* and *Elham Kashefi* and was published in Progress toward practical quantum cryptanalysis by variational quantum cloning." Physical Review A 105.4 (2022)[CDKK22]. Again, since the context of the research done in this paper is wider than the focus and interest of this thesis, we have only included the most relevant parts and main contributions of the current author.

2

Preliminaries

Man, he took his time in the sun
Had a dream to understand
A single grain of sand
He gave birth to poetry
But one day'll cease to be
Greet the last light of the library

– Nightwish, The Greatest Show on Earth

We start with a general remark regarding this chapter. This chapter attempts to cover all the necessary backgrounds, topics, concepts and tools used in this thesis. Some sections mainly provide general knowledge on the subject, and others, briefly introduce, sometimes in more detail, the definitions or tools used later on in the thesis. Throughout this preliminary chapter, whenever a specific notion or tool is introduced, we navigate the reader to the part of the thesis it is employed. However, for a reader with familiarity with the general topics covered here, we suggest skipping this chapter and returning to each subsection when referred to subsequently in the following chapters.

2.1 Quantum information and quantum computing

Let us begin the chapter by giving some background on quantum information and quantum computing. We assume some familiarity with quantum mechanics and although it will not be necessary for understanding the content of this thesis, it is encouraged for enjoying it. We also assume familiarity with linear algebra.

2.1.1 Quantum states and Hilbert space

The concept of *quantum states* and where they live, which is called *state space* comes from the first postulate of quantum mechanics. According to this postulate, any isolated physical system can be described (or be associated with) a *vector*, in a complex vector space with an inner product, known as *Hilbert space*. This vector, that completely describes the physical system, or sometimes the physical

quantum mechanical property of a physical system, is called the *state* of the system and is a normalised unit vector in the Hilbert space [NC10][1]. We denote a Hilbert space of dimension d by $\mathcal{H}$ or sometimes $\mathcal{H}^d$. Any d-dimensional Hilbert space is equipped with a set of d orthonormal vectors called a basis. The most important quantum systems in quantum information are 2-level quantum systems or quantum states living in a 2-dimensional Hilbert space. We call this special state a *qubit*. The following set of vectors is a complete basis for a qubit, referred to as the computational bases:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{2.1}$$

Here $|.\rangle$ (called 'ket') or $\langle .|$ (called 'bra') are known as *Dirac notation* and are the most common notation in quantum. Moreover, for any Hilbert space $\mathcal{H}$ we can define a dual space denoted by $\mathcal{H}^*$, where for any $|\psi\rangle \in \mathcal{H}$, there exists a dual $\langle\psi| \in \mathcal{H}^*$, that is its complex conjugate, such that $\langle\psi|\psi\rangle = 1$. Also the inner product between two vectors $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ is shown in the Dirac notation as $\langle\phi|\psi\rangle$.

Any qubit state can be written as a linear combination of the basis for instance: $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ where $|\alpha|^2 + |\beta|^2 = 1$ for any $\alpha, \beta \in \mathbb{C}$, since the state should be normalised. This linear combination of other quantum states (for instance any basis state) is called a *superposition* of those quantum states. According to quantum mechanics, any normalised superposition of the states is also a valid member of the Hilbert space, due to linearity and hence is another valid quantum state. The coefficients α and β are also called *amplitudes* and are complex numbers. One of the most useful superpositions, in the equal-weight superposition of the computational basis like the following:

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \qquad |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \tag{2.2}$$

one can see that the state $|+\rangle / |-\rangle$ are also orthonormal and hence form another basis for qubit Hilbert space. This basis is called *plus-minus basis* or *X-basis* (we will see why in 2.1.6.2). The generalisation of such uniform superposition basis states in higher dimension is also known as *Fourier basis*.

2.1.2 Mixed states and density matrices

Now, let us give a more general and complementary formalism for describing qubits and quantum states. When we describe a quantum system by a vector in the Hilbert space, we *deterministically* describe its state (however, as we will see, the process of revealing the state is itself probabilistic), in this case, we call it a *pure* quantum state. Nevertheless, not all the systems in nature are like that.

[1] There is an alternative version of the first postulate of quantum mechanics that has an additional statement that the other way is also assumed, meaning that giving a Hilbert space, where a physical system is described in it, any vector of the Hilbert space is also a potential state of a physical system. Sometimes this is inherent in the evolution postulate, however, it is interesting to think about it as the first postulate as well.

Some systems, are in fact, a probability distribution over different pure quantum states. We call these states *mixed states*, and we can represent them as follows:

$$\rho := \sum_s p_s |\psi_s\rangle \langle\psi_s| \tag{2.3}$$

where $|\psi_s\rangle \langle\psi_s|$ is the outer product of the two vectors. The above representation means that we prepare a pure state $|\psi_s\rangle$ with probability p_s. Also, one can see that ρ is no longer a vector, but a matrix. This kind of matrices, called *density matrices* [Fan57] are a more general way of describing all quantum systems, including the pure one, since if we have only one probability $p = 1$, then $\rho = |\psi\rangle \langle\psi|$, which is a pure state and the density matrix equivalent of $|\psi\rangle$.

In operator language, a density operator for a system is a positive semi-definite, Hermitian operator of trace one ($\mathsf{Tr}(\rho) = 1$) acting on the Hilbert space of the system. We denote the set of all the density matrices associated with the Hilbert space $\mathcal{H}$, as $\mathcal{S}(\mathcal{H})$. Geometrically, this is a convex set. Also a pure state always satisfy $\mathsf{Tr}(\rho^2) = 1$, while as for a mixed state $\mathsf{Tr}(\rho^2) < 1$. This is a good criterion for checking the purity of a density matrix.

2.1.2.1 Bloch sphere

For the 2-dimensional space of qubits, there is a simple and pleasant geometrical representation for all the possible pure and mixed states. It is described as a unit sphere, called *Bloch sphere*, as shown in Fig. 2.1.

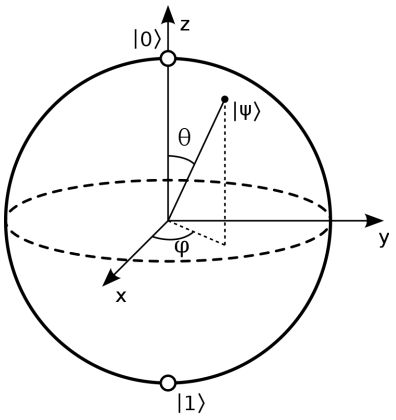

Figure 2.1: The Bloch sphere (figure from: [Wik22])

The surface of the Bloch sphere represents the pure state of a qubit and all the points inside of the sphere represent the set of the mixed states. A pure qubit state can be described in terms of the angles associated with its Bloch vector, as can be seen in the figure. For $0 \leqslant \theta \leqslant \pi$ and $0 \leqslant \phi < 2\pi$, the state of a qubit $|\psi\rangle$ is described as:

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle \tag{2.4}$$

A general qubit density matrix, can be written as [BL06]:

$$\rho = \frac{1}{2}(\mathbb{I} + x_1 X + x_2 Y + x_3 Z) \tag{2.5}$$

where $\mathbb{I}$ is the identity matrix and X, Y and Z are the following matrices known as *Pauli matrices*:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.6}$$

Thus, states of single qubits are characterized by a vector $(x1, x2, x3) \in \mathbb{R}^3$ taken from the unit ball, which is the Bloch sphere.

2.1.2.2 Composition of quantum systems and entanglement

The very original formulation of quantum mechanics talks about a single quantum system described by a vector in Hilbert space. However, if we want to describe the joint state of two quantum systems $\mathcal{H}_1$ and $\mathcal{H}_2$, we need to deal with their composition. The most common framework for the composition of quantum states is tensor product composition. In other words, the composite system $|\psi_{1,2}\rangle$ is described as a unit vector in the tensor product of the Hilbert spaces $\mathcal{H}_1 \otimes \mathcal{H}_2$.[2]

If a quantum state can be written as the tensor product of all its subsystems, we say that the state is *separable*, for example: $|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$. Nevertheless, not all the states in $\mathcal{H}_A \otimes \mathcal{H}_B$ can be written as such. The states that cannot be described in this tensor product form are called *entangled* states, and physically, they contain some non-classical correlation known as *entanglement*. The following defines the general definition of separable and entangled states for bipartite mixed states:

Definition 1 (Separable and entangled mixed states [BL06])**.** A mixed state ρ_{AB} is separable if and only if it can be represented as a convex combination of the product of projectors on local states in the form of the following equation. Otherwise, the mixed state is said to be entangled.

$$\rho_{AB} = \sum_{i=1}^{K} p_i \, |e_i\rangle \langle e_i| \otimes |f_i\rangle \langle f_i| \tag{2.7}$$

where $|e_i\rangle$ and $|f_i\rangle$ are a basis for subsystem A and B respectively.

We also note that the states in Eq. (2.7) describe the most general state of a class of states called LOCC, meaning the most general states that two parties, Alice and Bob can prepare using only *local operation* and *classical communication*.

Another advantage of the density matrix formalism is that it allows us to describe the quantum states of the subsystems of a joint quantum system, even

[2] This is usually considered as an (extended) axioms of the quantum mechanics, however, it is also possible not to assume it, and to derive it instead from general composition rules and physical evidence [AD78].

if the state is not separable [NC10]. For this we can take the partial trace of the other subsystem, to obtain the subsystems of interest, for instance:

$$\rho_A = \mathrm{Tr}_B(\rho_{AB}), \quad \text{and,} \quad \rho_B = \mathrm{Tr}_A(\rho_{AB}) \tag{2.8}$$

where Tr_B meaning taking the trace over subsystem B, using the basis of this subspace. ρ_A and ρ_B are called a *reduced density matrix* of the system.

Now, let us introduce the most important entangled bipartite states in quantum information, also known as *Bell states*, which are as follows:

$$\begin{aligned} |\Phi^{\pm}\rangle &:= \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle) \\ |\Psi^{\pm}\rangle &:= \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \end{aligned} \tag{2.9}$$

where $|00\rangle = |0\rangle_A \otimes |0\rangle_B$ (and similarly for the rest of the basis). These states contain the maximum amount of entanglement between all the pure bipartite states. Furthermore, they have an interesting property that their reduced density matrices are the state $\frac{\mathbb{I}}{2}$, which is a state known as *maximally mixed* state.[3] For instance, we have:

$$\rho_A = \rho_B = \mathrm{Tr}_A(|\Phi^+\rangle\langle\Phi^+|_{AB}) = \frac{\mathbb{I}}{2} \tag{2.10}$$

Looking at the reduced density matrix of joint quantum systems can in general give information about the amount of entanglement contained in these systems.

2.1.3 Quantum operations and measurements

So far we gave a brief introduction to quantum systems and some of their properties. Now it is time to talk about how quantum systems evolve and transform into other quantum systems.

The first form of quantum operation that we know, according to postulates of quantum mechanics, are unitary operators. A unitary matrix U, can transform a pure and mixed quantum state as follows:

$$|\psi'\rangle = U|\psi\rangle \quad \text{and,} \quad \rho' = U\rho U^{\dagger} \tag{2.11}$$

Recalling the Bloch sphere, the unitary operation of any pure qubit state is equivalent to a rotation of the vector on the surface of the Bloch sphere (up to a phase factor). However, unitary matrices are not the most general form of quantum operations. General quantum transformations are *Completely Positive Trace Preserving (CPTP or CPT)* maps which include also unitary matrices. These operations are also called a *quantum channel* and can map a general density matrix $\rho \in \mathcal{H}$ to another density matrix $\rho' \in \mathcal{H}'$ (where $\mathcal{H}'$ is often the same as $\mathcal{H}$, but not necessarily), as follows:

$$\mathcal{E} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}'), \qquad \rho' = \mathcal{E}(\rho) \tag{2.12}$$

[3] In general maximally mixed state for Hilbert space of dimension d is given as $\frac{\mathbb{I}_d}{d}$.

Quantum channels can take the following general from:

$$\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger \tag{2.13}$$

where E_k are operators that should satisfy the following criterion for the overall operation $\mathcal{E}$ to be trace-preserving:

$$\sum_k E_k^\dagger E_k = \mathbb{I} \tag{2.14}$$

This representation of quantum channels is called *operator-sum* formalism [NC10] and the decomposition of a quantum channel into such operators is also sometimes called *Kraus decomposition*. There also exists an alternative way of representing quantum channels from the point of view of system-environment interaction. This point of view is very interesting since it shows that all the operations can eventually be described via a unitary operation on a larger or expanded Hilbert space which also includes the environment. Let the quantum state ρ be entangled with a system $|E\rangle$ that describes the environment. If a unitary operation is applied to the joint state of the system-environment, the operation that is applied to the system alone can be described as follows:

$$\rho' = \mathcal{E}(\rho) = \mathrm{Tr}_E[U(\rho \otimes |E\rangle\langle E|)U^\dagger] \tag{2.15}$$

This operation is no longer a unitary but a CPTP map. We should also note that this later interpretation gives us a good intuition and toolkit to study the effect of noise on the quantum system in the same way as we describe any other transformations of them. A quantum *noise* is also described as a CPTP map and be studied with the same mathematical toolkits theoretically. Commonly we define specific classes of quantum channels that model the most common errors and noisy behaviour that happens to the actual devices. The most famous ones are *bit-flip noise channel*, *phase-flip noise channel*, *Pauli noise channel*, *depolarising noise channel*, *dephasing noise channel*, and *amplitude-damping noise channel* [BL06, NC10].

2.1.3.1 Measurements

We now introduce one of the most central types of operations in quantum mechanics *i.e.* measurements. Measurement operators offer a mathematical formalism for studying the process of observation and extracting the real values for the physical properties of a quantum system. These values are in some sense *the classical information* of the system given by expectation values of a Hermitian observable. Quantum measurements are described as a set of linear operators $\{M_m\}$ acting on the state where the index m refers to each measurement outcome. If $|\psi\rangle$ is the pure quantum state before the measurement, then the probability of obtaining result m, and the state of the system after the measurement is given as follows

[NC10]:

$$p(m) = Pr[\text{obtaining result } m] = \langle\psi| M_m^\dagger M_m |\psi\rangle$$
$$|\psi_m\rangle = \frac{M_m|\psi\rangle}{\sqrt{\langle\psi| M_m^\dagger M_m |\psi\rangle}} \tag{2.16}$$

Thus quantum measurements are probabilistic in nature, and since the probability should be preserved over the full set of measurement, they also satisfy the following completeness equation:

$$\sum_m \langle\psi| M_m^\dagger M_m |\psi\rangle = \sum_m p(m) = 1 \tag{2.17}$$

This probability rule for quantum systems, also known as *Born's rule*, for a general mixed system is given as follows:

$$p(m) = Pr[\text{obtaining result } m] = \mathsf{Tr}[M_m \rho M_m^\dagger] \tag{2.18}$$

The first type of measurement, and a very useful one, are *projective measurements*, which are given by a set of projective operators. A simple example is a set $\{M_0, M_1\}$ where $M_0 = |0\rangle\langle 0|$ and $M_1 = |1\rangle\langle 1|$. This is a qubit measurement which projects everything in the Z basis of the Bloch sphere, and it is also famously known as *measurement in the computational basis*.[4]

The most general class of measurement in the quantum world is given by a mathematical formalism known as *Positive Operator-Valued Measure (POVM)*. A POVM is described as a set of positive operators $\{E_m\}$ satisfying the relation $\sum_m E_m = \mathbb{I}$, and they obey the same Born's rule as we described earlier. This class also includes the projective measurements, however, one difference between the projective measurements and a POVM non-projective one is that the cardinality of the set of POVM measurements over a Hilbert space can be larger than the dimension as opposed to the projective ones. For instance, for qubits, we have seen that the full set of computational basis measurements, includes 2 projectors (which is the same for measuring on any arbitrary basis), but one can define the following valid set of POVM measurements on a qubit:

$$E_1 = (2-\sqrt{2})|1\rangle\langle 1|, \qquad E_2 = (2-\sqrt{2})|-\rangle\langle -|$$
$$E_3 = \mathbb{I} - E_1 - E_2 \tag{2.19}$$

The POVM can be interpreted physically in different ways. The first one is when we are applying a projective measurement on the joint state of our system within a larger system or correlated with another system. In this case, although the measurement on the larger Hilbert space is projective, the resulting measurement on the main systems that we are interested in is not, and it's instead a POVM. This scenario is similar to the case we have discussed before regarding the CPTP maps

[4] The computational basis measurement can be easily generalised to any dimension by only making the projective operator from the computational basis of that Hilbert space.

and there is a good reason for this resemblance due to the fact that POVMs are also CPTP maps and can be described in that formalism. Another way of physically interpreting the POVM is when the real measurement devices are not perfect and instead of performing a perfect projection, they perform a combination of a projection and another quantum operation. In that case, again the measurement can be mathematically described as a POVM. This last point is very important in distinguishing quantum states from each other as we will see in Section 2.2.

2.1.4 Distance measures

Distance measures are mathematical tools for comparing systems on different aspects, for instance, their information quantity. In the classical world, these comparisons are often straightforward. As an example, to compare classical bit strings, we can simply check their equality. But one can also define a better and more fine-grained distance for classical information, for example, by counting the number of places where two bitstrings are different. This distance is called *Hamming distance* in classical information theory and gives a good measure for quantifying classical information in many cases. But how about quantum information. As we know by now, the quantum information lives inside the state of a qubit, that is, a vector in a continuous vector space. And more importantly, while revealing this information (measurement) we are dealing with a probabilistic process. Hence comparing and quantifying the distance between quantum information and generally quantum systems are more tricky! Fortunately, the mathematical background of quantum mechanics and quantum information is strong enough to handle this more complicated situation, and a large variety of quantum distance measures have been defined in the literature, each of which, is useful for different problems that we face in this field [NC10, BL06, GR18, MPS+10, MPO22, BDS+18, GLN05]. Here we only introduce the very few most relevant distance measures for this thesis.

But before, that let us start with a classical distance, which has been incorporated in the quantum regime very similarly. This is the case when we want to compare two probability distributions $\{p(x)\}$ and $\{q(x)\}$. One very common and quite intuitive way of defining a notion of distance for them is as follows:

$$d(p(x), q(x)) := d_{\ell_1}(p(x), q(x)) = \frac{1}{2}\sum_x |p(x) - q(x)| \tag{2.20}$$

This distance is called *trace distance* or ℓ_1*-norm*. And the first quantum distance that we introduce is the generalisation of this distance. The trace distance is defined as follows:

Definition 2 (Trace distance)**.** For any general quantum state ρ and σ, their trace distance is defined as:

$$d_{\mathrm{Tr}}(\rho,\sigma) := \frac{1}{2}\mathrm{Tr}|\rho-\sigma| = \frac{1}{2}\sum_i |\lambda_i^{\rho\sigma}| \tag{2.21}$$

where $|\rho-\sigma| := \sqrt{(\rho-\sigma)^\dagger(\rho-\sigma)}$ is defined as the positive square root of the matrix, and $\lambda_i^{\rho\sigma}$ are eigenvalues of the Hermitian, but not necessarily positive, matrix $(\rho-\sigma)$.

The second measure of distance that is widely used all over quantum information, and this thesis included, is *fidelity*. Although fidelity is not a *metric* on the space of density matrices [NC10], it is one of the most useful measures of the 'closeness' of two quantum states. One of the reasons is that fidelity has an operational meaning: intuitively it expresses the probability that one state will pass a test to identify as the other one. The fidelity between two quantum states in the most general case is known as *Uhlmann's fidelity* and is defined as follows:

Definition 3 (Fidelity)**.** For any general quantum state ρ and σ, their Uhlmann's fidelity is defined as:

$$F(\rho,\sigma) := \left(\mathrm{Tr}\left[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right]\right)^2 \tag{2.22}$$

which is equal to the squared overlap $F(|\psi_\rho\rangle, |\psi_\sigma\rangle) := |\langle\psi_\rho|\psi_\sigma\rangle|^2$, for two pure quantum states $|\psi_\rho\rangle$ and $|\psi_\sigma\rangle$.

The third quantum distance that we introduce, is closely related to the fidelity. In fact, we first define a geometrical metric in the space of quantum states, known as *Bures angle*, as follows:

$$\Theta_{\mathrm{BA}}(\rho,\sigma) := \arccos\sqrt{F(\rho,\sigma)} \tag{2.23}$$

The Bures angle is itself a distance but it is also associated with another important metric in the quantum information, known as *Bures distance* which is also the quantum equivalent of *Fubini-Study* metric [Stu05].

Definition 4 (Fubini-Study/Bures distance)**.** For any general quantum state ρ and σ, their Bures/Fubini-Study distance is defined as follows:

$$d_{\mathrm{FS}}(\rho,\sigma) := \left(2(1-\sqrt{F(\rho,\sigma)})\right)^{\frac{1}{2}} = \sqrt{2(1-\cos\Theta_{\mathrm{BA}})} \tag{2.24}$$

There are several important and useful properties and features of these distances which we need to cover for the purpose of this thesis. First, we need to note that all these distances, including the trace distance and fidelity, should be

preserved under the unitary evolution of quantum states. This is because unitaries, preserve the inner product and hence should also preserve any notion of distance that we define over the space of density matrices. Thus we have the following central relations [NC10]:

$$\begin{aligned} d_{\mathrm{Tr}}(U\rho U^\dagger, U\sigma U^\dagger) &= d_{\mathrm{Tr}}(\rho,\sigma) \\ F(U\rho U^\dagger, U\sigma U^\dagger) &= F(\rho,\sigma) \end{aligned} \tag{2.25}$$

But how about the distance between quantum states, after a non-unitary general quantum channel is applied to them? It can be shown that quantum channels are *contractive*, meaning that they decrease the distance between quantum states. This is captured in the following theorem in terms of trace distance:

Theorem 1. *[Contractivity of quantum channels [NC10]] Suppose $\mathcal{E}$ is a CPTP map. Let ρ and σ be any two density operators. We have:*

$$d_{Tr}(\mathcal{E}(\rho),\mathcal{E}(\sigma)) \leqslant d_{Tr}(\rho,\sigma) \tag{2.26}$$

The same result can be reformulated in terms of fidelity leading to the fact that $F(\mathcal{E}(\rho),\mathcal{E}(\sigma)) \geqslant F(\rho,\sigma)$ under any CPTP operation, which is usually referred to as *Monotonicity of the fidelity* [NC10].

Another property of trace distance that will come in very handy in our proofs is what is known as *strong convexity* and is stated as follows:

Theorem 2. *[Strong convexity of the trace distance [NC10]] Let $\{p_i\}$ and $\{q_i\}$ be two probability distributions over the same index set, and ρ_i and σ_i be density matrices associated with the same index set. Then the trace distance satisfies the following:*

$$d_{Tr}\left(\sum_i p_i\rho_i, \sum_i q_i\sigma_i\right) \leqslant d_{\ell_1}(p_i,q_i) + \sum_i d_{Tr}(\rho_i,\sigma_i) \tag{2.27}$$

where $d_{\ell_1}(p_i,q_i)$ is the classical trace-distance between the two probability distribution.

Similarly, we have *strong concavity* for fidelity:

Theorem 3. *[Strong concavity of fidelity [NC10]] Let $\{p_i\}$ and $\{q_i\}$ be two probability distributions over the same index set, and ρ_i and σ_i be density matrices associated with the same index set. Then the trace distance satisfies the following:*

$$F\left(\sum_i p_i\rho_i, \sum_i q_i\sigma_i\right) \geqslant \sum_i \sqrt{p_i q_i} F(\rho_i,\sigma_i) \tag{2.28}$$

which also results in the weaker version called *concavity* of the fidelity:

$$F(\sum_i p_i \rho_i, \sigma) \geqslant \sum_i p_i F(\rho_i, \sigma) \tag{2.29}$$

Finally, it is important to be able to translate between fidelity and trace distance. The relation between the two, is given via the following inequality:

$$1 - \sqrt{F(\rho,\sigma)} \leqslant d_{\mathrm{Tr}}(\rho,\sigma) \leqslant \sqrt{1 - F(\rho,\sigma)} \tag{2.30}$$

One can also define distances and norms on the operator space. These distances are called *operator norms*. The first important example is the *operator infinity norm* or ℓ^∞-norm. In general, an operator norm ℓ^∞ is defined on a Banach space[5] or a bounded sequence of elements or vectors of that space as follows:

$$||x||_\infty = \sup_n |x_n|, \tag{2.31}$$

The operator norm on the Hilbert space is defined over the space of bounded linear operators as,

$$||O||_\infty = \sup ||Ox|| : \forall ||x|| \leqslant 1, \tag{2.32}$$

We also note that for the operator norms, $||.||_1$ is the dual norm of $||.||_\infty$ [HR16].

The final distance measures that we introduce, which is particularly beneficial distance in the quantum setting, is a distance called *diamond norm*, defined as follows:

Definition 5 (Diamond norm)**.** For any two CPTP map (quantum channel) Λ_1, Λ_2, their diamond norm is defined as,

$$\| \Lambda_1 - \Lambda_2 \|_\diamond := \max_\rho (\| (\Lambda_1 \otimes \mathbb{I})[\rho] - (\Lambda_2 \otimes \mathbb{I})[\rho] \|_1) \tag{2.33}$$

where $\| \, . \, \|_1$ is the ℓ_1-*norm*, and the maximum has been taken over all the density matrices ρ.

Operationally diamond norm quantifies the maximum probability of distinguishing operation Λ_1 from Λ_2 in a single-use, and it is a sensible measure to quantify the difference between unitary operators or other quantum channels.

2.1.5 Entropic uncertainty relations

In this section, we introduce a more advanced but very useful toolkit in quantum information that is also related to cryptography. It consists of a mathematical framework and several inequalities known as *conditional entropies* or *entropic uncertainty relations*, which have been used in the formal security proofs of several

[5] Banach space is a complete normed vector space. That is, a Banach space is a vector space with a metric that allows the computation of vector length and distance between vectors, and is complete in the sense that a 'Cauchy sequence' of vectors always converges to a well-defined limit that is within the space [AAD11]

quantum protocols, specifically, the Quantum Key Distribution (QKD) protocols [Ren08, TL17].[6] We have mostly exploited the content of this section in Chapter 6, Section 6.4.5.1 and 6.4.6 and Chapter 7 Section 7.2.1.

But first, we need to introduce the notion of entropy in quantum and classical information. In classical information theory, the entropy for a random variable is defined as follows, and it is called *Shannon's entropy*.

Definition 6 (Shannon's entropy)**.** Let X be a discrete random variable on a finite set $\mathcal{X} = \{x_1, \ldots, x_n\}$, with probability distribution function $p(x) = Pr(X = x)$. The entropy $H(X)$ of X is defined as:

$$h(X) = -\sum_{x \in \mathcal{X}} p(x) \log p(x) = \mathbb{E}[-\log p(X)] \tag{2.34}$$

while we consider $0 \log 0 = 0$ and logarithm is usually taken to the base 2, in which case the entropy is measured in bits.

Intuitively, this measure quantifies the amount of 'information' (or on the other side 'uncertainty') in a system. As we mentioned, quantum states also include a certain degree of classical information, which we can extract through the probabilistic procedure of measurements. As a result, we can also assign entropy to quantum states. The quantum version of Shannon's entropy is called *Von Neumann entropy*[7] which is defined as follows:

Definition 7 (Von Neumann entropy)**.** For a quantum-mechanical system described by a density matrix ρ, the von Neumann entropy is

$$S(\rho) = -\mathsf{Tr}(\rho \log \rho) = -\sum_i \lambda_i \log_2(\lambda_i) \tag{2.35}$$

where λ_i are the eigenvalues of ρ, we consider $0 \log 0 = 0$ and logarithm is taken to the base 2 or e. Furthermore $S(\rho)$ is zero if and only if ρ represents a pure state.

We are now ready to talk about uncertainty in quantum mechanics, in terms of entropy. Heisenberg's uncertainty principle is one of the most important fundamental properties of quantum mechanics which is mathematically speaking due to the non-commuting property of observables, like Pauli X and Z. Reformulating

[6] For this subsection we assume some familiarity with the concept of QKD protocols since we refer to it several times. However, the details of the protocol are not compulsory for understanding the tools that we introduce here. We do not intend to give a background on the QKD protocol(s) as it will make this preliminary even longer than it is. We refer the readers to [NC10] for a general overview of the protocol and to [TL17] for a comprehensive and advanced description and security proofs.

[7] In this thesis, we mostly use $H(.)$ to denote the general notion of entropy which can in some cases refer to Shannon's entropy and in some others to Von Neumann entropy. However, if we specifically want to emphasise Shannon's entropy, we use the notation $h(.)$

these relations in terms of entropic quantities has been very useful and the most well-known uncertainty relation for these operators was given by Deutsch [Deu83] and later improved [MU88]. The relation, for X and Z observables, is given as follows:

$$H(X) + H(Z) \geqslant \log_2(\frac{1}{c}) \tag{2.36}$$

where c denotes the maximum overlap between any two eigenvectors of X and Z. Since the entropy is defined with respect to a random variable, we need to see what are our random variables here. First, we consider a quantum system A where the state is described with the density matrix ρ_A on a finite-dimensional Hilbert space. We then assume a measurement is performed on a X and Z basis as projective operators that project the state into the subspace spanned by those bases. Thus the random variables are defined via measurements of the observers X and Y. In the most general case, the measurements are a set of POVM operators on system A denoted as $\{M^x\}_x$ and $\{N^z\}_z$ satisfying the Born rule for obtaining outcomes x and z to be as follows:

$$P_X(x) = \mathsf{Tr}[\rho_A M^x] \quad , \quad P_Z(z) = \mathsf{Tr}[\rho_A N^z] \tag{2.37}$$

In this case, the Eq. (2.36) still gives the generalised uncertainty relation with the difference that the c is defined as follows:

$$c = \max_{x,z} c_{zx}, \quad \text{and} \quad c_{xz} = \| \sqrt{M^x}\sqrt{N^z} \|^2 \tag{2.38}$$

where $\| \, . \, \|$ denotes the operator norm (or infinity norm) defined in Section 2.1.4. The above uncertainty relation can be extended to conditional entropy as well in the context of *guessing games* as has been defined in [CBTW17]. Assume two parties, Alice and Bob, where Bob prepares a state ρ_A and Alice randomly performs the X and Z measurements leading to a bit K. Then Bob wants to guess K given the basis choice $R = \{0, 1\}$. The conditional Shannon entropy is defined as follows:

$$H(K|R) := H(KR) - H(R) \tag{2.39}$$

Thus one can get the same uncertainty relation with the conditional entropy as:

$$H(K|R = 0) + H(K|R = 1) \geqslant \log_2(\frac{1}{c}) \tag{2.40}$$

Similar, to the classical case, for a bipartite system ρ_{AB} the conditional Von Neumann entropy is defined as follows:

$$H(A|B) := H(\rho_{AB}) - H(\rho_B) \tag{2.41}$$

Furthermore, this can be generalised to any tripartite quantum system with state ρ_{ABC}. An interesting property here is an inequality referred to as *data processing inequality* [CBTW17] which states that the uncertainty of A conditioned

on some system B never goes down if B performs a quantum channel on the system. In other words for any tripartite system ρ_{ABC} where system C will perform a quantum operation on the quantum state to extract some information, we have the following:

$$H(A|BC) \leqslant H(A|B) \tag{2.42}$$

Given the above inequality leads to the general uncertainty relations between any tripartite system. One of the most common scenarios is when we have two honest parties, Alice and Bob, and an *eavesdropper* or *adversary* called Eve, for example in the QKD protocol. In this case, the following entropic inequality holds:

$$H(K|ER) + H(K|BR) \geqslant \log_2(\frac{1}{c}) \tag{2.43}$$

Where K is the measurement output and R is the basis bit. This imposes a fundamental bound on the uncertainty in terms of von Neumann entropy, in other words, the amount of information that an eavesdropper can extract from the joint quantum systems shared between the three parties, is fundamentally bounded by quantum mechanics. These inequalities can also be extended to the case where n bits are encoded in n quantum states where R^n and K^n are bit-strings denoting the basis random choices for the qubits and measurement outputs respectively, and B^n denotes Bob's bit-string. Also, E denotes Eve's system which is a general quantum system operating on n-qubit messages and any arbitrary local system. We have the following inequality:[8]

$$H(K^n|ER^n) + H(K^n|B^nR^n) \geqslant n\log_2(\frac{1}{c}) \tag{2.44}$$

The amount of information shared between joint quantum systems can also be defined in terms of other informatic quantities such as *mutual information* or *accessible information*. Again, let us discuss these quantities in a two party scenario. Consider a scenario where Alice prepares a pure quantum state drawn from the ensemble $\{p_y, |\psi_y\rangle\}$ with the density matrix ρ_{AB}, where

$$\rho_{AB} = \sum_y p_y |y\rangle_A\langle y| \otimes |\psi_y\rangle_B\langle\psi_y|. \tag{2.45}$$

Bob knows the ensemble i.e., the mixed state ρ_{AB}, but not the particular state that Alice chose. He wants to acquire as much information as possible about y. Bob collects his information by performing a generalized measurement, the POVM $M_{\tilde{y}}$. Bob's state is of the form $\rho_B = \mathrm{Tr}_A(\rho_{AB})$ as it is the subsystem of the larger density matrix. If Alice's preparation choice was y, Bob will obtain the measurement outcome $\tilde{y}$ with conditional probability $p(\tilde{y}|y) = \langle\psi_y| M_{\tilde{y}} |\psi_y\rangle$. For this kind of classical-quantum state ρ_{AB}, the amount of information Bob can extract from this measurement is given by a quantity called *mutual information (MI)* $I(Y;\tilde{Y})_\rho$ between $Y, \tilde{Y}$ which is defined as follows.

$$I(Y:\tilde{Y}) := h(Y) - h(\tilde{Y}|Y), \tag{2.46}$$

[8] This is the main result we will use in Section 6.4.5.1 and Section 6.4.6 for our security proof.

Nonetheless, all the entropic quantities that we have discussed so far, work well in the asymptotic limit, while other similar quantities are more suited to capture the finite-size systems. *Min- and max-entropy* are the notion first proposed by Renner [Ren08] as the natural generalizations of what was known as *conditional Rényi entropies* [Ré61] to the quantum setting. The definition is as follows:

Definition 8 (Min- and max- entropy [KRS09])**.** Let $\rho = \rho_{AB}$ be a bipartite density operator. The min-entropy of A conditioned on B is defined by:

$$H_{\min}(A|B)_\rho := -\inf D_\infty(\rho_{AB} \,\|\, \mathbb{I}_A \otimes \sigma_B) \tag{2.47}$$

where the infimum ranges over all normalized density operators σ_B on subsystem B and $D_\infty(. \,\|\, .)$ is defined as follows:

$$D_\infty(\tau \,\|\, \tau') := \inf\{\lambda \in \mathbb{R} : \tau \leqslant 2^\lambda \tau'\} \tag{2.48}$$

and the max-entropy is defined as:

$$H_{\max}(A|B)_\rho := -H_{\min}(A|C)_\rho \tag{2.49}$$

where the min-entropy is evaluated for a purification ρ_{ABC} of ρ_{AB}.

The above entropies can be then parameterised by a parameter $\varepsilon \geqslant 0$ called the *smoothness* parameter. The smooth version of the min and max entropies can be defined as follows:

Definition 9 (ε-smooth min/max entropy [KRS09])**.** Let $\rho = \rho_{AB}$ be a bipartite density operator and let $\varepsilon \geqslant$ be a parameter. The ε-smooth min/max entropy of A conditioned on B are defined by:

$$\begin{aligned} H^\varepsilon_{\min}(A|B)_\rho &:= \sup_{\rho'} H_{\min}(A|B)_{\rho'}, \\ H^\varepsilon_{\max}(A|B)_\rho &:= \inf_{\rho'} H_{\max}(A|B)_{\rho'} \end{aligned} \tag{2.50}$$

where the supremum ranges over all density operators $\rho' = \rho'_{AB}$ which are ε-close to ρ.[a]

[a] The ε-closeness can be defined with respect to both trace distance and fidelity. However usually defining them in terms of fidelity is more suitable since it is invariant under purification.

The final relevant tool of information theory that we need to introduce in this section is called quantum *Asymptotic Equipartition Property (AEP)* defined in [Ren08]. This is the quantum equivalent of classical (AEP) that roughly speaking, talks about the probability of typical sets occurring in a random or stochastic process when having a series of many random variables. Here we need a special case of quantum AEP for a n-fold quantum-classical system, which we represent

as the following theorem.

Theorem 4 (quantum AEP for quantum-classical states [Ren08]). *Let $\rho_{XB} \in \mathcal{S}(\mathcal{H}_X \otimes \mathcal{H}_B)$ be density operator that is classical on $\mathcal{H}_X$ and let $N \in \mathbb{N}$. Then, for any $\varepsilon \geqslant 0$ we have,*

$$\frac{1}{N} H_{min}^{\varepsilon}(\rho_{XB}^{\otimes N} | \rho_B^{\otimes N}) \geqslant H(\rho_{XB}) - H(\rho_B) - \eta \tag{2.51}$$

where $\eta := (2H_{max}(\rho_X) + 3)\sqrt{\frac{\log(\frac{1}{\varepsilon})}{N} + 1}$, is a function of the smoothing parameter ε and N.

2.1.6 Quantum computing

We have so far given a brief and general background on quantum information. In this section, we will glance over the necessary tools and concepts from quantum computing that we will require for the rest of this thesis. Let us begin with this question: *What do we need to make a quantum computer?* This question was answered by DiVincenzo in [DiV00], where certain criteria have been proposed for constructing a quantum computer. The following are the seven proposed criteria (the last two are necessary for *quantum communication*).

DiVincenzo criteria [DiV00]

1. A scalable physical system with well-characterized qubit
2. The ability to initialize the state of the qubits to a simple fiducial state (quantum state preparation)
3. Long relevant decoherence times
4. A qubit-specific measurement capability
5. A *universal* set of quantum gates

* The ability to interconvert stationary and flying qubits

* The ability to faithfully transmit flying qubits between specified locations

In the previous sections, we have covered the first four criteria since we have introduced qubits and their transformations (as well as the notion of noise) and measurements. In this section, we will focus on the fifth one while we introduce quantum computation and quantum gates. To see why we need quantum gates, we need to have a look at different models of computation, especially in the quantum world.

2.1.6.1 Different models of quantum computation

The first abstract classical model for computation was a *Turing machine (TM)*. A Turing machine contains four main elements [NC10]: (a) a program, (b) a finite state control, co-ordinating the other operations of the machine; (c) a tape, (which is like a memory); and (d) a read-write tape-head, which points to the position on the tape which is currently readable or writable. This simple system is capable of capturing any classical algorithm. The *Quantum Turing machine (QTM)* gives the same type of abstraction for quantum computing. Quantum Turing Machine was firstly proposed by Paul Benioff in 1980 and 1982 [Ben80, Ben82], and then further formalised by David Deutsch in 1985 [Deu85] while the alternative model which we will talk about has also been introduced. Similar to the classical case, a QTM has also a finite set of states $Q = \{q_0, q_1, \dots\}$, a finite set of input and working alphabet and an infinite quantum tape that models the quantum memory and a single 'head'. QTM is usually initialised at a state $|\psi(0)\rangle$ and will perform the computation by applying unitary transformation to the state *i.e.* at every step $|\psi(i+1)\rangle = U|\psi(i)\rangle$. Finally, the process of reading will include quantum measurements as one can expect. Although the intuitive notion of QTM is quite simple, the formal definition is rather complicated and hence we skip introducing it here.[9]

The most common model of quantum computation (and the one used in this thesis) is the *quantum circuit model*, which is a quantum generalisation of the classical *circuit model*. Classically, a circuit consists of several inputs and outputs (bits), wires which describe these systems, and several logical gates [NC10]. A logical gate is a binary function $f : \{0,1\}^m \to \{0,1\}^n$ for example, `AND`, `OR` or `NOT` gates. In the quantum circuit model, on the other hand, our inputs are qubit (or more generally quantum states), and our logical gates are unitary transformations. In the classical circuit model, to be able to perform *any* classical computation we need a *universal gate set*. The quantum circuit model is no different. A set of quantum gate $\mathcal{G} = \{G_i\}$ is universal if *any general n-qubit unitary operation* can be approximated using a quantum circuits that uses this gate set, with an arbitrary accuracy [NC10].

The set of universal quantum gates is not *unique* and different options have been proposed with different theoretical and most importantly implementational advantages and disadvantages for implementing over different types of hardware [CGC+12]. All of them, however, require some *sing-qubit* gates and some *entangling gates*. In the next section, we introduce some of the most widely used quantum gates.

To conclude this section, let us briefly mention another model of quantum computing known as *Measurement-Based Quantum Computing (MBQC)* introduced in [RB01]. The reason for this naming is that in this model, the initial resource is an entangled state in a form of a graph or cluster (called *graph state* and *cluster state*), and each operation is performed by applying a measurement. This technique is also known as *gate teleportation* [GC99]. Moreover, MBQC

[9] Moreover this is not the model of computation that we use in this thesis.

has been shown to be equivalent to the circuit model. We do not go into further details about this model, since we will not use it in the thesis.

2.1.6.2 Quantum gates

The first set of single-qubit quantum gates that we introduce is *Pauli* gates represented by the Pauli matrices we have seen in Eq. (2.6). We first note that the computational basis, are the eigenvectors of Z, and the plus-minus basis, are the eigenvectors of the Pauli matrix X. The eigenvectors of the Y operator are also very similar to the X and are given as follows:

$$|+i\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \qquad |-i\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \tag{2.52}$$

Also one can easily check the action of X, Y and Z gate on the computational basis, by applying their matrix on the basis vectors.

$$\begin{aligned} X|0\rangle &= |1\rangle, \quad X|1\rangle = |0\rangle \\ Y|0\rangle &= -i|1\rangle, \quad Y|1\rangle = i|0\rangle \\ Z|0\rangle &= |0\rangle, \quad Z|1\rangle = -|1\rangle \end{aligned} \tag{2.53}$$

The X gate is the equivalent of the classical 'bit-flip' gate, and the Z gate is a 'Phase gate'. The Y gate is the combination of both since $Y = iXZ$.

The next gate is called *Hadamard gate*, denoted as H that acts as follows on the computational basis:

$$\begin{aligned} H|0\rangle &= |+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ H|1\rangle &= |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \end{aligned} \tag{2.54}$$

The Hadamard gate in fact transforms the computational basis to plus-minus basis. Also, the Hadamard gate creates the symmetric superposition of computational basis even in higher dimension if it is applied as a tensor product form $H^{\otimes n}$ over n qubits. The matrix representation of H is as follows:

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{2.55}$$

As mentioned before, the unitary transformation of a qubit is equivalent (up to a global phase) to rotation on the Bloch sphere, hence one can define the general following rotation single-qubit gates [NC10]:

$$\begin{aligned} R_X(\theta) &:= e^{-i\theta/2X} = \begin{pmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\cos\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \\ R_Y(\theta) &:= e^{-i\theta/2Y} = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ -\cos\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} \\ R_Z(\theta) &:= e^{-i\theta/2Z} = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix} \end{aligned} \tag{2.56}$$

The most useful 2-qubit gates are CNOT (also called CX or controlled-X) and CZ (or controlled-Z) gates. The importance of these gates is that they can create entanglement. Let us first give the matrix representation of these gates:

$$\mathrm{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad \mathrm{CZ} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{2.57}$$

In the above gates, the first qubit acts as a 'control' and the second qubit as a 'target'. CNOT applies a bit-flip or X gate on the second qubit if the control qubit is $|1\rangle$ (and does nothing if it is $|0\rangle$), and similarly, CZ applies a Z gate on the second qubit conditioned on the first one being $|1\rangle$. We also note that the CNOT together with the set of all single-qubit unitary gates form a universal gate set.

Other general *controlled-gates* can be also defined similarly as follows:

$$\mathrm{CU}_{12} = |0\rangle\langle 0|_1 \otimes \mathbb{I}_1 + |1\rangle\langle 1|_2 \otimes U_2 \tag{2.58}$$

where U is conditionally applied on the second qubit.

Finally, a quantum gate that we will use throughout this thesis is another 2-qubit (or generally multi-qubit gate) gate known as the SWAP gate. The SWAP gate on two quantum states with arbitrary dimensions acts as follows:

$$\mathrm{SWAP}\,|\psi\rangle\,|\phi\rangle = |\phi\rangle\,|\psi\rangle \tag{2.59}$$

This gate swaps between the Hilbert space of two quantum states. The qubit SWAP gate can be built from three CNOT gates, and is given with the following matrix:

$$\mathrm{SWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.60}$$

As a final remark, we note that any single-qubit gate may be approximated to arbitrary accuracy using a finite set of gates [NC10]. This is the result of one of the most important theorems in quantum computing, namely *Solovay–Kitaev theorem* [Kit97]. More precisely, the Solovay–Kitaev theorem states that for any single-qubit gate U and any $\varepsilon \geqslant 0$, it is possible to approximate U to a precision ε using $\Theta(log^c(1/\varepsilon))$ gates from a fixed finite set, where c is a small constant approximately equal to 2.

2.2 Distinguishability and verification of quantum states

An important difference between qubits and classical bits (and generally quantum and classical states) is that it is impossible to obtain the exact classical descrip-

tion[10] of a single given copy of a quantum system. This important limitation imposed by quantum mechanics is closely related to *no-cloning theorem*, which we will thoroughly introduce in Section 2.3, and we will also further discuss this fundamental connection in Section 3.2. However, having access to copies of the same quantum system allows for the extraction of the state's description. As a result, there exists a bound on how well one can derive the classical description of quantum states depending on their dimension and the number of available copies. This problem is known as the problem of *state estimation* in quantum information [Hra97, PR04]. Due to the experimental relevance, in addition to quantum information, this problem has been also widely studied in quantum optics [PR04, DJ99, FKF00, OIO+12].

Another tightly related yet different problem in quantum information is the problem of *state discrimination*. State discrimination refers to the task of distinguishing an unknown (pure or mixed) state ρ in a known set of states. More precisely, given an d-level quantum system ρ be one of the states from the set $\{q_i, \rho_i\}_{i=1}^{N}$, that the ensemble of states ρ_is, each happening with probability q_i, the goal is to determine ρ is which one of the states of the set, by performing the best possible POVM, leading to the minimum error discrimination probability.

For the case of two mixed states, the best probability of discrimination is given by a famous bound in quantum information theory known as *Holevo-Helstrom* bound:

$$Pr_{\text{guess}}^{opt} = \frac{1}{2} + \frac{1}{2}\|q_1\rho_1 - q_2\rho_2\|_1 = \frac{1}{2} + d_{\text{Tr}}(q_1\rho_1 - q_2\rho_2) \tag{2.61}$$

Also, for two pure quantum states $|\psi_1\rangle$ and $|\psi_2\rangle$, there exists a general optimal strategy for state discrimination with projective measurements. This result which we represent in the following theorem is an indirect consequence of *Neumark's theorem* (or Naimark's theorem) for general POVMs [BK15], and hence sometimes called Neumark's measurements.

Theorem 5. *The best discrimination strategy for two pure state $|\psi_1\rangle$ and $|\psi_2\rangle$ with projective measurements $\{|v_1\rangle, |v_2\rangle\}$, where $|v_1\rangle$ and $|v_2\rangle$ are in the span of $|\psi_1\rangle$ and $|\psi_2\rangle$such that $\langle v_1|v_2\rangle = 0$, is when they are symmetric with respect to the angle bisector of $|\psi_1\rangle$ and $|\psi_2\rangle$, and $|v_i\rangle$ is closer to $|\psi_i\rangle$ for $i = 1, 2$. On outcome "$|v_i\rangle$", we guess $|\psi_i\rangle$. Moreover, let the angle between $|\psi_1\rangle$ and $|\psi_2\rangle$ be defined as: $\theta = \arccos|\langle\psi_1|\psi_2\rangle|^2$. Then the success probability of this strategy is given by:*

$$Pr_{succ} = |\langle\psi_1|v_1\rangle|^2 = \cos^2\left(\frac{\pi/2 - \theta}{2}\right) = \frac{1}{2} + \frac{1}{2}\sin\theta \tag{2.62}$$

One can check that this optimal probability, can be obtained from Holevo-Helstrom bound in Eq. (2.61) as a special case.

We can also assume another quantum state discrimination scenario where

[10] Here by 'classical description' we mean the value of the amplitudes in a specific basis with arbitrary precision

we do not get any false results, while we allow the measurement outcome to be inconclusive. This means that if the measurement outcome indicates one of the states, we know, for sure, that it's the correct state. Although sometimes the measurement's outcome is: 'I don't know'! In the literature of quantum computing, this scenario is known as *unambiguous state discrimination* [Iva87, Die88, Per88]. We note that this problem is particularly of interest from an experimental point of view and while dealing with imperfect measurement devices and observers [Ber07, MSB04, Ber10]. In this problem again, the goal is to find the best set of POVM measurements for this problem. Then minimise the probability of an inconclusive outcome, finding the optimal POVMs. For the case of two pure quantum states with equal prior probability, the following optimal strategy has been given and proved optimal in [Iva87, Die88, Per88].

Theorem 6. *The best strategy to unambiguously discriminate two pure states $|\psi_1\rangle$ and $|\psi_2\rangle$ is to carry out the POVM $\{E_0, E_1, E_2\}$ where we guess $|\psi_2\rangle$ if the outcome is 1, we guess $|\psi_1\rangle$ if the outcome is 2, and the discrimination is inconclusive if the outcome is 0. The optimal probabilities for this case are given as follows for the angle between $|\psi_1\rangle$ and $|\psi_2\rangle$ is θ.*

$$\begin{aligned} Pr[\textit{outcome 1}] &= tr(E_1\rho) = 1-\cos\theta \\ Pr[\textit{outcome 2}] &= tr(E_2\rho) = 0 \\ Pr[\textit{outcome 0}] &= tr(E_0\rho) = \cos\theta \end{aligned} \tag{2.63}$$

Nevertheless, this is the simplest case of unambiguous state discrimination and the problem has been generalised to N linearly independent states in [CB98], and to mixed states in [BFH06, RST03] and the reader can also find one of the most recent developments on this topic in [Kar21].

Note that, in general, the discrimination problem is directly related to the trace distance between the given quantum states. More generally, distinguishing between different quantum states is related to their *quantum distance*, quantified by distance measures. We have introduced some of them in Section 2.1.4. While possibly the most common distance measure for this purpose is the trace distance, here we want to give a general definition of distinguishability with Uhlmann fidelity. This definition is one of the most acquainted definitions in the thesis.

Definition 10 (μ-distinguishability)**.** Let $0 \leqslant \mu \leqslant 1$ be the distinguishability threshold parameter. We say two quantum states ρ and σ are μ-distinguishable if $0 \leq F(\rho,\sigma) \leqslant 1-\mu$.

Note that two quantum states, ρ and σ, are *completely distinguishable* or 1-distinguishable ($\mu = 1$), if $F(\rho,\sigma) = 0$.

One can also define the ν-indistinguishability in the same manner:

Definition 11 (ν-indistinguishability)**.** Let $0 \leqslant \nu \leqslant 1$ the indistinguishability threshold parameter. We say two quantum states ρ and σ are ν-indistinguishable if $\nu \leq F(\rho, \sigma) \leq 1$.

2.2.1 Verifying quantum states

Due to the impossibility of perfectly distinguishing quantum states, checking the equality of two completely unknown states is a non-trivial task. The task of *equality testing* is a simple but an extremely important task and a building block for lots of complicated quantum protocols [BCWdW01, BBD+97, XAW+15]. The objective is to test whether two *unknown* quantum states are the same. First, we introduce the most well-known quantum algorithm for equality testing. The content of this section is widely used throughout the thesis in all the chapters (perhaps less used in Chapter 7). Thus, familiarity with the notions and notations used here is crucial.

2.2.1.1 SWAP test

Given a single copy of two unknown quantum states ρ and σ, is there a simple test to optimally determine whether the two states are equal or not? This question was answered affirmatively by Buhrman et al. [BCWdW01] when they provided a test called the SWAP test. This test was initially used by the authors to prove an exponential separation between classical and quantum resources in the simultaneous message passing model. Since then it has been used as a standard tool in the design of various quantum algorithms [BCMdW10, KDK17]. A SWAP test circuit takes as an input the two unknown quantum states ρ and σ and attaches an ancilla $|0\rangle$. A Hadamard gate is applied to the ancilla followed by the control-SWAP gate and again a Hadamard on the ancilla qubit. Finally, the ancilla is measured in the computational basis and we conclude that the two states are equal if the measurement outcome is '0' (labelled accept). Fig. 2.2 illustrates this test in the special case when the state σ is a pure state and shown by $|\psi\rangle$.

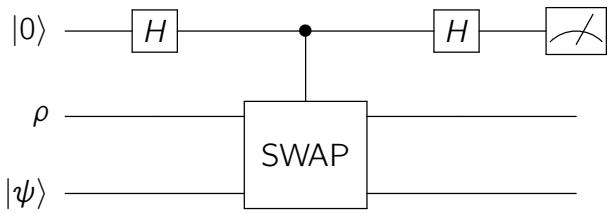

Figure 2.2: The SWAP test circuit

It can be shown that the probability the SWAP test accepts the states ρ and σ is [KMY03],

$$\Pr[\text{SWAP accept}] = \frac{1}{2} + \frac{1}{2}\text{Tr}(\rho\sigma) \tag{2.64}$$

In the special case of when at least one of the states (let's say σ) is a pure state $\sigma = |\psi\rangle\langle\psi|$, the probability of acceptance is,

$$\Pr[\text{SWAP accept}] = \frac{1}{2} + \frac{1}{2}|\langle\psi|\rho|\psi\rangle| = \frac{1}{2} + \frac{1}{2}F(\rho, |\psi\rangle\langle\psi|) \tag{2.65}$$

Thus when at least one of the two states is a pure state, the acceptance probability is related to the fidelity between the states. Implying that if the states are the same, the acceptance probability is 1. However, when the states are different, the SWAP test accepting the states implies an error. The error in the SWAP test, when the states are not identical (also called the one-sided error), is Pr[accept]. Nevertheless, this error can be brought down to any desired error $\varepsilon > 0$ by running multiple instances of the SWAP test circuit. Let M be the number of copies of both input states. Then the number of instances, required to bring down the error probability to a desired ε is,

$$\begin{aligned} \Pr[\text{SWAP error}] = \prod_{j=1}^{M} \Pr[\text{SWAP accept}]_j = (\frac{1}{2} + \frac{1}{2}F)^M = \varepsilon \\ \Rightarrow M(\log(1+F) - 1) = \log(\varepsilon) \Rightarrow M \approx \mathcal{O}(\log(1/\varepsilon)) \end{aligned} \tag{2.66}$$

where $F = F(\rho, |\psi\rangle\langle\psi|) = \langle\psi|\rho|\psi\rangle$ and we use the fact that fidelity is independent of ε.
Now let us introduce a generalisation of this equality test.

2.2.1.2 Generalised SWAP test

The above SWAP test is optimal in Equality testing (in a single instance) of two unknown quantum states when one has a single copy of the two states. However, there are certain quantum protocols where one has access to multiple copies of one unknown state $|\psi\rangle$ and only a single copy of the other unknown state ρ and the objective is to provide an optimal Equality testing circuit. Considering this scenario, Chabaud et al. [CDM+18] provided an efficient construction of such a circuit, a generalised SWAP (GSWAP) test circuit. A GSWAP circuit takes as an input a single copy of ρ, M copies of $|\psi\rangle$ and $\lceil \log M + 1 \rceil$ copies of the ancilla qubit $|0\rangle$. The generalised circuit is then run on the inputs, and the ancilla qubits are measured in the computational basis. Fig. 2.3 is a generic illustration of such a circuit. For more details on the circuit refer to the original work [CDM+18].

It can be shown that the probability the GWAP circuit accepts two quantum states ρ and $|\psi\rangle$ is,

$$\Pr[\text{GSWAP accept}] = \frac{1}{M+1} + \frac{M}{M+1}\langle\psi|\rho|\psi\rangle = \frac{1}{M+1} + \frac{M}{M+1}F \tag{2.67}$$

where $F = F(\rho, |\psi\rangle\langle\psi|)$. We note that in the special case of $M = 1$, the GSWAP test reduces to the SWAP test. Also in a single instance, GSWAP provides a better Equality test compared to the SWAP test since it reduces the one-sided

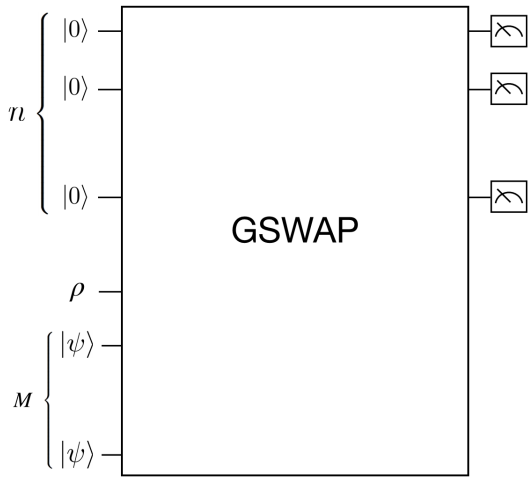

Figure 2.3: GSWAP: A generalisation of the SWAP test with a single copy of ρ and M copies of $|\psi\rangle$. The circuit also inputs $n = \lceil \log M + 1 \rceil$ ancilla qubits in the state $|0\rangle$. At the end of the circuit, the ancilla states are measured in the computational basis.

error probability. In the limit $M \to \infty$, we obtain the optimal acceptance probability of $\Pr[\text{accept}] = F = \langle\psi|\rho|\psi\rangle$. Another important feature of GSWAP is that it can achieve any desired success probability $\varepsilon (\geqslant F)$ in just a single instance which is impossible to achieve using SWAP circuit. However, the number of copies required is exponentially more than the number of instances that the SWAP circuit has to run to achieve the same error probability,

$$\begin{aligned} \Pr[\text{GSWAP error}] = \Pr[\text{GSWAP accept}] &= \frac{1}{M+1} + \frac{M}{M+1} F = \varepsilon \\ \Rightarrow M &\approx \mathcal{O}(1/\varepsilon) \end{aligned} \tag{2.68}$$

2.2.1.3 Abstract quantum test

We can also abstract the notion of quantum equality testing of quantum states. We introduce our own abstract version of such test algorithms by defining the necessary conditions for a general quantum test.

Definition 12 (Quantum Testing Algorithm)**.** Let $\rho^{\otimes\kappa_1}$ and $\sigma^{\otimes\kappa_2}$ be κ_1 and κ_2 copies of two quantum states ρ and σ, respectively. A Quantum Testing algorithm $\mathcal{T}$ is a quantum algorithm that takes as input the tuple $(\rho^{\otimes\kappa_1},\sigma^{\otimes\kappa_2})$ and accepts ρ and σ as equal (outputs 1) with the following probability

$$\Pr[1 \leftarrow \mathcal{T}(\rho^{\otimes\kappa_1},\sigma^{\otimes\kappa_2})] = 1 - \Pr[0 \leftarrow \mathcal{T}(\rho^{\otimes\kappa_1},\sigma^{\otimes\kappa_2})] = f(\kappa_1,\kappa_2,F(\rho,\sigma))$$

where $F(\rho,\sigma)$ is the fidelity and $f(\kappa_1,\kappa_2,F(\rho,\sigma))$ satisfies the following limits:

$$\begin{cases} \lim_{F(\rho,\sigma)\to 1} f(\kappa_1,\kappa_2,F(\rho,\sigma)) = 1 & \forall\,(\kappa_1,\kappa_2) \\ \lim_{\kappa_1=1,\kappa_2\to\infty} f(\kappa_1,\kappa_2,F(\rho,\sigma)) = F(\rho,\sigma) & \\ \lim_{\kappa_1\to\infty,\kappa_2=1} f(\kappa_1,\kappa_2,F(\rho,\sigma)) = F(\rho,\sigma) & \\ \lim_{F(\rho,\sigma)\to 0} f(\kappa_1,\kappa_2,F(\rho,\sigma)) = \mathtt{Err}(\kappa_1,\kappa_2) & \end{cases} \tag{2.69}$$

with $\mathtt{Err}(\kappa_1,\kappa_2)$ characterising the statistical error of the test algorithm.

As an example, for the GSWAP test where $\kappa_1 = 1$ and $\kappa_2 = M$, we obtain from Eq. (2.68) that the probability of acceptance in the limit $F(\rho,|\psi\rangle\langle\psi|) \to 1$ is 1, while it is $\frac{1}{M+1}$ in the limit $F(\rho,|\psi\rangle\langle\psi|) \to 0$. It can be inferred from the above definition that the quantum test can be idealized by forcing the $\mathtt{Err}(\kappa_1,\kappa_2)$ to be zero for any given number of copies. We discuss this last point later in Chapter 4, when we introduce an ideal version of such abstract tests.

2.3 Quantum cloning

In this section, we introduce one of the core concepts of this thesis: the no-cloning theorem and quantum cloning. This section (specifically subsection 2.3.1) is essential for Chapter 7 and not mostly used in other chapters, except for the general notion of no-cloning. First, we discuss the impossibility of *perfectly* cloning quantum states via the no-cloning theorem and then we discuss how we can step out of this limitation on the quantum world and go beyond this impossibility.

The no-cloning theorem states that it is not possible to perfectly clone an *unknown* quantum state. The proof of this theorem is most commonly known to be the work of Wootters and Zurek [WZ82], and also independently done by Dieks [Die82] in 1982. Although it seems that it has originally been discovered in 1970 by Park [Par70]. One formulation of the no-cloning theorem is as follows.

Theorem 7 (The no-cloning theorem)**.** *There exists no unitary transformation U_c that performs the following operation on an arbitrary unknown state $|\psi\rangle \in \mathcal{H}$, a blank (or reference) state $|0\rangle \in \mathcal{H}$ of the same Hilbert space and any arbitrary ancillary state $|a\rangle$:*

$$|\psi\rangle|0\rangle|a\rangle \xrightarrow{U_c} |\psi\rangle|\psi\rangle|a_\psi\rangle \tag{2.70}$$

There exist several proofs of this theorem. Here we give a simple one, similar to what can be found in [BL06]. The proof is by contradiction. Let us assume that such a unitary U_c exists. We note that all the states $|\psi\rangle$, $|0\rangle$ and $|a\rangle$ (whose dimension does not need to be specified) are normalized. Since the state $|\psi\rangle$ is arbitrary, the unitary should be working similarly for any state. Now we assume two non-orthogonal input states $|\psi\rangle$ and $|\phi\rangle$, for which the cloning transformation should be as follows.

$$\begin{aligned} U_c(|\psi\rangle|0\rangle|a\rangle) &= |\psi\rangle|\psi\rangle|a_\psi\rangle \\ U_c(|\phi\rangle|0\rangle|a\rangle) &= |\phi\rangle|\psi\rangle|a_\phi\rangle \end{aligned} \tag{2.71}$$

where $|a_\psi\rangle$ and $|a_\phi\rangle$ represent the output states of the ancilla after the cloning transformation. Regardless of the dimension of the ancillary states, we have that $|\langle a_\psi|a_\phi\rangle| \leqslant 1$. Also, the unitary transformation preserves the inner product. Now, let us inner product both sides of the Eq. (2.71), which leads to the following:

$$\langle\psi|\phi\rangle = \langle\psi|\phi\rangle^2 \langle a_\psi|a_\phi\rangle \Rightarrow \langle\psi|\phi\rangle = \frac{1}{\langle a_\psi|a_\phi\rangle} \tag{2.72}$$

which can clearly be never satisfied and hence the contradiction has been shown. There are two cases where one cannot reach such contradictions. The first one is the trivial case of $\langle\psi|\phi\rangle = 1$, and the other one is the two states are known to be orthogonal *i.e.* $\langle\psi|\phi\rangle = 0$. The latter case is intuitively very insightful since it corresponds to cloning a classical bit, which we know is possible. We will dig further into this in Chapter 3. Moreover, we note that here, no assumption has been made on the unitary, and the no-cloning has been only the result of the unitarity of the transformations in quantum mechanics.

It is also worth mentioning that no-cloning is such a fundamental aspect of nature that it extends to other areas of physics. In fact, cloning is also impossible if we consider the impossibility of superluminal signalling to be held, which we do, due to special relativity. This no-go theorem is known as *no-signalling* [Gis98, NC03] and it is known that if perfect cloning could be possible, it would lead to a contradiction with the fact that no signal can travel faster than the speed of light. No-cloning, is also deeply connected to many other no-go results in quantum information, such as no-broadcasting theorem [BCF+96, PHH08], no-deleting theorem [KPB00], no-superposing theorem [OGHW16, DKK17].

2.3.1 Cloning beyond the no-cloning theorem

Well, *the show must go on*! The no-cloning theorem has not been the end of the road for this part of quantum information. On the contrary, the beginning of a rich field of research. To go around the no-cloning limitation, we must lower our expectations from the cloning machine! Remember that we expected the cloning machine to be *deterministic* (meaning that we always want to get the clones with probability 1) and *exact* (which means that the output states should be both perfect copies of the initial state). It turned out that by relaxing each of these conditions, quantum cloning can be made possible. We call the transformation that achieves such tasks a quantum cloning machine (QCM).

Relaxing the first condition leads to a class of quantum cloning known as *probabilistic cloning*, originated by these works [DG97, DG98]. A probabilistic cloning machine produces *perfect* clones, but only *some times*, i.e. it succeeds with a certain probability. As one can guess, there are information-theoretic upper bounds on this success probability. In this thesis, we do not focus on probabilistic cloning, although we refer the interested readers to these reviews on quantum cloning, including the probabilistic cloning [SIGA05, FWJ+14, BL06].

The second condition, on the other hand, was historically the first one to be relaxed by Buzek and Hillery in [BH96], leading to the field of *Approximate quantum cloning*. In approximate cloning, the operation is deterministic, however, we allow the clones to be not perfect, i.e. have some distance from the original quantum state. The most important property of an approximate cloner is the quality of the output clones, which is measured by the fidelity between the clone and the original state. The next important factor is the family of states we require the cloning machine to clone. Each family of states leads to a class of approximate cloning machines. For instance, if the family of states we consider is *all* the possible states of a Hilbert space, the cloner is called *universal quantum cloner*. However, this set can be made more restricted, meaning that more prior information of the initial states is known, which in turn will lead to cloners with higher optimal fidelity as we will see.

Another property of a cloning machine is *symmetry*. When a cloning machine is symmetric, it means that both of the outputted clones should be the same, relative to the comparison measure. This symmetry is considered when the fidelity is to be optimised, which means that most of the time the symmetric cloners are optimised with respect to *local fidelity*, i.e. the fidelity of each clone state. While we can also have *asymmetric cloners* [Cer00b, Cer00a, IAC+05, DFC05] where the fidelity of one clone is higher than the other one. The quality (especially in the asymmetric case) of the cloner can also be measured with respect to the joint state of both of the clones in comparison to two perfect clones. This fidelity measured is referred to as *global fidelity*. In Chapter 7, where we introduce a machine learning algorithm for the task of cloning, this distinction between global and local fidelity becomes a subtle and important factor.

Finally, the generalisation of the cloning problem is when we have $M > 1$ copies of the input state and we require to create $N > M$ approximate clones (known as

$N \to M$ cloning). In this case enforcing *symmetry* would correspond to

$$F_{\mathrm{L}}^{j} = F_{\mathrm{L}}^{k}, \qquad \forall j,k \in \{1,\dots N\}. \tag{2.73}$$

Where F denotes the fidelity. Now let us briefly discuss three main classes of cloning machines.

2.3.1.1 Universal quantum cloning

The earliest result in approximate cloning was a universal symmetric cloning machine (UQCM) [BH96]. A UQCM is completely agnostic about the input state and is aimed to clone any given quantum state of a Hilbert space with a given dimension. For the case of a qubit, where we want to clone all the qubits on the Bloch sphere, it has been shown that this cloner can achieve the optimal cloning fidelity of $5/6 \approx 0.8333$. This optimal cloner, takes as input an initially unknown state, a blank state and an ancillary qubit, and maps them to a 2-qubit state (the ancillary state is traced out) where the fidelity of each subsystem, or in other words, each reduced density matrices, is optimised to the maximum value of fidelity, leading to the optimal local qubit fidelity of $F_{\mathrm{L,opt}}^{\mathrm{U},1} = F_{\mathrm{L,opt}}^{\mathrm{U},2} = 5/6$.

Now, in the generalised case, we can provide multiple (M) copies of a state to the cloner and request N output approximate clones which is referred to as $M \to N$ cloning [GM97, BEM98]. Generalizing the universal cloning fidelity ($F_{\mathrm{L,opt}}^{\mathrm{U},j}$) := $F_{\mathrm{L,opt}}^{\mathrm{U},j}(1,2)$) to the $M \to N$ scenario, the optimal local fidelity becomes:

$$F_{\mathrm{L,opt}}^{\mathrm{U},j}(M,N) = F_{\mathrm{L,opt}}^{\mathrm{U}}(M,N) = \frac{MN+M+N}{N(M+2)} \tag{2.74}$$

Here $N-M$ ancilla qubits are used to assist, so the initial state is $|\psi_A\rangle^{\otimes M} \otimes |0\rangle^{\otimes N-M}$. We also note that in the limit $M \to \infty$, an optimal cloning machine becomes equivalent to a quantum state estimation machine [SIGA05] for universal cloning. We will further discuss the intuitive meaning and relevance of this important result for our purpose in Chapter 3. In the context of cryptography, $M \to N$ cloning can be also modelled as having N adversaries, $E_1 \dots E_N$ who receive M copies of the state to be cloned.

2.3.1.2 Phase-covariant cloning

Phase-covariant (introduced by [BCMDM00]) states are equatorial states of Bloch sphere. It is common to choose the $\mathsf{X}-\mathsf{Y}$ plane to describe such states:

$$|\psi_{xy}(\eta)\rangle = \frac{1}{\sqrt{2}}\Big(|0\rangle + e^{i\eta}|1\rangle\Big) \tag{2.75}$$

It is known that a phase-covariant cloning machine (PCQCM) that clones one equatorial qubit to two clones has the optimal local fidelity $F_{\mathrm{L,opt}}^{\mathrm{PC}} \approx 0.85 > 5/6$, which is notably higher than the universal case.

These states are relevant since they are used in BB84 QKD protocols and also in universal blind quantum computation protocols [BB14, BFK09]. Here in the context of QKD, we have the following scenario: Assume that Alice wishes to transmit quantum information to Bob (here the information is the bits of the key) but the channel is subject to an eavesdropper, Eve, who wishes to adversarially gain knowledge on the message sent by Alice.

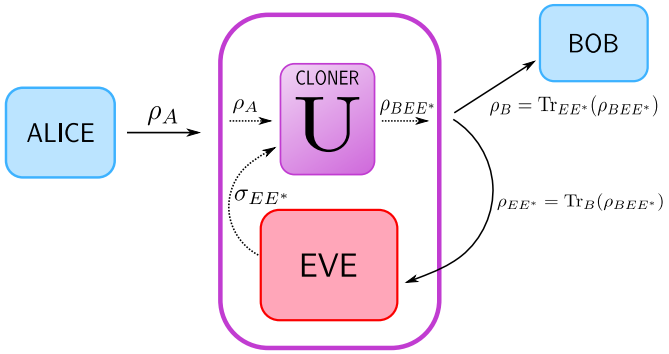

Figure 2.4: Cartoon illustration of an eavesdropping attack by Eve, trying to clone the state, ρ_A, Alice sends to Bob. Eve injects a 'blank' state (which can be a specific state or an arbitrary state depending on the scenario, which ends up as a clone of ρ_A) and an ancillary system (E^*). She then applies the cloning unitary, U. We can assume the cloner is manufactured by Eve to give her the greatest advantage. The state Bob receives will be the partial trace over Eve's subsystems, $\rho_B = \mathrm{Tr}_{EE^*}(\rho_{BEE^*})$, and Eve's clone will be ρ_E, where ρ_{BEE^*} is the full output state from the QCM.

We illustrate this in Fig. 2.4, for a single Eve. In this picture, Alice (A) sends a quantum state[11], ρ_A, to Bob (B). A cloning based attack strategy for Eve (E) could be to try and clone Alice's state, producing a second (approximate) copy which she can use later in her attack, with some ancillary register (E^*).

Interestingly, the cloning of phase-covariant states can be accomplished in an *economical* manner, meaning without needing an ancilla system for Eve, E^*[NG99]. However, as noted in [SIGA05], removing the ancilla is useful to reduce resources if one is *only* interested in performing cloning, but if Eve wishes to attack Alice and Bob's communication, it is more beneficial to apply an ancilla-based attack. Intuitively, this is because the ancilla also contains information about the input state which Eve can extract. Of interest to our purposes, is an explicit quantum circuit which implements the cloning transformation. A unified circuit [BBHB97, FWJ+14, FMWW01] for the above cases (universal and $\mathrm{X}-\mathrm{Y}$ phase covariant) can be seen in Fig. 2.5. The parameters of the circuit, $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \alpha_3\}$, are given by the family of states the circuit is built for [BBHB97, FWJ+14, FMWW01].

For phase-covariant cloning of $\mathrm{X}-\mathrm{Y}$ states, we explicitly have the following

[11] This will typically be a pure state, $\rho_A := |\psi\rangle\langle\psi|_A$, but it can also be generalised to include mixed states, in which the task is referred to as *broadcasting*[BCF+96, CC07, DF07]. The *no-broadcasting* theorem is a generalisation of the no-cloning theorem in this setting.

optimal angles:

$$\begin{aligned}\alpha_1^{XY} = \alpha_3^{XY} &= \arcsin\sqrt{(\frac{1}{2} - \frac{1}{2\sqrt{3}})} \approx 0.477,\\ \alpha_2^{XY} &= -\arcsin\sqrt{(\frac{1}{2} - \frac{\sqrt{3}}{4})} \approx -0.261\end{aligned} \tag{2.76}$$

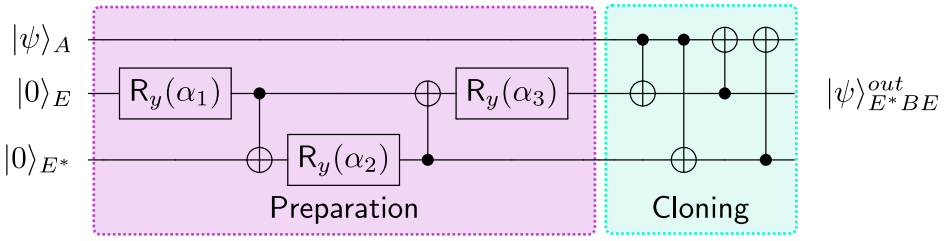

Figure 2.5: Ideal cloning circuit for universal and phase covariant cloning. The Preparation circuit prepares Eve's system to receive the cloned states, while the Cloning circuit transfers information. Notice that the output registers which contain the two clones of $|\psi\rangle_A$ to Bob and Eve in this circuit are registers 2 and 3 respectively.

For completeness, let us also have a look at the optimal local fidelity for the general $M \to N$ case, which has been studied in [FMWW01, DM03]. There is no unique expression for the optimal local fidelity as a function of N and M. However, for the $1 \to N$ case the optimal phase covariant fidelity is given by [DM03] as follows depending on N being odd or even:

$$F^{\mathrm{PC}}_{\mathrm{L,opt}}(1,N) = \begin{cases} \frac{1}{2}(1+\frac{N+1}{2N}) & \text{odd } N \\ \frac{1}{2}(1+\frac{\sqrt{N(N+2)}}{2N}) & \text{even } N \end{cases} \tag{2.77}$$

2.3.1.3 State-dependent cloning with fixed overlap

Now we introduce another class of cloning machines where we aim to clone two non-orthogonal unknown quantum states with a known fixed overlap[12]. This was one of the original scenarios studied in the realm of approximate cloning [BDE+98] but is difficult to tackle analytically. Let us consider the simplest case first, where one considers two states of the type:

$$\begin{aligned}|\psi_1\rangle &= \cos\phi|0\rangle + \sin\phi|1\rangle\\ |\psi_2\rangle &= \sin\phi|0\rangle + \cos\phi|1\rangle\end{aligned} \tag{2.78}$$

which have a fixed overlap, $s = \langle\psi_1|\psi_2\rangle = \sin 2\phi$. It has been shown in [BDE+98] that the optimal *local fidelity* for this scenario is the following:

$$\begin{aligned}F^{\mathrm{FO},j}_{\mathrm{L,opt}} &= \frac{1}{2} + \frac{\sqrt{2}}{32s}(1+s)(3-3s+\sqrt{1-2s+9s^2})\\ &\times\sqrt{-1+2s+3s^2+(1-s)\sqrt{1-2s+9s^2}},\ j\in\{1,2\}\end{aligned} \tag{2.79}$$

[12] This cloning is originally referred to as 'state-dependent cloning', so we herein use this term referring to this scenario.

It can be shown that the *minimum* value for this expression is achieved when $s = \frac{1}{2}$ and gives $F_{\text{L,opt}}^{\text{FO},j} \approx 0.987$, which is much better than the symmetric phase-covariant cloner.

Let us also have a look at the global fidelity in the general $M \to N$ case for this cloning machine which is given as [BL06, BDE+98]:

$$F_{\text{G,opt}}^{\text{FO}} = \frac{1}{2}(1 + s^{M+N} + \sqrt{1-s^{2M}}\sqrt{1-s^{2N}}) \tag{2.80}$$

Interestingly, it can be shown that the state-dependent quantum cloning machine (SDQCM) which achieves this optimal *global* fidelity, does not saturate the optimal *local* fidelity. Computing the local fidelity for the globally optimized SDQCM gives [BL06]:

$$\begin{aligned} F_{\text{L},*}^{\text{FO},j}(M,N) = \frac{1}{4}\Big(\frac{1+s^M}{1+s^N}\left[1+s^2+2s^N\right] + \frac{1-s^M}{1-s^N}\left[1+s^2-2s^N\right] \\ + 2\frac{1-s^{2M}}{1-s^{2N}}\left[1-s^2\right]\Big) \qquad \forall j \end{aligned} \tag{2.81}$$

In contrast, computing the optimal local fidelity for this scenario [BDE+98] (for $1 \to 2$ cloning) is:

$$\begin{aligned} F_{\text{L,opt}}^{\text{FO},j} = \frac{1}{2} + \frac{\sqrt{2}}{32s}(1+s)\left(3-3s+\sqrt{1-2s+9s^2}\right) \\ \times \sqrt{-1+2s+3s^2+(1-s)\sqrt{1-2s+9s^2}}, \ \forall j \end{aligned} \tag{2.82}$$

It can be shown that the *minimum* value for this expression is achieved when $s = \frac{1}{2}$ and gives $F_{\text{L,opt}}^{\text{FO},j} \approx 0.987$, which is also much better than the symmetric phase-covariant cloner. Nevertheless, comparing Eq. (2.82) and Eq. (2.81) reveals that $F_{\text{L},*}^{\text{FO},j}(1,2)$ is actually a *lower* bound for the optimal local fidelity, $F_{\text{L,opt}}^{\text{FO},j}$ in Eq. (2.82). This point is crucially relevant for us and we will go back to it in Chapter 7 where we will use this scenario as a case study for quantum coin flipping protocols and for the design of our variational cloning algorithm.

The state-dependent cloning has been studied concerning the security of QKD. Although, as we have discussed optimal cloning-based attacks for the BB84 protocol are given with the optimal phase covariant cloner. However, this type of cloning has not been widely used for the study of other cryptographic protocols, which is one of our main contributions in Chapter 7. We also note that, interestingly, state-dependent cloning has been recently used to demonstrate advantages related to quantum contextuality [LS20].

2.4 Haar measure and random matrix theory

In this section, we introduce some mathematical background for a concept that is the building block of quantum randomness. The notion of the Haar measure is a particularly important one, and consequently, we have also used it in almost all the chapters. However, other random matrix theory toolkits introduced in this section are used in Chapter 5. In mathematics, the *Haar measure* assigns an 'invariant volume' to subsets of a locally compact topological group. This measure was introduced by Haar in 1933 [Haa33], though its special case for Lie groups had been introduced earlier by Hurwitz as *invariant integral* [DF17]. This measure has been used in many fields such as group theory, representation theory, random matrix theory, ergodic theory and quantum information. Let us first introduce the mathematical definitions and then give an intuition on its application in quantum information.

A Haar measure is a non-zero measure on any locally compact group G such that $\mu : G \to [0, \infty)$ such that for all $X \subset G$ and $x \in G$ we have the following translation invariance property for $\mu(X) = \int_{x \in G} d\mu(x)$:

$$\mu(xX) = \mu(Xx) = \mu(X) \tag{2.83}$$

In particular, the Haar measure $d\mu(U)$ can be defined for a unitary group $U(d)$. Sampling unitaries from Haar measure on $U(d)$ is equivalent to geometrically uniform sampling from unitary groups of that dimension.

Let us take the 2-dimensional Hilbert space as an example. We recall that pure qubit states can be represented as a vector or a point on the surface of the Bloch sphere. Assume that we want to pick uniformly random qubits on the surface of the Bloch sphere. Since every point on the sphere is parameterised by θ and ϕ according to Eq. (2.4), one approach would be to uniformly sample values for these two parameters and the result will be random qubits on the Bloch sphere. Although, as it is shown in Fig. 2.6, by doing this, the resulting vectors will not be uniformly distributed on the surface of the Bloch sphere and they will be more concentrated around the poles. On the contrary, if one samples the vectors uniformly at random according to the Haar measure over $SU(2)$, they will be distributed uniformly over the Bloch sphere (Fig. 2.6 (b)). In practice, however, sampling from the Haar measure requires exponential (in d) resources [Kni95].

We are also interested in characterising the properties of the eigenvalues of Haar-random unitary matrices and their distributions. Problems of this sort have been widely studied in the field of *random matrix theory*. Here we introduce some of the important results in this field that we will use later on in Chapter 5.

The first result that we need, is known as *Weyl density formula* or *Weyl integration formula*, and is stated as follows:

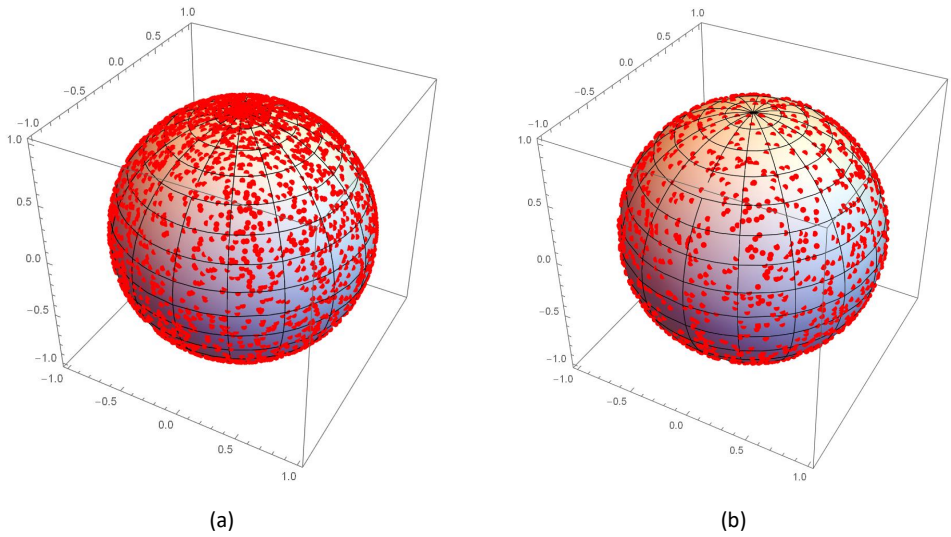

Figure 2.6: Haar measure and non-uniform sampling over the Bloch sphere. Figure (a) shows the case where the qubits have been sampled through a uniform sampling of the qubit parameters. This sampling leads to a geometrically non-uniform sampling over the sphere. Figure (b) shows sampling qubits according to Haar measure over $SU(2)$ which leads to a uniform sampling over the Bloch sphere.

Lemma 1 (Weyl integration formula on $U(n)$ [Mec19])**.** *Let $\{e^{i\theta_j}\}_{j=1}^{n}$ be the eigenvalues of $n \times n$ random unitary matrix. The unordered eigenvalues of a random unitary matrix have the following eigenvalue density*

$$\frac{1}{n!(2\pi)^n} \prod_{1 \leqslant j < k \leqslant n} |e^{i\theta_j} - e^{i\theta_k}|^2 \tag{2.84}$$

with respect to $d\theta_1 \dots d\theta_n$ on $(2\pi)^n$. That is, for any $g : U(n) \to R$ with

$$g(U) = g(VUV^*) \quad \text{for any } U, V \in U(n),$$

(i.e., g is a class function), if U is Haar-distributed on $U(n)$, then

$$\mathbb{E}[g(U)] = \frac{1}{n!(2\pi)^n} \int_{[0,2\pi)^n} \tilde{g}(\theta_1, \dots, \theta_n) \prod_{1 \leqslant j < k \leqslant n} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 \dots d\theta_n \tag{2.85}$$

where $\tilde{g} : [0, 2\pi)^n \to \mathbb{R}$ is the (necessarily symmetric) expression of $g(U)$ as a function of the eigenvalues of U.

As discussed in [Mec19], one consequence of the above lemma is that the eigenvalues of random unitary matrices want to spread out. For any given pair of eigenvalues labelled by (j, k), $|e^{i\theta_j} - e^{i\theta_k}|^2$ is zero if $\theta_j = \theta_k$, and is 4 if $\theta_j = \theta_k + \pi$ (and in that neighborhood if they are roughly antipodal). This produces the effect

alternatively known as 'eigenvalue repulsion'.

Another important tool in the study of the eigenvalues of random matrices is the *empirical spectral measure* defined as,

$$\tilde{\mu} = \frac{1}{n}\sum_{j=1}^{n} \delta_{e^{i\theta_j}} \tag{2.86}$$

where $e^{i\theta_j}$ are the eigenvalues of the unitary matrix and δ is the probability distribution function over the eigenvalues. The empirical spectral measure is a probability measure to encode the ensemble of eigenvalues which puts equal mass at each of the eigenvalues of U. This encoding is very useful for representing the spreading of the eigenvalues on the complex unit circle denoted by $\mathbb{S}^1 \subseteq \mathbb{C}$.

Next, we need the following important theorem by Diaconis-Shashahani [DS94], that shows the convergence of the eigenvalues of the Haar-random matrices to the uniform distribution over the unit circle:

Theorem 8 ([DS94]). *Let U be uniformly chosen from Haar-measure in $U(d)$, Let ν be the uniform distribution on $\mathbb{S}^1$. Then as $d \to \infty$, the $\tilde{\mu}_U$ converges, weakly in probability, to ν:*

$$\tilde{\mu}_U \xrightarrow{d\to\infty} \nu \tag{2.87}$$

Finally, we introduce the following result by Wieand [Wie02] which is very useful in working with the statistics of the eigenvalues of random unitaries.

Theorem 9 ([Wie02]). *Let U be a unitary matrix chosen from Haar measure in $U(d)$, and let $\{e^{i\theta_1}, \ldots, e^{i\theta_d}\}$ be the eigenvalues of U. Fix a finite number of intervals on the unit circle $I_1 = (e^{i\theta_{1j}}, e^{i\theta_{1l}}), \ldots, I_m = (e^{i\theta_{mj}}, e^{i\theta_{ml}})$. Define the random variables $N_{\theta_1}, \ldots, N_{\theta_m}$ to be the number of eigenvalues in each arc defined by the intervals. In the limit of large d, the mean and variance of N_{θ_k} are as follows:*

$$\mathbb{E}_d[N_{\theta_k}] = \frac{d(\theta_{kj} - \theta_{kl})}{2\pi} \tag{2.88}$$

and

$$Var(N_{\theta_k}) = \frac{1}{\pi^2}(\log(d) + 1 + \gamma + \log|2\sin(\frac{\theta_{kj} - \theta_{kl}}{2})|) + o(1). \tag{2.89}$$

where $\gamma \approx 0.577$ is the Euler's constant.

This theorem, gives a concrete formula for calculating the expectation value and variance of the random variable that represents the number of eigenvalues of a random unitary matrix, in each arc of the unit circle and hence can be used to study the distribution of eigenvalues of random matrices.

2.5 Quantum cryptography

Now we focus on another field of research that we tightly connected with this thesis, namely *quantum cryptography*. Quantum cryptography is almost as old as quantum computing itself and studies different cryptographic problems that involve, in several ways, quantum mechanical systems. These quantum systems can be employed by honest parties to perform a cryptographic task, or else can be exploited by an adversary, or a dishonest party, trying to attack the system. We have already seen an example of a secure protocol (*i.e.* QKD) where parties have some limited quantum capabilities like preparing and measuring qubit states in a specific basis. We will introduce another example of such protocols in this chapter that achieves a functionality which is impossible with only 'classical' cryptography (see 2.5.7). Yet another sub-field of quantum cryptography is *post-quantum* cryptography, dedicated to studying the security of 'classical' systems against quantum adversaries and the design of quantum-secure cryptographic schemes. In this field, the most vital aspect of a quantum adversary is its quantum computing capability, especially after the discovery of algorithms such as Shor's and Grover's where there exists (a potentially) significant *quantum-speedup* [WK19]. The main idea here is to keep the cryptographic schemes classical while designing them based on assumptions and mathematical problems that are also hard for quantum computers to solve. Due to the current technological challenges and inefficiency of quantum systems, as well as the incompatibility of many of the quantum protocols with today's existing classical cryptosystems, this idea is perhaps the most popular discipline today for achieving security in the quantum world [WK19, BBGP16]. Among the existing attempts in this field to guarantee 'quantum-resistant' with classical cryptographic schemes, one of the most successful ones is *lattice-based cryptography* [MR09]. However, there is a full spectrum between going towards fully quantum systems and keeping them fully classical. Numerous work has been done in this area, which is also of particular interest in this thesis. Therefore, in this section, we will introduce a handpick of concepts and protocols from different branches of quantum cryptography, which are either essential for the results we will establish later or will help the reader with an improved understanding of future topics. However, the introduction we give here is by no means exhaustive.

Before going over the more technical materials, let us settle on a few basic notations and terminologies that we will widely throughout the thesis.

We start with the notion of *security parameter*. The **security parameter**, which we denote as λ in this thesis, is a parameter that quantifies the security level of the systems, or in other words, the complexity of the problem based on which the cryptographic scheme has been designed. Roughly speaking, the security parameter measures how 'hard' it is for an attacker (which we call adversary from now on) to break the cryptographic scheme. As such, the objective is to design cryptographic schemes that *for any* adversary, the success probability of the adversary is 'small' relative to this parameter.

Now let us focus on the word *'small'* in the previous sentence and try to formalise that. The smallness, in the world of cryptography, is usually formalised

via a concept known as *negligible function*, defined as follows:

Definition 13 (Negligible function)**.** A function $\varepsilon : \mathbb{N} \to \mathbb{R}$ is a *negligible function*, if for every positive integer c, (or equivalently every polynomial function of the security parameter $poly(.)$) there exists an integer $N_c > 0$ (or $N_{poly} > 0$) such that for all $x > N_c$ ($x > N_{poly}$) the following holds:

$$|\varepsilon(x)| < \frac{1}{x^c} \qquad \left(\text{or } |\varepsilon(x)| < \frac{1}{poly(x)}\right) \tag{2.90}$$

Thus, we require the success probability of the adversary to be a negligible function of the security parameter λ, which we denote as either $negl(\lambda)$ or $\varepsilon(\lambda)$. We refer to [KL20] for properties of the negligible functions.

Another terminology that we need to introduce, is the notion of *One-Way Function (OWF)*. An OWF is a function that is 'easy' to compute on every input, but 'hard' to invert given the image of a random input. OWFs are usually considered as a *computational assumption* on cryptography, referring to the hardness complexity in the definition. More formally, an OWF is defined as follows [KL20]:

Definition 14 (One-way function (OWF))**.** A function $f : \{0,1\}^n \to \{0,1\}^l$ is a one-way function if:

1. f can be evaluated in polynomial time on every input.
2. for every *probabilistic polynomial time (PPT)* algorithm $\mathcal{A}$ there exists a negligible function ε such that:

$$\Pr_{X \leftarrow \{0,1\}^n}[\mathcal{A}(f(X), 1^n) \in f^{-1}(f(X))] \leqslant \varepsilon(n) \quad \forall n. \tag{2.91}$$

We also note that one-way-ness can be also defined over a family of functions. OWFs are an important part of modern cryptography, hence we refer the interested reader to [KL20] for more information about the topic. Now, let us introduce another important terminology in the next section.

2.5.1 Formal frameworks for cryptanalysis

Security proofs in cryptography are usually given in a formal mathematical framework that allows for careful analysis of the cryptographic tasks. The most widely-used framework in modern cryptography is the *game-based framework*. In the game-based framework, the cryptographic definition (or task) is formalised as a *game* played between the adversary (In this thesis, we usually denote the adversary as $\mathcal{A}$) and a hypothetical honest party called *challenger* (we usually denote the challenger by $\mathcal{C}$). Both parties can be (and usually are) probabilistic algorithms and processes. The game should be defined in a way that captures the task of interest, and it is defined over a probability space. The security then is proved

by performing a probabilistic analysis and measured in the success probability of the adversary in winning the game. This model is quite popular in cryptography as it is relatively easy to understand and utilise, while it is powerful enough to capture the security of any cryptographic primitive. Moreover, the framework can also be translated in the quantum regime, as has been previously done by several works such as [BZ13a, BZ13b, AMRS20]. In this thesis, we will also use this framework for our security proofs and will define our quantum definitions based on this paradigm.

However, for completeness, we will also briefly mention other security paradigms. The other famous security framework is *simulation paradigm* [GGM86].[13] Here the security is captured by comparing two scenarios (or two worlds): An ideal scenario which is secure by definition, and a real-world, where an adversary can interact with the real execution of the protocol. In this regime, the protocol is secure, if any adversary in the real model cannot do much better than if it was involved in the ideal model or, in other words, the two scenarios are *indistinguishable*.

The simulation-based framework, however, investigates the security in a standalone model, *i.e.* when we consider the execution of that protocol only, as a single instance. However, more complicated cryptographic protocols are often composed of smaller sub-protocols and components and one requires to ensure the security of the whole system is preserved if *secure protocols* are being composed together. Although this is not only non-trivial, there are several examples where this is not the case, *i.e.* composing secure protocols leads to a non-secure system [KL20]. A cryptographic framework that addresses this problem has been introduced by Canetti [Can01] and is called *Universal Composability (UC)* framework. This framework is closer in nature to the simulation paradigm. Despite being very powerful, often proving the security of protocols in this manner is more complicated, and there are considerably fewer cryptographic schemes that have been proven *composably secure*. Another similar framework that also captures the composability issue into account is a framework introduced by Maurer [Mau05] and is called *Abstract Cryptography (AC)* framework. Both of these frameworks have also been extended to quantum setting [MQR09, MR16].

2.5.2 Adversarial models in the quantum world

In this section, we will introduce different adversarial models in the quantum world, specifically the ones that we will mostly encounter in this thesis. Some of the adversaries that we discuss are more common in the quantum information literature (for instance against quantum protocols such as QKD). While some others are, generally speaking, translations of different classes of classical adversaries into the quantum world, and as a result, mostly adopted from cryptography literature. As in this thesis, we deal with various quantum adversaries we attempt to give a general and coherent overview of them in this section.

[13] Also called *simulation-based framework* and *real world/ideal world paradigm*

Let us first set up the definition of an adversary! An adversary is an algorithm, defined as a polynomial-time uniform family of quantum circuits, that can be either deterministic or probabilistic (the probabilistic ones are more common to consider due to generality), and their goal is to perform a task that leads to breaking a protocol, or more generally a cryptosystem, under specific assumptions. Hence, each class of adversaries is usually characterised by the set of assumptions that we consider for such algorithms.

In this thesis, we are interested in *quantum adversaries*, i.e. adversaries who also possess quantum capabilities, in addition to their usual classical computational power. As a result, the first assumption on our adversaries of interest is *quantum mechanics* itself! Thus, we assume that a quantum adversary is subject to the laws of quantum mechanics, which we also assume to be *correct* and *complete*[14].

If no additional assumptions have been made on the adversary, which inherently means that the adversary's computational power is *unbounded*, the adversary is often called *unbounded quantum adversary*. The security that is achieved against such adversaries is called *information-theoretic security*, as opposed to *computational security* where there are assumptions on the computational capabilities [PR21]. This is the strongest known notion of security that the marriage of quantum mechanics and information theory has been made possible [BS16]. The security of many quantum protocols, such as QKD, quantum money, quantum coin-flipping and so on, has been studied in this security model. The adversary sometimes appears in the form of an eavesdropper (usually called Eve) who wants to access some encoded information that is being exchanged through a channel controlled by this adversary (like in the case of QKD). In some other protocols, the adversary plays the role of a malicious party in a protocol who wants to cheat or deviate from the honest behaviour (like in the case of quantum money or quantum coin-flipping).

Nevertheless, this notion is usually too strong, and almost none of the classical cryptosystems that we have can resist unbounded quantum adversaries [WK19, Mos18]. Thus a more common and standard adversarial model is the class of *quantum polynomial-time* (QPT) adversaries. Here the computational power of the adversary has been limited to polynomial time (in the security parameter). The QPT adversary is, in fact, an *efficient* quantum adversary that, in terms of complexity theory, is allowed to run polynomial-time uniform family of quantum circuits. Despite being computationally bounded, this class of adversary is still very powerful, and it is known that many existing cryptosystems based on computational hardness assumptions are still broken against this adversary as well, due to the existence of efficient quantum algorithms such as Shor's algorithm [Sho94], and Grover's algorithm [Gro96] that the adversary can exploit[WK19, Son14]. A

[14] Although one might consider quantum mechanics as an established model that describes the nature, which is true to some extent and precision as all of the theories in physics, its correctness and completeness is still an assumption we (happily) carry along with ourselves throughout this thesis (and generally in quantum cryptography and quantum information). Despite debates on the *completeness* of quantum theory, as discussed in [PR21], this is still a very justifiable assumption to make

QPT adversary can also be given oracle access to the classical or quantum primitive. The oracle model is a common cryptographic technique that is widely used in security proofs since it facilitates modelling adversarial behaviours where some information about the scheme is gathered (in some earlier stages or by interacting with the scheme and observing its properties). We discuss such oracles in the next section as well in Section 3.3.3 in Chapter 3. Since a QPT adversary is polynomial bounded, it is also bounded in the oracle model to a polynomial number of queries to the given oracle [15]. The quantum adversaries that have this somewhat quantum communicative access to the primitives are also sometimes called *online* quantum adversaries, while an *offline* quantum adversary can only have classical information of the primitive, and later on, use a quantum algorithm together with the classical data to break the cryptosystem. The study of the security of classical cryptosystems against *offline* quantum adversaries, is famously known as post-quantum cryptography [BBGP16, BL17, Son14]. Moreover, the more technical term for this security model is the *standard security model*, while as when the oracle access to the primitive is considered quantum, the term *quantum security model* is used [BZ13a, BZ13b, GHS16]. One of the key elements of the quantum security model is the superposition queries, that is the adversary can query many classical values in one quantum query in the form of a superposition of those states. Superposition queries enables a broader range of non-trivial attacks [KLLNP16, SS17, BZ13a, GHS16] that are not possible in the classical regime, or the standard security model.

Although the quantum security model might seem too strong to be considered for classical schemes, there are well-justified reasons for considering it. First of all, we note that for a quantum scheme, the natural model to consider *is* the quantum security model, since any type of interaction with the primitives will be via quantum states and since classical primitives can also be generalised as quantum ones, this model is theoretically more general and hence interesting. Moreover, in the future, one can consider a world where classical computers have been replaced with quantum ones and hence even the classical routines and cryptographic algorithms are being run on a quantum computer. This scenario argument has been also given by Boneh and Zhandry in [BZ13b]. Nonetheless, from a more practical point of view, one argument against this model for classical primitives is that a possible countermeasure against *superposition attacks* is to forbid any kind of quantum access to the oracle through measurements. However, in such a setting the security relies on the physical implementation of the measurement tool which itself could be potentially exploited by a quantum adversary. Thus, and as it has previously been advocated in [BZ13b, BZ13a, KLLNP16, AMRS20], providing security guarantees in the quantum security model is crucial even practically.

Once the oracle's access to the primitive is considered, the adversarial models can be further categorised based on the assumptions of the access level of the adversary. These classes of adversaries are usually the adaptation of the usual

[15] In general, in the field of algorithm complexity, time complexity and query complexity has been considered separately in many cases, while as in cryptography we usually consider the QPT adversary to be polynomial in both.

classical models in the quantum regime. As usual, we start with the strongest case. The strongest access level is when an adversary can directly access the oracle and query any arbitrary quantum state of their choice. Also, the queries can be issued *adaptively*, meaning that the adversary can choose the next query depending on the responses of the oracle to previous queries. In classical cryptography, this attack model is called *Chosen Message Attack (CMA)* model. The quantum analogue of this model has been introduced by Boneh and Zhandry [BZ13b], and called *Quantum Chosen Message Attack (qCMA)*[16]. This is one of the main quantum adversarial models that we use in this thesis. However, we will carefully define our version of qCMA within our given security game in Chapter 3.

Another important note that is worth mentioning here is that although in CMA/qCMA models the queries are *adaptive*, the term 'adaptive' is also used in the literature for another security level, where the adversary has been given an extra learning phase, usually after receiving the main message. This level of adaptiveness is meaningful for some definitions for instance for encryption schemes one can consider such a model for ciphertext, known as CCA2. This model has also been brought into the quantum world [CEV20]. We will also briefly mention this model for our case studies.

Next, the adversary can be weakened if we restrict the direct access to the oracle and instead the adversary has access to a random set of queries, chosen from a certain distribution[17]. This is often referred to as *Random Message Attack (RMA)* and can also be translated in the quantum setting when the set of queries are quantum input and output samples of the quantum oracle. Sometimes this type of adversaries is also called *non-adaptive* or *weak* adversaries. As we will see in Chapter 3, and also later in Chapter 6, this adversarial model has close connections with the models that are considered in learning theory.

Bounding the computational power of the quantum adversary is not the only option to go to a weaker quantum adversarial model. It is also possible to remain in the information-theoretic security regime while making instead, some reasonable assumptions about the storage capabilities of the adversary. Making an assumption about the adversary's capability in storing quantum data is a technologically sensible assumption due to the difficulty of building quantum memories, despite the latest efforts and progress [LST09, WLZ+19, BRA+19, GI20, LRGR+21, BBFO+19, WMH+20, DKLP02]. This has given rise to a quantum adversarial model known as *bounded quantum-storage model*. This model has been introduced by [DFSS05] and inspired from its classical counterparts [Mau93, CM97]. In this model, we assume that a quantum adversary can only store a limited number of qubits, yet it is computationally unbounded. It is common for the protocols

[16] When used for encryption scheme, there are several attack models associated with this security level. If the adversary can choose the plaintext arbitrarily, it is referred to as *(quantum) chosen-plaintext attack ((q)CPA)* and if the adversary is allowed to choose the ciphertext, *(quantum) chosen ciphertext attack ((q)CCA)*. Also, due to the complications that exist in the quantum setting and different notions of oracles, there are several definitions for this security level in the literature [GHS16, CEV20, GKS21, CETU21]

[17] Which is usually the uniform distribution over all the possible set of messages

in this model to assume no quantum memory for the honest parties, while the adversaries can only store a small fraction of the qubits sent in the protocol by assumption. Several protocols that are impossible to achieve in the unbounded model have proven to be secure in the quantum bounded storage model, such as oblivious transfer [DFSS05] or bit commitment [Unr11]. Moreover, another realistic assumption to be made about the quantum memories is that they are noisy. This assumption has been considered in [WST08], leading to a model called *noisy quantum-storage model*. In this thesis, we do not consider the memory-restricted adversarial models, but an enthusiastic reader can find further readings in [BS16, PR21].

As the final note, since entanglement is also a precious quantum resource, one can also consider models where quantum adversaries are restricted in using this resource. This model has been studied in [BCF+14] in the context of position-based cryptography, where it has been assumed that the adversaries cannot share entanglement.

2.5.3 Quantum accessible oracles for classical functions

In the previous section, we have discussed different adversarial models and the role of oracles in them. In this section, we define quantum oracles for classical primitives. These oracles are also called *quantum accessible oracles*.

A quantum oracle is a unitary transformation $\mathcal{O}$ over a D-dimensional Hilbert space that can be queried with quantum states. The quantum oracle can grant quantum access to the evaluation transformation of a classical or quantum primitive *i. e.* a classical function. The quantum accessible oracle gives, in fact, a reversible quantum implementation of that function. The first way of doing so is what is referred to as *the standard oracle* [KKVB02, BZ13b, BZ13a, GHS16, GKS21, CEV20].

In the standard quantum-query model, the adversary $\mathcal{A}$ has black-box access to a reversible version of f, which is a classical-polynomial-time computable deterministic or randomised function of the evaluation $\mathcal{E}$, through an oracle $R\mathcal{O}_f^{\mathcal{E}}$ which is a unitary transformation. The evaluation oracle can be represented as:

$$R\mathcal{O}_f^{\mathcal{E}} : \sum_{m,y} \alpha_{m,y} |r\rangle_{\mathcal{O}} |m, y\rangle \rightarrow \sum_{m,y} \alpha_{m,y} |r\rangle_{\mathcal{O}} |m, y \oplus f(m; r)\rangle \tag{2.92}$$

Here m is the message, and y is the ancillary system required for unitarity. In general, the standard oracle can also capture randomised evaluations with a randomness r picked from $\mathcal{R} \subseteq \{0,1\}^l$ as the randomness space, although in this case, the oracle may not be a unitary transformation. The unitary representation of the standard oracle has been introduced in several works such as [GHS16, GKS21, CEV20] with slightly different approaches that lead to an equivalent adversary's state, which is a completely mixed density matrix with respect to the randomness subspace. Nevertheless, in this thesis, to emphasise that the adversary cannot gain access to the internal randomness register of the oracle directly and avoid some potential artificial entanglement attacks, we opt for the approach of [GKS21] and

consider the randomness as an internal state of the oracle which is re-initiated for each query with a new classical value r. This choice is also due to the fact that the oracle needs to output the randomness register as a separable state, otherwise, an unwanted entanglement will be created between the adversary's output state and the internal register of the oracle, as also mentioned in [GKS21]. Moreover, if the primitive requires that the randomness is returned to the adversary for each query (as a classical bit-string or a function of r), it can be recorded in the adversary's auxiliary state y that can be extended to also capture the randomness space. An example of such construction will be introduced later in Section 3.6. Finally, we specify that for deterministic primitives (denoted by $\mathcal{O}_f^{\mathcal{E}}$) the structure is similar, except that the randomness register is not used.

Now let us introduce another type of quantum accessible oracles. First, let's assume that the function $f : \{0,1\}^n \to \{0,1\}^n$ is a bijection. In this case, the following transformation is a unitary:

$$\mathcal{O}_f : \sum_m \alpha_m |m\rangle \to \sum_m \alpha_m |f(m)\rangle \tag{2.93}$$

This can be generalised for non-length-preserving functions as the following transformation:

$$\mathcal{O}_f : \sum_{m,y} \alpha_{m,y} |m,y\rangle \to \sum_{m,y} \alpha_{m,y} |\phi_{m,y}\rangle \tag{2.94}$$

where the length of the ancillary register $|y| = |f(m)| - |m|$ and $\phi_{m,0} = f(m)$ for every m. This type of oracles are called *minimal oracles*, and as one can see this is closer to a general unitary, and hence they can consider to be a more powerful oracle than the standard oracle. One main difference between minimal and standard oracle is that if the function f is an encryption scheme $Enc_k(.)$, then the adjoint of the minimal encryption oracle is the decryption oracle i.e. $\mathcal{O}_{Enc}^{\dagger} = \mathcal{O}_{Dec}$, while as this is not the case for the standard oracle [GHS16].

Another type of quantum oracles is the *Fourier oracle* or *Fourier phase oracle* which is defined as follows [KKVB02, Zha19]:

$$\text{Fourier}\mathcal{O}_f^{\mathcal{E}} : \frac{1}{\sqrt{2^{m2^n}}} \sum_{m,y} |m,y\rangle \to \frac{1}{\sqrt{2^{m2^n}}} \sum_{m,y} e^{2\pi i f(m).y/2^n} |m,y\rangle \tag{2.95}$$

It has been shown that the Fourier oracle and standard oracle are equivalent [KKVB02]. Also in [Zha19] sophisticated techniques have been developed to record the adversary's queries made to the oracle, which is a very challenging task in the quantum regime when the queries are quantum, due to properties such as unclonability. Being able to record quantum queries is needed in some of the proof techniques and reductions in cryptography, and it is specifically relevant in the quantum random oracle model (QROM) [BDF+11]. This issue has been addressed in [Zha19] by introducing a new type of oracle called *compressed oracle*, which can be both defined as a standard or Fourier oracle. We avoid introducing this technique here since we do not use oracle recording techniques in this thesis.

2.5.4 Classical pseudorandomness

Randomness is perhaps one of the most crucial elements in modern cryptography. However, it is folklore knowledge that achieving true randomness in the classical world is practically impossible. That is why the concept of *pseudorandomness* has been introduced in cryptography as an efficient and practical approximation of truly random objects. Generally, a pseudorandom object, should not be distinguishable from its truly random counterpart. To formally define this concept, first, we need to ask the following question: 'Indistinguishable to what?' Since this is a cryptographic concept, let's consider an adversarial scenario. It is rather obvious that an unbounded adversary who can cover all the possible objects of the set can always make this distinction, then pseudorandomness is an *computational* security assumption. Thus the answer to that question is to an efficient or computationally bounded adversary (or, more generally, distinguisher). The pseudorandomness can be generally defined as follows [KL20]:

Definition 15 (Pseudorandomness)**.** Let $\mathcal{D}$ be a distribution over n-bit strings.$\mathcal{D}$ is (t,ε)-pseudorandom if for all adversaries $\mathcal{A}$ running in time at most t, $\mathcal{A}$ cannot distinguish $\mathcal{D}$ with a uniformly random distribution U_n over the n-bit. In other words, the following holds:

$$|\Pr_{x\leftarrow\mathcal{D}}[\mathcal{A}(x)=1]-\Pr_{x\leftarrow U_n}[\mathcal{A}(x)=1]|\leqslant\varepsilon \tag{2.96}$$

Nevertheless, in the asymptotic case, the ε needs to be a negligible function in the security parameter and the pseudorandomness has been usually defined over a finite family of objects with a specified distribution. Now, we can define the first important pseudorandom object in modern cryptography, the Pseudorandom Generators (PRG):

Definition 16 (PRG [KL20])**.** Let ℓ be a polynomial and let G be a deterministic polynomial-time algorithm such that for any n and any input $x\in\{0,1\}^n$, the result $G(x)$ is a string of length $\ell(n)$. We say that G is a Pseudorandom Generator (PRG) if the following conditions hold:

- **(Expansion:)** For every n it holds that $\ell(n)>n$
- **(Pseudorandomness:)** For any PPT algorithm $\mathcal{A}$, there is a negligible function $negl$ such that:

$$|Pr[\mathcal{A}(G(x))=1]-Pr[\mathcal{A}(r)=1]|\leqslant negl(n) \tag{2.97}$$

where both probabilities are taken over the randomness of $\mathcal{A}$, the first one over uniform choice of $x\in\{0,1\}^n$ and, and the second one, over uniform choice of $r\in\{0,1\}^{\ell(n)}$.

Thus a PRG is an efficient and deterministic algorithm that expands a *short* uniformly random seed into a longer pseudorandom string.

Next, we define *Pseudorandom Functions (PRF)* which are a family (usually keyed-family) of functions that are computationally indistinguishable from the set of uniformly random functions from the same domain and range. PRFs are formally defined as follows:

Definition 17. [Pseudorandom Functions (PRF)] Let $\mathcal{K}, \mathcal{X}, \mathcal{Y}$ be the keyspace, the domain and range, all implicitly depending on the security parameter λ. A keyed family of functions $\{PRF_k : \mathcal{X} \to \mathcal{Y}\}_{k \in \mathcal{K}}$ is a pseudorandom function (PRF) if for any polynomial-time (PPT) algorithm $\mathcal{A}$, PRF_k with a random $k \leftarrow \mathcal{K}$ is indistinguishable from a truly random function $f \leftarrow \mathcal{Y}^{\mathcal{X}}$ in the sense that:

$$|\Pr_{k \leftarrow \mathcal{K}}[\mathcal{A}^{PRF_k}(1^\lambda) = 1] - \Pr_{f \leftarrow \mathcal{Y}^{\mathcal{X}}}[\mathcal{A}^{f}(1^\lambda) = 1]| = negl(\lambda). \tag{2.98}$$

PRFs can also be equivalently defined in a game-based fashion as an indistinguishability game [KL20]. At the beginning of the game, an honest challenger flips a coin and selects to be in a random or pseudorandom world. Then according to the selected world, the challenger picks either a function f from the truly random family of functions; or picks a random key, and consequently, a pseudorandom function F_k. Then every time the adversary issues a query, the challenger responds with f or F_k, depending on the random bit b. The adversary's objective is then to guess b, or in other words, distinguish between truly random and pseudorandom worlds.

PRFs are extremely practical tools for cryptography, and many classically secure constructions are based on them. It is also known that under the assumption of OWF, a PRF family can be constructed.

One can also translate this concept to the quantum setting, where the adversary/distinguisher is quantum. This notion is referred to as *quantum-secure Pseudorandom Functions (qPRF)* and is defined as follows:

Definition 18. [quantum-secure Pseudorandom Functions (qPRF) [Zha12]] Let $\mathcal{K}, \mathcal{X}, \mathcal{Y}$ be the keyspace, the domain and range, all implicitly depending on the security parameter λ. A keyed family of functions $\{PRF_k : \mathcal{X} \to \mathcal{Y}\}_{k \in \mathcal{K}}$ is a quantum-secure pseudorandom function (qPRF) if for any polynomial-time quantum oracle algorithm $\mathcal{A}$, PRF_k with a random $k \leftarrow \mathcal{K}$ is indistinguishable from a truly random function $f \leftarrow \mathcal{Y}^{\mathcal{X}}$ in the sense that:

$$|\Pr_{k \leftarrow \mathcal{K}}[\mathcal{A}^{PRF_k}(1^\lambda) = 1] - \Pr_{f \leftarrow \mathcal{Y}^{\mathcal{X}}}[\mathcal{A}^{f}(1^\lambda) = 1]| = negl(\lambda). \tag{2.99}$$

Similarly, it has been proven that qPRFs can exist under the assumption of quantum-secure OWFs [Zha12].

2.5.5 Quantum pseudorandomness

The definitions of this sections are directly used in Chapter 3, Section 3.6.2.2 and throughout Chapter 5. Like what we have seen in the previous section, *quantum* pseudorandom objects can also be defined. The notion of quantum pseudorandomness has been defined for the first time in [JLS18], by introducing *Pseudorandom Quantum States (PRS)* and *Pseudorandom Unitaries (PRU)* as a computational version of true quantum randomness. In Section 2.4 we have introduced the Haar measure as a measure for perfect and uniform randomness over the quantum states and unitary transformations. Informally, pseudorandom states/unitaries are a set of states/unitaries that are computationally indistinguishable from Haar-random states/unitaries to a quantum polynomial-time adversary.

More formally, PRS is defined as follows:

Definition 19. [Pseudorandom Quantum States (PRS) [JLS18]] Let $\mathcal{H}$ be a Hilbert space and $\mathcal{K}$ the key space. $\mathcal{H}$ and $\mathcal{K}$ depend on the security parameter λ. A keyed family of quantum states $\{|\phi_k\rangle \in S(\mathcal{H})\}_{k\in\mathcal{K}}$ is *pseudorandom*, if the following two conditions hold:

- **Efficient generation**. There is an efficient quantum algorithm G which generates the state $|\phi_k\rangle$ on input k. That is, for all $k \in \mathcal{K}, G(k) = |\phi_k\rangle$.
- **Pseudorandomness**. Any polynomially many copies of $|\phi_k\rangle$ with the same random $k \in \mathcal{K}$ is computationally indistinguishable from the same number of copies of a Haar random state. More precisely, for any efficient quantum algorithm $\mathcal{A}$ and any $m \in poly(\lambda)$,

$$|\Pr_{k\leftarrow\mathcal{K}}[\mathcal{A}(|\phi_k\rangle^{\otimes m}) = 1] - \Pr_{|\psi\rangle\leftarrow\mu}[\mathcal{A}(|\psi\rangle^{\otimes m}) = 1]| = negl(\lambda). \tag{2.100}$$

where μ is the Haar measure on $S(\mathcal{H})$.

And PRU that are the quantum equivalent of PRFs are also defined as follows:

Definition 20. [Pseudorandom Unitary Operators (PRU) [JLS18]] A family of unitary operators $\{U_k \in \mathcal{U}(\mathcal{H})\}_{k\in\mathcal{K}}$ is a pseudorandom unitary if two conditions hold:

- **Efficient computation**. There is an efficient quantum algorithm Q such that for all k and any state $|\psi\rangle \in S(\mathcal{H}), Q(k,|\psi\rangle) = U_k|\psi\rangle$.
- **Pseudorandomness**. U_k with a random key k is computationally indistinguishable from a Haar random unitary operator. More precisely, for any efficient quantum algorithm $\mathcal{A}$ that makes at most polynomially many queries to the oracle:

$$|\Pr_{k\leftarrow\mathcal{K}}[\mathcal{A}^{U_k}(1^\lambda)=1] - \Pr_{U\leftarrow\mu}[\mathcal{A}^{U}(1^\lambda)=1]| = negl(\lambda). \tag{2.101}$$

where μ is the Haar measure on $S(\mathcal{H})$. Note that here we focus on the Pseudorandomness condition of the PRU definition.

Another approximation of Haar-randomness in quantum information is the notion of *t-designs*. Although these objects are also often called 'pseudorandom' in the mathematical physics literature, they are analogous to t-wise independent random variables in theoretical computer science [JLS18]. Quantum state and unitary t-designs are informally approximating the Haar measure up to t-th order polynomials or tensor products. Thus one way of defining the quantum states t-design is as follows:

$$\sum_i p_i(|\phi_i\rangle\langle\phi_i|)^{\otimes t} = \int_{\mu_\psi}(|\psi\rangle\langle\psi|)^{\otimes t} d\mu_\psi \tag{2.102}$$

where p_i is the probability of each state $|\phi_i\rangle$ and the integration in the right hand side is over Haar measure [DCEL09, EA5]. Similarly a unitary t-design can be defined as follows:

$$\sum_i p_i U_i^{\otimes t}\rho(U_i^{\otimes t})^\dagger = \int_{\text{Haar}} U^{\otimes t}\rho(U^{\otimes t})^\dagger dU \tag{2.103}$$

On the right-hand side of the equation is the expectation for the t-fold tensor product of Haar measure is also denoted by $\mathbb{E}_H^t(\rho)$. Also, a t-design can be defined approximately. The ε-approximate t-design has been introduced by Brandão, Harrow and Horodecki [BHH16] as follows:

Definition 21 (ε-approximate t-design)**.** We say a family of unitary with distribution $\mathcal{D}$ given as the set $\{p_i, U_i\}$ forms an ε-approximate t-design if the following holds:

$$(1-\varepsilon)\mathbb{E}_H^t(\rho) \leqslant \sum_i p_i U_i^{\otimes t}\rho(U_i^{\otimes t})^\dagger \leqslant (1+\varepsilon)\mathbb{E}_H^t(\rho) \qquad \forall\rho\in\mathcal{S}(\mathcal{H}^{\otimes t}) \tag{2.104}$$

We conclude this section by mentioning some of the numerous applications of t-designs in different areas of quantum computing and quantum information, including quantum supremacy [BFNV19], verification and benchmarking [HFGW18, NZO+21, EHW+20], the physics of blackholes [HP07], cryptography [AM17, JLS18], and machine learning [MBS+18].

2.5.6 Unforgeability

Unforgeability is the desired security property for many primitives such as Message Authentication Codes (MACs) and digital signatures. Informally, unforgeability ensures that an adversary cannot produce valid input-output pairs of the evaluation function of the primitive with only limited access to its oracle, or in other words, from a previously learnt set of input and outputs of the function. The unforgeability of a classical primitive can be studied against classical or quantum adversaries in the different adversarial models that we have introduced in Section 2.5.2. In this section, we first introduce different levels of classical unforgeability, and then we also give some of the proposals for translating this notion to the quantum setting. Later in Chapter 3, we generalise the quantum unforgeability inside a formal and unified framework. Thus this section is mostly relevant for Chapter 3, Section 3.5. Unforgeability is also a central security property for quantum schemes such as quantum money, as we will further discuss in that chapter.

2.5.6.1 Classical Unforgeability

Goldwasser et al. [GMR88] define different notions of unforgeability for digital signatures. They consider various types of attacks including CMA where the adversary is allowed access to the signing oracle on a list of messages of their choice. They define *existential forgery* as the attack where the adversary can forge a valid signature for at least one new message; and the notion of *selective forgery* as an attack where the adversary can forge a valid signature with non-negligible probability for a particular message chosen by the adversary prior to accessing the signing oracle.

An et al. [ADR02] define a slightly stronger notion of unforgeability called *strong unforgeability* that requires the adversary not only to be unable to generate a valid signature on a 'new' message but also to be unable to generate even a valid 'new' signature on an already signed message. *Strong Existential Unforgeability (SEUf)*, also called *strong unforgeability*, has formally been defined in [BSW06] by Boneh et al.

Bellare *et al.* [BGR95] define the notion of Strong Existential Unforgeability under chosen message and chosen verification queries attack (SEUF-CMVA) for message authentication codes (MACs). In both of these attack models, the adversary is allowed a chosen message oracle access, as defined for digital signatures in [GMR88]. Although in the later attack model for message authentication codes, the experiment also allows verifying queries through oracle access. This model is justified for MACs as unlike digital signatures, where the verification algorithm is

public, the adversary cannot run the verification algorithm on their own. *(Weak) Existential Unforgeability (EUf) under chosen message attacks* is a natural definition for MACs defined by Bellare et al. [BKR00] and comes by extending the one for digital signatures [GMR88].

Moreover, Dodis et al. [DKPW12] define the notion of *selective unforgeability under adaptive chosen message and chosen verification queries (SelUF-CMVA)*.

A yet weaker notion called *universal unforgeability* requires the adversary to produce a fresh tag for a uniformly random message given as a challenge to the adversary [AHM+14]. This notion, again, can be considered against both attack models: chosen message and chosen verification query attack (UniUF-CMVA) and chosen message attack (UniUF-CMA). Table 2.1 summarizes all these different classical notions of unforgeability.

Attack Model / Def. level	CMVA	CMA
SEUf (strong)	-	[ADR02, BSW06, BGR95]
EUf (weak)	[DKPW12, BGM04]	[BSW06, BKR00]
SelUf (selective)	-	[DKPW12]
UniUf (universal)	[AHM+14]	[AHM+14]

Table 2.1: Classical unforgeability definitions from strongest to weakest. *CMVA* - adaptive chosen message queries and limited access to the verification oracle. *CMA* - (adaptive) chosen message attacks. In the cases marked with "-", no definition has been proposed yet to the best of our knowledge.

2.5.6.2 Unforgeability in the quantum world

In the quantum regime, the definition of unforgeability defined by Boneh and Zhandry [BZ13b, BZ13a] (denoted by BZ), is described as a quantum analogue of *strong existential unforgeability* and it is in the chosen message attack (CMA) model. The formal definition of BZ (EUF-qCMA) for digital signatures is as follows:

Definition 22. [BZ or (EUF-qCMA) [BZ13a]] A system S (Sign/Mac), is existentially unforgeable under a quantum chosen message attack (EUF-qCMA) if no adversary after issuing q quantum chosen message queries, can generate $q+1$ valid classical message-tag pairs with non-negligible probability in the security parameter.

Another definition of unforgeability against quantum adversaries called *blind unforgeability* was proposed in [AMRS20]. This more recent definition aims to capture some attacks that are not captured by BZ. This notion defines an algorithm to be forgeable if there exists an adversary who can use access to a 'partially blinded' oracle to validate responses of the messages that are in the blinded region and hence only respond to the queries that are not in this region. A blinded

operation for a function $f : X \to Y$ and a subset of messages $B \subseteq X$ is defined as:

$$Bf(x) = \begin{cases} \perp, & \text{if } x \in B \\ f(x), & \text{otherwise} \end{cases} \tag{2.105}$$

Where in particular for the definition of unforgeability, the elements of X are placed in B independently at random with a particular probability ε, denoted by B_ε. Then the security game of unforgeability has been defined as follows with the adversary having access to the blinded oracle.

Definition 23. [[AMRS20](Def.4&5)] Let $\Pi = (KeyGen, Mac, Ver)$ be a MAC with message set X. Let $\mathcal{A}$ be an algorithm, and $\varepsilon : \mathbb{N} \to \mathbb{R}_{\geqslant 0}$ an efficiently computable function. The blind forgery experiment $BlindForge_{\mathcal{A},\Pi}(n, \varepsilon)$ proceeds as follows:

1. Generate key: $k \leftarrow KeyGen(1^n)$
2. Generate blinding: select $B_\varepsilon \subseteq X$ by placing each m into B_ε independently with probability $\varepsilon(n)$.
3. Produce forgery: $(m, t) \leftarrow \mathcal{A}^{B_\varepsilon MAC_k}(1^n)$.
4. Outcome: output 1 if $Ver_k(m, t) = acc$ and $m \in B_\varepsilon$; otherwise output 0.

From this game blind-unforgeability is defined as follows.
A MAC scheme Π is blind-unforgeable (BU) if for every polynomial-time uniform adversary $(\mathcal{A}, \varepsilon)$

$$Pr[BlindForge_{\mathcal{A},\Pi}(n, \varepsilon(n)) = 1] \leqslant negl(n).$$

and the probability is taken over the choice of key, the choice of blinding set, and any internal randomness of the adversary.

Thus, in this definition, a forgery happens if the adversary can produce a valid tag for a message within the blinded region. We refer to this definition of unforgeability as BU. This definition imposes the challenge to be orthogonal to the previously queried messages.

We also recall the following useful theorem from [AMRS20]:

Theorem 10. *[from [AMRS20]] Let $\mathcal{A}$ be a QPT such that $supp(\mathcal{A}) \cap R = \emptyset$[a] for some $R \neq \emptyset$. Let $\mathtt{MAC}$ be a MAC, and suppose $A^{\mathtt{MAC}_k}(1^n)$ outputs a valid pair $(m, \mathtt{Mac}_k(m))$ with $m \in R$ with non-negligible probability. Then $\mathtt{MAC}$ is not BU-secure.*

[a] Here $supp(\mathcal{A})$ denotes the support of $\mathcal{A}$ that is defined as follows. Let $\mathcal{A}$ have oracle access to a classical function $f : \{0,1\}^n \to \{0,1\}^m$. Let $|\psi_i\rangle$ be the state of the the query i or equivalently the intermediate state after applying U_i in the sequence of $\mathcal{O}U_q\mathcal{O}\ldots U_1$ on an initial state $|0\rangle_{XYZ}$ where X denotes the input registers. Then $supp(\mathcal{A})$ is defined to be the set of input strings x such that there exists a function f with the respective oracle such that $\langle x|\psi_i\rangle_X \neq 0$ for at least one of the queries.

In addition to these two main definitions, another definition for quantum unforgeability has been given in [GYZ17] for *one-time* unforgeable schemes, which we will skip representing it due to the lack of generality and since it is less relevant for this thesis. Another related and interesting work is the study of *non-malleability* and its relation to authentication in the quantum regime which has been studied in [AM17].

2.5.7 Coin-flipping

In this section, we introduce *coin-flipping* which is a cryptographic task that allows two mutually distrustful parties to agree on a common random bit. We particularly need the familiarity with this cryptographic functionality for Chapter 7 (Section 7.2.2 and 7.4.2). This task has been first introduced by Blum [Blu83] and has been motivated in the following scenario: Alice and Bob need to agree on the output of a coin-flip over the phone for an important decision. However, they don't trust each other. The formal definition of a coin-flipping task is given as follows:

Definition 24 ((Strong) coin-flipping)**.** The task of coin flipping consists of two mutually distrustful players, Alice and Bob, and the goal is for both players to output the same random bit $c \in \{0,1\}$ such that the following properties hold:

- **Correctness:** if both Alice and Bob are honest then b is uniformly distributed: $p(c=0) = p(c=1) = 1/2$.
- **ε-secure:** neither player can force $p(c=0) \geqslant 1/2+\varepsilon$ or $p(c=1) \geqslant 1/2+\varepsilon$, where $p(c)$ is the probability that the honest player outputs a value c.

The smallest ε for which a protocol is ε-secure is called the **bias**.

It has also been shown in [Blu83] that unconditionally secure coin-flipping is impossible in the classical world, meaning that no classical coin-flipping protocol is secure, or no value of $\varepsilon < 1/2$ can be achieved for security. Nevertheless,

coin-flipping with computational assumptions is possible since there exists a coin-flipping protocol, assuming perfectly secure OWF exists. Also, Cleve [Cle86, CI93] extended the computational coin-flipping into r-rounds and showed an upper bound of $\Omega(1/r)$ for any two-party r-round coin-flipping protocol as the bias.

Historically, the first quantum coin-flipping protocol was introduced by [MSCK99] which has conjectured to achieve arbitrary small bias, although a full security proof has not been given in the paper. In Chapter 7 we will introduce this protocol and as one of our contributions, we show how it can be broken with cloning-based cryptanalysis.

Later a quantum coin-flipping protocol has been introduced by Aharonov et al. [ATSVY00] which provably achieves the bias 0.42 in an information-theoretic way. We will also closely study this protocol in Chapter 7, so we avoid repetition in here.

Another quantum coin-flipping protocol with qutrits has been introduced by Ambainis in [Amb04] which achieves a better bias than Aharonov's protocols with 0.25 bias. Given this improvement, an interesting question was whether one we achieve arbitrary small bias for coin-flipping using quantum information. This question has been answered negatively by Kitaev when given the following bound for the bias of any strong coin-flipping protocols:

$$\varepsilon_{min}^{qcf} = \frac{\sqrt{2}-1}{2} \approx 0.207. \tag{2.106}$$

Hence perfectly secure quantum coin-flipping is also impossible.

The requirement on the coin-flipping can we weakened if the choice of Alice and Bob is predetermined, say Alice always wants to bias the coin towards 0, and Bob is vice versa. This leads to the notion of *weak coin-flipping*. Weak coin-flipping has been studied in the literature for many years and it has been shown in 2007 [Moc07] that that weak quantum coin flipping, with arbitrarily small (but non-zero) bias, is possible. Although the protocol that achieved that arbitrary small bias is complicated and scales exponentially in $1/\varepsilon$ in the number of rounds. The long-standing problem of the weak coin-flipping has been finally solved by Arora et al. in 2019 [ARW19].

2.6 Quantum and classical learning

In this section, we enter the last field of research where we have adopted and exploited many of our toolkits and the concepts we have used in this thesis and which is rather different from the other fields of research we have talked about so far. Here, we will talk about 'learning' in a broad context which includes learning theory, machine learning and some areas of quantum information processing such as tomography. However, as each of these subjects has a very rich literature on its own, and entering each of them with enough precision and care requires a separate thesis, our introduction here will be very brief and we will mostly focus on the tools that we have directly used in the thesis.

Learning is the act of acquiring knowledge, but generally speaking, in physics and computer science, there are two types of learning. Either we want to learn a 'system' or an unknown property or feature of the system by interacting with it, or we want the 'machines' and computers to be able to do something of a similar nature. However, later we need to teach the machines to 'learn' first, which means adopting a methodical approach to leverage 'data' to improve the performance of a learning task. The first case is a problem usually studied in physics (and generally through experiments and simulations), and the second is a sub-field of computer science known as *learning theory*[18] and *machine learning*. Although seemingly very different, the latest progress in the field of machine learning, as well as different approaches to simulating physical systems, has brought these fields closer together. An example of this is the applications of machine learning in particle physics [RWR+18, AFFS19]. On the other hand, physics has also inspired machine learning models, for instance in development of *Boltzmann machine* [SK75] or *Born machine* [CCW18, CMDK20].

On the other hand, the potential advantage of quantum computing in bringing speedup to some problems has motivated researchers to design quantum machine learning techniques and algorithms. One of the earliest examples (perhaps the earliest one) was an algorithm developed for solving linear equations and similar problems in matrix algebra by Harrow, Hasidim, and Lloyd [HHL09], where exponential speedup for some operations has been shown. However, the field of *Quantum Machine Learning (QML)* has expanded fast over the very few years that have been past since its birth. For a great review of the field we refer the reader to [SSP15, MGL22].

We start on the physics side, explaining the existing notions of learning in quantum information processing, and we finish by introducing some of the tools that we require from classical and quantum learning theory and quantum machine learning.

2.6.1 Quantum state and process tomography

Quantum state tomography and quantum process tomography are the process of determining the state of a quantum state or describing quantum dynamics and are of great importance to quantum computation and quantum information, both theoretically and experimentally. However, they are both challenging and resource-intensive tasks [BCD+09].

In Section 2.2, we have explained the reason behind the difficulty of determining and distinguishing quantum states. Quantum process tomography is even more challenging since we aim to fully characterise the dynamics of a quantum system. For example, we want to characterise a quantum gate (a unitary transformation), or a quantum channel. Both of these tasks have applications in the verification and benchmarking of quantum systems and quantum computers.

Let us start with state tomography. The easiest way (and clearly the most

[18] Also called computational learning theory

inefficient) way to perform process tomography, is to simply use the Born's rule. This method also sometimes called *linear inversion* involves preparing (or acquiring in any way) many copies of an unknown state ρ, and repeatedly performing projective or POVM measurements, in order to extract the expectation values of the probability, of obtaining a histogram of the observations of the measurements, and finally, the amplitudes to fully describe the state. However, this method is highly infeasible and requires asymptotically many copies and measurements to achieve good precision. This duo to the fact that to reconstruct a d-dimensional density matrix, one needs to determine d^2-1 independent parameters and requires many measurements for each. (where d is already exponential in the input size *i.e.* the number of qubits). A better approach is to restrict the domain of the density matrices to a more 'likely' space. This method is called *Maximum Likelihood Estimation (MLE)* and involves searching for the density matrix that maximises the likelihood of giving the experimental results. The 'likelihood' is a probability function assigned to the observable that would most likely detect the state. Using this method a complexity of $O(d^4)$ has been achieved in the general cases [PR04, QHL+13]. Yet another method is to use Bayesian estimations such as *Bayesian Mean Estimation (BME)* for this task [BK10]. This method requires reasonable prior information about the systems however it can achieve a minimum estimation probability of $\frac{1}{N+d}$, with N being the number of observables.

Exploiting machine learning techniques in recent years has enabled tremendous improvements in this field. For instance a general complexity of $O(d^3)$ has also been achieved using neural networks [XX18], and the results have been improved in the following works [APJAD18, TMC+18, RKKN19].

Finally, the most recent breakthrough in this field has been made by Huang et. al in [HKP20], where they have shown that even though the best-known techniques for full state tomography are still exponential (and believed to remain so), one can still extract many useful properties of a quantum state, efficiently and without requiring exponential copies of that state. This discovery has a great deal of significance in very different areas of quantum information and quantum computing, as we will discuss further in the future chapters.

Going back to quantum process tomography, we first note that this process is closely related to quantum state tomography. Let us give a simple example. Assume you are given an unknown unitary gate (a single-qubit gate, for instance) and want to extract the full unitary matrix. To do so, you need to know the action of the unitary on a full set of basis (say, the computational basis). Thus you will prepare states in the computational basis and apply the unitary to them. However, the output state of the gate, $U|\psi\rangle$, is unknown and can be any state on the Bloch sphere in our specific example. Thus one needs to measure the state and repeat the process many times to get a good approximation of the action of U, only on the computational basis. Then for the state $|1\rangle$, the same process needs to be repeated. Nevertheless, similar to state tomography, this is the most naive and inefficient way. A variety of different disciplines exist here such as *Standard quantum-process tomography (SQPT)*, *ancilla-assisted process tomography (AAPT)*, and *direct characterization of quantum dynamics (DCQD)*.

We refer the reader to [MRL08] for a review of these different strategies and their required resources.

A point worth mentioning here is that for process tomography (maybe unlike state tomography), the full characterisation of the dynamics might not be needed for many problems and applications. An example of this is *randomised benchmarking* which is a method for testing the quality and capability of quantum hardware by estimating the average error rates of the gates [EA5]. The other example is the idea of *quantum emulation*. Since quantum emulation will have particular importance in our work, we will introduce it separately in the next section.

2.6.2 Quantum Emulation

We now describe the concept of *Quantum Emulation (QE)* and an algorithm called *universal quantum emulator* developed by Marvian and Lloyd in [ML16].[19] The *quantum emulation algorithm* is a quantum process learning tool that can outperform the existing approaches based on quantum tomography [DLP01]. Generally speaking, the goal of the quantum emulator is to mimic the action of an unknown unitary transformation on an unknown input quantum state[20] while having access to some 'data' in the form of input-output samples of the unitary[21]. We return to this algorithm in Chapter 3 Section 3.4 to further analyse the algorithm and repurpose it.

Before diving into the details of this algorithm, we make a small remark on the difference between 'emulation' and 'simulation'.

2.6.2.1 Emulation vs simulation for a quantum process

Emulating a quantum process, in the sense that is introduced here, is different from process tomography since the goal is not to extract the full description of a quantum process but instead to 'learn' enough from it to be able to mimic its behaviour. However, when one defines the notion of emulation in this manner, it suggests a similarity with another notion, that is also particularly important for physical systems *i.e.* 'simulation'. Especially because the 'unknown unitary' that we are trying to emulate here, can correspond to the Hamiltonian of a physical system. Despite similarities, here we want to emphasise that the notion of simulation and emulation for a quantum process have crucial differences.

In simulating a quantum dynamics, for instance, a unitary, we usually start from an initial state of a quantum system (for example, a many-body system of particles), and we use abstractly speaking a *'simulation machine'* that produces a

[19] Although this may not be the only possible algorithm for the purpose of quantum emulation, is the only one we are aware of by the time of writing this thesis, and hence the content of this section are mainly based on the mentioned work.

[20] The algorithm can be applied to any unknown state, however, the high fidelity performance is achieved when the target state is in the span of the data hence, not fully unknown.

[21] The term 'emulation', however, has been used in other meanings both in physics and cryptography. While throughout this thesis whenever we use the term, we refer to emulation in the context used in this section.

'simulation' or a mathematical approximation that mimics the required property of the system. Sometimes the method consists of letting the system evolve with a stochastic process [BCC+15].[22] The key point here, is that the simulation of the system needs to completely (however with some approximation) obey the same dynamics and transformations, otherwise, it will not be useful to study the original system.

An emulator, on the other hand, is not trying to completely recreate the transformation. Instead, it outputs what that transformation is 'supposed to do' on a new quantum state. Although it mimics the system's behaviour in an input-output manner, it does not need to obey the same dynamics. In that sense, an emulation algorithm is closer to a machine learning process, as we will also discuss later in Chapter 3. This distinction between simulation and emulation has been illustrated in Fig. 2.7. With these in mind, we now introduce the quantum emulation algorithm.

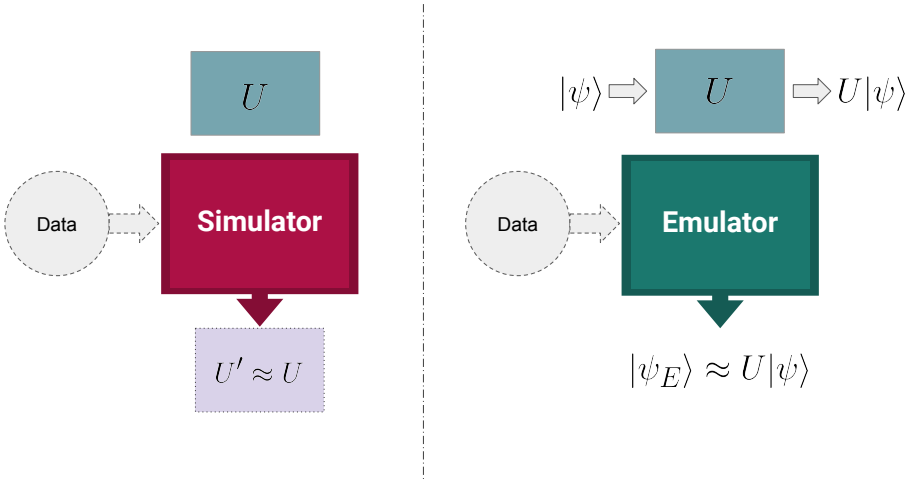

Figure 2.7: Illustration of the contrast between the notions of emulation and simulation. Particularly regarding a quantum process, denoted by a unitary U. An emulator, as opposed to a simulator, does not necessarily recreate the same dynamics but instead mimics the action of the unitary on a new quantum state.

2.6.2.2 QE: the circuit and description of the algorithm

The circuit of the quantum emulation algorithm is depicted in Fig. 2.8 (recreated from [ML16]) and works as follows: Let U be a unitary transformation on a D-dimensional Hilbert space $\mathcal{H}^D$, $S_{in} = \{|\phi_i\rangle ; i = 1, ..., K\}$ be a sample of input states and $S_{out} = \{|\phi_i^{out}\rangle ; i = 1, ..., K\}$ the set of corresponding outputs, i.e $|\phi_i^{out}\rangle = \mathrm{U}|\phi_i\rangle$. Also, let d be the dimension of the Hilbert space $\mathcal{H}^d$ spanned by S_{in} and $|\psi\rangle$, a challenge state. The goal of the algorithm is to find the output of U on $|\psi\rangle$, that is $\mathrm{U}|\psi\rangle$.

[22] This is usually referred to as quantum random walk

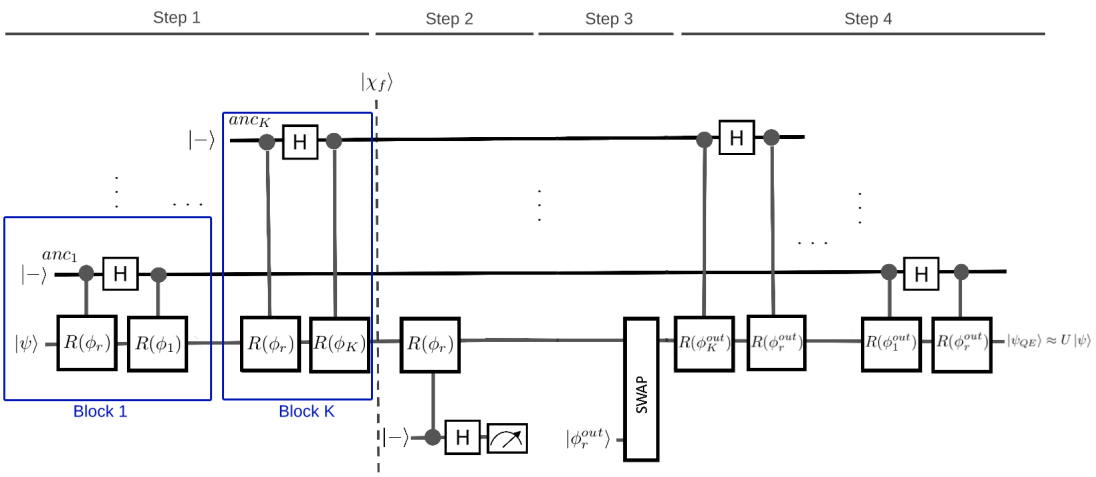

Figure 2.8: The circuit of the quantum emulation algorithm. $|\phi_r\rangle$ is the reference state and $|\phi_r^{out}\rangle$ is the output of the reference state. $R(*)$ gates are controlled-reflection gates. In each block of Step 1, a reflection around the reference and another sample state is being performed.

The main building blocks of the algorithm are controlled-reflection gates described as:

$$R_c(\phi) = |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes e^{i\pi|\phi\rangle\langle\phi|} \tag{2.107}$$

A controlled-reflection gate acts as the identity ($\mathbb{I}$) if the control qubit is $|0\rangle$, and as $R(\phi) = e^{i\pi|\phi\rangle\langle\phi|} = \mathbb{I} - 2|\phi\rangle\langle\phi|$ if the control qubit is $|1\rangle$. The circuit also uses Hadamard and SWAP gates and consists of four stages.

Stage 1. K number of sample states and a specific number of ancillary qubits are chosen and used through the algorithm. We assume the algorithm uses all of the states in S_{in}. The ancillary systems are all qubits prepared at $|-\rangle$. Let $|\phi_r\rangle \in S_{in}$ be considered as the reference state. This state can be chosen at random or according to some distribution. The first step consists of $K-1$ blocks wherein each block the following gates run on the state of the system and an ancilla:

$$W(i) = R_c(\phi_i)HR_c(\phi_r). \tag{2.108}$$

In each block represented by Eq. (2.108), a controlled-reflection around the reference state $|\phi_r\rangle$ is performed on $|\psi\rangle$ with the control qubit being on the $|-\rangle$ ancillary state. Then a Hadamard gate (H) runs on the ancilla followed by another controlled-reflection around the sample state $|\phi_i\rangle$. This repeats for each of the K states in S_{in}, such that the input state is being entangled with the ancillas, and also it is being projected into the subspace $\mathcal{H}^d$. By doing this, the information of $|\psi\rangle$ is encoded in the coefficients of the general entangled state. This information is the overlap of $|\psi\rangle$ with all the sample inputs. By reflecting around the reference state in each block, the main state is pushed to $|\phi_r\rangle$ and the probability of finding the system at the reference state increases. The overall state of the circuit after Stage 1 is:

$$[W(K)...W(1)]\,|\psi\rangle\,|-\rangle^{\otimes K} \approx |\phi_r\rangle\,|\Omega(anc)\rangle \tag{2.109}$$

where $|\Omega(anc)\rangle$ is the entangled state of K ancillary qubits. The approximation comes from the fact that the state is not only projected on the reference quantum

state but is also projected on other sample quantum states with some probability. We present a more precise formula in the next subsection.

Stage 2. In this stage, first a reflection around $|\phi_r\rangle$ is performed and after applying a Hadamard gate on an extra ancilla, that ancilla is measured in the computational basis $\{|0\rangle, |1\rangle\}$. Based on the output of the measurement, one can decide whether the first step was successful (i.e. the output of the measurement is 0) or not. If the first step is successful, the main state has been pushed to the reference state. In this case, the algorithm proceeds with Stage 3. If the output is 1, it implies that the projection was unsuccessful and that the input state remains almost unchanged. In this case, either the algorithm aborts or it goes back to the first stage and picks a new state as the reference. This stage has a post-selection role which can be skipped, to output a mixed state of two possible outputs.

Stage 3. The main state is swapped with $|\phi_r^{out}\rangle = \mathrm{U}|\phi_r\rangle$ that is the output of the reference state. This is done by employing a SWAP gate. At this point, the overall state of the system is:

$$(\mathrm{SWAP} \otimes I^{\otimes K})\,|\phi_r^{out}\rangle\,|\phi_r\rangle\,|\Omega(anc)\rangle = |\phi_r\rangle\,|\phi_r^{out}\rangle\,|\Omega(anc)\rangle\,. \tag{2.110}$$

By tracing out the first qubit, the state of the system becomes $|\phi_r^{out}\rangle\,|\Omega(anc)\rangle$.

Stage 4. The last stage is very similar to the first one except that all blocks are run in reverse order, and the reflection gates are made from corresponding output quantum states. The action of stage 4 is equivalent to:

$$W^{out}(i) = R_c(\phi_i^{out})HR_c(\phi_r^{out}) = (\mathrm{U} \otimes I)W(i)(\mathrm{U}^\dagger \otimes \mathbb{I}). \tag{2.111}$$

After repeating this gate for all the output samples, U is applied to the projected components of $|\psi\rangle$, and by restoring the information of $|\psi\rangle$ from the ancilla, the input state approaches $\mathrm{U}|\psi\rangle$. The overall output state of the circuit at the end of this stage is:

$$[W^{out}(1)...W^{out}(K)]\,|\phi_r^{out}\rangle\,|\Omega(anc)\rangle \approx \mathrm{U}\,|\psi\rangle\,|-\rangle^{\otimes K} \tag{2.112}$$

where equality is obtained whenever the success probability of Stage 2 is equal to 1.

The property of interest to measure the success or quality of the emulation algorithm is the fidelity of the output state $|\psi_{QE}\rangle$ (the output state of QE on the main register) and the intended output $\mathrm{U}|\psi\rangle$. In the original paper, the fidelity analysis is first provided for ideal controlled-reflection gates and later a protocol is presented to implement them efficiently using a technique called *quantum principal component analysis* introduced by Lloyd, Mohseni and Rebentrost [LMR14].[23] We now recall the following central theorem from [ML16]:

[23] For the purpose of this thesis as we are more interested in the theoretical bounds for the fidelity, all the gates including the controlled-reflection gates are assumed to be ideal keeping in mind that the implementation is possible.

Theorem 11. *[ML16] Let $\mathcal{E}_U$ be the quantum channel that describes the overall effect of the algorithm presented above. Then for any input state ρ, the Uhlmann fidelity of $\mathcal{E}_U(\rho)$ and the desired state $U\rho U^\dagger$ satisfies:*

$$F(\rho_{QE}, U\rho U^\dagger) \geqslant F(\mathcal{E}_U(\rho), U\rho U^\dagger) \geqslant \sqrt{P_{succ-stage1}} \tag{2.113}$$

where $\rho_{QE} = |\psi_{QE}\rangle\langle\psi_{QE}|$ is the main output state(tracing out the ancillas) when the post-selection in Stage 2 has been performed. $\mathcal{E}_U(\rho)$ is the output of the whole circuit without the post-selection measurement in Stage 2 and $P_{succ-stage1}$ is the success probability of Stage 1.

We point out that this algorithm with ideal controlled-reflection gates performs the emulation task with K total blocks and arbitrary precision ε given that,

$$T \geqslant \frac{d \times \log(8d\varepsilon^{-2})}{1 - |\lambda_{\mathcal{D}}|} \tag{2.114}$$

where $\lambda_{\mathcal{D}}$ is the eigenvalue of the overall channel with the second largest magnitude. Thus the algorithm (both sample and run time) complexity, in this case, is polynomial in d. Nevertheless, the spectral gap of the channel, namely $1 - |\lambda_{\mathcal{D}}|$, plays an important role as well, according to the Eq. (2.114), which is not very clear and easy to determine. This is why in Chapter 3 where we use this algorithm, we do not use these complexity results. Instead, we perform fidelity analysis based on the output of the algorithm.

Finally, in the imperfect setting and given the algorithm for approximating the controlled reflection gates, the overall algorithm can be implemented using the N_{tot} number of samples, and in time t_{tot} with precision $\varepsilon > 0$, as follows:

$$N_{tot} = \tilde{\mathcal{O}}\left(\frac{d^2 \times \varepsilon^{-1}}{(1 - |\lambda_{\mathcal{D}}|)^2}\right), \qquad t_{tot} = \tilde{\mathcal{O}}(N_{tot} \times \log D) \tag{2.115}$$

where $\tilde{\mathcal{O}}$ suppresses more slowly-growing terms.

2.6.3 Learning theory

In this section we discuss learning theory and we mainly use the definitions provided here in Chapter 3, Section 3.6.3.1. Learning theory is a theoretical subfield of machine learning or *computational learning* that provides the mathematical framework for studying and quantifying learning problems as well as the design and analysis of algorithms to solve them. The 'learner' is a classical/quantum algorithm (either deterministic or probabilistic), but the target of learning, can be very different objects, including functions, quantum states, quantum processes, distributions, etc. Also, depending on the learning task, one might not need to fully learn all the characteristics or a full description of the target, but perhaps some specific properties. Often the main question here is *how efficient* can the learning target be learned. This efficiency is measured commonly in time (time

complexity) or the size of the sample data (query complexity) [Ang92]. Another essential element of learning is data. We need data to learn from or train our machines to perform the desired task. The dataset, in learning theory, is also called *learning data*, or*training set*. Let of start by defining this sample data:

Definition 25 (Learning/sample/training dataset)**.** Let $\mathcal{T}$ be the learning object and let $\mathcal{X}$ be the domain or instance space of $\mathcal{T}$, and $\mathcal{Y}$ be the label set (for instance the range, if $\mathcal{T}$ is a function). The learning/sample/training data is a set $S = \{(x_i, y_i)\}_{i=1}^{K} \subseteq (\mathcal{X} \times \mathcal{Y})^K$, where $K = |S|$ is the size of the set. Furthermore, the datapoints x_i may have been chosen according a distribution $\mathcal{D}$.

Note that we have defined the above definition in a general way, such that it can be applied to different learning targets and different classical or quantum data. For instance, for a fully quantum data we have $\rho_i, \sigma_i \in \mathcal{H}$ and hence $\mathcal{X} = \mathcal{Y} = \mathcal{H}$.

Also, the learning problems are usually categorised in two types in the literature of learning theory, as follows [KV94]:

- **Unsupervised learning:** we require the learner to discover hidden structures in a set of unlabeled data points.
- **Supervised learning:** we want to learn a property or make a prediction based on a labelled dataset.

The next terminology that we need to introduce here is the notions of *concept class* and *hypothesis*. A 'concept' is a specific sample of the learning object. It is most commonly used for Boolean functions, so a concept is a specific $f : \mathcal{X} \to \mathcal{Y}$. A 'concept class' is the set of all the learning objects of interest *e.g.* for a function that would be a family of functions $\mathcal{F}$ such that $f \in \mathcal{F}$. A 'hypothesis' h, is the learner's output or guess for the function f. One ideally wants h to be as close as possible to f, for the learning process to be successful. We can also define a 'hypothesis class' H where all the $h \in H$, and by restricting the learner to choose h from H, we can bias the learner towards a particular set of solutions.

Before introducing the first formal learning definition for functions, we need to introduce one last concept. In Section 2.5.3 we have introduced oracles and their importance in cryptography. Here as well, we can assume the data has been obtained via interacting with an oracle. However, this is an specific oracle called *example oracle* or PEX [SG04, AdW17a]. An example oracle gets a concept f, a distribution $\mathcal{D}$, and when queried, outputs a sample data point as defined in Definition 25:

$$\mathrm{PEX}(f, \mathcal{D}) \to [(x, f(x)) : (x \leftarrow \mathcal{D})] \tag{2.116}$$

Now, we are ready to introduce the definition of *Probably Approximately Correct (PAC)* learning, introduced first by Valiant [Val84]. The name refers to the learner not being required to learn a function *exactly* ($h(x) = f(x)$), but rather approximately with a high probability. The definition has been given in different

ways with slight variations in the literature, however, for our purpose, we choose the one closer to [SG04, AdW17a].

Definition 26 (PAC-learnability)**.** A concept class $\mathcal{F}$ is (ε, δ)-PAC learnable, if a learning algorithm $\mathcal{L}$, given access to a PEX oracle, can generate a hypothesis h, for all distributions $\mathcal{D}$, and for any concept $f \in \mathcal{F}$, such that h is an ε-approximation of f under $\mathcal{D}$, with at least $1-\delta$ probability. *i.e.*

$$Pr[h \leftarrow \mathcal{L}^{\mathsf{PEX}} : \underset{x\in\mathcal{D}}{Pr}[h(x) \neq f(x)] \leqslant \varepsilon] \geqslant 1-\delta \tag{2.117}$$

Similar to what has been discussed in Section 2.5.3, the classical oracles can be translated to quantum setting, for modelling quantum access and to be able to leverage the quantum properties of a data that is encoded in quantum states, such as superposition. In this case, as well, the example oracle has been defined in the quantum setting in [BJ95]. A *Quantum Example Oracle (*QPEX*)* (also called *quantum random access oracle*), for a classical function (or concept) f, over a distribution $\mathcal{D}$ outputs the following state when queried:

$$\mathsf{QPEX}(f, \mathcal{D}) \rightarrow |\psi_{EX}\rangle := \sum_x \sqrt{\mathcal{D}(x)}\,|x, f(x)\rangle \tag{2.118}$$

Another way of translating the dataset queries into the quantum world, is by *Quantum Membership Query (*QMQ*)* oracles [SG04], that is much closer to the back-box oracles we have discussed in Section 2.5.3. A QMQ for a concept f, operates as follows:

$$\mathsf{QMQ}_f : |x, b, y\rangle \rightarrow |x, b \oplus f(x), y\rangle \tag{2.119}$$

which can also take as input superposition state of classical inputs. However, the QMQ is mostly used in the context of *exact* learning instead of PAC-learning. Using the QPEX, one can define a quantum variant of PAC-learnability[24] as follows:

Definition 27 (PAC-learnability with QPEX)**.** We say a concept class $\mathcal{F}$ is (ε, δ)-quantum-PAC learnable or (ε, δ)-PAC learnable with QPEX, if it is PAC-learnable according to Definition 26, with the difference that the learner has been given oracle access to QPEX.

We also point out that this notion of PAC-learnability is quite strong since it is over all the possible distributions. However, one can be interested in the learnability of a concept class over a specific distribution, for instance, a uniform distribution. We refer to this case as *PAC-learnability over D*, where we fix a distribution D from a larger set or all the possible distributions $\mathcal{D}$. We come back to this point in Chapter 3, Section 3.6.3.1.

[24] Although it is better to call this quantum-assisted PAC-learning or PAC-learning with QPEX oracle, to make a distinction with PAC-learning of quantum objects, for instance in this work [PM22a]

We conclude this section by mentioning that, as shown in [SG04], the notions of classical and quantum PAC-learning are equivalent, although not in terms of efficiency. More precisely, if a concept class $\mathcal{F}$ is quantum-PAC-learnable according to Definition 27, it is also classically PAC-learnable, while there exists a gap in terms of *efficient* PAC-learnability between quantum and classical case.

2.6.4 Quantum machine learning and variational algorithms

We have reached now the final section of the background and preliminary materials. In this section, we will give a very high-level introduction to *Quanutm Machine Learning (QML)*. This section is only needed for Chapter 7. As mentioned before, the QML is only about a decade old. However, due to the huge background brought into the field from classical machine learning and the particular interest and attention of the researchers in this field, it has grown quickly compared to its age. Here we only focus on a sub-filed quantum machine learning that we will need for this thesis, namely *Variational Quantum Algorithms (VQA)*. For a comprehensive review of quantum machine learning we refer the enthusiastic reader to [SSP15, BWP+17, MGL22, SP18b].

We also note that the term *quantum* in QML, refers to different classes of problems. In the first class, the learning algorithm is classical but the data is quantum (for instance, using neural networks to analyse measurement statistics from a quantum experiment), in the second one the learning algorithm is quantum, but the data is classical. However, we encode them in quantum states to enable the quantum algorithms to run on them (for instance [HHL09]). And lastly, both data and learning algorithms are quantum (we will see an example of this case in Chapter 7). For a better overview of each of these subfields and to see examples of each case in the NISQ era, we refer the reader to this thesis [Coy22]. Let us begin by introducing VQC.

2.6.4.1 Variational quantum algorithms

Generally speaking, a variational quantum algorithm is, in fact, a hybrid quantum-classical algorithm where the classical part is usually in charge of optimising parameters of a quantum object (a parameterised quantum state or a parameterised unitary circuit). The quantum part is the part of the algorithm that deals with interacting with a quantum input and producing a parametrised output such as $\rho(\boldsymbol{\theta})$, while $\boldsymbol{\theta}$ is the hyper-parameter that will be optimised [Coy22].

Since the first part of the systems that interact with the quantum data, has a quantum nature, and the outputted result are quantum states (while the classical part works with classical data), at some point, the data needs to be measured. This process is defined through a set of observable $\{O_i\}_{i=1}^K$, producing the set of measurement outcomes, or expectation values $\{\langle O_i\rangle_{\boldsymbol{\theta}}\}$, passed on to the classical optimiser. As one can guess, VQAs are quite heuristic methods, but lately, several works have been done to provide theoretical frameworks and guarantees for them, including [MRBAG16]. One of the key components of VQAs is the *cost function*

which defines the learning problem of interest, and we will briefly introduce it in what follows.

2.6.4.2 Cost function

A cost function is a function of the parameters of our model, which quantifies the quality of the learning algorithm. We define the cost function as follows:

Definition 28 (Cost function)**.** Let a learning problem $\mathcal{T}$ be parameterised and characterised by the hyperparamaeter $\boldsymbol{\theta}$, a *cost function* $C(\boldsymbol{\theta})$ is a function of $\boldsymbol{\theta}$, which a learner $\mathcal{L}$, attempts to find its global minimum in order to solve the learning task $\mathcal{T}$.

According to [CAB+21], a good cost function needs satisfy four main properties. **Faithfulness:** meaning that its minimum point should be a solution for the problem of interest in the parameter landscape; **efficient computability:** meaning that it should be reasonably feasible to estimate C in polynomial time. **trainability (or efficient optimisation property):** meaning that the cost function should be optimisable efficiently, *i.e.* it should be differentiable for calculating the gradients, or navigatable in the parameter landscape. And finally, it should have **operational meaning:** that is, a smaller cost function should correspond to closeness to the solution and quality of the learning.

Finding a suitable cost function for the learning problem of interest is one of the most crucial parts of a VQA, and it would contribute significantly to the success and efficiency of the VQA algorithm.

Finally, a cost function can be *local* or *global* [Coy22, CSV+21]. A local cost function corresponds to a local observable and a global cost function to a global one respectively. This distinction about the locality becomes very important in the quantum case, unlike classical cost functions, since local and global observables may differ in important quantum qualities such as entanglement. In Chapter 7, we will see that this property will become very relevant and theoretically interesting to study for our problem.

2.6.4.3 Ansätze

Another essential part of a VQA is its ansatz. This is the part of VQA that creates the parameterised states. Therefore, the form of the ansatz is relevant in the geometrical properties and the landscape of $\boldsymbol{\theta}$ [CAB+21]. More precisely, from an initial state $|\psi_0\rangle$ which is the input of the algorithm, the ansatz creates the parameterised state as follows:

$$|\psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})\,|\psi_0\rangle = U_L(\theta_L)\dots U_2(\theta_2)U_1(\theta_1)\,|\psi_0\rangle \tag{2.120}$$

As it is clear from the above equation, the parameterisation can be made by applying a series of parameterised unitaries $U(\boldsymbol{\theta})$, that is also referred to as *parameterised quantum circuits*. The structure of the ansätzecan be tailored

to the problem (*problem-inspired ansatz*), or it can be generic (which is also called *problem-agnostic ansatz*) [CAB+21]. One advantage of adopting problem-agnostic ansätzeis that they can be tailored to the specific hardware instead, or in other words, be made *hardware-native* or *hardware-efficient*. This type will be particularly desirable for NISQ devices where a limited set of native gates are available, such as Rigetti or QCIBM quantum machines. Finally, hardware-efficient ansätzeaim to reduce circuit depth in VQAs, which is yet another important point for NISQ machines.

2.6.4.4 Optimisation techniques

Finally, the last part of VQA is the classical optimisation method used to minimise the cost function. This part is essentially a classical machine learning subroutine. Hence, different classical optimisation methods can be used depending on the problem and training structure. One of the most exploited optimisation methods that are of particular interest for VQAs is the *gradient-based optimisation* technique. We will also use the same method for our purpose in this thesis.

The gradient of a parameterised cost function in the parametrised landscape is given by the following important result, famously known as the *parameter shift rule*. It states the following [CAB+21, SBG+19, MBK21]:

Theorem 12 (Parameter-shift rule)**.** *Let $C(\boldsymbol{\theta})$ be the cost function, generated using an ansatz of the form $U(\theta_i) = e^{-i\theta_i/2\sigma_i}$ where σ_i are the Pauli operators, the gradient of C with respect to parameter θ_i is given as follows:*

$$\begin{aligned}\frac{\partial C(\boldsymbol{\theta})}{\partial \theta_i} &:= \sum_k \frac{1}{2}(\mathrm{Tr}[O_k U^\dagger(\boldsymbol{\theta_i^+})\rho_k U^\dagger(\boldsymbol{\theta_i^+})] - \mathrm{Tr}[O_k U^\dagger(\boldsymbol{\theta_i^-})\rho_k U^\dagger(\boldsymbol{\theta_i^-})]) \\ &= \frac{1}{2}[C(\boldsymbol{\theta}_i^+) - C(\boldsymbol{\theta_i^-})]\end{aligned} \tag{2.121}$$

where $\boldsymbol{\theta}_i^\pm = (\theta_1, \ldots, \theta_i \pm \frac{\pi}{2}, \ldots, \theta_L)$

In general the above theorem can be generalised to a shift α, instead of $\frac{\pi}{2}$, in that case the coefficient will be $\frac{1}{2\sin\alpha}$, instead of $\frac{1}{2}$ [CAB+21]. Generally speaking, the above theorem shows that one can evaluate the gradient by shifting the parameter by some amount α, which makes the calculation of the cost function more efficient.

3

Unclonability, Unforgeability and Learnability

"I like crossing the imaginary boundaries people set up between different fields - it's very refreshing."

– Maryam Mirzakhani

3.1 Introduction

As mentioned, unclonability is one of the pillars of quantum information and perhaps one of the most fundamental sources of security in quantum cryptography. In this chapter, we make an effort to shine new lights on the meaning of unclonability by bringing it into a broader context. We try to understand a generalisation of no-cloning via two other equally fascinating notions: *Unforgeability* and *Learnability*. The former is a cryptographic notion which we have introduced in Section 2.5.6, and the latter is the subject of study in learning theory and machine learning, which we have briefly discussed in Section 2.6. We first set the scene for the rest of the thesis by expressing the intuitive meaning and connections among these concepts. As we move forward in the chapter, we attempt to formalise some of the presented ideas while also introducing the tools and definitions that we will need to establish our results in the upcoming chapters. In fact, in this chapter, we sketch a big picture, of which we manage to paint only some tiny sections in detail. However, we believe that the general and intuitive overview represents the idea that binds different chapters of this thesis together and will hopefully initiate further thought-provoking questions in this area of research.

Apart from introducing the concepts and notions of interest, we also discuss two main contributions in this chapter. The first one is a new cryptanalysis tool, namely quantum emulation, which serves as a new attack model, as well as a learning tool which builds a bridge between cryptography and learning and leads to several no-go results that we will introduce in this chapter and Chapter 4.

The second contribution is a new framework that formalises the notion of unforgeability in the quantum world. This framework generalises unforgeability

for both quantum and classical primitives and provides us with an accurate yet intuitive tool to study the notion of unclonability from a cryptographic point of view in our future chapters. We also show several case studies as well as no-go and positive results in this framework to demonstrate the broader applicability of the framework in cryptography, and outside the coverage of the cases studied in connection to unclonability.

3.1.1 Structure of the chapter

In Sections 3.2 and 3.3.1 we mostly provide intuitions on the relationships between several concepts including unclonability, unknown transformations, learnability, unforgeability and so on. These two sections can be viewed as extensions of the introduction since we introduce a few novel results. Nevertheless, the reader may find novelty in some of the arguments and more importantly, the way they have been described. In Section 3.4, we revisit the quantum emulation algorithm that we introduced in Section 2.6.2 (Chapter 2), and we give several attack examples based on this algorithm in 3.4.2. In Section 3.5, we introduce our generalised quantum framework for unforgeability and finally, in Section 3.6, we discuss the relevance of our framework through several no-go results, examples and new secure constructions.

3.2 Unclonability and Unknownness

Let us start with this rudimentary question: *Why quantum states are unclonable?* A straightforward answer to this question, and one often found in quantum mechanics and quantum computing textbooks is something along these lines: *Due to the unitarity of quantum mechanics.* Even though this answer is correct (as we have also seen in the proof of the no-cloning theorem in Chapter 2), we want to start this section by pondering whether perhaps there is a more fundamental answer to this question. To dive into this deeper level, we first need to ask the question in a manner that captures the more precise statement of the no-cloning theorem: *Why 'unknown' quantum states are unclonable?*

Allow us to discuss why the word 'unknown' bares such a great deal of significance here. First, a 'known' quantum state, *i.e.* a state which we precisely know its classical description, *is* perfectly clonable merely because by knowing the classical description, one can prepare as many copies as desired of that state. In fact, one can see the classical description as a recipe for making a unitary operation that generates that state over and over from a fixed state (for instance the computational basis $|0\rangle$). The second point is, as we have seen in Section 2.3, *not all* quantum states are unclonable! A known set of orthogonal quantum states is *clonable*. For the qubit case, if, for example, the state $|\psi\rangle \in \{|0\rangle, |1\rangle\}$ (but let's say we don't know which one) then there is a simple cloning machine using the CNOT gate as follows:

$$|\psi\rangle|0\rangle \xrightarrow{\text{CNOT}} |\psi\rangle|\psi\rangle \tag{3.1}$$

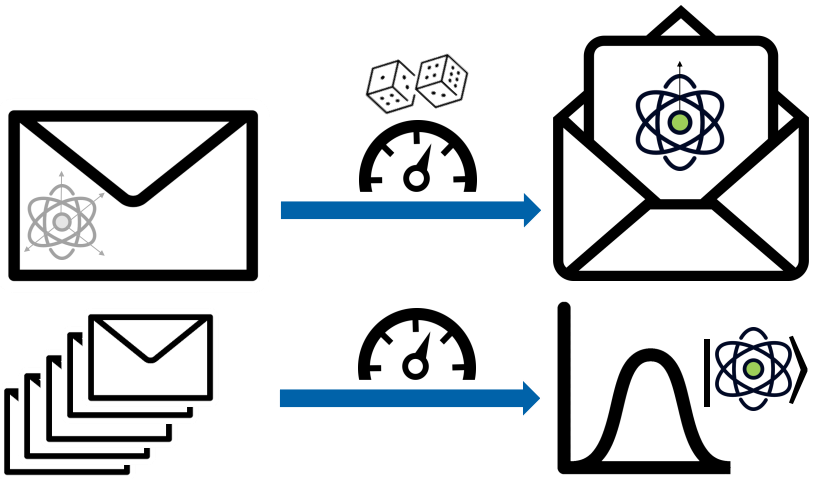

Figure 3.1: Illustration of 'unknownness' property of quantum states and act of measurement as a probabilistic and destructive operation. Here the quantum object, can be seen as a closed envelope that contains information (for instance about its spin, polarization, etc.). While to know this information one needs to open the envelope (measurement) which leads to a probabilistic outcome, and the envelope cannot be closed again, illustrating the destructiveness of the measurement. Moreover having multiple copies of the quantum states allows for extracting the classical description of the state in terms of the statistics of the measurement outcomes.

However, a set of orthogonal quantum states with a known basis, is not a piece of quantum information at all, but rather a classical one, even though carried by a quantum system. The reason is that one can always measure the quantum system on the given basis and *deterministically* know the state of the system. We note that in both examples, we are still in the regime of quantum mechanics, where all the transformations between pure quantum states are unitary, and we are still talking about a quantum mechanical property of a *physical* system, such as the spin of an electron. Therefore, it seems that unclonability, is an aspect of the *quantum information* that we believe to exist in a quantum mechanical system, not the system itself, or in other words, the *lack of information* about a quantum mechanical system, in the realm where we have uncertainty.

This 'unknownness' for quantum states comes from the two most fundamental qualities of measurement in quantum mechanics: Any measurement is inherently *probabilistic*, and it is usually *destructive*. That is, a quantum system, carries some information about its own physical property [1] (for instance, about its spin, polarization, etc.), while the physics would not allow us to know that information with certainty and without *leaving a mark* on the object, that is in a non-reversible way. We illustrate this in Fig. 3.1. On the other hand, imagine a world where perfect cloning of these unknown states is possible. In this world, an observer

[1] Whether this physical properties are 'carried' by the system in reality, or even whether such states bare existential reality, is a matter of philosophical and scientific debates to date. Despite the author's personal affection for the subject, it is very far from the topic of this thesis, and hence we avoid entering that realm, and we note that the sentence has been merely used with an illustrative purpose. We refer enthusiastic readers to some interesting references [Har13, PBR12, LJBR12, Bro19, Omn02, AK16]

who owns such a cloning machine, could make many copies of the quantum state and start measuring them one by one. This observer does not need to worry about destrying the states via measurement since they can keep increasing their knowledge of the quantum system, to the point of certainty, only forming a single copy[2]. Therefore, a perfect cloner would provide a *free source of information* for a quantum system, which is nonphysical.

Finally, we emphasise that as the amount of a priori information about the quantum state increases, cloning becomes more and more feasible. We recall the concept of approximate cloning that we have introduced in Section 2.3. As we have seen, restricting the family of states to be cloned to a specific family, and therefore have *partial prior information* about the state (for instance, phase-covariant or state-dependent cloning as opposed to universal cloning), lead to cloning machines that produce clones with higher optimal fidelity. We will go back to this point in Chapter 7 where we use cloning machines for cryptanalysis.

As we move along this thesis, we aim to show that the unknownness that is profoundly connected to the unclonability of quantum states, manifests itself in other forms of unclonability.

3.2.1 From unclonability of quantum states to unclonability of transformations

Now, we discuss another notion of quantum unclonability which is the unclonability of quantum transformations. Here, by quantum transformation, we mean either a unitary transformation or more generally, a CPTP map. First, we need to clarify what it means to clone a quantum process.

There are several ways to capture the unclonability of quantum transformations. The first one is as defined by Chiribella et. al. in [CDP08] where cloning a transformation Λ means exploiting a single use of Λ inside a quantum circuit, to perform the transformation $\Lambda \otimes \Lambda$ on bipartite states. In this context, a more general no-cloning theorem exists which subsumes the no-cloning of quantum states as a special case: Two black boxes $\mathcal{O}_1$ and $\mathcal{O}_2$ cannot be perfectly cloned by a single-use, unless they are the same, or they are perfectly distinguishable. Here the cloning is intertwined with the task of discrimination between two black boxes via a single-use and is characterised by the minimum of the worst-case error probability for discrimination.[3] An interesting consequence of this theorem is that not only quantum black-box operations, like unitary gates, are unclonable by a single use, but also classical transformations such as permutations of classical registers, are also unclonable by this definition. Let us highlight two key ingredients in this notion of unclonability: *black-box* and *single-use*. One can easily relate the 'black-box' to the concept of 'unknownness' that we have been discussing so far. But the notion of 'single-use' here reveals a yet more interesting fact about

[2]However, we need to emphasise that this does not mean that the uncertainty which exists due to quantum measurements disappears by having the classical deception of a quantum state

[3]The clonable cases then are when this discrimination probability is $p = 0$ (perfect discrimination) or $p = \frac{1}{2}$ (no discrimination).

the unclonability of quantum transformations, which shows an elementary relation between quantum cloning and quantum learning, as has been also noted in the paper. The reason is that the approximate version of the cloner that clones black boxes is allowed to slightly learn the transformation (up to a single interaction). We will elaborate on this point further and try to portray this connection more clearly in Section 3.3.

Another way one may look at the unclonability of quantum transformations is as a pre-processing step for the unclonability of quantum states. Assume we have again a black-box or unknown unitary U, which we use as a generator for an unknown state $|\psi\rangle$ *i.e.* we always input a known state (take state $|0\rangle$ as an example) and receive a state $|\psi\rangle = U|0\rangle$. Now, since the unitary is fully unknown, the output state is also unknown, therefore unclonable. Thus the unclonability of quantum states can also be viewed as the unclonability of its unknown generator. Nonetheless, in this case too, if the unitary is used repeatedly, the generated copies could be used to learn the state, as well as the unitary itself. Generalising this notion, we can think of the task of cloning a quantum transformation, as a general operation while having multiple time access to a fully unknown unitary, which can generate similar outputs of that unitary on different states (*e.g.*: on the full or partial set of bases). We pursue the understanding of this more general notion of unclonability, and its relation to cryptography.

Our first contribution to this end is to formalise the concept of unknown or black-box unitary in a way that would best suit our purpose. We introduce the notion of *Unknown Unitary (UU)*, which we will use throughout the thesis. Intuitively, an unknown unitary is a unitary that we have no prior information about prior to any interactions with it, *i.e.* querying it. However, we formalise this 'lack of prior information as a distinguishability problem. Let us elaborate on this: What we mean by knowing nothing about a unitary, in fact, means that the unitary can be any unitary matrix from the family of *all possible unitaries* that can operate over a Hilbert space (or the associated linear operator space) of a certain dimension. In other words, if this set was finite, the probability that we could guess which unitary has been selected from the set would be the completely random guessing probability over the uniform distribution, which would depend on the cardinality of the set. However, this set is infinite, but as we have seen in Section 2.4, the Haar measure, can describe a uniform measure of randomness over the space of quantum states and unitary transformations. Therefore, we can define the 'unknownness' of a unitary in terms of distinguishability from the Haar measure, as follows:

Definition 29 (Unknown Unitary (UU))**.** Let U^u be a family of unitary transformation where each $U \in U^u$ is a unitary over a D-dimensional Hilbert space, $\mathcal{H}^D$. Also, let λ be a parameter related to D.[a] We say U^u is a family of Unknown Unitaries (UU), if for all the quantum polynomial-time algorithms $\mathcal{A}$, there exists a negligible function $\varepsilon(\lambda) = negl(\lambda)$ such that the difference between the probability of estimating the output of any $U \in U^u$ on any randomly picked state $|\psi\rangle \in \mathcal{H}^D$ and the probability of estimating the output of a Haar random unitary operator on the same state is bounded by $\varepsilon(\lambda)$:

$$\mathop{\mathbb{E}}_{|\psi\rangle}\left[\left|\Pr_{U\leftarrow U^u}[F(\mathcal{A}(|\psi\rangle), U|\psi\rangle) \geqslant \delta(\lambda)] - \Pr_{U_\mu \leftarrow \mu}[F(\mathcal{A}(|\psi\rangle), U_\mu|\psi\rangle) \geqslant \delta(\lambda)]\right|\right] \leqslant \varepsilon(\lambda). \tag{3.2}$$

where μ denotes the Haar measure, F denotes fidelity, $\delta(\lambda) = non\text{-}negl(\lambda)$ is non-negligible value in λ.

[a] In cryptography, since this parameter is usually related to the security, it is called the security parameter.

In Section 2.5.5, we introduced the notion of quantum pseudorandomness and particularly pseudorandom quantum unitaries (PRUs). The definition of UU is, in its essence, very similar to that of PRUs. The reason is that, as discussed above, we have intuitively defined the unknownness with respect to *perfect randomness* in the quantum world, which is the Haar measure. Quantum pseudorandomness is an *approximation* of Haar measure or, in another perspective, a computational version of Haar measure. Additionally, we note that UU is a weaker notion, than PRU and can be considered as a *single-shot pseudorandomness*. We further explore this beautiful relationship between randomness/pseudorandomness, unknownness and unclonability in Chapter 5. As a final remark to conclude this section, we note that the relationship between unclonability and pseudorandomness has been discussed in [JLS18] as well, where the authors have demonstrated a cryptographic variant of the no-cloning theorem.

3.3 Unclonability and different notions of learning

In the previous section, we discussed the intuitive connection between unclonability and the lack of knowledge about a quantum mechanical object such as a quantum state or a quantum transformation, which is what we referred to as *unknownness*. As anything 'unknown' can eventually become 'known' through the process of interacting and learning,[4] our next favourite concept to study in this chapter would be 'learning'. Here, we look at learning from two perspectives: learning theory and cryptography. We seek to find similarities in what can be called learning in

[4] Although this may sound like a philosophical statement, it is also a scientific one! From a physics point of view, it simply refers to the fact that any quantum operation can be learned asymptotically.

these two different views, which will help us better connect them for the rest of the chapter and the following chapters.

3.3.1 Learning, forging and emulation

First, we start with learning. The term *learning* can correspond to numerous different definitions and meanings in different contexts covering purely fundamental to entirely practical spectrum. We are interested in learning in its most theoretical and fundamental sense, as it is often studied in learning theory. Therefore we consider learning as the operation of an algorithm (either deterministic or probabilistic) that targets to learn/predict an object (a function, quantum state, quantum process, distribution, etc.) from a given set of data, usually known as *learning data*, or sometimes *training set*, that includes information about that object. Moving into the quantum realm, both the learning algorithm and the sample data can be quantum, which has given rise to the massive field of quantum learning theory and quantum machine learning. Our principal interest lies in quantum objects such as quantum states and quantum operations.[5] We recall from Section 2.6.3 that the main question regarding learning is *How efficiently can we learn the object?* Where the question is usually answered in time complexity or sample complexity. In particular, the separation between classical and quantum algorithms in terms of the two above efficiency factors stands among the most challenging problems in the field.

Given this very general description of the learning task, the first notion of learning that we discuss is *learning an unknown quantum state*, which usually means learning the *classical description* or similar properties and characteristics of that state. Here the learning data is multiple physical copies of the same unknown state. The learning algorithm measures these copies and then post-processes the measurement outcomes to extract the classical description or the respective property.[6] In general, extracting a complete classical description with arbitrary precision requires an exponential number of samples. However, Huang *et.al.* introduced a technique known as *classical shadow* [HKP20], which allows the efficient learning of many properties of a quantum state. Another note worth mentioning here is that there is a well-known result in quantum information showing an equivalence between optimal universal $N \to \infty$ quantum cloning of pure states and optimal state estimation devices taking as input N copies of an unknown pure state [SIGA05, GM97]. More importantly, bounds on optimal cloning can be derived from this equivalence. This result establishes the deep connection between unclonability and unlearnability of quantum states.

The next notion is perhaps the most well-studied concept in learning theory: function learnability. We recall from Section 2.6.3, the scenario where given a family of functions, $\mathcal{F}$, we want to efficiently learn a representation of any function, $f \in \mathcal{F}$. The task of learning f can be considered to be *exact* or *approximate* where

[5] However, classical functions can also be represented via a quantum unitary (Section 2.5.3). Thus, they too are subjects of our investigation.

[6] This process is also known as *state tomography* [NC10, BCD+09], discussed in Section 2.6.1.

the latter, the function is learned approximately but with high probability, which is denoted as PAC-learning (Definition 26). In this learning model, the sample data is sample pairs $(x, f(x))$ where x has been sampled according to some distribution $\mathcal{D}$, given to the learning algorithm via an oracle, which outputs one such pair on each call. In the quantum version of these learning notions, one can also allow the oracle to give quantum access to the underlying classical function, producing superposition queries.

Next, assume a weaker notion of learning, where we do not expect the learner to be able to learn *all* the functions in a family (or a concept class), but instead given a function f, selected from a family $\mathcal{F}$, the learning algorithm needs to output a correct new pair $(x, h(x))$, from the set of inputs and outputs of f (which can also be given via interacting with an oracle). The validity of the pair is tested with a verification algorithm, which in many cases (but not all the times), checks whether $h(x) = f(x)$ or not. This scenario is a very well-known scenario in cryptography, known as *forgery*. In fact, in cryptography, one would ideally want to avoid all such cases where an efficient forger can exist for a function f. More precisely, in classical cryptography, the function f is usually selected from a keyed-family of functions (where the key is sampled uniformly at random) and the input x can be either selected from a distribution or by the forger itself arbitrarily. Here to ensure the security of a cryptographic scheme which employs function f, we require that no adversary is able to produce such valid pairs. In other words, the function f should be not easily learnable/predictable from a limited set of samples. Here the adversary, which is a probabilistic algorithm, runs an efficient learning process for this specific instance of a learning problem that we call forgery. In Section 3.3.3 we will further discuss this relationship and the specifications of the oracles.

Finally, learning a quantum transformation is yet another notion of learning, that shares similarities with the above. The most conventional notion is known as *quantum process tomography* in the literature. In process tomography, given an unknown unitary or quantum channel, we are interested in learning the classical description, or characteristics of the quantum process by interacting with it with many quantum states and measuring the outcomes. It is not hard to see that without any prior information about the unitary or channel, this task is highly inefficient [NC10, MRL08], given the fact that even learning a good approximation of an unknown quantum state is quite resource-intensive. Nevertheless, this is not the only notion of learning one can imagine for quantum processes. In Section 2.6.2.1, we introduced the notion of *emulation* and how it differs from *simulation*. As the simulation is probably closer to the concept of tomography (and even PAC-learning to some degree), we argue that emulation is very close to the notion of forging, where here, the emulation will forge an unknown unitary instead of a classical function. We recall that an emulation algorithm aims to produce a close approximation of the output of a unitary U, on a given state $|\psi\rangle$, from a set of input and output samples, which is similar to the forgery scenario presented above.

We keep these intuitions in mind as we move forward, and we focus on some

of these notions of learning such as emulation (Section 3.4).

3.3.2 Unforgeability and unclonability

Here, we look at the relation between unforgeability and unclonability. These two properties become particularly related for *quantum tokens* like quantum money [Wie83**?** , AC12], quantum coin [Gav12], and more generally what is known as *unforgeable tokens*. These are unique objects which can be produced and verified by an honest party, but no untrusted party can generate such valid tokens. Considering that the generators of such objects are classical functions or quantum processes, the role of the adversary would be to *forge* or learn them, as discussed in the previous section, and in the cryptography world, resisting forgery attacks is captured by the property of *unforgeability*. Despite its simple intuitive meaning, unforgeability is not very easy and trivial to capture formally, especially in the quantum world. In Section 2.5.6, we discussed different classical definitions of unforgeability and some candidates for unforgeability definitions in the quantum world. We will also present our formal framework for quantum unforgeability later, in this chapter. But for now, let us see how this property is related to unclonability.

Let us look at it through a very simple example regarding quantum money. Assume a bank (or the mint) producing some notes, *i.e.* the physical objects used as money. Each note has a unique serial number attached to it which provides a basis for the verification of the note when the user wants to use it for a transaction [Ver19]. These notes are distributed among untrusted users, who are willing to create more notes than they originally had in their possession. Thus unforgeability is an important property for a note, meaning that the dishonest party cannot come up with another unique serial number that would be also valid, and hence pass the verification. But on the other hand, an easy way of creating more notes is to simply copy the whole note, since the new one will also have a valid serial number. As a result, cloning a token is a simple but applicable forging attack.

In the classical world, nothing prevents a user with sufficient resources to forge the note physically. However, in 1983, Wiesner first introduced the idea of quantum money, based on the no-cloning theorem [Wie83]. Assuming your token to be a physical quantum system described by the quantum state $|\$_{\texttt{serial}}\rangle$ with the serial number `serial`, the unclonability of these states leads to the unforgeability of the quantum money scheme. This unforgeability stems from quantum mechanics, without any extra assumption. On the contrary, if a perfect quantum cloning machine could exist, no such quantum systems could satisfy unforgeability, irrespective of the scheme.

Wiesner's quantum money is not the only cryptographic primitives where unforgeability and unclonability are related. Another example is a public-key quantum money scheme known as *Quantum Lightning* [Zha21]. Similar to Wiesner's quantum money, quantum lightning also relies on the unclonability of quantum states to prevent forgery and duplication attacks. This scheme uses a one-time signature, and as discussed in the preliminaries, the main security property of a signature

scheme is unforgeability. Moreover, the public-key quantum money functionality has often been seen as a computational or cryptographic variant of the no-cloning theorem.

Another relevant cryptographic functionality that is worth mentioning in this section, is *Quantum copy-protection*, initially proposed by Aaronson [Aar09], and that has been developed throughout several works recently [ALP21, ALL+21, BJL+21, SW22]. The idea behind quantum copy-protection is to use quantum unclonability to achieve programs that cannot be copied. This functionality has many interesting applications such as *Secure software leasing* [ALP21, BJL+21]. Assume a program f is given not as a classical function but instead, in the form of a quantum state $|\psi_f\rangle$, such that from this state, f can be computed on any arbitrary input, yet it is infeasible to copy it or convert it into two other arbitrary states from which you can still compute f. Despite the clear connection between this functionality and the unclonability of quantum states, quantum copy protection is a stronger and more demanding requirement than simple unclonability. Also, quantum copy-protection does not directly link with unforgeability to the best of our knowledge. However, we argue that it is related to another concept that we have been discussing in this section, *i.e. learnability*. Aaronson shows that any 'learnable' program can be copy-protected. Here learnability means that the output of the function cannot be predicted from input and output behaviour, which is very similar to unforgeability. Thus it would not be surprising that 'forgeable' functions could also not be copy-protected. Nonetheless, as there are different definitions and levels of unforgeability, there are also several definitions for learnability in this context. Some of them have been introduced in [ALL+21]. More surprisingly, a recent result [ALP21] shows, that under certain computational assumptions, even certain unlearnable programs cannot be copy-protected. Therefore, the connection between learnability (as well as unforgeability) and copy-protection is still an intriguing open problem. We conjecture that the categorisation of copy-protectable functions from an unforgeability point of view can also be insightful.

Finally, we note that there are plenty of other cryptographic primitives of this sort, where unclonability is a core aspect which enables features that are infeasible classically. However, the careful study of every one of them, including the above examples, is a research field on its own and indeed, outside the scope of this thesis. Our goal in this section was to sketch the existence of this relation via some well-known examples to be able to highlight it later in the other forms of unclonability, such as the one we present in Chapter 4.

3.3.3 Unforgeability and learnability with quantum oracles

As we have discussed, unforgeability has a close and intuitive connection with the notion of learning. On the other hand, unforgeability is a cryptographic property we require for many different primitives and cryptographic schemes. We can therefore ask:

What does it mean to learn a cryptographic primitive?

First, we note that the core element of a cryptographic primitive is a classical function (or equivalently a quantum process in the quantum case) which maps the domain to the range, where the specific properties of this function (or operation) often are employed to achieve different cryptographic functionalities. In this context, learning a primitive means either learning the full underlying function (or quantum process) or alternatively learning the desired property. This learning process as we discussed in Section 3.3.1, can be used by an adversary who tries to break the cryptosystem. For this learning algorithm to be able to function, one needs the *learning data*. In the cryptography literature, similar to the learning theory, this learning data is usually obtained by interacting with an oracle that models an interactive platform for the full and perfect implementation of the evaluation function.

Here, we are mainly interested in quantum algorithms (and hence quantum adversaries) and the notion of quantum unclonability. Thus the relevant case for us is when the interaction with the primitive is also realised quantumly. For quantum primitives, this is a natural model to consider, as we are dealing with a quantum process that produces quantum outputs and often takes quantum inputs as well. For classical primitives on the other hand, according to the different adversarial models that we have introduced in Section 2.5.2, this brings us to the *quantum security regime*, where the interaction with the primitive is also considered to be quantum.

In the preliminaries (Section 2.5.3), we have already discussed how a quantum accessible oracle can be defined for a classical function. Here, we recall that notion and we also present the same model for quantum primitives. In this way, we can study both classes of primitives in an analogous model, *i.e.* the quantum oracle model, which we will use throughout the chapter and the thesis. We also discuss how learning from these quantum oracles, can essentially link to unforgeability or other similar cryptographic properties of interest.

3.3.3.1 Quantum oracle for classical vs quantum primitives

Let us first recall the *standard quantum oracle* for classical primitives which we introduced in Section 2.5.3. The standard oracle is a black-box unitary of a reversible version of a classical-polynomial-time computable function f, which can represent a deterministic or randomised primitive, defined as follows:

$$R\mathcal{O}_f : \sum_{m,y} \alpha_{m,y} |r\rangle_{\mathcal{O}} |m, y\rangle \rightarrow \sum_{m,y} \alpha_{m,y} |r\rangle_{\mathcal{O}} |m, y \oplus f(m; r)\rangle \tag{3.3}$$

Having access to this quantum operation, a quantum adversary $\mathcal{A}$ can get the outcome of the function f for several classical inputs m in one query, in a superposition form over the message set of their choice. However since the primitive can be randomised, the value of the function in each execution can also depend on the random value. This scenario is modelled via the randomised version of the oracle or the randomised oracle as shown above. In the case of

deterministic primitives, the adversary gets full query access to the unitary that maps all the computational basis to another combination of computational basis that represents the range of the function. In the randomised case, on the other hand, the adversary usually gets a more *limited* access to an extended unitary transformation over the joint Hilbert space of the function and the randomisation, or in other words, to the full Hilbert space of all the possible outcomes of all the functions, parameterised with the random value. In Section 2.5.3 we have briefly discussed some of the interesting cases and questions that arise regarding which level of access to the extended unitary can be permitted for the quantum adversary. Nonetheless, the key point that we want to emphasise is that the information about the primitives is accessible to the adversary via the *quantum* input and outputs of this unitary, which is, initially unknown to the adversary.

For quantum primitives, the modelling of this scenario is even more evident. Quantum schemes often work with families of quantum states which are usually unknown (or partially unknown) quantum states from the adversary's point of view, or similarly, an unknown quantum operation. In both cases, one can model the primitive with the evaluation between quantum states and thus can define the oracle as a general unitary transformation for a deterministic primitive, as follows:

$$\mathcal{O}_U : \sum_i \alpha_i |b_i\rangle \xrightarrow{U} \sum_i \beta_i |b_i\rangle \tag{3.4}$$

Here $\{|b_i\rangle\}$ are a basis (not necessary computational basis) for $\mathcal{H}^D$, the Hilbert space that the unitary operates upon. We note that quantum primitives can perform an arbitrary rotation of the bases. The analogue of this type of oracles for classical primitives, are *type-2* oracles (also called *minimal oracles*)[GHS16, GKS21]. We also note that in the non-randomised case, this oracle generalises the standard quantum oracle. In other words, the class of all the possible standard oracles of a certain dimension are a subclass of all the quantum oracles of the form Eq. (3.4) over the same dimension.

A randomised quantum primitive can also be defined similarly to the classical case. Here we give an abstract notation of a general randomised quantum primitive, but we further clarify the realisation of such oracles in the upcoming sections. We denote a general randomised unitary oracle for quantum primitives as follows:

$$\mathcal{O}_U : \sum_i \alpha_i |r\rangle_{\mathcal{O}} |m_i\rangle \xrightarrow{U} \sum_i \beta_i(r) |r\rangle_{\mathcal{O}} |m_i\rangle \tag{3.5}$$

Hence a $\mathcal{O}_U$ is a unitary over the joint space of the oracle's randomness register and the main input state, which consists of a family of smaller unitaries parameterised by a random internal parameter r.

Now, back to the problem of learning, we can see that in both cases, the learning data is a set of input and output quantum states $\{(\rho^{in}, \rho^{out})_i\}_{i=1}^{q}$ [7] of

[7] In general, these output states can be entangled across different queries. In some cases, where the query to the oracle was performed adaptively and sequentially, the number of quantum states and the mathematical structure of this quantum learning data might be different. However, we will argue that this generality can still hold. For additional technical discussion, see Section 3.5.1

an unknown unitary that is provided to the learning algorithm from interacting q-times with the respective quantum oracle. Let us point out a few aspects of this analogy between learning and cryptographic problems in the oracle model.

First, depending on whether the access to the oracle is direct or indirect via another agent or honest party, different quantum adversarial models emerge. From another perspective, the learning data can, in general, be sampled from a distribution, which can also impose restrictions or relax conditions on the learning problem. Second, the number of queries q (either in the worst case or average case) specifies the *query complexity* of the learning algorithm but also translates to the power of the quantum adversary in the cryptography language. Third, it is clear that in both of these models, any such unknown unitaries are eventually 'learnable' if the number of queries to the oracle is unbounded. The same situation in cryptography is commonly known as *brute-force attacks*. The lower the query complexity, the higher the efficiency of the algorithm and the more they become interesting attacks on cryptosystems. Therefore, at the intersection of cryptography and learning theory, we are usually interested in learning algorithms with polynomial [8] query complexity. Many of the learning algorithms, such as process tomography [DLP01, PR04, Hay05, KJ09, BCD+09, OW16], or certain quantum machine learning algorithms [BWP+17, WSK+21, AdW17b, PM22b] do not have such polynomial-size query complexity and thus despite being very useful in other areas, are not usually compelling toolkits for cryptanalysis. However, in recent years, there has been significant progress in the development of *efficient* algorithms for learning quantum states and quantum processes [ML16, Aar20, HKP20]. We believe these algorithms are powerful yet fairly unexplored tools while studying problems in the domain of cryptography. An example of this sort of learning algorithm is a technique called *shadow tomography* introduced by Aaronson [Aar20], which has also been studied in the context of quantum money.

In what follows, we study one candidate of such efficient learning algorithms. The algorithm that we study is an algorithm for quantum emulation which aims to generate the output of an unknown unitary to an unknown quantum state from learning samples of that unitary. We have sought to uncover the relevance of such methods in relation to unclonability and the cryptographic properties of unclonable objects. From the next section on, we use these two main ingredients to understand unclonability in a broader context: *quantum emulation* as a learning tool and *unforgeability* as a cryptographic characterisation. In order to make our intuitive arguments more precise, we will need to have a closer look at both of them.

3.4 Universal quantum emulator revisited

In the previous section, we discussed different notions of *learning* and their relation to unclonability, as well as cryptographic concepts like unforgeability. We

[8] In a given parameter that either quantifies the resource, or the security, or in some cases both. For instance, the number of input qubits of the algorithm.

have also seen that emulating an unknown quantum transformation is a way of *learning* with close proximity to *forging* that process. Previously in Section 2.6.2 of Chapter 2, we have seen a quantum algorithm from [ML16] that performs the task of emulating an unknown quantum unitary on an unknown input quantum state by having some of the input-output samples of the unitary. We have seen some properties of the universal quantum emulator algorithm, such as efficiency and complexity results. Re-purposing the algorithm from its original target in the context of tomography, here we take a new look at this algorithm, and more generally emulation, from a cryptanalysis perspective. In other words, since it seems the objectives of emulation and unforgeability are in the complete opposite direction of each other, we propose emulation as a general attack model against the notion of unforgeability in general. To make this statement more formal, we will first require a formal definition of *unforgeability* itself, for which we need to wait until the next section. While in this section, we revisit the quantum emulation algorithm from an adversarial point of view and as a cryptanalysis toolkit. For this purpose, we need to provide a new fidelity analysis of the algorithm exploiting a specific asymmetry in the algorithm that can be used effectively in adversarial scenarios. We then show a few examples of this new class of quantum attacks, which we call *quantum emulation attacks*.

3.4.1 Output fidelity analysis

We are interested in the fidelity of the output state $|\psi_{QE}\rangle$ of the algorithm and the intended output $\mathsf{U}|\psi\rangle$ to estimate the success. Here we are more interested in the explicit form of the output states and the theoretical bounds for the fidelity, rather than the complexity analysis provided in [ML16]. We note that for our calculations, all the gates including the controlled-reflection gates are assumed to be ideal keeping in mind that the implementation is possible with the technique of *quantum principal component analysis* developed in [LMR14], as also mentioned in [ML16]. We recall from Section 2.6.2 that the fidelity of the output is related to the success probability of the first stage of the algorithm in the following way:

$$F(\rho_{QE}, \mathsf{U}\rho\mathsf{U}^\dagger) \geqslant F(\mathcal{E}_{\mathsf{U}}(\rho), \mathsf{U}\rho\mathsf{U}^\dagger) \geqslant \sqrt{P_{succ-stage1}} \tag{3.6}$$

Also, from the proof of Theorem 11 from [ML16], it can be seen that the success probability of Stage 1 is calculated as follows:

$$P_{succ-stage1} = |\langle\phi_r|\,\mathsf{Tr}_{anc}(|\chi_f\rangle\langle\chi_f|)\,|\phi_r\rangle|^2 \tag{3.7}$$

where $|\chi_f\rangle$ is the final state of the circuit after Stage 1 and $\mathsf{Tr}_{anc}(\cdot)$ computes the reduced density matrix by tracing out the ancillas. The overlap of the resulting state and the reference state equals the success probability of Stage 1. Now we only use Eq. (3.7) for our analysis henceforward. For this section, we need a more precise and concrete expression for the output fidelity.

Here we point an important observation about the algorithm. The fidelity of the output state of the circuit highly depends on the choice of the reference

state Eq. (3.7) such that it may increase or decrease the success probability of the adversary in different security models as we will discuss. We establish the following recursive relation for the state of the circuit after the i-th block of Stage 1, in terms of the previous state:

$$|\chi_i\rangle = \frac{1}{2}[(I - R(\phi_r))\,|\chi_{i-1}\rangle\,|0\rangle + R(\phi_i)(\mathbb{I} + R(\phi_r))\,|\chi_{i-1}\rangle\,|1\rangle]. \tag{3.8}$$

Now by using this relation, we can prove the following theorem:

Theorem 13. *Let $|\chi_K\rangle$ be the output state of K-th block of the circuit (Fig. 2.8). Let $|\psi\rangle$ be the input state of the circuit, $|\phi_r\rangle$ the reference state and $|\phi_i\rangle$ other sample states. We have:*

$$\begin{aligned}|\chi_K\rangle = {} & \langle\phi_r|\psi\rangle\,|\phi_r\rangle\,|0\rangle^{\otimes K} + |\psi\rangle\,|1\rangle^{\otimes K} - \langle\phi_r|\psi\rangle\,|\phi_r\rangle\,|1\rangle^{\otimes K} \\ & + \sum_{i=1}^{K}\sum_{j=0}^{i}[f_{ij}2^{l_{ij}}|\langle\phi_r|\psi\rangle|^{x_{ij}}|\langle\phi_i|\psi\rangle|^{y_{ij}}|\langle\phi_r|\phi_i\rangle|^{z_{ij}}]\,|\phi_r\rangle\,|q_{anc}(i,j)\rangle \\ & + \sum_{i=1}^{K}\sum_{j=0}^{i}[g_{ij}2^{l'_{ij}}|\langle\phi_r|\psi\rangle|^{x'_{ij}}|\langle\phi_i|\psi\rangle|^{y'_{ij}}|\langle\phi_r|\phi_i\rangle|^{z'_{ij}}]\,|\phi_i\rangle\,|q'_{anc}(i,j)\rangle\end{aligned} \tag{3.9}$$

where l_{ij}, x_{ij}, y_{ij}, z_{ij}, l'_{ij}, x'_{ij}, y'_{ij} and z'_{ij} are integer values indicating the power of the terms of the coefficient. Note that f_{ij} and g_{ij} can be 0, 1 or -1 and $q_{anc}(i,j)$ and $q'_{anc}(i,j)$ output a computational basis of K qubits (other than $|0\rangle^{\otimes K}$).

We give an induction proof of this theorem in Appendix A.1.

Having a precise expression for $|\chi_f\rangle$ from Theorem 13, one can calculate $P_{succ-stage1}$ of Eq. (3.7) by tracing out all the ancillary systems from the density matrix of $|\chi_f\rangle\langle\chi_f|$. Also, now it is clear that if $|\psi\rangle$ is orthogonal to $\mathcal{H}^d$, the only term remaining in Eq. (3.9) is $|\psi\rangle\,|1\rangle^{\otimes K}$. So, the input state remains unchanged after the first stage and $P_{succ-step1} = 0$. For states projected in the subspace spanned by S_{in}, the overall channel describing the quantum emulation algorithm has always a fixed point inside the subspace [ML16]. Hence, Stage 1 is successful with probability close to 1 by assuming the gates to be ideal.

Let us see two simple examples of the above theorem, which we will use in the future. First, assume that we have only two sample states: one reference state $|\phi_r\rangle$, and another sample state $|\phi_1\rangle$, and their respective output states. Trying to run a quantum emulation algorithm with this database, will lead to an emulation algorithm that has only one block in the first stage. In this case, the output state $|\chi_1\rangle$ of the circuit after the first stage is given as a function of the following overlaps,

$$\langle\phi_r|\psi\rangle = \alpha, \quad \langle\phi_1|\psi\rangle = \beta, \quad \langle\phi_r|\phi_1\rangle = \gamma, \tag{3.10}$$

as follows:

$$|\chi_1\rangle = \alpha\,|\phi_r\rangle\,|0\rangle + |\psi\rangle\,|1\rangle - \alpha\,|\phi_r\rangle\,|1\rangle - 2\beta\,|\phi_1\rangle\,|1\rangle + 2\alpha\gamma\,|\phi_1\rangle\,|1\rangle \tag{3.11}$$

Expanding this two two-blocks emulation, using the same formula, we have:

$$\begin{aligned}|\chi_2\rangle &= \alpha\,|\phi_r\rangle\,|00\rangle + 2\gamma_1(\alpha\gamma_1-\beta_1)\,|\phi_r\rangle\,|01\rangle + |\psi\rangle\,|11\rangle \\ &+ (2\beta_1\gamma_1-\alpha-2\alpha\gamma_1^2)\,|\phi_r\rangle\,|11\rangle + 2(\alpha\gamma_1-\beta_1)\,|\phi_1\rangle\,|11\rangle \\ &+ 2(\alpha\gamma_2-\beta_2+2\beta_1\delta-2\alpha\delta\gamma_1-2\alpha\beta_1\gamma_1\gamma_2+2\alpha\gamma_1^2\gamma_2)\,|\phi_2\rangle\,|11\rangle\end{aligned} \tag{3.12}$$

where the coefficients are given by the following pair overlaps:

$$\begin{aligned}&\langle\phi_r|\psi\rangle=\alpha, \quad \langle\phi_1|\psi\rangle=\beta_1, \quad \langle\phi_2|\psi\rangle=\beta_2 \\ &\langle\phi_1|\phi_r\rangle=\gamma_1, \quad \langle\phi_2|\phi_r\rangle=\gamma_2, \quad \langle\phi_1|\phi_2\rangle=\delta.\end{aligned} \tag{3.13}$$

Although calculating these explicit forms seems repetitive, having them provides the ability to optimise the output fidelity with respect to the choice of states. We underline that for the initial purpose of the algorithm, this was not relevant as the sample states were assumed to be chosen randomly, while in some adversarial models, specifically when giving oracle access to the unknown target unitary, the adversary has the advantage of selecting the best possible states. Hence, having the explicit forms in terms of overlaps is a significant step towards using quantum emulation as a cryptanalysis toolkit.

3.4.2 Quantum Emulation Attacks

Now, let us see how we can use this algorithm with the new given picture and results in creative ways. The first quantum emulation attack that we present is the one that we will use several times in this thesis to show non-trivial impossibility results and also happens to be the simplest emulation attack. Here we present the most general format, without explicitly specifying the model or formal game in which it is used. Later in Section 3.6 and also Chapter 4 we will go back to this attack and use it in a formal way within game-based security models. The other two attacks we give in this section, mostly serve as toy examples to demonstrate the possibility of different attacks one can build given a small-size quantum emulator.

3.4.2.1 One-block quantum emulation attack

Imagine the scenario where an adversary $\mathcal{A}$ who has access to the following samples of an unknown unitary U:

$$\{|\phi_1\rangle, |\phi_r\rangle\}, \quad \{|\phi_1^{out}\rangle, |\phi_r^{out}\rangle\} \tag{3.14}$$

and tries to closely approximate the output $U|\psi\rangle$ for a target state $|\psi\rangle$. We assume the adversary $\mathcal{A}$, has the ability to select $|\phi_1\rangle$ and $|\phi_r\rangle$ and interact with the unitary to obtain the respective outcome.

Without loss of generality assume the case that $|\phi_1\rangle$ is a computational basis and the target state $|\psi\rangle$ is another computational basis (or more generally any states such that $\langle\psi|\phi_1\rangle=0$). Now the question will be how to choose the best $|\phi_r\rangle$

for this emulation. From Theorem 13 and Eq. (2.113) we already know that the reference state should have some overlap with the target state for the emulation to be successful in this case. For finding a general result, we parameterise the choice of the reference state in the amplitudes of the reference state, as follows:

$$|\phi_r\rangle = \sqrt{1-\alpha^2}\,|\phi_1\rangle + \alpha\,|\psi\rangle \tag{3.15}$$

Now $\mathcal{A}$ can run a one-block quantum emulation algorithm. We can directly use Eq. (3.11), noting that $\langle\phi_1|\psi\rangle = 0$ and $|\langle\psi|\phi_r\rangle|^2 = \alpha^2$ and $|\langle\phi_1|\phi_r\rangle|^2 = 1-\alpha^2$

$$|\chi_1\rangle = \alpha\,|\phi_r\rangle\,|0\rangle + |\psi\rangle\,|1\rangle - \alpha\,|\phi_r\rangle\,|1\rangle + 2\alpha(\sqrt{1-\alpha^2})\,|\phi_1\rangle\,|1\rangle\,. \tag{3.16}$$

By calculating $|\chi_1\rangle\langle\chi_1|$, tracing out the ancillary systems and using Theorem 11, we can bound the fidelity of the output state of the emulator, denoted by $|\psi_{QE}\rangle$, and the target which is $U|\psi\rangle$ as follows:

$$F(|\psi_{QE}\rangle\langle\psi_{QE}|\,, U^\dagger|\psi\rangle\langle\psi|\,\mathrm{U}) \geqslant \alpha^2[1+4(1-\alpha^2)^2] \tag{3.17}$$

We can see that this fidelity, is a considerable value as long as α is not too small (in a cryptographic sense, the fidelity is a non-negligible function of the security parameter as long as α is not a negligible function). A trivial case is where fidelity is 1, for $\alpha = 1$, which means the reference state and the target state are the same. But there is also a non-trivial case for the fidelity to become unity, and that happens for $\alpha = \frac{1}{\sqrt{2}}$. This is when the reference state is a uniform superposition of the target state and the other sample.

Thus, we can see that this freedom of carefully choosing the sample states of an unknown unitary, enables an adversary to perform a successful emulation with high fidelity. This is important in the context of cryptography as the unknown unitary can be a quantum oracle of a classical primitive, and the output of the emulation can leak important information about the underlying function. In fact, as we will see later, this sort of attack directly connects to the unforgeability property.

3.4.2.2 Quantum emulation attack on the inverse function

Let us see another example. Assume that we have a classical function f that is efficiently computable but hard or inefficient to invert (for instance a one-way function). However, we assume that the function is invertible *i.e.* the f^{-1} exists. Our goal is to show some information about the inverse of f that can be extracted, with only having a standard quantum oracle access to U_f, using a small quantum emulation attack.

Let the following unitary be the standard quantum oracle for classical function f:

$$U_f : \sum_{x,t} \alpha_{x,t}\,|x,t\rangle \to \sum_{x,t} \alpha_{x,t}\,|x, f(x)\oplus t\rangle \tag{3.18}$$

Since f is easy to compute the unitary U_f is also efficiently implemented since we have assumed the inefficiency of computing the inverse implies that in general, an

efficient implementation of $U_{f^{-1}}$ does not exists, or in other words, we can assume that $U_f^\dagger \neq U_{f^{-1}}$ and having access to U_f and its complex conjugate does not lead to a trivial implementation of $U_{f^{-1}}$. Nonetheless, the standard oracle form of $U_{f^{-1}}$ takes the following form:

$$U_{f^{-1}} : \sum_{y,t} \alpha_{y,t} |y,t\rangle \rightarrow \sum_{y,t} \alpha_{y,t} |y, f^{-1}(y) \oplus t\rangle \tag{3.19}$$

where $y = f(x)$ and thus $|y, f^{-1}(y) \oplus t\rangle = |f(x), f^{-1}(f(x)) = x \oplus t\rangle$.

We show that an adversary $\mathcal{A}$ can emulate $U_{f^{-1}}$ and extract information about f^{-1} without having any oracle access to $U_{f^{-1}}$, and by only querying U_f[9].

First, we show the analysis of the emulation's output if the sample states from $U_{f^{-1}}$ were available. Then we propose a method to translate the queries of U_f to the queries of $U_{f^{-1}}$ which are required for the emulation attack.

Sample states $\mathcal{A}$ needs the following queries from the unknown unitary $U_{f^{-1}}$:

$$\{|\phi_1\rangle = |y_1, 0\rangle\,, \quad |\phi_r\rangle = \frac{1}{\sqrt{2}}(|y_1, 0\rangle + |y_k, 0\rangle)\} \tag{3.20}$$

and their respective outputs:

$$\begin{aligned} &\{|\phi_1^{out}\rangle = |y_1, f^{-1}(y_1)\rangle = |f(x_1), x_1\rangle\,, \\ &|\phi_r^{out}\rangle = \frac{1}{\sqrt{2}}(|y_1, f^{-1}(y_1)\rangle + |y_k, f^{-1}(y_k)\rangle) = \frac{1}{\sqrt{2}}(|f(x_1), x_1\rangle + |f(x_k), x_k\rangle)\} \end{aligned} \tag{3.21}$$

where $y_1 = f(x_1)$ and $y_k = f(x_k)$ are classical outputs of the function f. Also note that $|\phi_1\rangle$, as well as $|\phi_1^{out}\rangle$, are a computational basis of $\mathcal{H}^D$ over which U_f and $U_{f^{-1}}$ operate. Thus, this query is equivalent to a classical query.

Target state Let the target state of emulation, and the expected outcome be $|y_k, 0\rangle$, and $|y_k, x_k\rangle$ respectively, where the second register $|x_k\rangle$ is the desired output of $U_{f^{-1}}$ (or the pre-image of y_k).

Given the one-block emulation attack from the previous section, we know that the output fidelity for this case ($\alpha - \frac{1}{\sqrt{2}}$) is equal to 1. Therefore $\mathcal{A}$ could perfectly extract the x_k with probability one if having access to $U_{f^{-1}}$ oracle to be able to get the above samples. Now let us see how one can obtain the required samples, by interacting with U_f instead.

Translating queries of U_f to queries of $U_{f^{-1}}$:

[9] We note that this is a forgery attack on such functions, however it clearly does not break the one-wayness property since to break this property, one needs a pre-image of a random y should be found.

Let the following be the sample states of $\mathcal{A}$ after querying U_f:

$$\begin{aligned} &|\phi_1^u\rangle = |x_1, 0\rangle, \quad |\phi_1^{u,out}\rangle = |x_1, f(x_1)\rangle \\ &|\phi_2^u\rangle = \frac{1}{\sqrt{2}}(|x_1, 0\rangle + |x_k, 0\rangle), \quad |\phi_2^{u,out}\rangle = \frac{1}{\sqrt{2}}(|x_1, f(x_1)\rangle + |x_k, f(x_k)\rangle) \end{aligned} \tag{3.22}$$

Our goal is to make the transformation in a non-destructive way, instead of simply measuring the states which would lead them to collapse. Obviously, $|\phi_1\rangle$ can be prepared classically and the output, $|\phi_1^{out}\rangle$ can be easily obtained from $|\phi_1^{u,out}\rangle$ by simply performing a SWAP gate between first and second part of the register *i.e.* $|\phi_1^{out}\rangle = \text{SWAP}_{12}|\phi_1^{u,out}\rangle$. Similarly, $|\phi_r^{out}\rangle$, can be obtained by swapping the first and second register of $|\phi_2^{u,out}\rangle$, *i.e.* $|\phi_r^{out}\rangle = \text{SWAP}_{12}|\phi_2^{u,out}\rangle$.

It remains to obtain $|\phi_r\rangle$ from $|\phi_2^{u,out}\rangle$ which is more complicated as there is an entanglement between the registers that need to be removed. We give a small sample algorithm, Algorithm 1, to perform this task for the qubit case, although it is easily generalizable to n-qubits as well.

Algorithm 1 Mini sample converter algorithm

Description: Translating $|\phi_2^{u,out}\rangle$ to $|\phi_r\rangle$: We assume that $|x_1\rangle$, $|x_k\rangle$ and respectively $|y_1\rangle$ and $|y_k\rangle$ are qubit. If $x_1 \neq x_k$, then the state $|\phi_2^{u,out}\rangle$ is an entangled state which can be written as: $|\phi_2^{u,out}\rangle = \frac{1}{\sqrt{2}}(|0, y_1\rangle + |1, y_k\rangle)$. The algorithm proceeds as follows:

- Add an ancillary qubit $|0\rangle$, and perform a SWAP gate in the first qubit and the ancillary qubit leading to:

$$\begin{aligned} \text{SWAP}_{a1}|0_a\rangle|\phi_2^{u,out}\rangle &= \text{SWAP}(\frac{1}{\sqrt{2}}(|0\rangle|0\rangle|y_1\rangle) + |0\rangle|1\rangle|y_k\rangle)) \\ &= \frac{1}{\sqrt{2}}(|0\rangle|0\rangle|y_1\rangle) + |1\rangle|0\rangle|y_k\rangle) \end{aligned}$$

- Rewrite the first qubit in $|\pm\rangle$ basis:

$$|+\rangle(\frac{|0\rangle|y_1\rangle) + |0\rangle|y_k\rangle}{2}) + |-\rangle(\frac{|0\rangle|y_1\rangle) - |0\rangle|y_k\rangle}{2})$$

- Measure the first qubit in $|\pm\rangle$ basis.
- One of the two outcomes $\frac{|0\rangle|y_1\rangle) + |0\rangle|y_k\rangle}{\sqrt{2}}$ or $\frac{|0\rangle|y_1\rangle) - |0\rangle|y_k\rangle}{\sqrt{2}}$ are outputed with probability $\frac{1}{2}$.
- Apply SWAP gate on two registers, leading to the following states:

$$\frac{1}{\sqrt{2}}(|y_1\rangle|0\rangle) + |y_k\rangle|0\rangle) \quad \text{or} \quad \frac{1}{\sqrt{2}}(|y_1\rangle|0\rangle) - |y_k\rangle|0\rangle)$$

Note that, even though the algorithm's outcome is probabilistic, both of the output states have the desired superposition form, albeit with different relative phases[10]. This relative phase, will not affect the success probability of the emulation algorithm. In other words, emulation samples can be obtained deterministically, from the given states. This case is similar to the technique used in [DKK17] where superposition is created with the desired weight but with different phases.

Examples of this section, although simple, showcase the applicability of quantum algorithms, such as quantum emulation, that exploit the power of quantum data (quantum queries) as new attacks on cryptosystems. As shown through these toy examples, many such algorithms, when used in an adversarial scenario, can be adjusted to perform even stronger attacks, compared to the cases made for their original purposes, such as tomography or general learning. Hence the study of this class of algorithms from a cryptanalysis perspective bares outstanding importance in the field.

3.5 A unified framework for quantum unforgeability

After establishing the intuitive relations between unclonability, unforgeability and learning, and introducing the emulation cryptanalysis toolkit, now it is time to formalize those intuitions in the form of a formal framework for unforgeability as a cryptographic property for quantum and classical primitives. The game-based security framework is a standard model for formally defining security properties of cryptographic primitives such as encryption algorithms, digital signature schemes, physical unclonable functions or quantum money [BZ13a, GHS16, AMRS20, AMSY16, SJS16, AC12, Aar09]. Classical cryptographic primitives have also widely been studied in a quantum game-based framework, where parties are quantum (are able to run quantum circuits) [BZ13a, GHS16, AMRS20, SJS16]. Inspired by these works, we generalise the quantum game-based framework to formalize quantum unforgeability, in a way that it is compatible with the notion of learnability and unclonability and is useful for the rest of the thesis. Additionally, our framework unifies different levels of unforgeability as well as capturing quantum and classical primitives. First, we show the abstract and formal version of the definition and then we show how it can naturally cater for quantum primitives and different adversarial levels.

3.5.1 Framework and Formal definitions

Let $\mathcal{F} = (\mathcal{S}, \mathcal{E}, \mathcal{V})$ be a classical or quantum primitive with $\mathcal{S}$, $\mathcal{E}$, and $\mathcal{V}$ being the setup, evaluation, and verification algorithms respectively. We specify unforgeability as a game between a challenger $\mathcal{C}$ (that models the honest parties) and an adversary $\mathcal{A}$ (that captures the corrupted parties). The adversary's goal is

[10] There is a simple way, however to achieve a deterministic outcome. We can add an additional step to correct the output conditioned on the measurement outcome, as it is usually done in MBQC.

to *closely approximate* the output of the evaluation algorithm $\mathcal{E}$ on a *new challenge* such that it passes the verification with high probability. As we work in the quantum regime, we give adversary full quantum query access to the primitive, either classical or quantum. For classical primitives, as we assume the adversary has quantum oracle access to the primitive, we adopt the technique of quantum oracles defined in [BZ13a, BDF+11] for formalizing quantum query-response interaction between the adversary and the challenger.

The security game considered here consists of several phases. First, $\mathcal{C}$ runs the setup algorithm $\mathcal{S}$ to generate the parameters required throughout the game and instantiates the evaluation oracle $\mathcal{O}^{\mathcal{E}}$, the verification oracle $\mathcal{O}^{\mathcal{V}}$, and the message space $\mathcal{M}$. The learning phase defines the threat model. For now, we only consider the quantum equivalent of the chosen-message attack model for coherence and simplicity. Nevertheless, we show the extension to other attack models and types of adversaries, such as weak adversaries, in Section 3.5.3. The challenge phase determines the security notion captured by the game. The formal specification of our quantum games is presented in Game 1. But let us first go informally over each phase of the game and clarify the differences between quantum and classical primitives.

Setup. In the setup phase, $\mathcal{C}$ generates the parameters required in subsequent phases by running the setup algorithm of the primitive $\mathcal{F}$ on input λ (the security parameter), and the oracles are being instantiated accordingly.
Quantum case: For quantum primitives, the evaluation oracle is defined according to Eq. (3.5) for deterministic and randomised primitives respectively and the verification oracle implements a quantum test algorithm as defined in Definition 12.

Learning phase. In the learning phase, the adversary interacts with the evaluation oracle. For now, we only focus on chosen-message attack (cma) security. $\mathcal{A}$ requires the oracle evaluation on any input state ρ_i^{in}. The oracle evaluations are handled by $\mathcal{C}$ who issues the requests on ρ_i^{in} to $\mathcal{O}^{\mathcal{E}}$ and forwards to $\mathcal{A}$ the respectively received outputs ρ_i^{out}, where $i = \{1, \ldots, q = poly(\lambda)\}$. We also note that $\mathcal{A}$ can have an internal register σ and we allow for creating entanglement between $\mathcal{A}$'s register and output queries. Specifically for classical primitives, each $\rho_i^{in} = |\phi_i^{in}\rangle\langle\phi_i^{in}|$ where $|\phi_i^{in}\rangle = \sum_{m_i,y_i} |m_i, y_i\rangle$ is usually a pure state with m_i being the message and y_i the ancillary system. If the queries are being generated by $\mathcal{A}$, in most cases it can be assumed that they have the classical information underlying them, while output queries need to be considered as unknown quantum states to the adversary. In Section 3.5.3 we also represent the model for *weak adversary* which is the quantum equivalent of random-message attack (rma) in the classical world, as well as an alternative way of capturing adaptive adversaries. We also note that this phase for quantum primitives is similar to the classical ones, where $\{\rho_i^{in}\}_{i=1}^q$ represents input chosen message queries and $\{\rho_i^{out}\}_{i=1}^q$ is the respective outputs after the interaction with the oracle sent to $\mathcal{A}$ by the challenger.

Challenge phase. In this phase, the challenge that the adversary has to respond to is chosen in three different ways, each corresponding to a specific level of unforgeability. Similar to classical notions of unforgeability, the strongest notion is *existential unforgeability* denoted by qEx in the game, and whereby the adversary picks the message for which it will produce a forgery. On the other hand, in *selective unforgeability*, denoted qSel, the adversary picks the challenge but needs to commit to it before interacting with the oracle. Hence in Game 1 the selective challenge phase happens before the learning phase. A further way of weakening the unforgeability notion is when the challenge message is chosen by the challenger $\mathcal{C}$ uniformly at random from the set of all the messages. In any case, a classical message $m \in \mathcal{M}$ is selected (for classical primitives) where $\mathcal{M}$ is the set of classical messages.
Quantum case: If the primitive is quantum, the main difference is that $\mathcal{M} = \mathcal{H}^D$ (or $\mathcal{M} = \mathcal{S}(\mathcal{H}^D)$ for density operators) is a Hilbert space and $m = |\psi_m\rangle \in \mathcal{H}^D$ (or equivalently, $m = \rho_m \in \mathcal{S}(\mathcal{H}^D)$) is a quantum challenge in the D-dimensional Hilbert space. In the qUni challenge phase where the message is chosen by the challenger $\mathcal{C}$ uniformly at random from the set of all the messages, for quantum primitives it should be selected uniformly according to the Haar measure over $\mathcal{H}^D$. We also need to mention that for qSel challenge phases, $\mathcal{A}$ is required to submit the (efficient) classical description of the quantum state ρ_m. This is a technicality related to the verification phase, as it allows the challenger to prepare the required number of copies of the correct output for the most general form of verification.

We impose different conditions on the challenge phases which will be formalized later in the guessing phase. These conditions prevent the adversary from mounting trivial attacks.

Guess phase. In this phase, the adversary submits their forgery t for the challenge m. They win the game if the output pair (m, t) passes the verification algorithm with high probability. In addition, for qSel, the message m should be the same as the message submitted in the challenge phase. Here the condition in the challenge phase that we have mentioned is formally checked. The quantum challenge phase needs to be carefully specified to avoid capturing trivial attacks such as sending one of the previously learnt states as the challenge of the adversary. As a result, we have introduced the notation $m \notin_\mu \rho^{in}$ denoting μ-distinguishability from all the input learning phase states. When m is a classical bit-string the same condition should hold for the quantum encoding of m into a computational basis *i.e.* $|m\rangle$ (or $|m, 0\rangle$). Note that the case $\mu = 1$ implies the challenge quantum state has no overlap with any of the quantum states queried in the learning phase.

Important note: We emphasise, that we do not specify how the challenger could check whether the adversary meets the condition or not. Implementing this check is not crucial for our security analysis, where we only need to be able to characterise the instances that might present a security violation. The key point to note is that this can effectively be checked given a run against a given

adversary. Indeed, then ρ_i^{in} and ρ_i^{out} can be characterised by the probability analysis allowing proofs of security or exhibition of attacks.[11]

Regarding the verification oracle, for classical primitives the forgery pair (m, t) is classical and the verification oracle $\mathcal{O}_f^{\mathcal{V}}$ runs the classical verification algorithm $\mathcal{V} = \mathtt{Ver}(k, m, t, r)$. Here r is the randomness if the primitive is randomised.

Quantum case: This phase is similar to the classical case. Here, it can be seen that this is the most natural way of characterising a forgery for quantum primitives since the difference between quantum states is usually measured by their indistinguishability and with quantum distance measures. The main difference in this phase is the difference between the classical and quantum verification procedures. The verification is fairly straightforward for classical primitives since the equality test can be easily performed whereas for quantum primitives, both message and forgery are quantum states, and the verification oracle $\mathcal{O}_U^{\mathcal{V}}$ should call a quantum test algorithm $\mathcal{T}$ that checks the equality of quantum states as in Definition 12. Note that the challenger can prepare copies of correct outputs locally.

[11] This argument is a matter of debate in different areas of cryptography. Some researchers believe imposing any condition within the formal game needs to be done via an efficient and specified process while others, including the author, believe that the conditions only need to specify instances with calculatable probability for the purpose of the proofs. A similar case has been discussed in [CGK+16] (Section VI) regarding the definitions of verifiability in e-voting protocols, where some very widely-accepted definitions such as the one proposed in [KZZ15a, KZZ15b] have verifying subroutines that do not necessarily run in polynomial time. However, we note that imposing such conditions in the definition leads to the fact that extra care is needed when definitions such as this one are used to prove security. Since some reductions may not carry over if the conditions are not being executed, the security proofs can be more complicated. As we will see in what follows, this has been taken into account in our security proofs.

Formal definition of Generalised Quanutm Unforgeability (qGU)

Game 1. Formal definition of the quantum games $\mathcal{G}^{\mathcal{F}}_{q,c,\mu}(\lambda,\mathcal{A})$ where λ is the security parameter, q the number of queries issued to the evaluation oracle in the learning phase, μ the overlap allowed between the challenge and previously queries messages, and c the level of unforgeability. *The game* $\mathcal{G}^{\mathcal{F}}_{q,c,\mu}(\lambda,\mathcal{A})$[a]

Setup phase:

- $\texttt{param} \leftarrow \mathcal{S}(\lambda)$
- The oracles $\mathcal{O}^{\mathcal{E}}$ and $\mathcal{O}^{\mathcal{V}}$ and the message space $\mathcal{M}$ are instantiated given $\texttt{param}$.

Selective challenge phase:

- if $c = \mathsf{qSel}$: $\mathcal{A}$ picks $m \in \mathcal{M}$ and sends it to $\mathcal{C}$.

First learning phase:

- $\mathcal{A}$ (adaptively) issues queries $\rho_1^{in},\ldots,\rho_q^{in}$ (where $q = poly(\lambda)$) to $\mathcal{C}$. To each query ρ_i^{in} the challenger $\mathcal{C}$ queries $\mathcal{O}^{\mathcal{E}}$ on ρ_i^{in}, and forwards the received respective output ρ_i^{out} to $\mathcal{A}$. The adversary can also have an internal register σ which may be entangled with the output queries.

Challenge phase:

- if $c = \mathsf{qEx}$: $\mathcal{A}$ picks $m \in \mathcal{M}$ and sends it to $\mathcal{C}$.
- if $c = \mathsf{qUni}$: $\mathcal{C}$ picks $m \xleftarrow{\$} \mathcal{M}$ uniformly at random and sends m to $\mathcal{A}$

Second learning phase: As the *first learning phase*
Guess phase:

- if $c = \mathsf{qEx}$ OR $c = \mathsf{qSel}$: continue if $m \notin_{\mu} \rho^{in}$[b], else aborts.
- $\mathcal{A}$ generates the forgery t, and outputs to $\mathcal{C}$ the pair: $(m,t) \leftarrow \mathcal{A}(\{\rho_i^{in},\rho_i^{out}\}_{i=1}^{q},\sigma)$
- $\mathcal{C}$ queries the verification oracle: $b \leftarrow \mathcal{O}^{\mathcal{V}}(m,t)$
- $\mathcal{C}$ outputs b

[a] $c \in \{\mathsf{qEx},\mathsf{qSel},\mathsf{qUni}\}$; $0 < \mu \leq 1$.

[b] $\notin_{\mu} \rho^{in}$ denotes at least μ-distinguishability from all the ρ_i^{in}. For the classical message $m \in \{0,1\}^n$, the condition should hold for $|m\rangle$, *i.e.* the quantum encoding of m in computational basis.

We omit the parameter q when we consider arbitrarily polynomially many queries to the evaluation oracle issued by $\mathcal{A}$. We can now formally define *Existential*, *Selective* and *Universal Unforgeability* of primitives as instances of our game as follows.

Definition 30 (μ-qGEU)**.** A cryptographic primitive $\mathcal{F}$ provides μ-quantum existential unforgeability if the probability of any QPT adversary $\mathcal{A}$ of winning the game $\mathcal{G}^{\mathcal{F}}_{\mathsf{qEx},\mu}(\lambda,\mathcal{A})$ is at most negligible in the security parameter,

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qEx},\mu}(\lambda,\mathcal{A})] \leqslant negl(\lambda). \tag{3.23}$$

We also define a stronger security notion for existential unforgeability which considers any overlap μ.

Definition 31 (qGEU)**.** A cryptographic primitive $\mathcal{F}$ provides quantum existential unforgeability if it provides μ-quantum existential unforgeability for all non-negligible μ.

Definition 32 (μ-qGSU)**.** A cryptographic primitive $\mathcal{F}$ provides μ-quantum selective unforgeability if for any q the advantage of any QPT adversary $\mathcal{A}$ of winning the game $\mathcal{G}^{\mathcal{F}}_{q,\mathsf{qSel},\mu}(\lambda,\mathcal{A})$ over $P_{ov}(q,\mu)$ is at most negligible in the security parameter,

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{q,\mathsf{qSel},\mu}(\lambda,\mathcal{A})] \leqslant P_{ov}(q,\mu) + negl(\lambda). \tag{3.24}$$

We call $P_{ov}(q,\mu)$ the "overlap probability" describing the probability for trivial attacks via the overlap allowed by the parameter μ.[a]

[a] Note that by definition $\mathcal{A}$ can always achieve $P_{ov}(q,\mu)$, hence $\mathcal{A}$'s winning probability is always lower-bounded by this value.

The need for allowing an adversary to win with probability $P_{ov}(q,\mu)$ is similar to the classical definitions where the adversary is required to boost the success probability from some trivial value such as a random guess. Here, by allowing the adversary to create an overlap between the learning phase space and challenge, some unavoidable attacks exist which are independent of the actual primitive at hand, and as such needs to be extracted to characterise the gap between trivial and effective adversaries and hence precisely define a proper distance-based definition.

Definition 33 (P_{ov} for classical primitives)**.** For all q and for all μ, for a classical primitive where the evaluation oracle is a standard oracle $\mathcal{O}^{\mathcal{E}}_{f}$, the overlap probability for q-query games is equal to $P_{ov}(q,\mu) = 1-\mu^q$.

The expression $P_{ov}(q,\mu) = 1-\mu^q$ that is chosen in the above definition for the overlap probability for classical primitives, is the probability of a trivial attack performed via simply measuring the superposition queries. A straightforward calculation of this measurement probability for q queries with the same degree of overlap leads to the expression $1-\mu^q$.

A similar notion can be defined for quantum primitives. In this case, it is clear

that the adversary's success probability in finding the output by measurement strategy is almost zero and hence defining the P_{ov} as defined by Definition 33 leads to zero overlap probability. However, in this case, as well, there is another scenario that may lead to trivial attacks, which is due to the error produced by the quantum test algorithm in distinguishing the states with certain overlap. An example of this is the SWAP test which has a one-sided error of $\frac{1}{2}$ even for perfectly distinguishable states. This is a fundamental difference between the quantum world and classical primitives where equality can be checked deterministically. To have a general characterisation of P_{ov} for quantum primitives, this probability needs to be defined with respect to the test algorithm as follows.

Definition 34 (P_{ov} for quantum primitives)**.** Let ρ_{max} be the input learning phase query with the maximum overlap with the challenge state $|\psi\rangle$, allowed by the μ-distinguishability condition. Let the $\mathcal{O}_U^{\mathcal{E}}$ be the unitary oracle for the quantum primitive applying $\mathrm{U}_{\mathcal{E}}$ to the quantum inputs and let $\mathcal{O}^{\mathcal{V}}$ implement a quantum test algorithm $\mathcal{T}$. Then $\rho_{max}^{out} = \mathrm{U}_{\mathcal{E}}\rho_{max}\mathrm{U}_{\mathcal{E}}{}^{\dagger}$ is the output of the query from the oracle and $\rho^{out} = |\psi^{out}\rangle\langle\psi^{out}| = \mathrm{U}_{\mathcal{E}}|\psi\rangle\langle\psi|\mathrm{U}_{\mathcal{E}}{}^{\dagger}$ is the correct output of the challenge $|\psi\rangle$. We define the P_{ov} as the error probability of the test algorithm $\mathcal{T}$ on distinguishing ρ_{max}^{out} and ρ^{out} as follows:

$$P_{ov} = Pr[1 \leftarrow \mathcal{T}((\rho_{max}^{out})^{\otimes\kappa}, (\rho^{out})^{\otimes\kappa})] \tag{3.25}$$

This definition also implies an intuitive and practical approach to determine the desired $\mu < 1$ for quantum primitives, as it states that for any specific quantum primitive or the protocols based on that primitive, the μ should not allow for the above overlap attacks with a probability larger than the required security threshold. Nevertheless, if one assumes a reasonably good quantum test algorithm, this probability for quantum primitives is usually less than the classical ones due to quantum state distinguishability and lack of adversary's knowledge over the transformation of the output bases.

When selective unforgeability holds for any overlap μ we say that the primitive is quantum selective unforgeable.

Definition 35 (qGSU)**.** A cryptographic primitive $\mathcal{F}$ provides quantum selective unforgeability if it provides μ-quantum selective unforgeability for all non-negligible μ.

Now we give yet a weaker definition, namely *Universal Unforgeability*. Note that the μ-distinguishability condition is not necessary for universal unforgeability, as the challenge is chosen by the challenger, independently of the adversary's queries and the probability is taken over all the choices of the challenge state hence it is no longer meaningful to count for possible overlaps as trivial attacks.

Definition 36 (qGUU)**.** A cryptographic primitive $\mathcal{F}$ is quantum universally unforgeable if the probability of any QPT adversary $\mathcal{A}$ of winning the game $\mathcal{G}^{\mathcal{F}}_{\mathsf{qUni}}(\lambda, \mathcal{A})$ is negligible in the security parameter λ,

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qUni}}(\lambda, \mathcal{A})] \leqslant negl(\lambda). \tag{3.26}$$

3.5.2 Hierarchy and relationship to other definitions

In this section, we formally establish the hierarchy between the different levels of Generalised Unforgeability captured by our framework. Furthermore, for completeness, we also investigate how our definitions formally relate to the previously proposed ones for classical primitives. In particular, we show this relationship between 1-qGEU and the definitions of BZ and BU introduced in Section 2.5.6.2 in the preliminaries. In Fig. 3.2, we map out the results presented in this section.

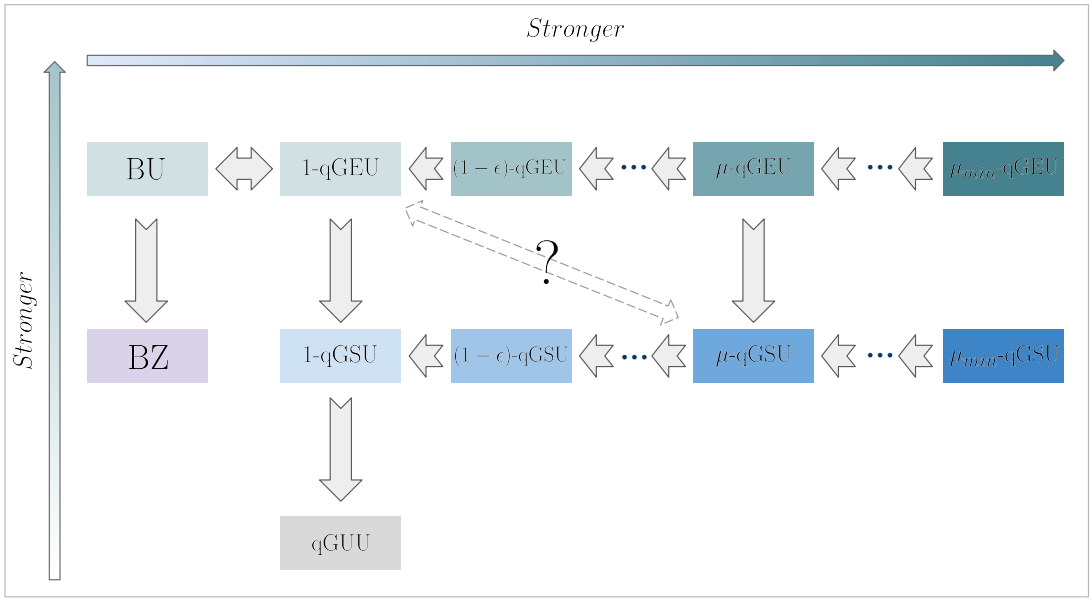

Figure 3.2: Relationship between different definitions of *Generalised Quantum Unforgeability*, BU and BZ. From down to up and left to right the definitions become stronger. $\varepsilon = \varepsilon(\lambda)$ is a negligible function in the security parameter and $\mu_{min} = non\text{-}negl(\lambda)$ is the smallest valid degree for μ. It is unknown whether μ-qGSU with smaller μ, implies μ-qGEU with bigger μ.

First, we establish the relationship between different instances of our game-based definition. We show that as expected for both existential and selective unforgeability, the definitions become stronger when decreasing the μ parameter from 1 and hence μ-qGEU implies 1-qGEU.

Theorem 14. *If $\mu_1 \leqslant \mu_2$ then μ_1-qGEU (resp. μ_1-qGSU) implies μ_2-qGEU (resp. μ_2-qGSU)*

Proof. The proof is straightforward for qGEU. Let $\mathcal{A}$ win against μ_2-qGEU, Let $\mathcal{A}'$ be the adversary who wants to attack μ_1-qGEU. $\mathcal{A}'$ queries the same learning phase queries as $\mathcal{A}$ and then calls $\mathcal{A}$. Since $\mu_1 \leqslant \mu_2$ any two states that are μ_2-distinguishable are also μ_1-distinguishable, then the challenge of $\mathcal{A}$ will necessarily satisfy the condition for μ_1-qGEU. Then $\mathcal{A}'$ can also win the game with non-negligible probability. For μ-qGSU the distinguishability argument is similar, although there is also the P_{ov} probability that is a function of μ. Thus we need to show the following:

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qSel},\mu_2}(\lambda,\mathcal{A})] - P_{ov}(\mu_2) \geqslant Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qSel},\mu_1}(\lambda,\mathcal{A})] - P_{ov}(\mu_1)$$

Which is also equivalent to showing the following statement:

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qSel},\mu_2}(\lambda,\mathcal{A})] - Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qSel},\mu_1}(\lambda,\mathcal{A})] \geqslant P_{ov}(\mu_2) - P_{ov}(\mu_1)$$

The LHS of the inequality is always positive due to the above distinguishability argument, and the P_{ov} is always a non-increasing function of μ for both types of primitives (also the negligible factor is omitted from both sides). Take the P_{ov} for the classical primitives for instance, which is equal to $1-\mu^q$. Therefore, the RHS of the inequality will be equal to $\mu_1^q - \mu_2^q$ which is always a non-positive value as $\mu_1 \leqslant \mu_2$. Then the above inequality holds and the theorem has been proved. □

Furthermore, it is easy to observe that for any given μ, μ-qGEU implies μ-qGSU. This is due to the fact that if the adversary wins the game by committing to their favourite message before the learning phase, they will necessarily win when picking the message after the learning phase.

Universal unforgeability is also intuitively weaker than existential unforgeability similarly to their classical counterpart. The same thing holds despite the winning condition for these two instances being very different. In universal unforgeability, the adversary wins only if they win the game on average over all the randomly picked messages. In our case, we are only interested in QPT adversaries, and as the universal definition is not parameterised by μ, it is not evident whether *qGUU* is weaker than μ-qGSU. The following theorem formally establishes the implication. We prove the theorem for 1-qGSU which, in turn, implies μ-qGSU for any μ.

Theorem 15. *μ-qGSU implies qGUU.*

Proof Sketch. The full proof can be found in the Appendix A.2. Here we present the key ideas of the proof. We show if there exists an adversary $\mathcal{A}$ that wins the qGUU game then 1-qGEU (1-qGSU) also breaks and the implication to μ-qGEU (μ-qGSU) is straightforward. First, we show that the distinguishability condition for $\mu = 1$ can be satisfied. Thus we write the winning probability of $\mathcal{A}$ as the combination of probabilities of winning for the selected message being orthogonal

to the learning phase or not:

$$\begin{aligned} \Pr_{x\in\mathcal{M}}[1 \leftarrow \mathcal{A}(x)] &= \Pr_{x\in\mathcal{M}'}[1 \leftarrow \mathcal{A}(x)]Pr[x \in \mathcal{M}'] + \Pr_{x\notin\mathcal{M}'}[1 \leftarrow \mathcal{A}(x)]Pr[x \notin \mathcal{M}'] \\ &= non\text{-}negl(\lambda) \end{aligned} \tag{3.27}$$

where $\mathcal{M}'$ is the set of all the challenges with no overlap with the learning-phase states. By calculating this probability we show that $\Pr_{x\in\mathcal{M}'}[1 \leftarrow \mathcal{A}(x)]$ is also non-negligible. In the second part of the proof we show that as long as the previous average probability holds, we can always construct an efficient adversary $\mathcal{A}'$ that uses $\mathcal{A}$ to win the selective unforgeability game. We prove this by partitioning the space of $\mathcal{M}'$ into equal polynomial-size subspaces and show that if the average probability over $\mathcal{M}'$ is non-negligible, then $\mathcal{A}'$ can always win the 1-qGEU game by randomly picking one of the subsets to pick the message from, as there will exist at least one message that allows $\mathcal{A}$ to win the game with non-negligible probability. As a result, $\mathcal{A}'$ wins the game with non-negligible probability. □

Now, we show an equivalence between an instance of our existential unforgeability definition and BU. Since this result is mainly for the sake of completeness and will not be directly related to the rest of the thesis, again we give a proof sketch here, and we give the full proof in the Appendix A.3.

Theorem 16. *1-qGEU is equivalent to BU.*

Proof Sketch. We show that 1-qGEU implies BU and vice versa. First, we show that if a scheme is not BU unforgeable against a QPT adversary then it is not 1-qGEU unforgeable either. Let $\mathcal{A}$ be a QPT adversary who forges a scheme $\mathcal{F} = (\mathcal{S}, \mathcal{E}, \mathcal{V})$ with message set $\mathcal{M} = \{0,1\}^n$ in the BU definition. Following the definition of BU, if $\mathcal{A}$ can win against the BU game, there exists a non-empty set $\mathcal{B}_\varepsilon$ for which $\mathcal{A}$ interacts with the blinded oracle associated with it and outputs a pair (m^*, t^*) where $t^* = f(m^*)$ (where f is the classical function of the evaluation $\mathcal{E}$) such that $\mathcal{V} = Ver_k(m^*, t^*) = acc$, and also the $m^* \in \mathcal{B}_\varepsilon$ with non-negligible probability in $\lambda = poly(n)$. By rewriting a general query state to the blinding oracle in orthogonal and non-orthogonal sub-spaces to the main forgery state, we can show that there exists a unitary non-blinding oracle that generates equivalent queries for this scenario. We then show that this new unitary oracle can be queried equivalently by an adversary who satisfies all the conditions of 1-qGEUand therefore can also generate a forgery that passes the test algorithm with also non-negligible probability. Hence we have shown that 1-qGEU implies BU.

For the other way of implication, we show that a QPT adversary $\mathcal{A}$ who wins the 1-qGEU, has also a non-empty support, $supp(\mathcal{A}) \cap R = \emptyset$, for some $R \neq \emptyset$, and can output a valid pair $(m^*, f(m^*))$ with $m^* \in R$ with non-negligible probability. Intuitively, this is due to the orthogonality condition that is required to be satisfied in the 1-qGEU game between the learning subspace and the forgery

state. According to the Theorem 10 in Chapter 2, this implies that the primitives against such an adversary are not BU-secure. This concludes the proof. □

From the above theorem and the equivalence of BU and BZ against classical adversaries, we derive the following corollary.

Corollary 1. *1-qGEU ≡ BU ≡ BZ against classical adversaries.*

3.5.3 Unforgeability against weak vs adaptive adversaries

Another variant of the unforgeability definition appears when we weaken the adversary in choosing the queries in the learning phase freely and adaptively. One way of imposing such limitations on the adversary is to assume that the adversary has no direct oracle access and instead has only access to a random set of queries (random input-output samples from the oracle) being selected at random from a specific distribution by the honest party. This attack model is commonly called the *random message attack* model in classical cryptography. Moreover, in the practical sense, this type of limited adversary represents *network adversarial model*, *i.e.*, when the adversary has only access to the communication channel. Network adversaries can get the input and output samples of a primitive, only by intercepting the messages that are exchanged between the two honest parties during a protocol.

The adversary we have considered so far, which we refer to as *adaptive adversary*, can query the evaluation function of the primitive through oracle access adaptively, in the sense that not only the queries have been chosen by the adversary, but they can depend on the previously received responses in general. We also introduce an alternative way of capturing full adaptive quantum adversaries in the Appendix A.4. On the other hand, a weak non-adaptive adversary, cannot choose the queries and instead receives q queries (where $q = poly(\lambda)$) of $\mathcal{E}$, in the form of input and output pair. Most commonly, the queries are being picked at random from a uniform distribution by an honest party, but more generally, one can consider the case where queries are being selected from a distribution $\delta_{\mathcal{D}}$ over the message space.

In what follows, we restate a new instance of our Game 1, which captures weak adversaries as well as adaptive ones. We also note that weak adversaries are mostly of interest regarding universal unforgeability, thus we only restate the game for universal unforgeability.

Universal Unforgeability with weak and adaptive adversaries

Game 2. Formal definition of the quantum games $\mathcal{G}_q^{\mathcal{F}}(\lambda, \mathcal{A})$ where λ is the security parameter, q the number of queries issued to the evaluation oracle in the learning phase.

Setup phase:

- same as Game 1

Learning phase:

- If the adversary is adaptive, $\mathcal{A} = \mathcal{A}_{ad}$ (same as Game 1):
  - $\mathcal{A}_{ad}$ selects any desired query $c_i \in \mathcal{M}$, and issues to $\mathcal{C}$ (up to q queries).
  - $\mathcal{C}$ queries $\mathcal{O}^{\mathcal{E}}$ on c_i, and sends the respective output r_i back to $\mathcal{A}_{ad}$.
- If the adversary is weak (non-adaptive), $\mathcal{A} = \mathcal{A}_{weak}$:
  - $\mathcal{C}$ selects a challenge $c_i \in \mathcal{M}$ uniformly at random from $\mathcal{M}$ and independent of i.
  - $\mathcal{C}$ queries the $\mathcal{O}^{\mathcal{E}}$ on c_i and produces the response $r_i = \mathcal{E}(c_i)$.
  - $\mathcal{C}$ issues to $\mathcal{A}_{weak}$ the set of random challenges and their respective responses $\{(c_i, r_i)\}_{i=1}^{q}$.

Challenge phase:

- same as Game 1 for universal challenge phase (qUni)

Guess phase:

- same as Game 1

Note that in the above game all the queries c_i and the responses r_i have been abstracted for simplicity, but can capture both quantum and classical queries. We will widely use this variant on the unforgeability game later in Chapter 6, Section 6.4.

Let us define universal unforgeability with weak adversaries as follows:

Definition 37 (Universal Unforgeability against weak Adversary)**.** A cryptographic primitive $\mathcal{F}$ is quantum universally unforgeable against a 'weak (non-adaptive) adversary' if the success probability of any weak QPT adversary $\mathcal{A}_{weak}$ in winning the game $\mathcal{G}^{\mathcal{F}}(\mathcal{A}_{weak}, \lambda)$ is at most a negligible function, $\varepsilon(\lambda)$, in the security parameter.

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}(\mathcal{A}_{weak}, \lambda)] \leqslant \varepsilon(\lambda) \tag{3.28}$$

3.5.3.1 A note on weak vs strong unforgeability

We also note that there is another way of characterising the strength of the unforgeability definition in the literature. We have formally defined our different instances of unforgeability as a quantum analogue of *weak unforgeability*. However, the same definition albeit with a small modification can be applied to capture *strong unforgeability*. First, we note that the difference between strong and weak unforgeability is only relevant to randomised primitives. For non-randomised primitives, these definitions are equivalent. In the classical world, for strong unforgeability, it is sufficient for the adversary to output a new pair to win the game and hence the adversary is allowed to pick one of the learning phase messages as the challenge and produce a new output with fresh randomness. In our definition, it is sufficient to expand the μ-distinguishability condition to the overall input of the oracle including the randomness, *i.e.* adversary's challenge state $|r^*\rangle\langle r^*| \otimes \rho_m$ needs to be μ-distinguishable from all the learning phase states with their randomness registers which can be written as $|r_i\rangle\langle r_i| \otimes \rho_i^{in}$. Once again for $\mu = 1$, this will capture the same definition as is expected.

3.6 Applications of qGU: possibility and impossibility results

In this section, we study the unforgeability of general classical and quantum primitives under the lens of our generalised unforgeability framework. We start with the strongest level of unforgeability in our framework, *i.e.* existential unforgeability, and we try to give examples for both classical and quantum primitives, all the way to the weakest unforgeability notion. We will see how this framework allows us to establish general possibility and impossibility results on different levels. It will also help us design non-trivial cryptographic primitives that satisfy a high level of quantum unforgeability against any quantum adversaries.

3.6.1 Generalised existentially unforgeable schemes

In this section, we turn our attention to 1-qGEU. First, we show a general and intuitive, yet important no-go result for μ-qGEU that is, no classical primitive (deterministic or randomized) can satisfy this level of unforgeability for any $\mu \neq 1$. This result states that 1-qGEU, (which is also equivalent to BU according to Theorem 16), is the strongest notion of existential unforgeability that any classical primitive can possibly achieve.

Theorem 17 (No classical primitive $\mathcal{F}$ is μ-qGEU secure)**.** *For any classical primitive $\mathcal{F}$ and for any μ such that $\mu \leqslant 1 - \frac{1}{D}$, where D is the dimension of the Hilbert space on which the evaluation oracle operates, there exists a QPT adversary $\mathcal{A}$ such that*

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qEx},\mu}(\lambda, \mathcal{A})] = non\text{-}negl(\lambda). \tag{3.29}$$

Proof. There exists a simple superposition attack that breaks μ-qGEU. Let $\mathcal{A}$ issue only one query which is the uniform superposition of all the inputs, which leads to an output of the form $\frac{1}{\sqrt{2^n}}\sum_m |r\rangle_{\mathcal{O}} |m, f(m;r)\rangle$, where we have taken $D = 2^n$. Then by measuring the first part of the register in the computational basis, the state will collapse to one of the basis and the adversary is able to produce a valid message-tag pair for a classical message with a negligible overlap with the learning phase. Hence $\mathcal{A}$ can always win the game for any $\mu \leqslant 1 - \frac{1}{2^n}$. □

Nevertheless, it is still possible to have schemes that are 1-qGEU secure through the following positive result:

Theorem 18. *qPRFs are 1-qGEU (1-qGSU) unforgeable.*

Proof. This is a straightforward result via equivalence of 1-qGEU to BU and a corollary from [AMRS20], where it is shown that *qPRF*s are *BU* secure. □

Although the general no-go result for classical primitives does not directly apply to quantum primitives, in Chapter 4, we show that most quantum primitives that we are interested in do not satisfy this definition either. However, it will not be surprising given the general no-go result we provide for selective unforgeability.

Another interesting positive result that we can demonstrate, is the peer of Theorem 18 for quantum primitives. We show that pseudorandom unitaries (PRU), which are the quantum counterpart of pseudorandom functions in the quantum world, can also satisfy 1-qGEU (and 1-qGSU).

Theorem 19. *PRU quantum primitives are 1-qGSU (1-qGEU) secure.*

Proof. We prove this by contradiction. Let $\mathcal{A}$ be an adversary who wins the 1-*qGSU* game with non-negligible probability (Note that here $P_{ov} = 0$). $\mathcal{A}$ selects a message m before (or after) the learning phase and then outputs the respective t such that it passes the verification test with non-negligible probability. Also by definition of 1-qGSU, $m \notin_{\mu} \rho^{in}$ for $\mu = 1$ and hence the message ρ_m is completely orthogonal to all ρ_i^{in}. Now we construct an adversary $\mathcal{A}'$ who is playing the PRU game. Let $\mathcal{A}'$ first query all the learning phase states of $\mathcal{A}$ and then also issue one more query which is ρ_m. Then $\mathcal{A}'$ calls $\mathcal{A}$ and receives the input-output pair of (m, t) such that ρ_t is non-negligible close to the actual output, *i.e.*

$$F(\rho_t, \mathsf{U}_{\mathcal{E}} \rho_m \mathsf{U}_{\mathcal{E}}{}^{\dagger}) = non\text{-}negl(\lambda) \tag{3.30}$$

Now $\mathcal{A}'$ can use this last query as a distinguisher between PRU and a unitary picked from the Haar measure since $\mathcal{A}'$ can estimate the output with non-negligible fidelity if the U_k had been picked from the family. Let $\mathcal{A}'$ runs a quantum equality test as described in Definition 12 on the $U_k|\psi\rangle$ obtained in the learning phase and ρ_t. Also note that if U is picked from the Haar measure family, the probability of producing the output is negligible by definition. Thus whenever the test shows equality, $\mathcal{A}'$ can conclude that the unitary has been picked from PRU. Thus for $\mathcal{A}'$, we have:

$$\Pr_{U\leftarrow U_k}[\mathcal{A}'^{U}(1^\lambda)=1]-\Pr_{U_\mu\leftarrow\mu}[\mathcal{A}'^{U_\mu}(1^\lambda)=1]=\textit{non-negl}(\lambda) \tag{3.31}$$

Which is a contradiction and the proof is complete. □

3.6.2 Generalised selectively unforgeable schemes

In this section, we establish results for μ-qGSU which restricts the adversary in two ways. First, by requiring the adversary to commit to the challenge before the learning phase, we prevent the adversary from picking any post-measurement state as their forgery challenge. Second, by subtracting the probability of any potential trivial attack, especially for classical primitives, from the winning probability of the game, we make the probability bounds tighter for the adversary. We also discuss why defining unforgeability in such a way leads to non-trivial results and establishes a separation between randomised and non-randomised constructions, therefore motivating the usefulness of the given definition.

3.6.2.1 Non-randomised schemes

Let us start with non-randomised schemes. To establish our result, we now take advantage of our proposed cryptanalysis toolkit, namely the *quantum emulation attack* (QEA), which we introduced earlier on in Section 3.4. Here we only show this no-go result for classical non-randomised primitives to avoid repetitions, but the same result holds for quantum constructions.

Theorem 20 (No classical (or quantum) non-randomised primitive $\mathcal{F}$ is μ-qGSU secure)**.** *For any classical/quantum deterministic primitive $\mathcal{F}$ and for any μ, in the range $\frac{1}{4}+\textit{non-negl}(\lambda)\leqslant\mu\leqslant 1-\textit{non-negl}(\lambda)$, there exists an effective QPT adversary $\mathcal{A}$ such that*

$$Pr[1\leftarrow\mathcal{G}^{\mathcal{F}}_{q(\lambda),\mathsf{qSel},\mu}(\lambda,\mathcal{A})]-P_{ov}(q(\lambda),\mu)=\textit{non-negl}(\lambda). \tag{3.32}$$

Proof. We show the proof for classical primitives but the same attack and results also holds for quantum primitives. We show that there exists a QPT adversary $\mathcal{A}$ who can win the game with non-negligible probability for any μ except when it is negligibly close to 0 or 1. The attack is the one-block emulation attack from Section 3.4.2.1 with the following setting. First $\mathcal{A}$ picks any two messages

$m, m' \in \mathcal{M}$ and sets m as the challenge. Then $\mathcal{A}$ queries the states $|\phi_1\rangle = |m', 0\rangle$ and $|\phi_r\rangle = \sqrt{1-\gamma^2}\,|m', 0\rangle + \gamma\,|m, 0\rangle$ by interacting with $\mathcal{O}_f^{\mathcal{E}}$, where γ is a real value such that $0 \leqslant \gamma \leqslant \sqrt{1-\mu}$ and such that the distinguishability condition of the μ-qGSU game is satisfied. After the learning phase, $\mathcal{A}$'s output state can be written as $\sigma_{out} = |\phi_1^{out}\rangle \otimes |\phi_r^{out}\rangle$ where $|\phi_1^{out}\rangle = \mathrm{U}_{\mathcal{E}}\,|\phi_1\rangle$ and $|\phi_r^{out}\rangle = \mathrm{U}_{\mathcal{E}}\,|\phi_r\rangle$. Followed by the fidelity analysis given in Section 3.4.2.1, we show that the success probability of $\mathcal{A}$ in producing the output of m *i.e.* $f(m)$ is bounded by $\gamma^2(1+4(1-\gamma^2)^2)$. This is because we have: $\langle\phi_1|\psi\rangle = 0$ and $|\langle\psi|\phi_r\rangle|^2 = \gamma^2$ and $|\langle\phi_1|\phi_r\rangle|^2 = 1-\gamma^2$, which gives us the following bound on the fidelity:

$$F(|\omega\rangle\langle\omega|, \mathrm{U}_{\mathcal{E}}^{\dagger}\,|\psi\rangle\langle\psi|\,\mathrm{U}_{\mathcal{E}}) \geqslant \gamma^2(1+4(1-\gamma^2)^2) \tag{3.33}$$

In general, γ^2 which is the overlap between the challenge state and the learning phase state can be as large as $1-\mu$ allowed by the definition, thus we set the maximum allowed value of overlap which is $\gamma = \gamma_{max} = \sqrt{1-\mu}$. Now we need to also determine P_{ov} and to show whether the adversary can boost the success probability by a non-negligible value. Here one of the queries is orthogonal to the challenge and there is only one query ($|\phi_r\rangle$) with overlap, thus according to Theorem 33 we have $P_{ov}(2,\mu) = 1-\mu^2$. As a result

$$\begin{aligned} Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qSel},\mu}(\lambda, \mathcal{A})] - P_{ov}(2,\mu) &= (1-\mu)[1+4(1-(1-\mu))^2] - (1-\mu^2) \\ &= \mu(1-\mu)(4\mu-1) \end{aligned} \tag{3.34}$$

Since $\frac{1}{4} + \textit{non-negl}(\lambda) \leqslant \mu \leqslant 1 - \textit{non-negl}(\lambda)$, then all the terms are non-negligible in the security parameter and this concludes the proof. □

The above theorem has a direct consequence which we represent as the following corollary:

Corollary 2. *No deterministic classical or quantum primitive $\mathcal{F}$ is qGSU (Definition 35) secure.*

Let us now discuss the intuitive meaning of it. First, we note that despite the above no-go theorem, qPRFs still provide 1-qGSU security (Theorem 18). However, this no-go result shows a fundamental vulnerability of any non-randomised classical primitive against forgeries, since the only way to ensure the security of primitives against such effective attacks is to guarantee that the adversary's forgery message is orthogonal to their learning subspace. Practically this guarantee can only be given by relying on the device implementation, which is arguably in contradiction with the whole motivation of obtaining security against more powerful quantum adversaries, to begin with [BZ13a]. Let us consider a non-randomised MAC scheme such as HMAC and NMAC. According to Theorem 20, these schemes do not satisfy existential nor selective unforgeability except for $\mu = 1$ and hence are always vulnerable against more powerful quantum adversaries

implementing superposition attacks. Nevertheless, one might argue that such definitions might be too strong, and the proposed attack might not demonstrate an intuitive forgery. To better demonstrate this potential vulnerability, let us show a slightly different example from what is used in the proof of the above theorem to argue there are instances of the game and attacks that can demonstrate an intuitive forgery situation.

Example 1. *Let $\mathcal{A}$'s state after the learning phase be $\sigma_{in} = |\phi_1^{in}\rangle \otimes |\phi_r^{in}\rangle^{\otimes 2}$ and $\sigma_{out} = |\phi_1^{out}\rangle \otimes |\phi_r^{out}\rangle^{\otimes 2}$ where the query states have been chosen as follows:*

$$|\phi_1\rangle = |m_1, 0\rangle \quad |\phi_r\rangle = \delta\,|m_1, 0\rangle + \gamma\,|m_2, 0\rangle + \gamma\,|m_3, 0\rangle \tag{3.35}$$

Where due to normalisation $|\delta|^2 + 2|\gamma|^2 = 1$, although we pick the $\delta = \sqrt{1-2\gamma^2}$ and γ to be real values for simplicity, thus $\gamma^2 \leqslant \frac{1}{2}$. Also note that $\mathcal{A}$ has two identical copies of $|\phi_r^{out}\rangle$. The attack consists of running two separate emulations for $|m_2, 0\rangle$ and $|m_3, 0\rangle$.

Let $|\phi_r\rangle$ be the reference state for the emulation, and the target state to be $|\psi\rangle = |m_2, 0\rangle$ or $|\psi\rangle = |m_3, 0\rangle$. Note that as $|\phi_1\rangle = |m_1, 0\rangle$ is orthogonal to both states and the reference state is symmetric with respect to them, the emulation's fidelity will be the same for both these states. Relying on Theorem 13, the output state of the QE algorithm with only one block will be:

$$\begin{aligned} |\chi_f\rangle = {} & \langle\phi_r|\psi\rangle\,|\phi_r\rangle\,|0\rangle + |\psi\rangle\,|1\rangle - \langle\phi_r|\psi\rangle\,|\phi_r\rangle\,|1\rangle - 2\,\langle\phi_1|\psi\rangle\,|\phi_1\rangle\,|1\rangle \\ & + 2\,\langle\phi_r|\psi\rangle\,\langle\phi_r|\phi_1\rangle\,|\phi_1\rangle\,|1\rangle\,. \end{aligned} \tag{3.36}$$

Note that $|\langle\phi_1|\psi\rangle| = 0$ and $|\langle\psi|\phi_r\rangle|^2 = \gamma^2$ and $|\langle\phi_1|\phi_r\rangle|^2 = 1-2\gamma^2$. Then according to Theorem 11, the fidelity of the emulation for both states is:

$$F(|\omega\rangle\langle\omega|, \mathsf{U}_{\mathcal{E}}\,|\psi\rangle\langle\psi|\,\mathsf{U}_{\mathcal{E}}{}^{\dagger}) \geqslant \gamma^2(1 + 4(1-2\gamma^2)^2) \tag{3.37}$$

Now we need to compare this probability with the P_{ov} probability which is $P_{ov}(3,\mu) = 1-\mu^3$ since the size of the learning phase includes 3 queries. We write the effective success probability of the adversary as:

$$\begin{aligned} Pr_{forge}[\mathcal{A}(m_2)] = Pr_{forge}[\mathcal{A}(m_3)] &= Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{3,\mathsf{qSel},\mu}(\lambda, \mathcal{A})] - P_{ov}(3,\mu) \\ &= \gamma^4(1 + 4(1-2\gamma^2)^2)^2 - (1-\mu^3) \end{aligned} \tag{3.38}$$

Finally, we do a functional analysis of the above probability to see in which cases it becomes non-negligible. First, we note that the success probability of the emulation attack is not greater than the trivial success probability for all the values of μ which shows that if we allow for too much overlap, the trivial attack already has a very high probability which is higher than the emulation's fidelity in this case. Next, since the highest allowed overlap is achieved when $1-\mu = \gamma^2$, we substitute the variable μ with $1-\gamma^2$ to find the degrees of μ for which an effective adversary exists. Hence we rewrite the winning probability of the Eq. (3.38) as follows:

$$\begin{aligned} Pr_{forge}[\mathcal{A}(m_2 \vee m_3)] &= \gamma^2(1 + 4(1-2\gamma^2)^2) - (1 - (1-\gamma^2)^3)) \\ &= \gamma^2(2 - 5\gamma^2 + 3\gamma^4) \end{aligned} \tag{3.39}$$

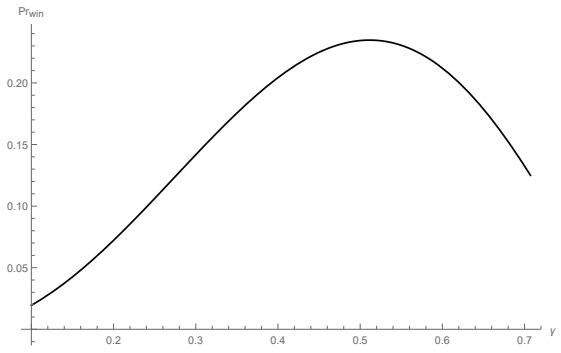

Figure 3.3: The winning probability of $\mathcal{A}$ to forge classical messages $\{m_2, m_3\}$ with the emulation attack. γ represents the overlap between the learning phase query and the target message.

Noting that the valid range for γ is $0 \leqslant \gamma \leqslant \frac{\sqrt{2}}{2}$, we plot the above function as it is shown in Fig. 3.3 and we can see that there is exist a valid range for μ such that the above forgery attack happens with non-negligible probability.

But more importantly, now having access to two copies of the reference state, the adversary can actually run the emulation attack twice, and produce the outputs of both m_2 and m_3 at the same time, with non-negligible probability. Thus for these values of μ, we have presented an adversary who can produce effective forgery for three classical messages m_1, m_2 and m_3 (Note that the first learning phase query is $|m_1, 0\rangle$ which is basically a classical query and as a result, $\mathcal{A}$ will always have the output for m_1) from a classical query, and two copies of the same quantum state which shows an intuitive forgery, especially that the presented attack is independent of the size of the messages and the dimensionality of the Hilbert space of the oracle. This sort of attack cannot be captured in the definitions of unforgeability that count the queries, such as BZ. Nevertheless, our approach to defining the notion of unforgeability is capable of showing such vulnerabilities against strong quantum adversaries.

3.6.2.2 Randomised schemes:

We have seen so far that by letting quantum adversaries exploit the power of superposition queries, they can mount effective attacks to break selective unforgeability in almost all the cases (most valid ranges of μ). A relevant question here would be whether there exists any scheme that can satisfy this quite strong level of unforgeability. Since it is an impossibility for non-randomised primitives, the only possible road ahead would be to employ randomisation of the primitive. In this section, we explore how to defend against general superposition adversaries. We show that this task is possible via randomisation. Concretely, we present randomised constructions for both classical and quantum cases, which satisfy qGSU, *i.e.* μ-qGSU for any μ. The key ingredient that allows this construction to be secure is that the randomisation has been used in an effective way such that the adversary is prevented from creating a known subspace for a specific unitary, even though they can query the challenge message in superposition. First, we start with classical primitives.

Let us first define the desired characteristic for the family of the classical functions used in our construction.

Definition 38 (Inter-function independent family:)**.** Let $F_k : \mathcal{K} \times \mathcal{X} \to \mathcal{Y}$ be a keyed family of functions with domain $\mathcal{X}$ and range $\mathcal{Y}$, where $\mathcal{X} = \{0,1\}^n$ and $\mathcal{Y} = \{0,1\}^m$. We say F_k is an *inter-function (pairwise) independent family* if for any efficient PPT adversary $\mathcal{A}$ and any two functions $F(k,.)$ and $F(k',.)$ picked uniformly at random from F_k, the probability of $\mathcal{A}$ finding an $x \in \mathcal{X}$ such that $F(k,x) = F(k',x)$, is negligible in the security parameter, *i.e.* the following condition should hold:

$$\Pr_{k,k' \leftarrow \mathcal{K}}[x \leftarrow \mathcal{A}(1^\lambda) \wedge F(k,x) = F(k',x)] = negl(\lambda) \tag{3.40}$$

The next step, is to show that a PRF family satisfies the above condition.

Lemma 2. *A PRF is an inter-function independent family.*

Proof. We want to show that any two randomly selected functions from a PRF family, satisfy the required pairwise-independency property of Definition 38. Let $F_k : \mathcal{K} \times \mathcal{X} \to \mathcal{Y}$ be a PRF family of functions where $|\mathcal{X}| = 2^n$ and $|\mathcal{Y}| = 2^m$. We want to show that there is no efficient adversary that can find an x such that $F(k,x) = F(k',x)$ for any two different, randomly picked keys k, k'. We prove by contradiction. We assume that F_k is a PRF but there exist an efficient adversary $\mathcal{A}$ that can find at least one $x \in \mathcal{X}$ such that for any two randomly picked functions from F_k we have:

$$\Pr_{k,k' \leftarrow \mathcal{K}}[x \leftarrow \mathcal{A}(1^\lambda) \wedge F(k,x) = F(k',x)] = non\text{-}negl(\lambda). \tag{3.41}$$

Now we construct a new family of functions from F_k which is a PRF. Let $F'_{k,k'} : \mathcal{K}^2 \times \mathcal{X} \to \mathcal{Y}$ be constructed as follows:

$$F'((k,k'),x) = F(k,x) \oplus F(k',x) \tag{3.42}$$

It is a well-known example in the literature that if F_k is a PRF, then $F'_{k,k'}$ is also a PRF. Now we show that if the Eq. (3.41) holds, then there also exist an adversary who can distinguish $F'((k,k'),x)$ form truly random function. Let $\mathcal{A}'$ query the same x' that has been found by $\mathcal{A}$. If $\mathcal{A}'$ queries $F'((k,k'),x)$, since $F(k,x') = F(k',x')$ with non-negligible probability, then the queries to $F'((k,k'),x)$ on x' should return 0^m. On the other hand the queries to the truly random function will return random bit-strings. As a results, $\mathcal{A}'$ can distinguish $F'((k,k'),x)$ from a truly random function which is a contradiction and hence we have proved that PRF satisfies the Definition 38. □

We can now give our construction based on PRFs or more generally, based on any family of classical functions satisfying the Definition 38.

Construction 1. Let $F : \mathcal{K} \times \mathcal{X} \to \mathcal{Y}$ be a PRF (or any other family satisfying Definition 38). Let $\mathcal{R} = \mathcal{K} = \{0,1\}^l$ be the randomness space. And let λ be the security parameter and l be polynomial in λ. The construction is defined by the following key generation algorithm, keyed evaluation algorithm, and keyed verification algorithm:

- **Key generation:** The secret key is picked uniformly at random from $\mathcal{K}$: $k \xleftarrow{\$} \mathcal{K}$
- **Evaluation:** The evaluation under key k on input m picks randomness r and applies $F(k \oplus r, \cdot)$ to m. Note that when responding to a quantum query, the same randomness is used for all the states of the superposition.
  - On input $m \in \mathcal{X}$:
  - $r \xleftarrow{\$} \mathcal{R}$
  - Return $F(k \oplus r, m) || r$
- **Verification:** The verification under key k of a pair $(m, (t, r))$, runs the evaluation algorithm on m under k with randomness r, and checks equality with t.
  - On input $(m, (t, r)) \in \mathcal{X} \times (\mathcal{Y} \times \mathcal{R})$:
  - If $F(k \oplus r, m) = t$ return $\top$, otherwise return $\bot$

We show that this construction satisfies μ-qGSU security.

Theorem 21. *Construction 1 is qGSU secure.*

Proof. We prove by contraposition. Let us assume there exists a QPT adversary $\mathcal{A}$ who plays the μ-qGSU game where the evaluation is according to Construction 1 and wins with non-negligible probability in the security parameter *i.e.* $\mathcal{A}$ wins the game by producing a valid tag t^* for their selected message m^* and randomness r^* with the following probability:

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{q,\mathsf{qSel},\mu}(\lambda, \mathcal{A})] - P_{ov}(q_r, \mu) = non\text{-}negl(\lambda) \tag{3.43}$$

Where the verification algorithm checks if $F(k \oplus r^*, m^*) = t^*$. We introduce the following games:

- `Game 0`. This game is the μ-qGSU for Construction 1, where $F(k \oplus r, .)$ is picked from F.
- `Game 1`. This game is similar to `Game 0`, except that $\mathcal{A}$ needs to produce forgery for a r^* which is one of the previously received random values of $\{r_i\}_{i=1}^{q}$ in the learning phase.

First, it is straightforward that the probability of the adversary winning μ-qGSU in `Game 0`, is at most negligibly higher than winning `Game 1`. Since r_i in both cases have been picked independently and uniformly at random and the probability of producing a forgery for a specific function with no query is negligible. Thus `Game 0` and `Game 1` are indistinguishable.

Now we recall the quantum oracle for this randomised construction. Let $R\mathcal{O}_c^{\mathcal{E}}$ be the random oracle for both games:

$$R\mathcal{O}_c^{\mathcal{E}} : \sum_{m,y} \alpha_{m,y} |r\rangle_{\mathcal{O}} |m, y\rangle \rightarrow \sum_{m,y} \alpha_{m,y} |r\rangle_{\mathcal{O}} |m, y \oplus (F(k \oplus r, m)||r)\rangle \tag{3.44}$$

Note that in each query a new function has been picked from F, but it is the same for all the messages in the superposition for that query.

Now we use the inter-function (pairwise) independent property of the family F. The construction requires the F to be a PRF family which is inter-function independent according to Definition 38, for two randomly selected keys. Now we need to also show that $F(k \oplus r, .)$ is a PRF as well, with a key k and any randomly selected randomness r, and as a result, we can use the inter-function independent property. This is clearly the case as the key k and any randomness r have been picked independently at random and if there exists a non-negligible advantage for the adversary to distinguish a $F(k \oplus r, .)$ from a truly random function for a value of r, there also exists an equivalent non-negligible advantage to distinguish a $F(k', .)$ where $k' = k \oplus r$ is a key selected uniformly at random. This is still the case even if the value r becomes public after the experiment. This is in contrast with the assumption that the family is PRF, hence we conclude that $F(k \oplus r, .)$ is a PRF. Now we can rely on the Lemma 2 that $F(k \oplus r, .)$ also satisfies the inter-function independent property and the following holds for each of the two functions drawn in any of the two queries:

$$\Pr_{i,j(i \neq j)}[x \leftarrow \mathcal{A}(1^{\lambda}) \wedge F(k \oplus r_i, x) = F(k \oplus r_j, x)] = negl(\lambda) \tag{3.45}$$

As a result, we show that the adversary can at most span a one-dimensional subspace of each $U_{k \oplus r}$. To show this we will calculate the probability of $\mathcal{A}$ in spanning at least a 2-dimensional common subspace from two different queries. This means that $\mathcal{A}$ needs to find at least two bases mapping to the same 2-dimensional subspace in the output Hilbert space. Moreover, we exclude that part of $\mathcal{A}$'s register that contains the classical value of the randomness to only capture the Hilbert space of each $U_{k \oplus r}$. Thus let the input bases be denoted by $|b\rangle = |m, z\rangle$ where z is a subset of y excluding the space for the randomness, for a specific m. Let $|e_i\rangle = U_{k \oplus r_i} |b\rangle = |z \oplus F(k \oplus r_i, m)\rangle$ and $|e_j\rangle = U_{k \oplus r_i} |b\rangle = |z \oplus F(k \oplus r_j, m)\rangle$ be the output states from two different queries. For these output bases to have some overlap, the two functions $F(k \oplus r_i, .)$ and $F(k \oplus r_j, .)$ need to return the same classical output with high probability. Although from Eq. (3.45), we have that the probability of finding such inputs that leads to a common basis is negligible:

$$\begin{aligned}
&\underset{i,j(i\neq j)}{Pr}[\{|e_i\rangle, |e_j\rangle\} \leftarrow \mathcal{A}(1^\lambda) \wedge \langle e_i|\, e_j\rangle \neq 0] \\
&\quad = \underset{i,j(i\neq j)}{Pr}[|b\rangle \leftarrow \mathcal{A}(1^\lambda) \wedge \langle b|\, U^\dagger_{k\oplus r_i} U_{k\oplus r_j}\, |b\rangle \neq 0] \\
&\quad = \underset{i,j(i\neq j)}{Pr}[|b\rangle \leftarrow \mathcal{A}(1^\lambda) \wedge \langle z\oplus F(k\oplus r_i,m)|\, z\oplus F(k\oplus r_j,m)\rangle \neq 0] \\
&\quad = \underset{i,j(i\neq j)}{Pr}[m \leftarrow \mathcal{A}(1^\lambda) \wedge F(k\oplus r_i,m) = F(k\oplus r_j,m)] = negl(\lambda)
\end{aligned} \tag{3.46}$$

This means that finding an even 2-dimensional common subspace between the different unitaries of the set is computationally hard for $\mathcal{A}$. Also since unitaries are distance preserving operators, this property holds for any sets of orthonormal basis, not necessarily the computational basis. As a result, by selecting a uniformly random function for each query, we have shown that no more than a one-dimensional subspace can be spanned for each specific unitary.

Now we calculate the upper-bound of $\mathcal{A}$'s probability from a single query to a fixed unitary $U_{k\oplus r^*}$ which we denote by U^* for simplicity. We recall that this query should be μ-distinguishable with the quantum encoding of m^*. Without loss of generality, let us write $\mathcal{A}$'s selected query for r^* as follows:

$$\begin{aligned}
&|\phi_{r^*}\rangle = \alpha\,|m^*, z, 0\rangle + \beta\,|\Omega\rangle\,|0\rangle\,, \\
&|\phi^{out}_{r^*}\rangle = (\alpha\,|m^*, z\oplus F(k\oplus r^*, m^*)\rangle + \beta U^*\,|\Omega\rangle)\,|r^*\rangle
\end{aligned} \tag{3.47}$$

where $|\Omega\rangle$ is a normalised state that includes a superposition of a set of messages $m \neq m^*$ and as a result, $\langle m^*, z|\,\Omega\rangle = 0$ and $\mathcal{A}$ sets the second part of the register to 0, such that the output randomness is a separable state and it can be excluded in the rest of the proof. Due to the fact that U^* is unitary, we know that $\langle m^*, z\oplus F(k\oplus r^*, m^*)|\, U^*\,|\Omega\rangle = 0$ and hence the probability of outputting $F(k\oplus r^*, m^*)||r^*$ from $|\phi^{out}_{r^*}\rangle$ is at most the probability of measuring it in the computation basis which is $|\alpha|^2$. This probability is maximum when $|\alpha| = |\alpha_{max}|$ which is when $\mathcal{A}$ uses the maximum allowed overlap of size $\sqrt{1-\mu}$. Hence we have:

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{q_r,\mathsf{qSel},\mu}(\lambda,\mathcal{A})] \leqslant 1-\mu \tag{3.48}$$

But on the other hand we have $P_{ov}(1,\mu) = 1-\mu$, which is the lower bound for $P_{ov}(q,\mu)$, and also since there is only one query to each function selected by each r, and Eq. (3.43) states that this probability is negligibly higher than $1-\mu$. Thus we have reached a contradiction that concludes our proof. □

We point out that for this construction to be secure, we did not need to use quantum secure PRFs (qPRF) as an assumption, and the PRF assumption plus the randomisation would bring the quantum security as a byproduct. This is in contrast with most quantum-secure unforgeable schemes in the quantum world [BZ13a, Zha15, AMRS20]. Nevertheless, qPRFs can also be used in Construction 1.

Now, we shall study the same problem for quantum primitives. Similar to the classical constructions, for quantum primitives too, we can use randomisation to

effectively secure them. The main idea is to select a new unitary transformation for each query using a classical randomness register. In this case, we need to clarify how such randomised quantum oracles can be implemented in a way that the overall transformation remains a specific unitary.

By recalling the abstract representation of the randomised quantum oracle that we gave in Section 3.3.3, the input state $|\psi_b\rangle = \sum_i \alpha_i |b_i\rangle$ (where $\{|b_i\rangle\}$ is a set of orthonormal bases) is mapped to a state $U(r)|\psi_b\rangle = \sum_i \beta_i(r)|b_i\rangle$ where $U(r)$ depends on the randomness and is different for each query *i.e.* the oracle uses its internal register $|r\rangle_{\mathcal{O}}$ to activate different $U(r)$ unitaries. However, for many constructions this randomness value r or a function of it like $g(r)$, will be necessary for verification and hence need to also be outputted. On the other hand, the register $|r\rangle_{\mathcal{O}}$ is the internal register of the oracle re-initiated for each query and some problems may arise if the adversary gets access to this register (see 3.3.3), thus in order to be able to output this value we expand the query space and we allow the input queries to be $|0\rangle \otimes |\psi_b\rangle$. We formulate the oracle as follows:

$$R\mathcal{O}_U^{\mathcal{E}} : |r\rangle_{\mathcal{O}} \otimes |0\rangle \otimes |\psi_b\rangle \rightarrow [\mathcal{I} \otimes \mathcal{I} \otimes U(r)]\,|r\rangle_{\mathcal{O}}\,|r\rangle\,|\psi_b\rangle \tag{3.49}$$

Note that for the purpose of our construction, in what follows, we assume that the ancillary state is initiated as a separable state $|0\rangle$ for simplicity, although if the adversary's ancillary register has not been initiated to zero, the randomness can be XORed to that value. The above oracle can be realised in several different ways but for a better demonstration, we give an explicit example in the circuit model, shown in Fig. 3.4. The input to the unitary evaluation of the oracle consists of two parts; one part includes the query and the second part is the internal randomness register which is initiated to a new value or equivalently to a new basis, for each query. This part in general acts as control qubits for the gates in the other part of the register that leads to applying a new overall unitary on the main query state. We note that the randomness register itself will remain untouched throughout the evaluation and finally its value is recorded in the $|0\rangle$ part of the input query. Here, $|r\rangle_{\mathcal{O}}$ is always on the computational basis. We also emphasize that for our construction we do not use, nor need to use, any explicit construction for the randomised oracle and we only rely on the specified assumption.

As follows from the above discussion, in quantum primitives with such randomised oracles, the security lies in the assumptions we consider on the family of $U(r)$s generated for each r. For instance, it is intuitive that a primitive where $U(r)$ are Haar random unitaries can be secure since the overall adversary's state after issuing polynomial queries to the oracle is almost indistinguishable from a totally mixed state. However, this assumption might be too strong. Hence we give a construction based on PRUs which is also the quantum analogue of PRFs that we used in our previous classical construction.

Construction 2. Let $\mathcal{F} = (\mathcal{S}, \mathcal{E}, \mathcal{V})$ be a quantum primitive with the evaluation unitary $\mathsf{U}_{\mathcal{E}} : \mathcal{H}^{\mathcal{R}} \otimes \mathcal{H}^{D} \to \mathcal{H}^{\mathcal{R}} \otimes \mathcal{H}^{D}$ where D is the overall dimension of the query and $\mathcal{H}^{\mathcal{R}}$ is a 2^{l} dimensional Hilbert space for the randomness. And let λ be the security parameter and l and $\log(D)$ be polynomial in λ. Also, let $\mathcal{U}_{PRU} = \{U_r\}_{r=0}^{L}$ be a PRU family with a cardinality L to be at least 2^{l}. The construction is defined as follows:

- **Setup:** The required parameters $\mathtt{param}$ is generated to instantiate the oracles.
- **Evaluation:** The evaluation picks randomness $r \xleftarrow{\$} \mathcal{R}$ uniformly, initialises the randomness register to $|r\rangle_{\mathcal{O}}$ and applies the following unitary, on each input query $|\psi_b\rangle = \sum_i \alpha_i |b_i\rangle$ where each $U(r) = U_r \in \mathcal{U}_{PRU}$

$$R\mathcal{O}_{U}^{\mathcal{E}} : |r\rangle_{\mathcal{O}} |0\rangle |\psi_b\rangle \xrightarrow{U_{\mathcal{E}}} [\mathcal{I} \otimes \mathcal{I} \otimes U(r)] |r\rangle_{\mathcal{O}} |r\rangle |\psi_b\rangle \tag{3.50}$$

- **Verification:** The verification oracle calls a quantum test algorithm $\mathcal{T}$ as defined in Definition 12 on $U(r) |\psi_b\rangle \langle\psi_b| U(r)^{\dagger}$ and the tag state ρ_t:
  - If $F(\rho_t, U(r) |\psi_b\rangle \langle\psi_b| U(r)^{\dagger}) = 1 - negl(\lambda)$ return $\top$ with a probability $1 - negl(\lambda)$
  - and $Pr[1 \leftarrow \mathcal{T}[(\mathsf{U}_{\mathcal{E}} \rho_{\delta} \mathsf{U}_{\mathcal{E}}{}^{\dagger})^{\otimes \kappa_1}, (\mathsf{U}_{\mathcal{E}} \rho_m \mathsf{U}_{\mathcal{E}}{}^{\dagger})^{\otimes \kappa_2}]] = negl(\lambda)$ for any state ρ_{δ} with δ^2-indistinguishable from ρ_m.

Now let us prove the selective unforgeability of the above construction.

Theorem 22. *Construction 2 is μ-qGSU secure for any $\mu \geqslant 1 - \delta^2$.*

Proof. We prove by contraposition. Let $\mathcal{A}$ be a QPT adversary who plays the μ-qGSU game where the evaluation oracle is as shown in the Eq. (3.50), and wins with non-negligible probability in the security parameter *i.e.* $\mathcal{A}$, wins the game by producing a valid tag ρ_t for their selected message ρ_m and randomness r^* with the following probability, after interacting with the oracle in the learning phase:

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qSel},\mu}(\lambda, \mathcal{A})] - P_{ov} = non\text{-}negl(\lambda) \tag{3.51}$$

Where the $P_{ov} = Pr[1 \leftarrow \mathcal{T}(\rho_{max}^{out})^{\otimes \kappa_1}, (\mathsf{U}_{\mathcal{E}} \rho_m \mathsf{U}_{\mathcal{E}}{}^{\dagger})^{\otimes \kappa_2}]$ according to Definition 34, and ρ_{max}^{out} is query with maximum allowed overlap from μ-distinguishability condition. Since the construction implies that $P_{ov} = negl(\lambda)$, this means:

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qSel},\mu}(\lambda, \mathcal{A})] = non\text{-}negl(\lambda) \tag{3.52}$$

Consequently, $\mathcal{A}$ can produce an output ρ_t with non-negligible fidelity with the actual output $\mathsf{U}(r^*) \rho_m \mathsf{U}(r^*)^{\dagger}$, for a $\mathsf{U}_{r^*} \in \mathcal{U}_{PRU}$. Now we consider two cases.

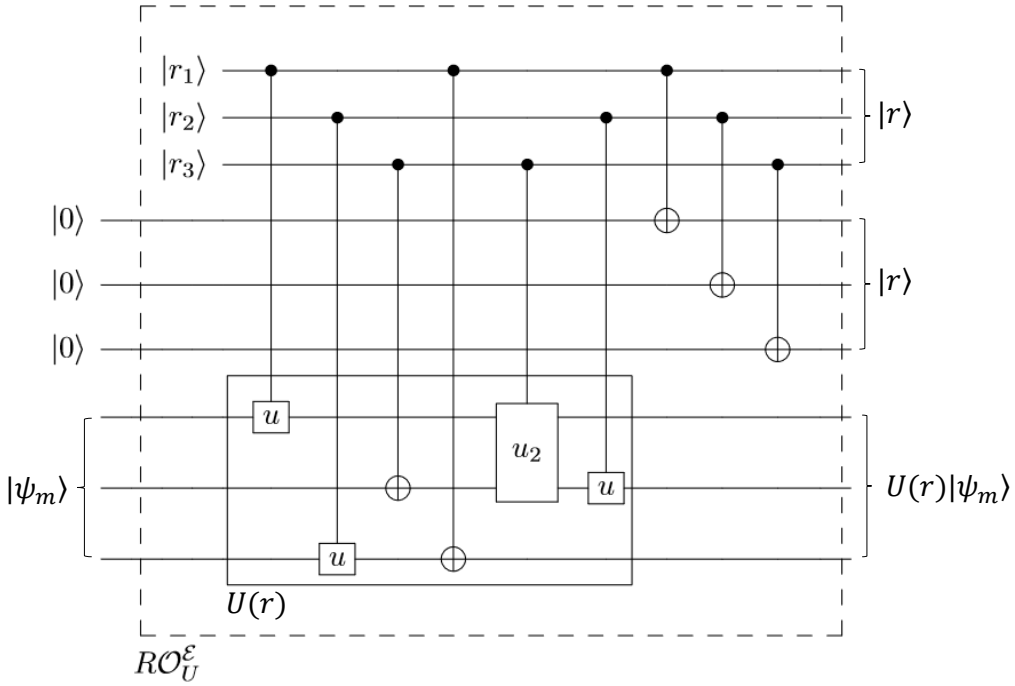

Figure 3.4: A sample circuit for randomised quantum oracle for quantum primitives. On each input query $|0\rangle|\psi_m\rangle$, a new randomness is initialised and the random unitary $U(r)$ acts on $|\psi_m\rangle$. The random unitary $U(r)$ consists of single and 2-qubit unitary gates selected at random in the setup phase, from a gate set required to construct any unitary $U(r)$ in the family $\mathcal{U}$ specified by the construction. These single and two-qubit gates are controlled by the randomness values $|r\rangle = |r_1, r_2, r_3\rangle$. In the last step, the classical value of randomness is recorded in the ancillary qubits of the query to be returned for verification.

Either r^* is one of the randomnesses that $\mathcal{A}$ has received during the learning phase, which means $\mathcal{A}$ can closely approximate the output of a random unitary $U(r^*)$ from a single query, or r^* is a new randomness value, for a new random unitary $U(r^*)$ where $\mathcal{A}$ has no query on it. We will show that each case leads to a contradiction.

First, we show that $\mathcal{A}$'s output state after the learning phase, *i.e.* σ_{out} cannot include more than a one-dimensional subspace of each of the $U(r)$ unitaries. To cover a subspace with a dimension of at least two, $\mathcal{A}$ needs to find a common output basis from two different queries. On the other hand, we note that as shown in [JLS18], any PRUs are generators of PRS that are a family of quantum states computationally indistinguishable from Haar measure. Hence the joint output states σ_{out} is also indistinguishable from Haar random states for $\mathcal{A}$ who is a QPT adversary. Now if $\mathcal{A}$ can find a common output subspace, it means that there are at least two states, corresponding to the bases of the 2-dimensional subspace, that are indistinguishable (or 0-distinguishable according to Definition 10), and hence $\mathcal{A}$ can use those queries to distinguish the distribution of states σ_{out}

and a Haar random distribution which contradicts the fact that the oracle will generate a PRS set of states after q queries. Now we show that each case will lead to a contradiction. We start with the second case where if $\mathcal{A}$ produces an indistinguishable (concerning $\mathcal{T}$) output for a random unitary with no query, then $\mathcal{A}$ can perform the learning phase locally without any interaction with the oracle and hence produce the output of any unitary picked from a family indistinguishable to Haar measure, which is a clear contradiction. For the first case, relying on the previous argument, we rewrite the learning phase states of the $\mathcal{A}$ after q queries, as follows:

$$\sigma_{in} = |\phi_{r^*}\rangle\langle\phi_{r^*}| \otimes \sigma_{in}^{q-1}, \quad \sigma_{out} = U_{r^*}|\phi_{r^*}\rangle\langle\phi_{r^*}|U_{r^*}^{\dagger} \otimes \sigma_{out}^{q-1} \tag{3.53}$$

where $|\phi_{r^*}\rangle$ is the query associated to U_{r^*} for which $\mathcal{A}$ produces a forgery and σ_{in}^{q-1} and σ_{out}^{q-1} are the input and output states of the remaining $q-1$ query respectively. We note that σ_{out}^{q-1} consists of $q-1$ quantum states with a distribution δ over a D'-dimensional Hilbert space s.t. δ is Haar-indistinguishable. Furthermore, the ancillary register where the r is encoded consists of q independent random values. Now let us construct an adversary $\mathcal{A}'$ who is a PRU distinguisher. Let $\mathcal{A}'$ interact with a unitary U either selected from $\mathcal{U}_{PRU}$ or from Haar measure, and query a state $|\phi_{r^*}\rangle$ as described above, and returns $U|\phi_{r^*}\rangle$ together with an ancillary register $|r\rangle$ where r picked uniformly at random. Then $\mathcal{A}'$ also locally creates $q-1$ Haar-random states and returns to $\mathcal{A}$ as the σ_{out}^{q-1}. Then $\mathcal{A}'$ also queries ρ_m from the oracle. Now $\mathcal{A}'$ uses the same test algorithm $\mathcal{T}$ to check the output of $\mathcal{A}$ *i.e.* ρ_t with the the oracle's output for the last query which is $U\rho_m U^{\dagger}$. From Eq. (3.52), we know that this probability is non-negligible, while as for a Haar random unitary the probability is negligible, thus can conclude that

$$|\Pr_{r\leftarrow\mathcal{R}}[\mathcal{A}'^{U_r}(1^{\lambda}) = 1] - \Pr_{U\leftarrow Haar}[\mathcal{A}'^{U}(1^{\lambda}) = 1]| = non\text{-}negl(\lambda). \tag{3.54}$$

which is a contradiction and the theorem has been proved. □

We conclude that even though generalised quantum selective unforgeability is too strong to be attained by deterministic schemes, one can come up with randomised constructions that satisfy even this strong level of unforgeability in the quantum world.

3.6.3 Generalised universally unforgeable schemes

We now draw our attention to the weakest notion of unforgeability in the hierarchy of our definitions and provide results for the universal unforgeability of different schemes. We recall that here the adversary receives a challenge picked by the challenger uniformly at random from the full message space. We need to emphasise that universal unforgeability is the most useful notion of unforgeability for our purpose, despite being the weakest. From now on and throughout the thesis we will mainly use this definition and explore its close relation to unclonability in

the same context as we have discussed in this chapter. We will also study the universal unforgeability of different primitives and protocols in future chapters.

But for now, for the sake of completeness and to complete the investigation of unforgeable schemes in our framework, we give two straightforward results for classical and quantum primitives.

Corollary 3. *qPRFs are qGUU secure.*

Proof. This is a direct implication of Theorem 18 where we have proved that qPRFs are 1-qGSU secure and Theorem 15 showing that 1-qGSU implies qGUU. □

We can also show that quantum PRU primitives are generally qGUU secure.

Corollary 4. *Deterministic quantum primitives based on PRU are qGUU secure.*

Proof. From Theorem 18 we know that PRU primitives are 1-qGSU secure. Also from Theorem 15, we have shown that qGUU is weaker than 1-qGSU. Thus any PRU primitive is qGUU secure. □

In Game 1, we have also introduced a second learning phase, after the challenge phase to capture universal unforgeability against stronger adaptive attack models. Here we also give a general no-go result for qGUU security of quantum primitives against such adversaries. This attack model is stronger than the usual chosen-message attack considered for universal unforgeability and is particularly interesting for quantum primitives. This is because for a quantum primitive, the adversary receives an unknown quantum state from the challenger and enabling the second learning phase does not lead to a trivial attack. We call this attack model, an adaptive-universal attack (*aua*). Nevertheless, we can show that a quantum adversary who can use entanglement can break the qGUU security of any deterministic primitive if the second learning phase is allowed. We show this specific instance of the game as $\mathcal{G}^{\mathcal{F}}_{\mathsf{qUni}-aua,\mu}(\lambda,\mathcal{A})$ and we note that again this instance should be parameterised with μ since a trivial attack can be mount if $\mathcal{A}$ tries to query the challenge phase again in the second learning phase. We present the result in the following theorem. However, we leave the proof for Appendix A.5.

Theorem 23 (No quantum non-randomised primitive $\mathcal{F}$ is aua-qGUU secure)**.** *For any deterministic quantum primitive $\mathcal{F}$ and for any μ such that $0 \leqslant \mu \leqslant 1 - non\text{-}negl(\lambda)$), there exists a QPT adversary $\mathcal{A}$ such that*

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qUni}-aua,\mu}(\lambda,\mathcal{A})] = non\text{-}negl(\lambda). \tag{3.55}$$

3.6.3.1 Relationship between universal unforgeability and learnability

Finally, we show a connection between universal unforgeability and the notion of function learnability, which we discussed in sections 2.6.3 and 3.3.1. More precisely, we show that universal unforgeability implies unlearnability in the PAC-learning setting. To do so, first, we need to clarify some technical remarks concerning the universal unforgeability to be able to link it with PAC-learnability.

First, we note that in Game 1 the learning phase is characterised by interaction with a general oracle $\mathcal{O}^{\mathcal{E}}$, which is included in the primitive $\mathcal{F}$. Here, to establish our result we assume that the oracle for the primitives of interest is a *quantum example oracle* (QPEX) as defined in Eq. (2.119). However, unlike Definition 27, in order to capture the chosen-message attack model that we consider in Game 1, we assume that the adversary gets to choose the distribution $\mathcal{D}$ (but not the selected function f from the concept class $\mathfrak{C}$). We argue that since the QPEX returns a quantum state of the superposition of all the inputs m with uniform weight over the distribution $\mathcal{D}$, for this type of oracles, choosing the distribution will be the equivalent of choosing the input quantum state (or its efficient classical description) of the oracle and receiving the respective quantum output. Therefore, we do not need to make a significant change in the learning phase of our game in order to capture this scenario. Also, in the qUni challenge phase, the message is not chosen uniformly, but from the distribution $\mathcal{D}$. We refer to this variant of universal unforgeability as *universal unforgeability under distribution* $\mathcal{D}$. Now, we can establish the following theorem:

Theorem 24. *Any family of universally unforgeable functions $\mathfrak{C}$, over distribution $\mathcal{D}$, is not PAC-learnable over $\mathcal{D}$.*

Proof. Let $f \in \mathfrak{C}$ be the evaluation function of a primitive $\mathcal{F}$ that is universally unforgeable under distribution $\mathcal{D}$, then by the definition of universal unforgeability, for any QPT adversary $\mathcal{A}$ who can make up to polynomial copy to the oracle, we have:

$$\Pr_{m\in\mathcal{D}}[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qUni}}(\lambda,\mathcal{A})] \leqslant negl(\lambda). \tag{3.56}$$

Let the verification algorithm check the equality of adversary's forgery, *i.e.* $t = h(m)$ with precision ε, that is the verification algorithm will pass the forgery if the following holds:

$$\Pr_{m\in\mathcal{D}}[h(m) \neq f(m)] \leqslant \varepsilon. \tag{3.57}$$

Thus we can rewrite the universal unforgeability of f as follows:

$$\Pr_{m\in\mathcal{D}}[\mathop{\mathbb{E}}_{m\in\mathcal{D}}[h(m) \neq f(m)] \leqslant \varepsilon] \leqslant negl(\lambda). \tag{3.58}$$

Now we assume a learner $\mathcal{A}_p$. We note that according to the definition of PAC-learnability with a QPEX oracle, the adversary gets samples from an unknown distribution $\mathcal{D}$. While as in the universal unforgeability, the adversary gets to choose a new desired distribution $\mathcal{D}_i$ for every query, where i denotes the index

number of the query. We note that for each selected function $f \in \mathfrak{C}$, the learner $\mathcal{A}_p$, is weaker than $\mathcal{A}$. We denote the hypothesis of $\mathcal{A}_p$ as h_p, we then have:

$$\Pr_{m\in\mathcal{D}}[\mathop{\mathbb{E}}_{m\in\mathcal{D}}[h(m) \neq f(m)] \leqslant \varepsilon] \leqslant \Pr_{m\in\mathcal{D}}[\mathop{\mathbb{E}}_{m\in\mathcal{D}}[h_p(m) \neq f(m)] \tag{3.59}$$

From Eq. (3.58) we have that the success probability $\mathcal{A}$ is bounded by a negligible value, thus, the probability of learner $\mathcal{A}_p$, in successfully outputting a hypothesis $h_p(m)$ such that $\Pr_{m\in\mathcal{D}}[h_p(m) \neq f(m)] \leqslant \varepsilon$ is also $negl(\lambda)$. For $\mathfrak{C}$ to be PAC-learnable, this probability needs to be $1-\delta$ for every function $f \in \mathfrak{C}$, while here the δ can only be negligibly close to 1. This concludes that $\mathfrak{C}$ is not PAC-learnable. □

We have shown a link between PAC-learning (with QPEX oracle) and quantum universal unforgeability. One can see from the above result that even if one of the functions in the family (concept class) is universally unforgeable, it is enough to show that the family is *not* PAC-learnable since PAC-learning requires the learner to learn *all* the functions in the concept class with the specified conditions.

3.7 Discussion and conclusions

We have seen in this chapter, how unclonability and especially the unclonability of quantum operations is related to the lack of information, which we characterise with the notion of unknownness. Looking at unclonability from this angle allowed us to expand our horizons into the realm of quantum randomness, cryptography and learning theory. We have discussed the connection between unclonability and unforgeability and between unforgeability and different other notions of learning. We have also talked about emulation as a learning mechanism that can be used as a new class of attacks. On the same note, we have studied a quantum emulation algorithm and developed some simple attacks based on the freshly provided analysis of the algorithm. More importantly, we have developed a universal and generalised framework for unforgeability in the quantum world. Unforgeability will become one of the principal components of this thesis and we will use the definitions and results of our framework in all the remaining chapters (except Chapter 7). Additionally, our case studies on different quantum and classical primitives in this chapter have shown that the generalised quantum unforgeability has shown the applicability of our framework and has also led us to propose both quantum and classical primitives that are secure against powerful quantum adversaries in a strong quantum security model. The first interesting future direction to this work would be to construct efficient and practical constructions for selective and universal unforgeability. These constructions can serve as quantum-secure MACs. Also, building efficient randomised oracles for quantum primitives using random quantum circuits or t-designs is an interesting future research direction.

We have also discussed the close relation between unforgeability and unclonability in the context of quantum money. A potentially attractive application of our framework would be for quantum money schemes. A question that can be of interest is whether one can design quantum money schemes with different levels

of unforgeability, as captured in our framework, and to what extent they can be practical?

As our final contributions, we have formalised our intuitive arguments about the correspondence between unforgeability and learning theory by showing a connection between PAC-learning with a quantum example oracle and a slightly different variant of universal unforgeability. Since in some cases proving universal unforgeability might be easier than PAC-learnability, this result can potentially give some criteria or cryptographic measures for checking the learnability of concept classes. Moreover, we conjecture that a similar result can potentially be obtained for quantum primitives, using the definitions of fully quantum PAC-learnability that exist in the literature, such as [HPS21, PM22b]. Nonetheless, we leave the investigation of this problem as future work.

Lastly, the link between cryptography and learning theory, especially in the quantum world and in the presence of fascinating phenomena such as unclonability, is an appealing and, in some ways fundamental area of research that we could only slightly touch upon in this chapter. Exploiting more novel quantum learning techniques such as shadow tomography and classical shadow [Aar20, HKP20] for cryptanalysis would be the next step in this line of research.

4

Quantum Physical Unclonable Functions

"In everything truth surpasses the imitation and copy."
– Marcus Tullius Cicero

4.1 Introduction

In the previous chapter we discussed cryptographic properties that are related to unclonability and we defined a new security framework, as well as attack tools to expand the study of unclonability of quantum processes from a cryptographic point of view. In this chapter, we introduce a different form of unclonability which is, neither restricted to quantum systems nor originated from quantum mechanics. Yet, this notion of unclonability is also a natural property of certain physical systems which emerges from uncontrollable imperfections, randomness and physical disorders. We refer to this type of unclonability as *Physical Unclonability* adopted from the term *Physical Unclonable Function (PUF)* originally introduced in hardware security. We bring the physical unclonability to the quantum world, and we study it as an abstract mathematical notion concerning the previously introduced concepts in cryptography and quantum information.

But first, let us intuitively describe physical unclonability. Imagine a factory that produces crystals for optical laboratories. The manufacturer intends to produce many copies of the same crystal over and over, with common specific optical properties such as scattering factors. Nevertheless, often no matter how good the product line is and how accurate the devices are, no two crystals produced by this factory are exact clones of each other. This is due to the fact that many uncontrollable parameters and physical randomnesses are involved in the process of crystal formation. Hence on some small scale, the internal structures of two crystals, even though sharing the same large-scale properties, are very different (Fig. 4.1), which makes each crystal unique at that level. Now, if such unique features can be somehow detected (for instance, by shining light on the crystal, which results in producing a unique scattering pattern), one can use each of such

crystals as a physical key. This key has a notable feature that not even the key-maker can have a copy of. Our example crystal or any similar physical device for that matter is called a physical(ly) unclonable function or a physical unclonable key.

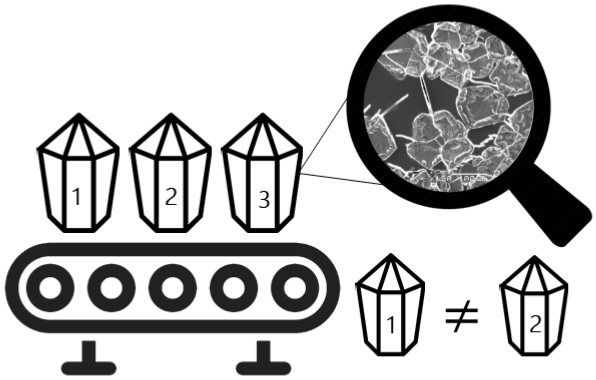

Figure 4.1: Illustration of the concept of a physical unclonable function showing that the underlying microscopic structure and the randomness appearing in the manufacturing process make each PUF device a unique one.

Thus PUFs are hardware structures designed to utilize this random and uncontrollable physical disorder that appears in any physical device during the manufacturing process. The behaviour of a PUF is usually equivalent to a set of Challenge-Response Pairs (CRPs) which are extracted through physically querying the PUF and measuring its responses (In our previous example, the optical parameters of the light are the challenge and the scattering pattern produced by the crystal is the response). The PUF's responses depend on its physical features and are assumed to be fundamentally unpredictable, *i.e.* even the manufacturer of the PUF, with access to many CRPs, cannot easily predict the response to a new challenge [RH14]. This property makes PUFs different from other hardware tokens in the sense that the manufacturer of a hardware token is usually completely aware of the behaviour of the token they have built [BFSK11].

In classical cryptography, physical unclonability is often considered as a *hardware assumption*. Considering hardware assumptions in cryptography, originated from an impossibility result by Canetti and Fischlin [CF01] on the impossibility of achieving secure cryptographic protocols without any setup assumptions. This result has motivated a rich line of research investigating the advantages of making hardware assumptions in protocol design. The idea was first introduced by Katz in [Kat07] and attracted the attention of researchers and developers as it adopts physical assumptions and eliminates the need to trust a designated party or to rely on computational assumptions. Among different hardware assumptions, PUFs have hugely impacted the field [BKOV17].

So far, the cryptographic literature has mainly considered what we will call classical PUFs (or cPUFs/CPUF). This includes, on an abstract level, the physical systems modelled by a classical function and restricted to classical CRPs. Most common cPUFs are electronic devices such as Arbiter PUFs [GCvDD02], Ring-

Oscillator based PUFs [SD07, DMM16] and SRAM PUFs [GKST07]. Optical PUFs were also introduced as cPUFs by Pappu *et al.* [PRTG02]. For a comprehensive overview of existing PUF structures, we refer the reader to [Mae13, Hal18]. Even though cPUF as an unclonable token is highly appealing and of interest for several real-world applications [CZZ17, HYKD14, DMAM17, ADM19, MBM+18, LZZ+19, Muk16], most cPUFs suffer from two major problems. First, most cPUFs generate only a finite and usually very limited number of completely independent CRPs [CZZ17] which is not ideal for many of the mentioned applications. Second, most of them are vulnerable against different attacks like side-channel [TPI19, CZZ17] and machine-learning [GTFS16, RH14, RSS+10, KG19]. In other words, they are not as unpredictable and unclonable as they were initially assumed to be. The aforementioned shortcomings of classical PUFs, on one hand, and their importance as a hardware security primitive in practice, on the other hand, call for investigating other potential physical unclonability in other areas of physics and cryptography. The quantum realm specifically, is one of the best areas to look for such phenomena for several reasons. First, the fundamental unclonability of quantum systems brings forward a potential advantage for achieving a stronger notion of physical unclonability. Second, from the point of view of physics, usually many random disorders that lead to physical unclonability, happen on the atomic and subatomic scale, ruled by the laws of quantum mechanics. Thus having a framework for the study of PUFs as a quantum objects seems to be much more informative. Third, quantum operations can usually generalise classical operations, and if defined carefully, the quantum analogue can encompass classical PUFs, leading to a better understanding of physical unclonability in general. And finally, from a cryptanalysis point, the recent advances in quantum technologies give rise to the question that whether quantum technologies can boost the security of cPUFs or if they, on the contrary, threaten their security. As we will argue later, some more promising classical PUFs such as optical PUFs, *are* quantum devices and can be attacked by a quantum adversary who exploits the power of quantum states and quantum algorithms. Hence to achieve any PUF-based application in the quantum era, the security needs to be properly analysed in a setting that includes quantum adversaries. To conclude, this is one of the few areas of research that lies in the intersection of fundamental physics and cryptography, and human curiosity calls for its theoretical exploration.

In the current chapter, we address the general and formal treatment of PUFs in the quantum world by defining quantum PUFs (qPUFs) as a quantum token/process that can be challenged with quantum states and output quantum states as a response. Our mathematical framework for qPUFs as a new quantum primitive is inspired by the theoretical literature of classical PUF, while we take into account the full capabilities of a quantum adversary. Similar to cPUF, not any function and process can be considered as a PUF and several requirements need to be satisfied. We identify the requirements a qPUF needs to meet to provide the main security property required for most of the qPUF-based applications, that is

unforgeability[1]. One of the main breakthroughs here is to show that the requirements of qPUFs are more restricted than their classical counterparts to achieve the same functionality promises, *i.e.* for qPUFs, the unpredictability is satisfied on a more fundamental level and under fewer assumptions. However, it is worth mentioning that in this chapter we do not focus on the practical constructions for qPUF as designing and implementing concrete qPUFs that satisfy our proposed level of security, remains a challenging task that we will slightly touch upon in the last chapter.

4.1.1 Structure of the chapter

We begin by giving a short background on classical PUF, in Section 4.2. Then in Section 4.3 we define qPUFs as general quantum channels and formalize the standard requirements of robustness, uniqueness and collision-resistance for qPUFs guided by their classical counterparts. We will show that given all the requirements, black-box unitary transformations are perfect candidates for qPUFs. We then formally define the notion of Unitary Quantum PUF (UqPUF). We also discuss the importance of the notion of *unknownness*, as defined in Definition 29 in Chapter 3, as the minimal assumption that leads to the unclonability of qPUFs.

In Section 4.4, we use our unforgeability game-based framework to study the security or unforgeability of general qPUFs. Also using the quantum emulation attacks and techniques that we have introduced in the previous chapter, we demonstrate successful attacks on qPUFs for some security levels. This leads to a general impossibility result for qPUFs. In doing so we establish several possibility and impossibility results. On the other hand, we formally prove that any qPUF provides *quantum universal unforgeability*, *i.e.* no QPT adversary can, on average, generate the response of a qPUF to random challenges. This is the main possibility result of this chapter, which shows a promising direction for research on quantum PUFs.

We conclude the chapter with a discussion and conclusion in Section 4.5. More specifically, we discuss the relevance of our definitions and security framework for other related types of PUF, including classical PUFs. We will argue how our proposed attacks can threaten the security of some of the existing PUF proposals and, we suggest a solution for making them secure by employing our framework and results.

4.1.2 Related works

The concept of Physical Unclonable Functions was first introduced by Pappu *et al.* [PRTG02] in 2001, devising the first implementation of an Optical PUF. Optical PUFs were subsequently improved to generate an independent number of CRPs [MAK+18].

[1] *Unpredictability* and *unclonability* are other equivalent terms for this notion used often in the literature.

More recently, the concept of *quantum read-out of PUF (QR-PUF)* was introduced in [Sko10] to exploit the no-cloning feature of quantum states to solve the spoofing problem in remote device identification protocols. The QR-PUF-based identification protocol has been implemented in [GHM+14]. In addition to the security analysis of this protocol against intercept-resend attack in [Sko10], its security has also been analysed against other special types of attacks targeting extracting information from an unknown challenge state [SMP13, YGLZ16]. In another work, [ND17], the continuous variable encoding is exploited to implement another practical QR-PUF based identification protocol. The security of this protocol has also been analysed only against prepare-and-resend attacks [Nik18, FNAF19]. Moreover, some other applications of QR-PUFs have been introduced in [SPM17] and [UWG+19]. However, all these prior similar works can be considered special cases of qPUFs and in a restricted security setting. (for more discussion, see Section 4.5)

In another independent and parallel recent work, Gianfelici et al. have presented a common theoretical framework for both cPUFs and QR-PUFs [GKB20]. They quantitatively characterise the PUF properties, particularly robustness and unclonability. They also introduce a generic PUF-based identification scheme and parameterise its security based on the experimental implementation of PUF.

4.2 Background on classical Physical Unclonable Functions

In this section, we briefly present the formal definition of PUFs as found in the classical literature [AMSY16, RS14, BFSK11]. Let a $\mathcal{D}$-family be a set of physical devices generated through the same manufacturing process. Due to unavoidable variations during manufacturing, each device has some unique features that are not easily clonable. A PUF is an operation making these features observable and measurable by the holder of the device.

As in [AMSY16, BFSK11], we formalize the manufacturing process of a PUF by defining the Gen algorithm that takes the security parameter λ as input and generates a PUF with an identifier **id**. Note that each time the Gen algorithm is run, a new PUF with new **id** is built. So, we have:

$$\mathbf{id} \leftarrow \mathrm{Gen}(\lambda). \tag{4.1}$$

Also, we define the Eval algorithm that takes a challenge x and a PUF **id** as inputs and generates the corresponding response $y_{\mathbf{id}}$ as output:

$$y_{\mathbf{id}} \leftarrow \mathrm{Eval}(\mathbf{id}, x). \tag{4.2}$$

Due to variations in the environmental conditions, for any given PUF with the identifier **id** (Let us call it $\mathrm{PUF}_{\mathbf{id}}$ from now on for a more intuitive notation), the Eval algorithm may generate a different response to the same challenge x. It is required that this noise be bounded as follows; if $\mathrm{Eval}(\mathbf{id}, x)$ is run several times,

the maximum distance between the corresponding responses should at most be δ_r. This requirement is termed the *robustness requirement*.

Consider a family of PUF generated by the same Gen algorithm, and assume the algorithm Eval is run on all of them with a single challenge x. To be able to distinguish each $\text{PUF}_{\mathbf{id}}$, it is required that the minimum distance between the corresponding responses be at least δ_u. This requirement is termed the *uniqueness requirement*.

The other requirement considered in [AMSY16] is *collision-resistance*. This imposes that whenever the Eval algorithm is run on $\text{PUF}_{\mathbf{id}}$ with different challenges, the minimum distance between the different responses must be at least δ_c. The parameters δ_r, δ_u, δ_c are determined by the security parameter λ. Robustness, uniqueness and collision-resistance are crucial for correctness of cryptographic schemes built on top of PUFs. The conditions $\delta_r \leq \delta_u$ and $\delta_r \leq \delta_c$ must be satisfied to allow for distinguishing different challenges and PUFs [AMSY16].

According to the above, a $(\lambda, \delta_r, \delta_u, \delta_c)$-PUF is defined as a pair of algorithms: Gen and Eval that provides the robustness, uniqueness and collision-resistance requirements. We call a $(\lambda, \delta_r, \delta_u, \delta_c)$-PUF a Classical PUF (cPUF), if the Eval algorithm runs on classical information such as bit strings. We also recall that a cPUF's Eval as a classical function $f : \{0,1\}^n \to \{0,1\}^m$, can be represented as a unitary transformation as follows (see Section 2.5.3 in the preliminaries):

$$\forall x \in \{0,1\}^n, \forall y \in \{0,1\}^m : U_f |x, y\rangle := |x, f(x) \oplus y\rangle \tag{4.3}$$

and thus if physically possible, a quantum adversary can query U_f on any desired quantum states such as the superposition of all the classical inputs.

4.3 Quantum Physical Unclonable Functions

In this section, we define a general notion for quantum PUFs. We consider a set of quantum devices that have been created through the same manufacturing process. These devices produce a general quantum state when challenged with a quantum state. Similar to the previously presented classical setting, we formalize the manufacturing process of qPUFs by defining a QGen algorithm:

$$\mathbf{id} \leftarrow \text{QGen}(\lambda) \tag{4.4}$$

where $\mathbf{id}$ is the identifier of $\text{qPUF}_{\mathbf{id}}$ and λ the security parameter.

We also need to define the QEval algorithm mapping any input quantum state $\rho_{in} \in \mathcal{S}(\mathcal{H}^{d_{in}})$ to an output quantum state $\rho_{out} \in \mathcal{S}(\mathcal{H}^{d_{out}})$ where $\mathcal{H}^{d_{in}}$ and $\mathcal{H}^{d_{out}}$ are the domain and range Hilbert spaces of $\text{qPUF}_{\mathbf{id}}$, denoted as:

$$\rho_{out} \leftarrow \text{QEval}(\text{qPUF}_{\mathbf{id}}, \rho_{in}). \tag{4.5}$$

For now, we allow QEval to be a general trace-preserving quantum map. We have:

$$\rho_{out} = \Lambda_{\mathbf{id}}(\rho_{in}) \tag{4.6}$$

Apart from these common algorithms that are analogue to the classical setting, we also require qPUFs as a primitive, to include an efficient test algorithm $\mathcal{T}$ as we have formally define in Definition 12 to test the equality between two unknown quantum states. We will also need the concept of quantum state distinguishability, which can be defined with different quantum distance measures such as trace distance or fidelity. Here we use the fidelity-based definitions of Definition 10 and Definition 11. We can now define a *Quantum Physical Unclonable Function (qPUF)* as follows.

Definition 39 (Quantum Physical Unclonable Function)**.** Let λ be the security parameter, and $\delta_r, \delta_u, \delta_c \in [0,1]$ the robustness, uniqueness and collision resistance thresholds. A $(\lambda, \delta_r, \delta_u, \delta_c)$-qPUF includes the algorithms: QGen, QEval and $\mathcal{T}$ satisfying Requirements 1, 2, and 3[a].

[a] In Requirements 1 and 3 the probabilities have been taken over the states of the domain Hilbert space, picked from any arbitrary distribution. In Requirement 2 the probability is over the family of CPTP maps between same input and output Hilbert spaces picked from an arbitrary distribution.

Requirement 1 (δ_r-**Robustness**)**.** [2] *For any* $\text{qPUF}_{\mathbf{id}}$ *generated through* QGen(λ) *and evaluated using* QEval *on any two input states* ρ_{in} *and* σ_{in} *that are* δ_r*-indistinguishable, the corresponding output quantum states* ρ_{out} *and* σ_{out} *are also* δ_r*-indistinguishable with overwhelming probability,*

$$\Pr[\delta_r \leq F(\rho_{out}, \sigma_{out}) \leq 1] = 1 - negl(\lambda). \tag{4.7}$$

Requirement 2 (δ_u-**Uniqueness**)**.** *For any two* qPUF*s generated by the* QGen *algorithm,* i.e. $\text{qPUF}_{\mathbf{id}_i}$ *and* $\text{qPUF}_{\mathbf{id}_j}$*, the corresponding CPTP map models,* i.e. Λ_i *and* Λ_j *are* δ_u*-distinguishable with overwhelming probability,*

$$\Pr[\ \| (\Lambda_i - \Lambda_j)_{i \neq j} \|_{\diamond} \geq \delta_u\ ] = 1 - negl(\lambda). \tag{4.8}$$

Requirement 3 (δ_c-**Collision-Resistance (Strong)**)**.** *For any* $\text{qPUF}_{\mathbf{id}}$ *generated by* QGen(λ) *and evaluated by* QEval *on any two input states* ρ_{in} *and* σ_{in} *that are* δ_c*-distinguishable, the corresponding output states* ρ_{out} *and* σ_{out} *are also*

[2] We should note that this requirement is satisfied for any qPUF, by definition, due to the contractivity of quantum channels, as we have defined the evolution algorithm as a CPTP map. However, since this is a crucial requirement for classical PUFs and an important property required for PUFs in general, we have decided to include it as a requirement for the completeness of the framework and for comparison's sake. Also, one might use a framework similar to the one presented in this chapter but with a PUF that is not necessarily a CPTP map, in which case the requirement is not always satisfied and needs to be checked.

δ_c-distinguishable with overwhelming probability,[3]

$$\Pr[0 \leq F(\rho_{out}, \sigma_{out}) \leq 1 - \delta_c] = 1 - negl(\lambda). \tag{4.9}$$

In many PUF-based applications such as authentication and identification, it is necessary that there be a clear distinction between different qPUF instances generated by the same QGen algorithm running on the same parameters λ [AMSY16]. To this end, the following conditions need to be satisfied: $\delta_c \leqslant 1 - \delta_r$ and $\delta_u \leqslant 1 - \delta_r$. We also note that Requirements 1 and 3 impose conditions on the evaluation algorithm of each of the qPUFs in the family while Requirement 2 is a property of the family of physical unclonable functions. In the majority of this chapter, we are interested in the security properties of each one of such functions in the family thus our results mostly concern the QEval algorithm.

We note that if qPUF is a general noisy quantum channel, the δ_c parameter can allow for some specific noise models. More specifically, the weak-collision resistance parameter *i.e.* the ratio of δ_c^o/δ_c^i is directly related to the channel parameters of the qPUF evaluation. Since we are interested in the cryptographic properties of qPUFs, and the collision-resistance is an important requirement for security, we choose the strong collision-resistance as the main requirement for quantum PUFs. We specify that the strong collision-resistance parameter can allow for noisy PUF evaluation under the coherent noise models. Such noise models preserve distances between the input and output states of the qPUF and this property makes them suitable candidates for quantum PUF. Also, it has been shown in [GD18] that a general noise can be modelled as a combination of coherent and incoherent noises. In other words, only the class of noise model with a close to zero incoherent factor can be considered to satisfy the δ_c (strong) collision resistance. Hence for the rest of this work, aiming to formalise the first general security framework, we consider this restricted noise setting that allows for an ideal qPUF and we leave further investigation that would depend on particular constructions for future works.

We have initially allowed for any CPTP map as QEval algorithm. Now, we let the QEval algorithm be a channel with the same dimension of domain and range Hilbert space, *i.e.* $d_{in} = d_{out}$. We show that under this assumption, only unitary transformations and CPTP maps that are highly close to unitary class, can simultaneously provide the (strong)collision-resistance and robustness requirements of qPUFs.

[3] A weaker variant of Collision-Resistance, with separate input/output bound can be also defined in a similar fashion where the responses generated by QEval on any two δ_c^i-distinguishable input states ρ_{in} and σ_{in}, should be at least δ_c^o-distinguishable. In fact, if $\delta_c^i = \delta_c^o = \delta_c$ we call the requirement a strong collision-resistance. Note that this equality holds up to a negligible value in the security parameter, *i.e.* if $\delta_c^i = \delta_c^o \pm negl(\lambda)$, the strong collision-resistance requirement has still been satisfied. If $\delta_c^o < \delta_c^i$ (the difference is non-negligible) then this is referred to as weak collision-resistance.

Theorem 25. *Let $\mathcal{E}(\rho)$ be a completely positive and trace-preserving map described as follows:*

$$\mathcal{E}(\rho) = (1-\varepsilon)U\rho U^\dagger + \varepsilon\tilde{\mathcal{E}}(\rho) \tag{4.10}$$

where U is a unitary transformation, $\tilde{\mathcal{E}}$ is an arbitrary (non-negligibly) contractive channel and $0 \leqslant \varepsilon \leqslant 1$. Then $\mathcal{E}(\rho)$ is a valid qPUF's evaluation algorithm (with equal domain and range dimensionality) for any λ and δ_c (up to a negligible factor), if and only if $\varepsilon = negl(\lambda)$.

Proof. First, we recall the contractive property of trace-preserving operations [NC10], the robustness is trivially satisfied. Hence the robustness is generally satisfied. As a result, the proof of the theorem reduces to proving for collision-resistance. Let ρ and δ be two δ_c-distinguishable challenge with fidelity $F(\rho,\sigma) \leqslant 1-\delta_c$. Again with the above argument the fidelity of the outputs cannot be smaller than $F(\rho,\sigma)$. Thus the δ_c requirement is satisfied if the fidelity of the response density matrices are equal up to a negligible value.

Now let $\rho_1 = U\rho U^\dagger$, $\sigma_1 = U\sigma U^\dagger$, $\rho_2 = \tilde{\mathcal{E}}(\rho)$, and $\sigma_2 = \tilde{\mathcal{E}}(\sigma)$. We use the joint concavity of the fidelity [NC10] to obtain the following relation for the channel's output fidelity:

$$\begin{aligned} F(\mathcal{E}(\rho),\mathcal{E}(\sigma)) &= F((1-\varepsilon)\rho_1 + \varepsilon\rho_2, (1-\varepsilon)\sigma_1 + \varepsilon\sigma_2) \\ &\geqslant (1-\varepsilon)F(\rho_1,\sigma_1) + \varepsilon F(\rho_2,\sigma_2) \end{aligned} \tag{4.11}$$

Since the first part of the channel is unitary which is distance preserving, we have $F(\rho_1,\sigma_1) = F(\rho,\sigma)$. Also due to contractivity we know that $F(\rho_2,\sigma_2) \geqslant F(\rho,\sigma)$. We then have:

$$F(\mathcal{E}(\rho),\mathcal{E}(\sigma)) - F(\rho,\sigma) \geqslant \varepsilon(F(\rho_2,\sigma_2) - F(\rho,\sigma)) \tag{4.12}$$

Now since the channel $\tilde{\mathcal{E}}$ is non-negligibly contractive, the value $F(\rho_2,\sigma_2) - F(\rho,\sigma)$ is not necessarily negligible and in order for the LHS of Eq. (4.11) to be always negligible, ε has to be negligible. So we have proved that CPTP maps of the form Eq. (4.10) can be δ_c collision resistance qPUFs only if $\varepsilon = negl(\lambda)$.

Now we show that all channels of the form of Eq. (4.10) where ε is negligible satisfy the strong collision resistance property up to a negligible value. To show that we recall the relation between fidelity and trace distance, that is $d_{\mathrm{Tr}}(\rho,\sigma) \leqslant \sqrt{1-F(\rho,\sigma)}$. We use this inequality to relate the distance between the states $\mathcal{E}(\rho)$ and $\mathcal{E}(\sigma)$ and the original distance between ρ and σ. By subtracting both sides, we get the following inequality:

$$\begin{aligned} F(\mathcal{E}(\rho),\mathcal{E}(\sigma)) - F(\rho,\sigma) &\leqslant d^2_{\mathrm{Tr}}(\rho,\sigma) - d^2_{\mathrm{Tr}}(\mathcal{E}(\rho),\mathcal{E}(\sigma)) \\ &\leqslant (d_{\mathrm{Tr}}(\rho,\sigma) - d_{\mathrm{Tr}}(\mathcal{E}(\rho),\mathcal{E}(\sigma)))(d_{\mathrm{Tr}}(\rho,\sigma) + d_{\mathrm{Tr}}(\mathcal{E}(\rho),\mathcal{E}(\sigma))) \\ &\leqslant 2(d_{\mathrm{Tr}}(\rho,\sigma) - d_{\mathrm{Tr}}(\mathcal{E}(\rho),\mathcal{E}(\sigma))) \end{aligned} \tag{4.13}$$

Next, we show the following inequality stating that the difference between the trace distance of the input and output for channels described as Eq. (4.10), is bounded by $\varepsilon d_{\mathsf{Tr}}(\rho,\sigma)$,

$$d_{\mathsf{Tr}}(\rho,\sigma) - d_{\mathsf{Tr}}(\mathcal{E}(\rho),\mathcal{E}(\sigma)) \leqslant \varepsilon d_{\mathsf{Tr}}(\rho,\sigma) \tag{4.14}$$

First, we note that the first part of the channel $\mathcal{E}$, which outputs density matrix $U\rho U^\dagger$ with probability $(1-\varepsilon)^2$, is a unitary and preserves the distance. As a result, for a fixed value of ε and any fixed arbitrary states ρ and σ, the difference between the trace distances of the output of $\mathcal{E}$ and the input states increases as $\tilde{\mathcal{E}}$ becomes more contractive. As the maximum contractivity of $\tilde{\mathcal{E}}$ occurs when $\tilde{\mathcal{E}} = \frac{I}{d}$, then the maximum difference between the output and input trace distances is satisfied for this instance of the channel. Let $\mathcal{E}'(\rho) = (1-\varepsilon)U\rho U^\dagger + \varepsilon\frac{I}{d}$. Then for a fixed ε we will have:

$$d_{\mathsf{Tr}}(\rho,\sigma) - d_{\mathsf{Tr}}(\mathcal{E}(\rho),\mathcal{E}(\sigma)) \leqslant d_{\mathsf{Tr}}(\rho,\sigma) - d_{\mathsf{Tr}}(\mathcal{E}'(\rho),\mathcal{E}'(\sigma)) \tag{4.15}$$

Now we calculate $d_{\mathsf{Tr}}(\mathcal{E}'(\rho),\mathcal{E}'(\sigma))$ using the definition of the trace distance:

$$\begin{aligned} d_{\mathsf{Tr}}(\mathcal{E}'(\rho),\mathcal{E}'(\sigma)) &= \frac{1}{2}\mathsf{Tr}[|\mathcal{E}'(\rho),\mathcal{E}'(\sigma)|] \\ &= \frac{1}{2}\mathsf{Tr}[|(1-\varepsilon)U\rho U^\dagger + \varepsilon\frac{I}{d} - (1-\varepsilon)U\sigma U^\dagger - \varepsilon\frac{I}{d}|] \\ &= (1-\varepsilon)(\frac{1}{2}\mathsf{Tr}[|U\rho U^\dagger - U\sigma U^\dagger|]) \\ &= (1-\varepsilon)d_{\mathsf{Tr}}(U\rho U^\dagger, U\sigma U^\dagger) \\ &= (1-\varepsilon)d_{\mathsf{Tr}}(\rho,\sigma) \end{aligned} \tag{4.16}$$

Substituting this back to Eq. (4.15), we get

$$d_{\mathsf{Tr}}(\rho,\sigma) - d_{\mathsf{Tr}}(\mathcal{E}(\rho),\mathcal{E}(\sigma)) \leqslant d_{\mathsf{Tr}}(\rho,\sigma) - (1-\varepsilon)d_{\mathsf{Tr}}(\rho,\sigma) = \varepsilon d_{\mathsf{Tr}}(\rho,\sigma) \tag{4.17}$$

Thus we have:

$$F(\mathcal{E}(\rho),\mathcal{E}(\sigma)) - F(\rho,\sigma) \leqslant 2\varepsilon d_{\mathsf{Tr}}(\rho,\sigma) \tag{4.18}$$

Now if and only if $\varepsilon = negl(\lambda)$ and since $0 \leqslant d_{\mathsf{Tr}}(\rho,\sigma) \leqslant 1$, we conclude that the difference between the fidelity is also negligible and hence the δ_c collision-resistance is satisfied up to a negligible value, and the proof is complete. □

The above theorem shows that only unitary or more generally, *ε-disturbed unitary maps* where ε is small, are suitable candidates for qPUFs, especially when strong collision resistance is required. In the rest of the chapter, we choose the QEval algorithm to be a unitary map. We call this type of qPUFs, Unitary qPUFs (or simply UqPUFs) and formally define them in Definition 40. Nevertheless, we believe studying more general non-unitary qPUFs will be interesting future research directions in this field (see Section 4.5).

So far, we concluded that in terms of the mathematical model, unitary quantum transformations are best suited to describe qPUFs. Now it is time to formalize

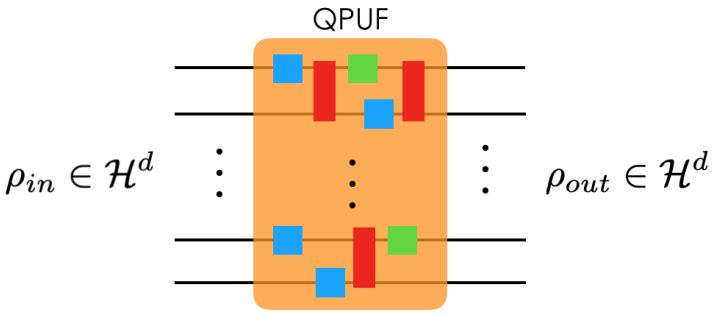

Figure 4.2: Illustration of qPUF as a unitary operation with input and output quantum states in $\mathcal{H}^d$. The blue and green boxes are single-qubit gates, and the red boxes demonstrate two-qubit and entangling gates. These are the abstract building blocks for the d-dimensional UqPUF, while to a QPT adversary, the unitary and the inherent structure is initially unknown.

the main hardware assumption of our qPUFs. We recall that in the classical setting it is assumed that the PUF behaviour is unknown even to the manufacturer. We also require UqPUF transformations to be initially unknown or in other words, behave as a unitary black-box of it which is exponentially hard to recover the full description. In the previous chapter, we have discussed the notion of *unknownness* and its relation to the unclonability and learnability of quantum processes. Now we invoke the same definition to formalize the hardware assumption of physical unclonability by requiring the unitary matrix of a qPUF to satisfy *unknownness* according to Definition 29 from Chapter 3, which formalises single-shot indistinguishability of the unitary from the family of Haar-random unitaries. An illustration of a unitary qPUF is given in Fig. 4.2. Let us formally define Unitary qPUFs (UqPUFs).

Definition 40 (Unitary qPUF (UqPUF))**.** A Unitary qPUF ($(\lambda, \delta_r)-$UqPUF) is a $(\lambda, \delta_r)-$qPUF where the QEval algorithm is modelled by an unknown unitary transformation $\mathsf{U}_{\mathbf{id}}$ over a D-dimensional Hilbert space, $\mathcal{H}^D$ according to Definition 29, such that for any quantum challenge ρ_{in} the respective response ρ_{out} is given as follows,

$$\rho_{out} = \mathrm{QEval}(\mathrm{UqPUF}_{\mathbf{id}}, \rho_{in}) = \mathsf{U}_{\mathbf{id}} \rho_{in} \mathsf{U}_{\mathbf{id}}^{\dagger}. \tag{4.19}$$

For simplicity and practical reasons, usually, the challenge is a pure quantum state denoted as $|\psi_{in}\rangle$, and the response of a UqPUF is simply given by $|\psi_{out}\rangle = U_{\mathbf{id}}|\psi_{in}\rangle$. Also, due to the distance-preserving property of UqPUFs, we drop δ_r from the notation and simply characterise UqPUF as λ-UqPUFs.

There are a couple of notes that are worth mentioning concerning this requirement. First, from the theoretical point of view, this requirement is a minimal and pre-challenge assumption, and considerably weaker than the assumptions needed for classical PUFs. A common requirement needed for classical PUF is *min-entropy* that informally captures the minimum extractable information about a cPUF from subsets of CRPs [AMSY16]. This requirement is morally closely related to the un-

predictability, or unforgeability of PUFs. Nevertheless, in the quantum setting, we aim to characterize the unpredictability as a byproduct of rather simpler hardware assumptions. Our proposed requirement intuitively requires the information of a UqPUF to be obtained only through querying it. One straightforward construction for UqPUF is to sample a unitary from a Haar-random Unitary family, but we believe there are more efficient ways to do this sampling [DCEL09, AE07] (see also Section 4.5, for more discussion about subsequent works and constructions of quantum PUF).

From a construction point of view, this condition may not seem easily achievable, but again practically, it is a reasonable assumption considering limited fabrication capabilities or the fact that simulating an arbitrary unitary on a quantum computer is not technologically easy due to noise and accumulated errors in each gate, even when the structure of the unitary is known. Moreover, there are promising constructions such as the family of optical schemes implemented using crystals or optical scattering media [ND17], where usually even the manufacturer does not know the underlying unitary unless querying it. On the other hand, in gate-based construction, one cannot avoid the fact that the manufacturer knows the underlying unitary. Hence this type of construction cannot provide security against an adversarial manufacturer. Nevertheless, if predicting the evolution of a quantum state is difficult this is enough for security under the usual PUF assumptions. As a result, such devices are still useful and practical for many applications as they can still provide security against any malicious adversary other than the manufacturer. The security framework that we will propose, on the other hand, covers both adversarial models where the manufacturer could be trusted or not.

The final deserving remark, before we move to the cryptanalysis of UqPUFs is that they also satisfy another natural notion of unclonability, known as no-cloning of unitary transformation [CDP08], discussed in Section 3.2.1. We recall that under this notion, two black-box unitary transformations $\mathcal{O}_1$ and $\mathcal{O}_2$ cannot be perfectly cloned by a single use, apart from the trivial cases of perfect distinguishability or when $\mathcal{O}_1 = \mathcal{O}_2$. Thus, two UqPUFs, as long as they correspond to different black-box unitaries, satisfied by the uniqueness requirement and our proposed assumption, are unclonable by quantum mechanics via a single use. Specifically, in the following section, we show how this unclonability property, can be expanded to the multiple-shot case by introducing the formal notion of unforgeability for quantum PUFs.

4.4 Cryptanalysis of Quantum Physical Unclonable Functions

Using the tools and framework that we have established in the previous chapter, for the study of unclonability and unpredictability via the cryptographic notion of unforgeability, we can now formally define this security notion for quantum PUFs and study the extent to which this property is satisfied for general UqPUFs as we have defined them. Other than the fundamental relationship that we aim

to establish between unclonability and unforgeability, in terms of cryptographic applications, the security of most PUF-based protocols relies on the unforgeability of PUFs [AMSY16]. In the context of PUF, unforgeability informally means that given a subset of challenge-response pairs of the target PUF, the probability of correctly guessing a new challenge-response pair shall be considerably small.

In the literature of classical cryptography and hardware security, the unforgeability of PUFs as classical functions is often studied in a game-based framework [AMSY16, BZ13b, DMAM17]. However, there have been studies of PUFs in the UC framework as well [BFSK11, OSVW13].

We recall the definition of unforgeability from the unified framework that we have defined in Chapter 3 and Game 1. Since the framework is defined to capture both quantum and classical primitives, we can easily adapt it for qPUFs. Here we only elaborate on what each of the stages means in the context of qPUFs.

In the **setup phase**, the necessary public and private parameters and functions are shared between the adversary and the challenger and the qPUF is generated.

The **learning phase** models the knowledge that the adversary can gain over a qPUF through queries. We consider chosen-input attacks model where the quantum adversary can choose any arbitrary (and potentially adaptive) query from the domain Hilbert space. Due to the quantum nature of queries, and to be able to fully characterize adversary's database, they have to prepare two copies of each challenge query, keep one in their database, and send the other one to the challenger.

The **challenge phase** captures the intended security notion. For qPUFs, we consider two types of challenge phase: existential and universal[4] as defined before. In the universal case, since we are in the regime of quantum unforgeability for quantum schemes, the uniform selection of the challenge is equivalent to choosing the challenge uniformly at random according to the Haar measure.

Finally, in the **guess phase**, the adversary outputs his guess of the response corresponding to the challenge chosen in the challenge phase. The equality of the adversary's response to the correct response is being tested by a quantum test algorithm as we have abstracted in Definition 12. The adversary wins the game if the output of the test algorithm is 1. Game 3 is adapted directly from Game 1 and formalises the unforgeability of qPUFs.

[4] This level of security is usually known as 'selective unforgeability' in the context of PUFs. Nevertheless, to avoid confusions, with the similar term used in Chapter 3 and for consistency, here we keep the term of universal unforgeability

Formal game-based unforgeability of qPUF

Game 3. Let $\text{qPUF} = (\text{QGen}, \text{QEval}, \mathcal{T})$ and $\mathcal{T}$ be defined as Definition 39 and Definition 12, respectively. We define the following game $\mathcal{G}^{\text{qPUF}}c\mu(\mathcal{A}, \lambda)$ running between an adversary $\mathcal{A}$ and a challenger $\mathcal{C}$:

Setup. The challenger $\mathcal{C}$ runs $\text{QGen}(\lambda)$ to build an instance of the qPUF family, $\text{qPUF}_{\mathbf{id}}$. Then, $\mathcal{C}$ reveals to the adversary $\mathcal{A}$, the domain and range Hilbert space of $\text{qPUF}_{\mathbf{id}}$ respectively denoted by $\mathcal{H}_{in}$ and $\mathcal{H}_{out}$ as well as the identifier of $\text{qPUF}_{\mathbf{id}}$, **id**. The challenger initialises two empty databases, S_{in} and S_{out} and shares them with $\mathcal{A}$.

Learning. For $i = 1 : k$

- $\mathcal{A}$ prepares two copies of a quantum state $\rho_i \in \mathcal{S}(\mathcal{H}_{in})$, appends one to S_{in} and sends the other to $\mathcal{C}$;
- $\mathcal{C}$ runs $\text{QEval}(\text{qPUF}_{\mathbf{id}}, \rho_i)$ and sends ρ_i^{out}, to $\mathcal{A}$;
- $\mathcal{A}$ appends ρ_i^{out} to S_{out}.

Challenge.

- If $c = \text{qEx}$: $\mathcal{A}$ picks a quantum state $\rho^* \in \mathcal{S}(\mathcal{H}_{in})$ at least μ-distinguishable from all the states in S_{in} and sends it to $\mathcal{C}$;
- If $c = \text{qUni}$: $\mathcal{C}$ chooses a quantum state ρ^* uniformly at random from the Haar-measure over the Hilbert space $\mathcal{H}_{in}$. The challenger keeps copies of ρ^* if necessary and sends one copy of ρ^* to $\mathcal{A}$.

Guess.

- $\mathcal{A}$ sends his guess ρ' to $\mathcal{C}$;
- $\mathcal{C}$ runs $\text{QEval}(\text{qPUF}_{\mathbf{id}}, \rho^*)$, and gets ρ_{out};
- $\mathcal{C}$ runs the test algorithm $b \leftarrow \mathcal{T}(\rho_{out}, \rho')$ where $b \in \{0, 1\}$ and outputs b. The adversary wins the game if $b = 1$.[a]

[a] Note that the learning phase queries include any general separable or entangled state.

We follow the same definitions of *quantum existential unforgeability* defined in Definition 30 and *quantum universal unforgeability* in Definition 36 for qPUFs based on the above game. Nevertheless, for the sake of completeness in the study of physical unclonability, we add here another level of security, which is against an exponential (or computationally unbounded) adversary instead of the usual QPT adversary considered in the unforgeability framework. We call this *quantum exponential unforgeability*, *quantum existential unforgeability*, and we define it as follows:

Definition 41 (Quantum Exponential Unforgeability)**.** A qPUF provides quantum exponential unforgeability if the success probability of any *exponential* adversary $\mathcal{A}$ in winning the game $\mathcal{G}^{\text{qPUF}}\text{qEx}\mu(\mathcal{A},\lambda)$ or $\mathcal{G}^{\text{qPUF}}_{\text{qUni}}(\mathcal{A},\lambda)$ is negligible in λ

$$Pr[1 \leftarrow \mathcal{G}^{\text{qPUF}}_{\text{qEx/qUni}}(\mathcal{A},\lambda)] = negl(\lambda) \tag{4.20}$$

4.4.1 Impossibility of exponential unforgeability for UqPUFs

After formalizing all the security games and definitions, it is time to derive general possibility and impossibility results regarding the quantum unforgeability of UqPUFs.

We start with the most powerful setting which is against the exponential quantum adversary. In the classical setting, cPUFs can be fully described by the finite set of CRPs, and this suffices for breaking unforgeability. More precisely, an unbounded or exponential adversary can extract the entire set of CRPs by querying the target cPUF with all possible challenges [CZZ17]. If the challenges are n-bit strings, the number of possible challenges is 2^n. However, in the quantum setting, a UqPUF can generate an infinite number of quantum challenge-response pairs such that extracting all of them is hard, even for exponential adversaries. This point, combined with limitations such as no-cloning and the limits on state estimation [BEM98], raise the question whether UqPUFs could satisfy unforgeability against exponential adversaries. Nevertheless, we answer this question negatively by proving that no UqPUF provides quantum exponential unforgeability as defined in Definition 41.

Theorem 26. (No UqPUF provides quantum exponential unforgeability) *For any λ-UqPUF, there exists an exponential quantum adversary $\mathcal{A}$ such that*

$$Pr[1 \leftarrow \mathcal{G}^{\text{UqPUF}}_{\text{qEx/qUni}}(\lambda,\mathcal{A})] = non\text{-}negl(\lambda) \tag{4.21}$$

Proof. The key idea of the proof is based on complexity analysis of unitary tomography and implementation of a general unitary by single and double qubit gates, since for an exponential quantum adversary, it will be feasible to extract the unitary matrix by tomography and then build the extracted unitary by general gate decomposition method. By using the Solovay-Kitaev theorem [NC10], we then show that the adversary can build the unitary matrix of the UqPUF performing on n-qubits, within an arbitrarily small distance ε using $O(n^2 4^n \log^c(n^2 4^n))$ gates and hence win the game with any test algorithm $\mathcal{T}$. Let $\text{UqPUF}_{\mathbf{id}}$ operate on n-qubit input-output pairs where $n = \log(D)$. In the learning phase, $\mathcal{A}$ selects a complete set of orthonormal basis of $\mathcal{H}^D$ denoted as $\{|b_i\rangle\}_{i=1}^{2^n}$ and queries $\text{UqPUF}_{\mathbf{id}}$ with each base 2^n times. So, the total number of queries in the learning phase is $k_1 = 2^{2n}$.

Then, $\mathcal{A}$ runs a *unitary tomography* algorithm to extract the mathematical description of the unknown unitary transformation corresponding to the $\text{UqPUF}_{\mathbf{id}}$, say $\mathsf{U}_{\mathbf{id}}$. It has been shown [NC10] that the complexity of this algorithm is $\mathcal{O}(2^{2n})$ for n-qubit input-output pairs. This is feasible for an exponential adversary. It is clear that once the mathematical description of the unitary is extracted, $\mathcal{A}$ can simply calculate the response of the unitary to any known quantum challenge. Nonetheless, we want to show the exponential adversary wins the weaker notion of the security, *i.e.* quantum universal unforgeability, where they have only one copy of the challenge state.

To win the game with the universal challenge phase, the adversary needs to implement the unitary. It is known that any unitary transformation over $\mathcal{H}^{2^n}$ requires $O(2^{2n})$ two-level unitary operations or $O(n^2 2^{2n})$ single qubit and CNOT gates [NC10] to be implemented. However, according to Solovay-Kitaev theorem [NC10], to implement a unitary with an accuracy ε using any circuit consisting of m single qubit and CNOT gates, $O(m\log^c(m/c))$ gates from the discrete set are required where c is a constant approximately equal to 2. Thus, an arbitrary unitary performing on n-qubit can be approximately implemented within an arbitrarily small distance ε using $O(n^2 4^n \log^c(n^2 4^n))$ gates.

Finally, $\mathcal{A}$ implements the unitary $\mathsf{U}'_{\mathbf{id}}$ with error ε. Let $\mathcal{A}$ get the challenge state $|\psi\rangle$ in the qUni Challenge phase. The adversary queries $\mathsf{U}'_{\mathbf{id}}$ with $|\psi\rangle$ and gets $|\omega\rangle = \mathsf{U}'_{\mathbf{id}}|\psi\rangle$ as output. Since the ε can be arbitrary small, then $F(\mathsf{U}_{\mathbf{id}}|\psi\rangle, \mathsf{U}'_{\mathbf{id}}|\psi\rangle) \geq 1 - negl(\lambda)$. So, $\mathcal{A}$'s output $|\omega\rangle$ passes any test algorithm $\mathcal{T}(|\psi^{out}\rangle^{\otimes\kappa_1}, |\omega\rangle^{\otimes\kappa_2})$ with probability close to 1. Again, an unbounded adversary wins the game $\mathcal{G}^{\text{UqPUF}}_{\text{qUni}}(\lambda, \mathcal{A})$ with probability 1. Also, since the existential challenge phase is a stronger definition, if $\mathcal{A}$ wins in game $\mathcal{G}^{\text{UqPUF}}_{\text{qUni}}(\lambda, \mathcal{A})$ they will also when $\mathcal{G}^{\text{UqPUF}}\text{qEx}\mu(\lambda, \mathcal{A})$ as well. Therefore we have:

$$Pr[1 \leftarrow \mathcal{G}^{\text{UqPUF}}_{\text{qEx/qUni}}(\lambda, \mathcal{A})] = 1. \tag{4.22}$$

that concludes the proof. □

We note that this result is expected as any qPUF (same as a classical PUF) can, in principle, be simulated with enough computational resources. Therefore no physical unclonability exists against an adversary with an unbounded quantum power since there is no level of *unknownness*, as in an exponentially powerful setting. We point out that similar relation exists in the relation between asymptotic state estimation and approximate quantum cloning (see Section 2.3 and Section 3.2). That is why the reasonable and achievable security model is usually against a qPUF in the hands of the adversary for a limited time or limited query such as QPT adversaries. It is also worth mentioning that from an engineering point of view, limiting the adversary to a certain number of queries on a hardware level can depend on the construction and, it might be possible in some qPUF implementations, while might not be feasible with some others. While this is an interesting problem to consider in qPUF implementations, from a cryptanalysis point, our given security analysis against a quantum adversary who is given

polynomial time in the security parameter of the qPUF is independent of the construction.

4.4.2 Impossibility of existential unforgeability for UqPUFs

Exploiting the quantum emulation tools introduced in Chapter 3 for cryptanalysis, we now turn to quantum existential unforgeability and show that no UqPUF provides quantum existential unforgeability for any $\mu \neq 1$ as defined in Definition 30. Note that the case $\mu = 1$ corresponds to the existential challenge state being orthogonal to all the queried states in the learning phase. With $\mu = 1$, the adversary is prevented from taking advantage of its quantum access to the qPUF to win the game.

Theorem 27. (No UqPUF provides quantum existential unforgeability) *For any λ-UqPUF, and $0 \leqslant \mu \leqslant 1 - non\text{-}negl(\lambda)$, there exits a QPT adversary $\mathcal{A}$ such that*

$$Pr[1 \leftarrow \mathcal{G}^{\mathsf{UqPUF}}\mathsf{qEx}\mu(\lambda, \mathcal{A})] = non\text{-}negl(\lambda). \tag{4.23}$$

Proof. We show there is a QPT adversary $\mathcal{A}$ who wins the game $\mathcal{G}^{\mathsf{UqPUF}}\mathsf{qEx}\mu(\lambda, \mathcal{A})$ with non-negligible probability in λ. We use a similar emulation attack presented in Section 3.4.2.1, which uses only one block of emulation algorithm. The learning phase queries are as follows where $|\phi_1\rangle$ can be any quantum state in $\mathcal{H}^D$, and $|\phi_3\rangle$ is any orthogonal state to $|\phi_1\rangle$ in the domain Hilbert space:

$$|\phi_2\rangle = \begin{cases} \frac{1}{\sqrt{2}}(|\phi_1\rangle + |\phi_3\rangle) & if\ 0 \leqslant \mu \leqslant \frac{1}{2} \\ \sqrt{\mu}|\phi_1\rangle + \sqrt{1-\mu}|\phi_3\rangle & if\ \frac{1}{2} < \mu \leqslant 1 - non\text{-}negl(\lambda) \end{cases} \tag{4.24}$$

Then, $\mathcal{A}$ sets $|\phi_3\rangle$ as his chosen challenge in the existential challenge phase. Note that $|\phi_3\rangle$ satisfies the μ-distinguishability condition with both $|\phi_1\rangle$ and $|\phi_2\rangle$. In the guess phase, to estimate the output of UqPUF to $|\phi_3\rangle$, the adversary $\mathcal{A}$ runs the QE with the reference state $|\phi_r\rangle = |\phi_2\rangle$.

Relying on Theorem 13, the output state of Stage 1 of the QE algorithm is:

$$\begin{aligned} |\chi_f\rangle = \langle\phi_2|\phi_3\rangle|\phi_2\rangle|0\rangle + |\phi_3\rangle|1\rangle - \langle\phi_2|\phi_3\rangle|\phi_2\rangle|1\rangle \\ - 2\langle\phi_1|\phi_3\rangle|\phi_1\rangle|1\rangle + 2\langle\phi_2|\phi_3\rangle\langle\phi_2|\phi_1\rangle|\phi_1\rangle|1\rangle. \end{aligned} \tag{4.25}$$

Having $\langle\phi_1|\phi_3\rangle = 0$ and we setting $\langle\phi_2|\phi_3\rangle = \alpha$ and $\langle\phi_2|\phi_1\rangle = \beta$, the final fidelity in terms of the success probability of Stage 1 is given as follows according to Eq. (3.17)

$$P_{succ-stage1} = |\alpha^2(1 + 4\alpha^2\beta^2)|^2. \tag{4.26}$$

We have different choices for the reference state depending on the distinguishability parameter μ. For cases where the adversary is allowed to produce a new state

with at least overlap half with all the states in the learning phase, by choosing the uniform superposition of the states where $\alpha = \beta = \frac{1}{\sqrt{2}}$, the output fidelity will be:

$$F(|\phi_3^{out'}\rangle\langle\phi_3^{out'}|, |\phi_3^{out}\rangle\langle\phi_3^{out}|) \geqslant \sqrt{P_{succ-stage1}} = 1. \tag{4.27}$$

where $|\phi_3^{out'}\rangle$ and $|\phi_3^{out}\rangle$ are the output of the QE algorithm and UqPUF to $|\phi_3\rangle$, respectively. According to the calculated fidelity, these two states are completely indistinguishable So, the success probability of $\mathcal{A}$ for any test according to Definition 12 is:

$$Pr[1 \leftarrow \mathcal{G}^{\text{UqPUF}}\text{qEx}\mu(\lambda, \mathcal{A})] = Pr[1 \leftarrow \mathcal{T}(|\psi^{out}\rangle, |\omega\rangle)] = 1 \tag{4.28}$$

which is the optimal choice of the reference. On the other hand, for the cases where the adversary is restricted to produce a challenge more than half distinguishable, we can still create a superposed state with $\alpha = \sqrt{1-\mu}$ and $\beta = \sqrt{\mu}$ and end up with the following fidelity of the emulation by setting $\mu = 1 - \textit{non-negl}(\lambda)$

$$\begin{aligned} F(|\phi_3^{out'}\rangle\langle\phi_3^{out'}|, |\phi_3^{out}\rangle\langle\phi_3^{out}|) &\geqslant |\alpha^2(1+4\alpha^2\beta^2)| \\ &= |(1-\mu)(1+4\mu(1-\mu))| \\ &= \textit{non-negl}(\lambda). \end{aligned} \tag{4.29}$$

Thus for any $\frac{1}{2} < \mu \leqslant 1 - \textit{non-negl}(\lambda)$:

$$Pr[1 \leftarrow \mathcal{G}^{\text{UqPUF}}\text{qEx}\mu(\lambda, \mathcal{A})] = Pr[1 \leftarrow \mathcal{T}(|\phi_3^{out}\rangle, |\phi_3^{out'}\rangle)] = \textit{non-negl}(\lambda) \tag{4.30}$$

And the proof is complete. □

This theorem implies that the adversary can always generate the correct response to his chosen challenge provided that he can query it in superposition with other quantum states during the learning phase in terms of the parameter μ. Note that since output quantum states in the learning phase are unknown to the adversary, the more straightforward strategy of superposing the learnt output quantum states cannot be efficiently performed. More precisely, the adversary cannot prepare the precise target superposition of the output states that are completely unknown [OGHW16, DKK17]. Therefore, the proposed attack in the above proof is general yet non-trivial.

4.4.3 Universal unforgeability of UqPUFs

We now show a positive result for the security of UqPUFs in general by further relaxing the level of security and considering quantum universal unforgeability. We will show that any UqPUF can provide this notion. This result also establishes the relationship between unforgeability, unknownness and physical unclonability in the quantum regime. Furthermore, note that in most PUF-based applications, the universal unforgeability is sufficient. We will discuss this further in Chapter 6.

We start by proving the following lemma which is a crucial step towards our proof. The lemma establishes the average probability of any state in $\mathcal{H}^D$ to be

projected in a subspace $\mathcal{H}^d$ where $d \leqslant D$. Based on this lemma, we calculate the probability of a state chosen uniformly at random from $\mathcal{H}^D$ (according to Haar-measure) to belong in the orthogonal subspace to the adversary's subspace. We recall that since quantum emulation is successful with a high probability when the target state has enough overlap with the sample subspace, the idea is to exploit the randomness involved in selecting the target state to prevent quantum emulation or similar attacks.

Lemma 3. *Let $\mathcal{H}^D$ be a D-dimensional Hilbert space and $\mathcal{H}^d$ a subspace of $\mathcal{H}^D$ with dimension d. Also, let Π_d be a projector, projecting any quantum state in $\mathcal{H}^D$ into $\mathcal{H}^d$. The average probability that any state, chosen uniformly at random from $\mathcal{H}^D$ from a Haar distribution, to be projected into $\mathcal{H}^d$ is equal to $\frac{d}{D}$*

$$\Pr_{|\psi\rangle,\Pi_d}\left[|\langle\psi|\Pi_d|\psi\rangle| = 1\right] = \frac{d}{D} \tag{4.31}$$

Proof. The proof is mainly based on the symmetry of the Hilbert space and the fact that the probability of falling into each subspace is equal for any state uniformly picked at random.

Note that any state $|\psi\rangle \in \mathcal{H}^D$ can be written in terms of the orthonormal bases of $\mathcal{H}^D$ denoted by $|b_i\rangle$, as follows:

$$|\psi\rangle = \sum_{i=0}^{D-1} \alpha_i |b_i\rangle \quad \text{with} \quad \sum_{i=0}^{D-1} |\alpha_i|^2 = 1 \tag{4.32}$$

where α_i are complex coefficients. A projection into a smaller subspace consists of choosing d bases of $\mathcal{H}^D$ in the form of $\sum_{j=0}^{d-1} |b_j\rangle\langle b_j|$. Without loss of generality, we can assume $D = md$ where m is an integer. This assumption is always correct for qubit spaces. This means that the larger Hilbert space can be divided into m smaller subspaces each with dimension d. Let $\{|e_i\rangle\}_{i=0}^{d-1}$ be a subset of $\mathcal{H}^D$ which makes a complete set of bases for one of the d-dimensional subspaces. A projector projects $|\psi\rangle$ into one of the subspaces. As $|\psi\rangle$ has been picked at random and the subspaces are symmetric, the probability of falling into each subspace is the same and equal to $\frac{1}{m}$ which is $\frac{d}{D}$. Otherwise either the sum of all probabilities would not be 1 or the $|\psi\rangle$ has not been picked uniformly at random from $\mathcal{H}^D$. This shows that on average the probability of projecting a state ψ is $\frac{d}{D}$. This can also be seen by the fact that the sum of all projectors in a complete set of projectors is equal to one. In this case, we have

$$\sum_{i=0}^{D-1} \Pi_i = \mathbb{I} \tag{4.33}$$

By sandwiching $|\psi\rangle$ on both sides we have:

$$\sum_{i=0}^{D-1} \langle\psi|\Pi_i|\psi\rangle = 1. \tag{4.34}$$

Each $\langle\psi|\Pi_i|\psi\rangle$ is itself equal to $\sum_{j=0}^{d-1}|\langle\psi|d_{ij}\rangle|^2$ where $|d_{ij}\rangle$s are the bases associated to the subspace that the projector Π_i projects into. This corresponds to all the permutations of d number of the coefficient $|\alpha_i|^2$ which will be $\frac{1}{d}$ on average. Since we have $\sum_{i=0}^{D-1}\frac{Pr_{\Pi_i}}{d}=1$, we can conclude that the average probability for all the projectors will be $\frac{d}{D}$. □

We need another small technical toolkit which allows us to derive our next result. We define another abstraction of the test algorithm of Definition 12 in direct relation to fidelity. We formalize the ideal test $\mathcal{T}_\delta^{ideal}$ as follows:

Definition 42 ($\mathcal{T}_\delta^{ideal}$ Test Algorithm)**.** We call a test algorithm according to Definition 12, a $\mathcal{T}_\delta^{ideal}$ Test Algorithm when for any two state $|\psi\rangle$ and $|\phi\rangle$ the test responds as follows:

$$\mathcal{T}_\delta^{ideal}=\begin{cases}1 & F(|\psi\rangle,|\phi\rangle)\geqslant\delta\\ 0 & otherwise\end{cases} \tag{4.35}$$

To establish our positive result, we first present a preliminary theorem which demonstrates the unforgeability of the UqPUF considering an ideal test algorithm which asymptotically satisfies the fidelity as defined in Definition 42.

Theorem 28. *For any unitary qPUF characterised by* $\mathsf{UqPUF}=(\mathsf{QGen},\mathsf{QEval},\mathcal{T}_\delta^{ideal})$*, with dimension* D*, and any non-zero* δ*, the success probability of any QPT adversary* $\mathcal{A}$ *in the game* $\mathcal{G}_{\mathsf{qUni}}^{\mathsf{UqPUF}}(\lambda,\mathcal{A})$ *is bounded as follows:*

$$Pr[1\leftarrow\mathcal{G}_{\mathsf{qUni}}^{\mathsf{UqPUF}}(\lambda,\mathcal{A})]\leqslant\frac{d+1}{D} \tag{4.36}$$

where $0\leqslant d\leqslant D-1$ *is the dimension of the largest subspace of* $\mathcal{H}^D$ *that can be spanned by* $\mathcal{A}$*'s sample database, during learning phase.*

Proof. Let $\mathcal{A}$ be a QPT adversary playing the game $\mathcal{G}_{\mathsf{qUni}}^{\mathsf{UqPUF}}(\lambda,\mathcal{A})$ where UqPUF is defined over $\mathcal{H}^D$. Let S_{in} and S_{out} be the input and output database of the adversary after the learning phase respectively, both with size k. Also, Let $\mathcal{H}^d$ be the d-dimensional Hilbert space spanned by elements of S_{in} where $d\leqslant k$ and $\mathcal{H}_{out}^d$ be the Hilbert space spanned by elements of S_{out} with the same dimension. $\mathcal{A}$ receives an unknown pure quantum state $|\psi\rangle$ as a challenge in the qUni challenge phase and tries to output a state ρ_ω as close as possible to $|\psi^{out}\rangle$. For the simplicity in the proof, we assume the adversary's forgery state is either a pure state or has a purification in the form of $|\omega\rangle$. We are interested in calculating the average probability for the fidelity of $\mathcal{A}$'s output state $|\omega\rangle$ and $|\psi^{out}\rangle$ be larger or equal to δ. We calculate this probability on average over all the possible states chosen uniformly at random according to the Haar measure over $\mathcal{H}^D$.

$$Pr[1\leftarrow\mathcal{G}_{\mathsf{qUni}}^{\mathsf{UqPUF}}(\lambda,\mathcal{A})]=\underset{|\psi\rangle\in\mathcal{H}^D}{Pr}[F(|\omega\rangle,|\psi^{out}\rangle)\geqslant\delta] \tag{4.37}$$

We aim to show, that for any $\delta \neq 0$, the success probability of $\mathcal{A}$ is negligible in λ.

Following the game definition, as the adversary selects states of the learning phase, the classical description of these states is usually known for them while the corresponding responses are unknown quantum states. Let $\mathcal{A}'$ be an adversary who also receives the classical description of the outputs or the complete set of bases of $\mathcal{H}^d$ and $\mathcal{H}^d_{out}$. Thus $\mathcal{A}'$ has a complete description of the map in the subspace; and as a result, has necessarily a greater success probability than $\mathcal{A}$.

$$Pr[1 \leftarrow \mathcal{G}^{\mathsf{UqPUF}}_{\mathsf{qUni}}(\lambda, \mathcal{A})] \leqslant Pr[1 \leftarrow \mathcal{G}^{\mathsf{UqPUF}}_{\mathsf{qUni}}(\lambda, \mathcal{A}')] \tag{4.38}$$

Therefore from now on throughout the proof, we calculate the success probability of $\mathcal{A}'$ who has full knowledge of the subspace, and we bound the success probability of $\mathcal{A}$ via $\mathcal{A}'$. We also note that the adversary cannot enhance their knowledge of the subspace by entangling their local system to the challenges of the learning phase since the reduced density matrix of the challenge/response entangled state lies in the same subspace $\mathcal{H}^d$ and $\mathcal{H}^d_{out}$. Hereby, upper-bounding the success probability of $\mathcal{A}$ with the success probability of $\mathcal{A}'$ who has the full knowledge of the subspace we have also included the possible entangled queries.

Now, we partition the set of all the challenges into two parts: the challenges that are completely orthogonal to the $\mathcal{H}^d$ subspace, and the rest of the challenges that have non-zero overlap with $\mathcal{H}^d$. We denote the subspace of all the states orthogonal to $\mathcal{H}^d$ as $\mathcal{H}^{d^\perp}$. We analyse the average success probability of $\mathcal{A}'$ in terms of the following partial probabilities:

$$\underset{|\psi\rangle \in \mathcal{H}^{d^\perp}}{Pr}[F(|\omega\rangle, |\psi^{out}\rangle) \geqslant \delta] \quad \text{and} \quad \underset{|\psi\rangle \notin \mathcal{H}^{d^\perp}}{Pr}[F(|\omega\rangle, |\psi^{out}\rangle) \geqslant \delta]. \tag{4.39}$$

We denote $F(|\omega\rangle, |\psi^{out}\rangle)$ as F_ω for simplicity. Since the probability of $|\psi\rangle$ belonging to any particular subset is independent of the adversary's learnt queries, the success probability of $\mathcal{A}'$ can be written as:

$$\begin{aligned} Pr[1 \leftarrow \mathcal{G}^{\mathsf{UqPUF}}_{\mathsf{qUni}}(\lambda, \mathcal{A}')] = &\underset{|\psi\rangle \in \mathcal{H}^{d^\perp}}{Pr}[F_\omega \geqslant \delta] \times Pr[|\psi\rangle \in \mathcal{H}^{d^\perp}] \\ &+ \underset{|\psi\rangle \notin \mathcal{H}^{d^\perp}}{Pr}[F_\omega \geqslant \delta] \times Pr[|\psi\rangle \notin \mathcal{H}^{d^\perp}] \end{aligned} \tag{4.40}$$

where $Pr[|\psi\rangle \in \mathcal{H}^{d^\perp}] = 1 - Pr[|\psi\rangle \notin \mathcal{H}^{d^\perp}]$ denotes the probability of the randomly selected $|\psi\rangle$ being projected into the subspace of $\mathcal{H}^{d^\perp}$ or in other words, have zero support in $\mathcal{H}^d$. From Lemma 3, we know that this probability for any subspace, is equal to the ratio of the dimensions. Here $\mathcal{H}^{d^\perp}$ is a $D-d$ dimensional subspace, thus $Pr[|\psi\rangle \in \mathcal{H}^{d^\perp}] = \frac{D-d}{D}$ and respectively $Pr[|\psi\rangle \notin \mathcal{H}^{d^\perp}] = \frac{d}{D}$. Also the probability is upper-bounded by the cases that the adversary can always win the game for $|\psi\rangle \notin \mathcal{H}^{d^\perp}$[5]. So, we have,

$$Pr[1 \leftarrow \mathcal{G}^{\mathsf{UqPUF}}_{\mathsf{qUni}}(\lambda, \mathcal{A}')] \leqslant \underset{|\psi\rangle \in \mathcal{H}^{d^\perp}}{Pr}[F_\omega \geqslant \delta] \times (\frac{D-d}{D}) + \frac{d}{D} \tag{4.41}$$

[5]This is one of the main reasons that our obtained upper-bound for the universal unforgeability

Finally, the only remaining term to be calculated is $\Pr_{|\psi\rangle\in\mathcal{H}^{d^\perp}}[F_\omega \geqslant \delta]$.
We write the expansion of $|\psi\rangle \in \mathcal{H}^D$ in an orthonormal basis for $\mathcal{H}^D$ as $|\psi\rangle = \sum_{i=1}^{D} c_i |e_i\rangle$. For any $|\psi\rangle \in \mathcal{H}^{d^\perp}$, the set of $\{|e_i\rangle\}_{i=1}^{D}$ can be the a union of the bases of $\mathcal{H}^d$, *i.e.* $\{|e_i^{in}\rangle\}_{i=1}^{d}$ and the bases of $\mathcal{H}^{d^\perp}$, *i.e.* $\{|e_i'\rangle\}_{i=d+1}^{D}$. Note that any state in $\mathcal{H}^{d^\perp}$ is orthogonal to all the $|e_i^{in}\rangle$ states. Thus, we can rewrite as follows

$$|\psi\rangle = \sum_{i=1}^{d} c_i^{in} |e_i^{in}\rangle + \sum_{i=d+1}^{D} c_i' |e_i'\rangle \tag{4.42}$$

Recall the case of interest is when $|\psi\rangle \in \mathcal{H}^{d^\perp}$, and , $\langle\psi| e_i^{in}\rangle = 0$. As a result, $c_i^{in} = 0$ and we have,

$$|\psi\rangle = \sum_{i=d+1}^{D} c_i' |e_i'\rangle \tag{4.43}$$

Similarly for the output state $|\psi^{out}\rangle = \sum_{i=1}^{d} c_i^{out} |e_i^{out}\rangle + \sum_{i=d+1}^{D} \alpha_i |b_i\rangle$, as the unitary preserves the inner product, $c_i^{out} = \langle e_i^{out}|\psi^{out}\rangle = \langle e_i^{in}| U^\dagger U|\psi\rangle = \langle e_i^{in}|\psi\rangle = 0$, and the correct output state can be written as

$$|\psi^{out}\rangle = \sum_{i=d+1}^{D} \alpha_i |b_i\rangle \tag{4.44}$$

where $\{|b_i\rangle\}_{i=1}^{D-d}$ are a set of bases for $\mathcal{H}_{out}^{d^\perp}$.
Finally, the adversary $\mathcal{A}'$ can produce a forgery written as follows

$$|\omega\rangle = \sum_{i=1}^{d} \beta_i |e_i^{out}\rangle + \sum_{i=d+1}^{D} \gamma_i |q_i\rangle \tag{4.45}$$

where the first part is spanned by the basis of learnt output subspace and the second part has been produced in $\mathcal{H}_{out}^{d^\perp}$ with $\{|q_i\rangle\}_{i=1}^{D-d}$ being a set of bases for $\mathcal{H}_{out}^{d^\perp}$. Based on unitarity argued above, the first part of the state $|\omega\rangle$ always gives a 0 fidelity, and for $\mathcal{A}'$ to optimise the probability all β_i should be zero. WHich makes $\sum_{i=1}^{D-d} \gamma_i |q_i\rangle \in \mathcal{H}_{out}^{d^\perp}$ where the normalization condition is $\sum_{i=1}^{D-d} |\gamma_i|^2 = 1$.

Since $|\psi\rangle$ is an unknown state selected uniformly at random and independent of the adversary, there are infinite choices for a set of bases orthogonal to $\{|e_i^{out}\rangle\}_{i=1}^{d}$, there is no unique way for $\mathcal{A}'$ choose or obtain the rest of the bases to complete the set. As a result, the choice of the $|q_i\rangle$ bases are also independent of $|e_i'\rangle$ or $|b_i\rangle$. In other words, knowing a matching pair of $(|q_i\rangle, |b_i\rangle)$ increases the dimension of the known subspace by one meaning the adversary has more information than it is assumed to have.

So, for each new challenge, $\mathcal{A}'$ produces a state $|\omega\rangle = \sum_{i=1}^{D-d} \gamma_i |q_i\rangle$ with a totally independent choice of bases. Without loss of generality we can fix the bases $|q_i\rangle$

is not tight, as this assumes the cases where adversary always wins with probability 1 if the state has any non-zero overlap with the sample subspace. Nevertheless, this upper-bound is enough for our purpose to show the unforgeability. Yet, obtaining tight upper-bounds for universal unforgeability is an interesting open question.

for different $|\omega\rangle$. To calculate the success probability of $\mathcal{A}'$, we calculate the fidelity averaging over all the possible choices of $|\psi\rangle$. The unitary transformation also preserves the distribution, in this case Haar distribution. This leads to a uniform distribution of all the possible $|\psi^{out}\rangle$. As a result, the average probability taken over all possible $|\psi\rangle$ is equal to the average probability over all possible $|\psi^{out}\rangle$,

$$\underset{|\psi\rangle\in\mathcal{H}^{d^\perp}}{Pr}[F_\omega \geqslant \delta] = \underset{|\psi^{out}\rangle\in\mathcal{H}^{d^\perp}_{out}}{Pr}[F_\omega \geqslant \delta]. \tag{4.46}$$

We now show that $\mathcal{A}'$ also needs to output $|\omega\rangle$ according to the uniform Haar distribution to win the game in the average case with the highest probability. Let $\mathcal{A}'$ output the states according to a probability distribution $\mathfrak{D}$ which is not uniform. Then, by repeating the experiment asymptotically many times, the correct response $|\psi^{out}\rangle$ covers the whole $\mathcal{H}^{d^\perp}_{out}$ while $|\omega\rangle$ covers a subspace of $\mathcal{H}^{d^\perp}_{out}$. This decreases the average success probability of $\mathcal{A}'$. So, the best strategy for $\mathcal{A}'$ is to generate the states $|\omega\rangle$ such that they span the whole $\mathcal{H}^{d^\perp}_{out}$, *i.e.* generating them according to the symmetric Haar uniform distribution.

Based on the above argument, and the fact that all the $|\omega\rangle$s are produced independently, we show that the average fidelity over all the $|\psi^{out}\rangle$ is equivalent to the average fidelity over all the $|\omega\rangle$. There are different methods for calculating the average fidelity over Hilbert spaces [ZS05], a common approach is to integrate over the symmetric measure such as Haar. In our case, the average fidelity can be formulated as $\int_{|\psi^{out}\rangle\in\mathcal{H}^{d^\perp}_{out}} |\langle\omega|\psi^{out}_x\rangle|^2 d\mu_x$ where $d\mu$ is the Haar measure based on which the reference state has been parameterized. Note that $|\omega\rangle$ can be different for any new challenge. Now we rewrite the above average with the new parameters:

$$\begin{aligned}
\int_{|\psi^{out}\rangle\in\mathcal{H}^{d^\perp}_{out}} F(|\omega\rangle, |\psi^{out}_x\rangle) d\mu_x &= \int_{|\psi^{out}\rangle\in\mathcal{H}^{d^\perp}_{out}} |\langle\omega|\psi^{out}_x\rangle|^2 d\mu_x \\
&= \int_{|\psi^{out}\rangle\in\mathcal{H}^{d^\perp}_{out}} |\sum_{i=1}^{D-d} \overline{\gamma_i}\langle q_i|\psi^{out}_x\rangle|^2 d\mu_x \\
&= \int_{|\psi^{out}\rangle\in\mathcal{H}^{d^\perp}_{out}} |\sum_{i=1}^{D-d} \overline{\gamma_{i_x}}\langle q_i|\psi^{out}\rangle|^2 d\mu_x \\
&= \int_{|\omega\rangle\in\mathcal{H}^{d^\perp}_{out}} |\langle\omega_x|\psi^{out}\rangle|^2 d\mu_x \\
&= \int_{|\omega\rangle\in\mathcal{H}^{d^\perp}_{out}} F(|\omega_x\rangle, |\psi^{out}\rangle) d\mu_x
\end{aligned} \tag{4.47}$$

We used the fact that fidelity is a symmetric function of two states and the measure of integral is the same for both cases where either of $|\psi^{out}\rangle$ or $|\omega\rangle$ are smoothly parametrized according to the symmetric measure. We use this equality

for averaging all the possible outputs for one $|\psi^{out}\rangle$. We wanted to calculate the probability of this average fidelity being greater than δ. To this end, we first calculate more generally, the probability that the average fidelity is non-zero. since we have $\underset{|\omega\rangle\in\mathcal{H}_{out}^{d\perp}}{Pr}[F_\omega \neq 0] + \underset{|\omega\rangle\in\mathcal{H}_{out}^{d\perp}}{Pr}[F_\omega = 0] = 1$, we calculate the probability of the zero fidelity as follows,

$$\begin{aligned}\underset{|\omega\rangle\in\mathcal{H}_{out}^{d\perp}}{Pr}[F_\omega = 0] &= \underset{|\omega\rangle\in\mathcal{H}_{out}^{d\perp}}{Pr}[|\langle\omega|\psi^{out}\rangle|^2 = 0] \\ &= Pr[\int |\sum_{i=1}^{D-d} \overline{\gamma_{i_x}} \langle q_i|\psi^{out}\rangle|^2 d\mu_x = 0] \\ &= \underset{x}{Pr}[(\sum_{i,j=1}^{D-d} \overline{\gamma_{i_x}}\alpha_j \langle q_{i_x}|b_j\rangle)^2 = 0]\end{aligned} \tag{4.48}$$

Based on the Cauchy–Schwarz inequality we obtain the following inequality:

$$[\sum_{i,j=1}^{D-d} \overline{\gamma_{i_x}}\alpha_j \langle q_i|b_j\rangle]^2 \geqslant \sum_{i,j=1}^{D-d} |\overline{\gamma_{i_x}}\alpha_j|^2 |\langle q_i|b_j\rangle|^2 \tag{4.49}$$

where,

$$\sum_{i,j=1}^{D-d} |\overline{\gamma_{i_x}}\alpha_j|^2 |\langle q_i|b_j\rangle|^2 = \sum_{i,j=1}^{D-d} |\overline{\gamma_{i_x}}\alpha_j|^2 |\langle q_i|b_j\rangle\langle b_j|q_i\rangle| = \sum_{i,j=1}^{D-d} |\overline{\gamma_{i_x}}\alpha_j|^2 |\langle q_i|\Pi_j|q_i\rangle| \tag{4.50}$$

Overall, we have,

$$\underset{|\omega\rangle\in\mathcal{H}_{out}^{d\perp}}{Pr}[F_\omega = 0] \geqslant \underset{x}{Pr}[\sum_{i,j=1}^{D-d} |\overline{\gamma_{i_x}}\alpha_j|^2 |\langle q_i|\Pi_j|q_i\rangle| = 0] \tag{4.51}$$

While the RHS is the probability of $|\omega\rangle$ being projected into the orthogonal subspace of a space that only includes $|\psi^{out}\rangle$ averaging over all the projectors. We use again Lemma 3. Here the dimension of the orthogonal subspace is equal to $D-d-1$, since the target subspace is one-dimensional and thus dimension of the orthogonal subspace needs to be subtracted by 1. We then have,

$$\begin{aligned}\underset{|\omega\rangle\in\mathcal{H}_{out}^{d\perp}}{Pr}[F_\omega = 0] &\geqslant \underset{x}{Pr}[(\sum_{i,j=1}^{D-d} |\overline{\gamma_{i_x}}\alpha_j|^2 |\langle q_i|\Pi_j|q_i\rangle|) = 0] \\ &\geqslant \frac{D-d-1}{D-d}\end{aligned} \tag{4.52}$$

And as a result,

$$\underset{|\psi^{out}\rangle\in\mathcal{H}_{out}^{d\perp}}{Pr}[F_\omega \neq 0] = \underset{|\psi^{out}\rangle\in\mathcal{H}_{out}^{d\perp}}{Pr}[|\langle\omega|\psi^{out}\rangle|^2 \neq 0] \leqslant \frac{1}{D-d} \tag{4.53}$$

Which holds for any non-zero δ as well. By substituting back into Eq. (4.41), we conclude that the success probability of $\mathcal{A}'$ is

$$Pr[1 \leftarrow \mathcal{G}_{\mathsf{qUni}}^{\mathsf{UqPUF}}(\lambda, \mathcal{A}')] = \frac{1}{D-d} \times (\frac{D-d}{D}) + \frac{d}{D} = \frac{d+1}{D} \tag{4.54}$$

And the success probability of $\mathcal{A}$ is also bounded by the same bound,

$$Pr[1 \leftarrow \mathcal{G}_{\mathsf{qUni}}^{\mathsf{UqPUF}}(\lambda, \mathcal{A})] \leqslant \frac{d+1}{D} \tag{4.55}$$

which completes the proof. □

Using this theorem that establishes a bound on the success probability of a QPT adversary in terms of fidelity, we can now prove a similar result for a general test algorithm.

Theorem 29. (Any UqPUF provides quantum universal unforgeability) *Let the test algorithm $\mathcal{T}$ be defined according to Definition 12 and satisfy the condition $\mathtt{Err}(\kappa_1, \kappa_2) = negl(\kappa_1, \kappa_2)$. Then any* UqPUF $= (\mathsf{QGen}, \mathsf{QEval}, \mathcal{T})$ *satisfies quantum universal unforgeability as for any QPT adversary, the following holds,*

$$Pr[1 \leftarrow \mathcal{G}_{\mathsf{qUni}}^{\mathsf{UqPUF}}(\lambda, \mathcal{A})] = negl(\lambda). \tag{4.56}$$

Proof. Let $|\psi\rangle$ be the challenge chosen in the universal challenge phase. Also, let $|\psi^{out}\rangle$ and $|\omega\rangle$ be the correct output of the UqPUF and the forgery state of adversary $\mathcal{A}$, respectively. Also, we assume there exists one copy of $|\omega\rangle$, thus $\kappa_1 = 1$, but the challenger may have κ_2 copies stored for verification. The success probability of $\mathcal{A}$ in the game $\mathcal{G}_{\mathsf{qUni}}^{\mathsf{UqPUF}}(\lambda, \mathcal{A})$ is equal to the probability of the test algorithm in outputting 1:

$$Pr[1 \leftarrow \mathcal{G}_{\mathsf{qUni}}^{\mathsf{UqPUF}}(\lambda, \mathcal{A})] = Pr[1 \leftarrow \mathcal{T}(|\psi^{out}\rangle^{\otimes \kappa_1}, |\omega\rangle^{\otimes \kappa_2})] \tag{4.57}$$

We simplify the notation of $Pr[1 \leftarrow \mathcal{T}(|\omega\rangle^{\otimes \kappa_1}, |\psi^{out}\rangle^{\otimes \kappa_2})]$ by substituting with $Pr[1 \leftarrow \mathcal{T}]$. We also note that all the probabilities are being taken on average over the uniform choice of the challenge, though we omit the notation. To calculate this probability, we consider two independent cases that leads to output 1. We introduce the parameter δ as the threshold for $F(|\omega\rangle, |\psi^{out}\rangle)$ which helps us to write the $Pr[1 \leftarrow \mathcal{T}]$ as sum of two terms, *i.e.* the probability of $\mathcal{T}$ outputting 1 while $F(|\omega\rangle, |\psi^{out}\rangle) \geq \delta$ and the probability of $\mathcal{T}$ outputting 1 while $F(|\omega\rangle, |\psi^{out}\rangle) < \delta$:

$$Pr[1 \leftarrow \mathcal{T}] = Pr[1 \leftarrow \mathcal{T}, F(|\omega\rangle, |\psi^{out}\rangle) \geqslant \delta] + Pr[1 \leftarrow \mathcal{T}, F(|\omega\rangle, |\psi^{out}\rangle) < \delta] \tag{4.58}$$

Let $\delta = negl(\lambda)$. We have,

$$\begin{aligned} Pr[1 \leftarrow \mathcal{T}] = {} & Pr[1 \leftarrow \mathcal{T} | F(|\omega\rangle, |\psi^{out}\rangle) \geqslant negl(\lambda)] Pr[F(|\omega\rangle, |\psi^{out}\rangle) \geqslant negl(\lambda)] \\ & + Pr[1 \leftarrow \mathcal{T} | F(|\omega\rangle, |\psi^{out}\rangle) < negl(\lambda)] Pr[F(|\omega\rangle, |\psi^{out}\rangle) < negl(\lambda)] \end{aligned} \tag{4.59}$$

From Theorem 28, we concluded that

$$Pr[F(|\omega\rangle, |\psi^{out}\rangle) \geqslant negl(\lambda)] \leqslant \frac{d+1}{D} \tag{4.60}$$

where d is the dimension of the subspace spanned by the learnt queries and $D = 2^n$ is the dimension of the domain and range Hilbert spaces and n is the number of qubits in each input/output state. Since the adversary is QPT, the number of learnt queries and as a result the value of d should be polynomial in n, *i.e.* $d = poly(n)$. Also, according to Definition 12, we have,

$$Pr[1 \leftarrow \mathcal{T} | F(|\omega\rangle, |\psi^{out}\rangle) < negl(\lambda)] = Err(\kappa_1, \kappa_2) \tag{4.61}$$

And,

$$Pr[1 \leftarrow \mathcal{T} | F(|\omega\rangle, |\psi^{out}\rangle) \geqslant negl(\lambda)] \leqslant F(|\omega\rangle, |\psi^{out}\rangle) \tag{4.62}$$

Considering the equality cases and due to the fact that $Pr[F(|\omega\rangle, |\psi^{out}\rangle) < negl(\lambda)] = 1 - Pr[F(|\omega\rangle, |\psi^{out}\rangle) \geqslant negl(\lambda)]$, the following equation is obtained

$$Pr[1 \leftarrow \mathcal{T}] = Err(\kappa_1, \kappa_2)(1 - \frac{d+1}{D}) + negl(\lambda)\frac{d+1}{D} \tag{4.63}$$

Recall that $\mathtt{Err}(\kappa_1, \kappa_2) = negl(\kappa_1, \kappa_2)$, $d = poly(n)$ and $D = 2^n$ and hence $\frac{d+1}{D} = negl(n)$ and the probability that the test algorithm outputs 1 is computed as

$$\begin{aligned} Pr[1 \leftarrow \mathcal{T}] &= negl(\kappa_1, \kappa_2)(1 - negl(n)) + negl(\lambda)negl(n) \\ &= negl(\kappa_1, \kappa_2) + negl(\lambda)negl(n) \end{aligned} \tag{4.64}$$

Let $\lambda = f(\kappa_1, \kappa_2, n)$, therefore we have

$$Pr[1 \leftarrow \mathcal{G}^{\mathsf{UqPUF}}_{\mathsf{qUni}}(\lambda, \mathcal{A})] = Pr[1 \leftarrow \mathcal{T}] = negl(\lambda) \tag{4.65}$$

and the proof is complete. □

We have shown that general UqPUFs together with a reasonably good quantum test algorithm, always satisfy universal unforgeability. We note that in deriving this result, we have mostly used the symmetries and geometry of Hilbert spaces and the randomness of the selected challenge according to the Haar measure, emphasising that in the quantum case, unlike the classical regime, the unpredictability of the qPUF can be proven given its unknownness or single-shot indistinguishability as a hardware assumption. One can also intuitively infer that any initially unknown unitary is hard to learn on average, given efficient-size oracle access.

4.4.4 A note on the unforgeability of quantum PUFs with public database

As we discussed in the previous section, the randomness of the challenge state, and the fact that an unknown single copy of it is available for the adversary, plays an important role in the unforgeability property of qPUFs. We specifically point

out the close relation to the fundamental limitation of the adversary in copying a single unknown quantum state with high fidelity. It might appear that using the no-cloning property of the challenge state, is enough to provide universal unforgeability, and one might even be able to achieve unforgeability if the unitary transformation is fully or partially public. In fact, this was one of the core ideas in early proposals for using quantum transformation as physical unclonable functions [Sko10, ND17]. In these works, it has been conjectured that an efficient adversary (QPT) is still incapable of providing a good estimation or guess for the response state, even if the unitary, or essentially the classical description of all the potential challenges are known to such adversaries. The given argument is simply the consequence of the challenge state being unknown and provided as a single copy, which presumably makes it hard for the adversary to determine the correct response. In this case, the best strategy would be to measure and estimate the challenge state, which will lead to a small success probability, and as a result, guessing the response state based on such measurements has accordingly low success probability. Making a PUF's database public is motivated since it can clear the need for securely storing big classical or quantum data.

Nevertheless, in the light of our new cryptanalysis tool, namely the QE, we note that such observations and conjectures are not in general correct and only apply to specific attack models such as cases where the adversary is restricted to only prepare and measure single-qubit quantum states. Against a general QPT adversary, we show that no unitary qPUF with a public or partially public database can provide universal unforgeability. This new result is a direct byproduct of our cryptanalysis of quantum emulation in Chapter 3, thus we present it as the following corollary.

Corollary 5. *Let* UqPUF $= (\mathsf{QGen}, \mathsf{QEval}, \mathcal{T}_{\delta}^{ideal})$*, be a unitary qPUF with unitary* U *of dimension* D*. Let* $\mathcal{A}$ *be a QPT adversary and let* $\mathcal{S} = (S_{in}, S_{out})$ *be an efficient-size* $(polylog(D))$ *sample set of* UqPUF *including challenge and response pairs respectively, known to* $\mathcal{A}$*. Let* $\mathcal{H}^d$ *be the subspace fully spanned by* $\mathcal{S}$*. For any challenge state* $|\psi\rangle$ *selected from any arbitrary distribution over* $\mathcal{H}^d$*,* $\mathcal{A}$ *can produce a state* $|\omega\rangle$*, with a very high fidelity compared to* $U|\psi\rangle$*. Therefore* UqPUF*s in this setting do not provide universal unforgeability.*[a]

[a] We note that the use of the term *universal unforgeability* here is slightly informal and different from the universal unforgeability as formally defined in Game 3 since here the challenge is selected from an arbitrary distribution over a subspace rather than being selected from Haar measure over the full space. However, we deliberately use the same expression, as it captures the same notion as the universal unforgeability where the challenge is being selected by the challenger and not the adversary. Only here, the selection of the challenge state is from a different space and distribution.

Proof. Let the database include q quantum input-output query pairs. Let $\mathcal{A}$ run a q-block QE using $\mathcal{S}$. Since, $|\psi\rangle$ is fully spanned by $\mathcal{S}$ or alternatively by the full basis of $\mathcal{H}^d$, then the quantum emulation algorithm can emulate the output of

$|\psi\rangle$, *i.e.* $U_{\mathsf{qPUF}}|\psi\rangle$ with an almost 1 fidelity according to Theorem 11. Thus this output passes any test algorithm with an overwhelming probability. □

As a result of the above corollary, it is fairly obvious that an efficient adversary can always successfully emulate the response of the UqPUF if the database is publicly known since they can locally build the set $\mathcal{S}$, from which the challenge state is selected, and win the universal unforgeability game with probability almost close to 1. Note that a universal quantum emulator is an efficient quantum algorithm, hence can be run by a QPT adversary. More importantly, the above result states that, in all cases where the adversary has a considerable amount of knowledge over an efficient subspace from which the challenge is selected, the quantum emulation attack can be performed on the unknown challenge, leading to a high fidelity forgery state and breaking unforgeability, even though the challenge state is unclonable.

4.5 Discussion and conclusions

We have formally defined quantum physical unclonable functions, as a new notion of unclonability. We established the minimum requirements and conditions to be satisfied at a hardware level such that a unitary transformation qualifies as a qPUF. In doing so, we have also studied the connection between unknownness, physical unclonability and no-cloning of unitary transformations. We have then analysed the unforgeability of qPUFs as their property of interest, both for understanding them as cryptographic primitives and from the application point of view. We proved that even though no qPUF can be exponentially or existentially unforgeable, our proposed general notion of unitary qPUF, provided the unknownness, always satisfies universal unforgeability that has a close connection with the unlearnability of these primitives efficiently. We now briefly discuss the relationship between our proposal and other types of PUFs, as well as the open questions and direction for future works.

First, we briefly discuss the relevance of our framework and result for cPUFs. As mentioned before, input-output pairs of a cPUF are bit-strings. Most of the available cPUF structures use digital encoding as their inputs and outputs. As a result, they can easily be integrated with other functionalities in Integrated Circuits (ICs). Considering encoding of such bit strings in the computational basis of a Hilbert space, the cPUFs can be considered as special types of UqPUFs. Given a quantum oracle access to such PUFs, they can be studied under the *quantum security model* as discussed in Chapter 3, queried with quantum states. In this model, our no-go result stating that no UqPUF provides quantum existential unforgeability can be extended to cPUFs, showing that they are also unable to provide this level of unforgeability for $\mu \neq 1$.

Another interesting point of comparison is to compare the assumptions that lead to unforgeability for qPUFs and cPUFs. According to [AMSY16], the min-entropy requirement (which imposes that the cPUF responses are linearly independent) is the main requirement of a cPUF which leads to existential unforgeability

[AMSY16] against classical adversaries with no quantum access to the cPUF. However, this requirement cannot be achieved with most of the common cPUF structures as shown in [GTFS16, RS14, RSS+10, KG19]. On the other hand, qPUFs only need the basic assumption on PUFs that let the behaviour of PUF be unknown to anyone [RH14]; and under that assumption, they can achieve a slightly weaker notion of unforgeability, yet against much stronger adversaries with quantum capabilities.

It is worth mentioning that in theory, UqPUFs (and as a result, cPUFs in the quantum security model) can still achieve existential unforgeability for $\mu = 1$. Nevertheless, finding physical structures and systems that provide this level of unforgeability is still an open question. To the best of our knowledge, there is no study on the quantum security of cPUFs in the literature. We emphasise that given the speedy progress in quantum technology the investigation of the security of cPUFs against quantum adversaries is crucial. The security of silicon cPUFs and the other types of cPUFs that cannot be queried by quantum states can be explored in the *post-quantum (or standard) security model* where the quantum adversary has only classical interaction with the primitive while they are equipped with a quantum computer. However, for the other types of cPUF structures like optical PUFs that can naturally be queried with quantum states, the security of cPUFs needs to be analysed in the quantum security model with quantum access to the cPUF oracle.

Another main category of PUFs is Quantum Read-out PUFs (QR-PUFs). Since they are also modelled via unitary transformations, they can be naturally compared to UqPUFs. The original definition of QR-PUFs considered quantumly-encoded challenge-response pairs. [Sko10, Sko12]. The security of QR-PUF-based identification protocols has been investigated in specific and limited security models, such as prepare-and-resend adversaries in [Sko10, Sko12, ND17, GHM+14, SMP13, Nik18, FNAF19] where either the full unitary transformation or equivalently the classical description of QR-PUF responses for any known challenge, is assumed to be public knowledge. The security of such PUF-based protocols relies on the bounds for estimating an unknown quantum challenge sent by the verifier.

Although our current framework as it is, will not be directly applicable to all sorts of protocols and scenarios in which QR-PUFs are defined and used due to specific sets of assumptions and adversarial models considered in these scenarios, we believe that QR-PUFs as a stand-alone primitive can be studied in our proposed framework. Following section 4.4.4, we discuss an extended class of such qPUFs which we call Public-Database PUFs (or PDB-PUFs). They include any PUF that can be queried with quantum (or quantumly encoded) challenges, producing quantum responses and are modelled by a publicly known unitary transformation or a public database equivalently. Our framework provides security notions against general and quantum adversaries in the standard game-based model. Hence we can also investigate the security of PDB-PUFs, by relaxing the unknownness condition for this class. Corollary 5 shows that no PUF in this class can provide universal unforgeability (and existential unforgeability). Yet

interestingly, the constructions proposed for such PUFs can be potential candidates for secure qPUFs (such as the optical qPUF presented in [ND17]), by removing the assumption of the public database while ensuring that the challenge subspace is unknown to the adversary. It is worth mentioning that the feasibility of other quantum attacks with current technologies has been discussed in [Sko10, Sko12, GHM+14, SMP13, Nik18, FNAF19]. However, it remains an interesting inquiry whether the quantum emulator attack can also be demonstrated on NISQ quantum devices.

Furthermore, we note that given that our bounds for universal unforgeability are not tight, it leaves a space for exploring the possibility of universal unforgeability under more efficient challenge sets. We will show an example of this in the next chapter. Finding tight bounds for the unforgeability of qPUFs and more generally, quantum primitives is yet another attractive and challenging open problem. Later on, in Chapter 6, we will see that toolkits from quantum information theory such as entropic uncertainty relations and data processing inequalities are useful and powerful tools to prove similar results for specific constructions and we, therefore, suggest that the proposed problem and supervised learning can be both attacked using similar tools. Information-theoretic bounds for quantum vs classical machine learning has been also studied in [HKP21]. Considering the close relationship between unforgeability and learning problems, we believe this work can assist uncover tighter universal forgery bounds for qPUF and other general quantum primitives.

An important complementary question that we left open is the design of concrete qPUF constructions based on the proposed formal framework. Developing sufficiently secure constructions for quantum PUF would be much more complicated than their classical counterparts as one needs to deal with many complications of the quantum world such as noise and decoherence. One of the important steps in this route is to study non-unitary qPUFs while relaxing the strong collision-resistance requirement to a weaker version. Such qPUFs will allow for more general noise models, and if proven to be secure, they will push the practical construction of qPUFs one step further towards experimental realisation. Another challenge in the way of industrialising qPUFs is the need for quantum memory for some of the qPUF-based protocols. It is an interesting question to what degree this resource can be reduced or even removed in different protocols. We will try to address this question to some extent in Chapter 6.

Finally, certification of qPUFs brings up compelling questions that are both of theoretical and practical interest. One example is to develop new efficient techniques for certifying the *effective dimensionality* of quantum black-box primitives such as UqPUFs. These techniques will not only be beneficial certification tools for qPUFs in practice but can also be new toolkits for certification more generally.

4.5.1 Subsequent works

To round off this chapter, let us briefly mention a few related works which appeared after the completion of the work presented in this chapter. Firstly, in [KMK21] a circuit-based construction has been proposed for qPUFs. The con-

struction uses t-designs to accomplish the functionality of UqPUFs as well as showing the requirements of qPUFs are satisfied. Although this specific construction does not provide physical unclonability against the manufacturer due to the gate-based construction, it can be used against third-party adversaries (and not the manufacturer).

Another proposal for achieving physical unclonability via quantum devices is given in [PSA+21]. In this work, the concept of a Classical-Readout QPUF (CR-QPUF) has been introduced where a quantum device is queried classically. Such PUFs aimed to utilize the noisy behaviour of quantum devices as the main source of physical unclonability and to remove the requirement for a quantum memory in the qPUF-based authentication protocols. In the proposed protocol, the challenge is a classical description of a parameterized unitary, which runs on a quantum computer. Then, the mean value of the measurement outcome over the qubits in the computational basis has been taken to be the response. In the proposed protocols both challenges and responses are communicated over a classical channel. Even though this construction seems practically feasible and interesting, one can debate that considering the mean-value as the response, will most probably substantially reduce the security as a consequence of shrinking the response space to a high degree. Later in [PPS21], the authors will show that this observation is correct and such PUFs can be efficiently learned using machine learning methods. Another appealing aspect of the work in [PPS21] is that it formalises the class of Classical Readout Quantum PUFs (CR-QPUFs) using the statistical query (SQ) model that seems another promising approach to studying non-unitary PUFs with underlying quantum properties. The modelling attack proposed in this work explicitly shows the insufficiency of this class by successfully implementing the attack on the QCIBM quantum machine.

The research on authentication protocols using optical PUFs also continued in the two following works [WCL+21, Nik21] demonstrating experimental realisation of photonic PUFs in specific authentication protocols.

5

Connection Between Quantum Pseudorandomness and Quantum Hardware Assumptions

"Unforeseen surprises are the rule in science, not the exception. Remember: Stuff happens."

– Leonard Susskind

5.1 Introduction

In the previous chapter, we have thoroughly studied the concept of physical unclonability, and quantum physical unclonable functions as hardware assumptions. We have also proved several results about their main cryptographic properties. Furthermore, in the course of the last two chapters, we have attempted to disclose the fundamental connection between the notions of unclonability and randomness. In this chapter, we turn into another stimulating key concept in cryptography, namely *pseudorandomness*, and we show the relationship between this notion and physical unclonability, or even more generally, hardware assumptions.

Pseudorandomness is one of the most fundamental concepts in cryptography and complexity theory. In contrast to true randomness, it captures the notion of primitives that behave randomly to the computationally-bounded observers [Yao82, Sha83, BM84]. Pseudorandom objects like pseudorandom number generators (PRGs) and pseudorandom functions (PRFs) play a crucial role in designing classical symmetric key cryptographic protocols for secure communications [GGM86, HILL99, LR88, Rom90]. These pseudorandom objects can be designed by exploiting the algebraic properties of families of keyed functions like keyed hash functions. Nevertheless, constructing these pseudorandom objects is challenging and usually relies on some computational assumptions. Recently Ji, Liu, and Song [JLS18] introduced the concept of quantum pseudorandomness as a quantum analogue of this concept by introducing pseudorandom quantum states (PRS, Definition 19) and pseudorandom unitaries (PRU, Definition 20). These are

families of states or unitary transformations indistinguishable from Haar measure (true random measure) to any quantum computationally-bounded observer. Even though quantum pseudorandomness is a very new field of research, it has already found many applications in cryptography [JLS18, AMR20, BS20, MY21, AQY22], complexity theory [Kre21, BCHJ+21], learning theory [HBC+21], and high energy physics [BFV19, KTP20]. The existing PRS schemes are constructed under computational assumptions such as quantum-secure PRFs or quantum secure one-way functions [BS20, JLS18]. An interesting question arises here: *Whether quantum pseudorandomness can be achieved under different sets of assumptions, for instance, hardware assumptions?* In this chapter, we mainly try to address this question. Given our specific interest in qPUFs, as a well-defined hardware assumption, we mostly focus on them and for the first time, we show the construction of quantum pseudorandom unitaries from quantum PUFs and vice-versa. We also point out that in the classical world, the relationship between PUFs and pseudorandomness has also been studied [RSS09] and it would be interesting to see if such a relationship also exists in the quantum setting.

Understanding this connection not only provides a deeper understanding of quantum pseudorandomness itself but also can substantially improve the construction of qPUFs and qPUF-based applications as well. Let us give an example. In the previous chapter, we have seen that a Haar-random family of unitaries can, by definition, be a family of secure qPUFs. Nonetheless, sampling Haar-random unitaries and states requires exponential resources [Kni95, NZO+21] and hence is experimentally challenging [CHS+15]. Moreover, the challenge distribution for universal unforgeability is required to also be Haar. If PRSs, can be used within the framework of universal unforgeability as a challenge set, a considerable improvement will occur in the practicality of any universally unforgeable scheme, including quantum PUFs.

In this work, we make substantial progress in the challenges mentioned above. Firstly, we show that PRS can replace the Haar-random assumption in the challenge state's selection for universal unforgeability. We further show that PRUs can be used as a viable candidate for qPUFs. This result provides yet another novel and efficient technique for constructing qPUFs.

Concerning our first question about alternative ways of constructing quantum pseudorandomness, we show that a qPUFs family can also be a family of PRUs. This, in turn, makes them special physical generators of pseudorandom quantum states.

Later, we give a novel construction of PRUs by exploring yet another hardware requirement, *i.e.* the uniqueness property. This result is shown generally for any family of unitary matrices with a certain specified degree of uniqueness, not only qPUFs. And as long as the uniqueness property can be assumed at a hardware level, it relates a hardware assumption to quantum pseudorandomness. Informally, we prove that any family of unitary transformations over d-dimensional Hilbert space satisfying almost-maximal uniqueness in the diamond norm is also a PRU family for sufficiently large d. Hence any PUF family satisfying this degree of uniqueness, is also a PRU.

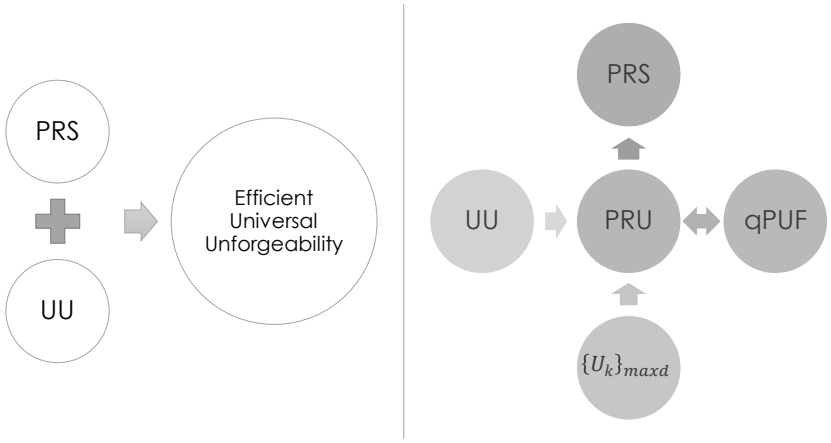

Figure 5.1: Pictorial summary of the results: The left-hand figure demonstrates Theorem 30 stating that universal unforgeability of unknown unitaries can be achieved efficiently using PRSs. The right figure depicts the relationship between unknown unitaries (UUs), quantum physical unclonable functions (qPUFs), pseudorandom unitaries (PRUs) and families of almost maximally-distanced unitaries ($\{U_k\}_{maxd}$) proved in Theorem 31, Theorem 32, Theorem 33, and Theorem 34. It also shows that they can be used as generators for PRSs.

Our investigation in this chapter helps establish a close connection between these two new fields and gives us novel insights into both physical unclonability and quantum pseudorandomness. A summary of our results is shown in Fig. 5.1. We are optimistic that the connections we foster here will enrich both fields.

5.1.1 Structure of the chapter

We begin the chapter with a question: Is it possible to have universal unforgeability with a PRS challenges set instead of a Haar-random state without losing any security guarantee? In Section 5.2, we give a positive answer to this question with a formal security proof. In section 5.3, we show that one can construct a family of unknown unitaries from PRUs, which gives a potentially efficient proposal for constructing a family of qPUFs. Finally, in section 5.4, we show that given hardware assumptions such as uniqueness and unknownness, one can achieve quantum pseudorandomness, which completes our picture.

5.2 Efficient unforgeability with PRSs

In this section, we investigate the problem of *universal unforgeability* with efficiently producible pseudorandom quantum states. As specified in the universal unforgeability framework in Chapter 3, the challenge states should be picked at random from Haar measure by the challenger. This is an important condition for the unforgeability of unknown unitary transformations. Since producing Haar random state is a challenging and resource-intensive task, to take the first step

towards the realization of universally unforgeable schemes, we attempt to replace this condition with its computational equivalent, *i.e.* the notion of PRS, introduced in Chapter 2 (Section 2.5.5). We first relax this condition by defining a variant of the universal unforgeability game, namely quantum *Efficient Universal Unforgeability (qEUU)* where the challenger picks the challenge states from a pseudorandom family of quantum states. Then we formally prove that unknown unitaries satisfy this notion of unforgeability. Furthermore, we briefly discuss how such pseudorandom quantum states can be efficiently generated using classical pseudorandom functions.

We define efficient universal unforgeability as follows:

Definition 43 (quantum Efficient Universal Unforgeability (qEUU))**.** Let Game $\mathcal{G}^{eff}_{\mathsf{qUni}}$ be the same as Game 3, except that in the challenge phase, the challenge states are being picked from the PRS family of states with a generation algorithm $G(k)$ with a key $k \in \mathcal{K}$, run in the setup phase. A primitive provides *efficient quantum universal unforgeability* if the success probability of any QPT adversary $\mathcal{A}$ in winning the game $\mathcal{G}^{eff}_{\mathsf{qUni}}$ is negligible in the security parameter λ,

$$Pr[1 \leftarrow \mathcal{G}^{eff}_{\mathsf{qUni}}(\lambda, \mathcal{A})] = negl(\lambda) \tag{5.1}$$

For the purpose of our proof, we also rewrite the pseudorandomness property of the PRS as a game which we formalized in the following:

PRS distinguishability game

Game 4. Let $\mathcal{H}$ be a Hilbert space and $\mathcal{K}$ the key space. The dimension of $\mathcal{H}$ and size of $\mathcal{K}$ depend on the security parameter λ. Let $\{|\phi_k\rangle \in S(\mathcal{H})\}_{k\in\mathcal{K}}$ be a keyed family of quantum states with efficient generation algorithm $G(k) = |\phi_k\rangle$ on input k. We define the following distinguishability game between an adversary $\mathcal{A}$ and a challenger $\mathcal{C}$:

Setup phase. The challenger $\mathcal{C}$ selects $k \xleftarrow{\$} \mathcal{K}$ and $b \xleftarrow{\$} \{0,1\}$ at random.

Challenge phase.

- If $b = 0$ (PRS world): $\mathcal{C}$ prepares m copies of $|\phi^0\rangle = |\phi_k\rangle$ by running $G(k)$.
- If $b = 1$ (Random world): $\mathcal{C}$ prepares m copies of a Haar-random state $|\phi^1\rangle = |\psi\rangle$.
- $\mathcal{C}$ sends $|\phi^b\rangle^{\otimes m}$ to $\mathcal{A}$.

Guess phase. $\mathcal{A}$ guesses b.

We now establish our main result regarding efficient unforgeability of unknown

unitary primitives.

Theorem 30. *Any unitary transformation U selected from a family of unknown unitaries satisfies* **quantum efficient universal unforgeability** *against any QPT adversary.*

Proof. We prove this theorem by contraposition in a game-based setting. We want to show that starting from the assumption of pseudorandomness of PRS in the efficient universal unforgeability game, if there exists a QPT adversary who succeeds to win this game, with non-negligible probability, there will also exist an adversary who can efficiently distinguish between PRS and Haar random states, which is in contrast with the initial assumption and as a result show a contradiction. First, we need to specify the following games:

- `Game 1`: This is the universal unforgeability game as specified in Game 1, with the only difference that the challenge state $\rho^* = |\phi_{k^*}\rangle\langle\phi_{k^*}|$ is chosen from a PRS family.

- `Game 2`: This is the PRS distinguishability game as specified in Game 4.[1]

- `Game 3`: This is a variation of Game 4 where $\mathcal{C}$ in addition to initial resources, has also access to a publicly known and implementable unitary U. In the challenge phase, $\mathcal{C}$ does the following: Generates m copies of $|\phi^0\rangle = |\phi_k\rangle$ using $G(k)$, or m copies of Haar random states $|\phi^1\rangle = |\psi\rangle$ depending of b, then on each copies applies the public unitary U and sends $(U|\phi^b\rangle)^{\otimes m}$ to $\mathcal{A}$. The rest of the game is similar to `Game 2`.

- `Game 4`: This game is similar to `Game 3`, except that $\mathcal{C}$ publicly chooses an l and l' such that $l + l' = m$ and sends l copies of the generated state and l' copies of the state after applying the unitary U, *i.e.* sends $|\phi^b\rangle^{\otimes l} \otimes (U|\phi^b\rangle)^{\otimes l'}$ to $\mathcal{A}$.

- `Game 5`: This game is similar to `Game 4` except the public unitary has been replaced by an unknown unitary $\tilde{U}$ of the same dimension. Hence in this game, similar to Game 1, we also assume a learning phase for $\mathcal{A}$ before the challenge phase. The learning phase is as follows: $\mathcal{A}$ issues $q = poly(\lambda)$ queries $\{\rho_i\}_{i=1}^q$ to $\mathcal{C}$, on each query $\mathcal{C}$ generates $\rho_i^{out} = \tilde{U}\rho_i\tilde{U}^\dagger$ by applying the unitary on the query state and sends ρ_i^{out} to $\mathcal{A}$. Then the rest of the game is similar to `Game 4` and at the end of the challenge phase $\mathcal{A}$ receives $|\phi^b\rangle^{\otimes l} \otimes (\tilde{U}|\phi^b\rangle)^{\otimes l'}$

[1] One small remark is that in PRS game, the unitary is picked inside the game, while the universal unforgeability game takes the unitary as part of the primitive and hence applies to any selected unitary. However, since we will show that our result applies to *all* unknown unitary matrices, it will also hold in the average case. Thus we drop this distinction in the course of the proof to avoid confusion.

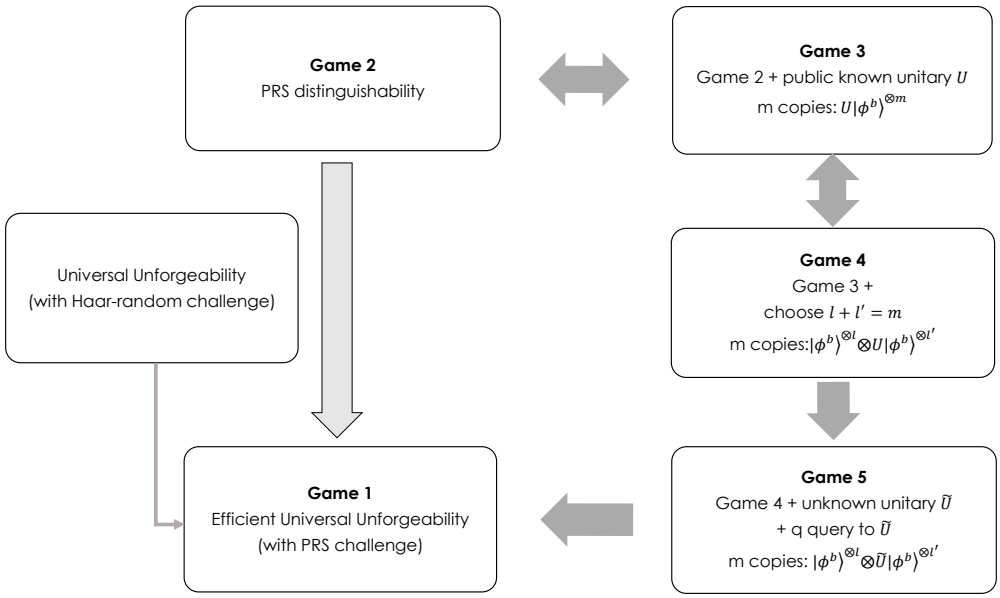

Figure 5.2: Proof sketch of Theorem 30 with the intermediate games.

Figure 5.2 illustrates the sketch of the proof. We first show that `Game 2`, `Game 3` and `Game 4` are equivalent. We note that unitary transformations are distance invariant and hence they also preserve the distribution of states, as a result applying a unitary to the state will not affect the distribution and the distinguishability of the quantum states, and as a result `Game 2` and `Game 3` are equivalent. Furthermore, in `Game 4`, since the unitary is public, $\mathcal{A}$ can either apply U on the first l copies $|\phi^b\rangle^{\otimes l}$ and end up with m copies of $(U|\phi^b\rangle)^{\otimes m}$ or alternatively apply $U^\dagger$ on the next l' copies $(U|\phi^b\rangle)^{\otimes l'}$ and get m copies of $|\phi^b\rangle^{\otimes m}$, and hence be reduced to either `Game 2` or `Game 3`. As a result, we have

$$\texttt{Game 2} \equiv \texttt{Game 3} \equiv \texttt{Game 4} \tag{5.2}$$

Now we show that `Game 4` implies `Game 5` *i.e.* if an adversary wins distinguishability in `Game 5` with probability p, they will also win in `Game 4` with the same probability.

The proof is straightforward as highlighted here. Let $\mathcal{A}$ be an adversary who wins `Game 5`, which means after the learning phase leading to a polynomial-size database of input-outputs of the unknown unitary $\tilde{U}$, and receiving $|\phi^b\rangle^{\otimes l} \otimes (\tilde{U}|\phi^b\rangle)^{\otimes l'}$, they can guess b with non-negligible probability better than random guess:

$$\Pr_{|\phi^b\rangle}[b \leftarrow \mathcal{A}(|\phi^b\rangle^{\otimes l} \otimes (\tilde{U}|\phi^b\rangle)^{\otimes l'})] = \frac{1}{2} + \textit{non-negl}(\lambda). \tag{5.3}$$

Now let's assume an adversary $\mathcal{A}'$ who plays `Game 4` and has to guess b by receiving the state $|\phi^b\rangle^{\otimes l} \otimes (U|\phi^b\rangle)^{\otimes l'}$ can guess b with same l and l' where U is a public unitary. Now $\mathcal{A}'$ can run $\mathcal{A}$ as a subroutine and $\mathcal{A}'$ sends to $\mathcal{A}$ the response to the same learning phase states from U. Since U is public $\mathcal{A}'$ can run it locally and

produce the required queries. Then $\mathcal{A}'$ also sends the state $|\phi^b\rangle^{\otimes l} \otimes (U|\phi^b\rangle)^{\otimes l'}$ to $\mathcal{A}$ and since $\mathcal{A}$ guesses the b with a probability non-negligibly better than half, so does $\mathcal{A}'$. As a result, we have shown that:

$$\texttt{Game 4} \Rightarrow \texttt{Game 5} \tag{5.4}$$

Finally, we show that `Game 5` implies `Game 1`. By contradiction, we assume there exist an adversary $\mathcal{A}$ who wins the unforgeability game with non-negligible probability. Let $\tilde{U}$ be the unknown unitary and $\mathcal{A}$'s forgery state be $|\omega\rangle$ and let the challenge state of `Game 1` be a PRS state $|\phi_k\rangle$. We have:

$$\begin{aligned} Pr[1 \leftarrow \mathcal{G}^{eff}_{\mathsf{qUni}}(\lambda, \mathcal{A})] &= \Pr_k[1 \leftarrow \mathcal{T}(|\omega\rangle, (\tilde{U}|\phi_k\rangle)^{\otimes\kappa})] \\ &= \Pr_k[F(|\omega\rangle, \tilde{U}|\phi_k\rangle) = \textit{non-negl}(\lambda)] \\ &= \textit{non-negl}(\lambda). \end{aligned} \tag{5.5}$$

Now we construct an adversary $\mathcal{A}'$ playing an instance of `Game 5` where $l = 1$ and $l' = m-1$. In the learning phase $\mathcal{A}$ interacts with the unknown unitary $\tilde{U}$ with the same learning phase states required for $\mathcal{A}$ and sends the query states $\{\rho_i^{out}\}_{i=1}^q$ together with the challenge state $|\phi^b\rangle$ to $\mathcal{A}$. Then $\mathcal{A}$ produces the forgery $|\omega\rangle$ as his guess for $\tilde{U}|\phi^b\rangle$. Now $\mathcal{A}'$ verifies $|\omega\rangle$ with the same test algorithm $\mathcal{T}$ where $\kappa = m-1$, since $\mathcal{A}'$ has $m-1$ copies of $\tilde{U}|\phi^b\rangle$ to check with. Then $\mathcal{A}'$ outputs the same b as outputted by the $\mathcal{T}$. The success probability of $\mathcal{A}'$ is as follows. If $b = 0$, the state is a PRS and the contradiction assumption is satisfied. Hence $\mathcal{A}$'s forgery state will pass the test algorithm with high probability. On the other hand if $b = 1$, the state has been picked from Haar measure and as a result of Theorem 29, the success probability of $\mathcal{A}$ winning the forgery game and producing a state to pass the test is negligible. Since guessing b in Game 5 with probability better than random guess is equivalent to the difference between the success probability of $\mathcal{A}'$ in winning the game in the two different scenarios, we have:

$$\begin{aligned} &|\Pr_{k\leftarrow\mathcal{K}}[\mathcal{A}'(|\phi_k\rangle \otimes (\tilde{U}|\phi_k\rangle)^{\otimes m-1}) = 1] - \Pr_{|\psi\rangle\leftarrow\mu}[\mathcal{A}'(|\psi\rangle \otimes (\tilde{U}|\psi\rangle)^{\otimes m-1}) = 1]| \\ &= |\Pr_{k\leftarrow\mathcal{K}}[\mathcal{A}(|\phi_k\rangle) = 1] - \Pr_{|\psi\rangle\leftarrow\mu}[\mathcal{A}(|\psi\rangle) = 1]| \\ &= \textit{non-negl}(\lambda) - \textit{negl}(\lambda) = \textit{non-negl}(\lambda) \end{aligned} \tag{5.6}$$

Here, as a concrete example, we can consider the GSWAP to be the equality test and, we show how this check can efficiently be performed to show the gap and hence the implication of the two later games. Let us denote the adversary's purified forgery state as $|\omega_b\rangle$. According to Eq. (2.67), the probability of the GSWAP accepting this state given $m-1$ copies of reference state $\tilde{U}|\phi^b\rangle$, has the following relation with the fidelity of the forgery state:

$$\Pr[\text{GSWAP accept}] = \frac{1}{m} + \frac{m-1}{m} F(\tilde{U}|\phi^b\rangle, |\omega_b\rangle)^2 \tag{5.7}$$

Assuming $\mathcal{A}$ wins the unforgeability game for PRS state with non-negligible probability implies that this fidelity is a non-negligible value in the security parameter,

hence $F(\tilde{U}|\phi^0\rangle, |\omega_0\rangle) = \delta = non\text{-}negl(\lambda)$. On the other hand, for Haar-random state this fidelity is always a negligible value and we have that $F(\tilde{U}|\phi^1\rangle, |\omega_1\rangle) = negl(\lambda)$. As a result the difference between $\mathcal{A}$'s success probability in the two cases is as follows:

$$\begin{aligned}
&|\Pr_{k\leftarrow\mathcal{K}}[\mathcal{A}'(|\phi_k\rangle \otimes (\tilde{U}|\phi_k\rangle)^{\otimes m-1}) = 1] - \Pr_{|\psi\rangle\leftarrow\mu}[\mathcal{A}'(|\psi\rangle \otimes (\tilde{U}|\psi\rangle)^{\otimes m-1}) = 1]| \\
&= \frac{1}{m} + \frac{m-1}{m} F(\tilde{U}|\phi^0\rangle, |\omega_0\rangle) - \frac{1}{m} + \frac{m-1}{m} F(\tilde{U}|\phi^1\rangle, |\omega_1\rangle) \qquad (5.8) \\
&= \frac{m-1}{m}(\delta - negl(\lambda)) \approx \frac{m-1}{m}\delta = non\text{-}negl(\lambda)
\end{aligned}$$

As a result, we have shown that there exist a non-negligible gap and hence $\mathcal{A}'$ can also win the `Game 5`. In conclusion, we have shown the following relation:

$$\texttt{Game 2} \equiv \texttt{Game 3} \equiv \texttt{Game 4} \Rightarrow \texttt{Game 5} \Rightarrow \texttt{Game 1} \qquad (5.9)$$

This means that an adversary winning the unforgeability game, with the challenge being picked from a PRS family, can also distinguish PRS states from Haar random states which is a contradiction and concludes the proof. □

We have formally shown that PRS states are enough to achieve quantum universal unforgeability. For completeness let us briefly discuss the construction of these states with the existing proposals. Ji, Liu, and Song [JLS18] propose several constructions for generating a PRS family using classical quantum-secure PRFs. Hence, they show that PRS can be constructed under the assumption that a quantum-secure one-way function exists. Another similar notion called Asymptotically Random State (ARS) has also been introduced in [BS19]. In both works, first, oracle access to a classical random function is given to efficiently construct a PRS, indistinguishable from Haar random states even for exponential adversaries. Then by relying on the existence of quantum-secure one-way functions, they replace the truly random function, with a post-quantum secure PRF to achieve security against polynomial adversaries. With this approach, one can construct computationally secure n-qubit PRS, which is also desired for the unforgeability security property. However, as discussed in [BS20], these methods are not scalable and an n-qubit PRS generator cannot necessarily be employed to produce a random state for k-qubit where $k < n$. For these reasons, in [BS20] the authors introduce a scalable construction for PRSs which, unlike prior works, relies on randomising the amplitudes of the states instead of the phase. The authors use Gaussian sampling methods to efficiently achieve PRSs.

5.3 From pseudorandom unitaries to UU and UqPUFs

We prove that a family of unitaries satisfying the computational assumption of PRU is also a family of unknown unitary transformations. As a result of this implication, efficient constructions such as PRU or t-design can also satisfy the notion of universal unforgeability. Moreover, this result establishes for the first time, a

link between a computational assumption of PRU with a hardware assumption such as unknownness.

Theorem 31. *A family of PRUs, $\mathcal{U} = \{U_k\}_{k\in\mathcal{K}}$ is also a family of unknown unitary (UU).*

Proof. We prove this by contradiction. Let $\mathcal{U}$ be a family of PRUs but not a family of UU which means that there is a quantum polynomial-time (QPT) adversary $\mathcal{A}$ who can estimate the output of a randomly picked $U \leftarrow \mathcal{U}$ where $\mathcal{U}$ is a family of UU, on a state $|\psi\rangle$, non-negligible better than the output of a $U \leftarrow \mu$ picked from a Haar-random unitary μ over a d-dimensional Hilbert space. Thus for $\mathcal{A}$ the following holds:

$$|\Pr_{U\leftarrow\mathcal{U}}[F(\mathcal{A}(|\psi\rangle), U|\psi\rangle) \geqslant \textit{non-negl}(\lambda)] - \Pr_{U_\mu\leftarrow\mu}[F(\mathcal{A}(|\psi\rangle), U_\mu|\psi\rangle) \geqslant \textit{non-negl}(\lambda)]| = \textit{non-negl}(\lambda). \tag{5.10}$$

Let $\mathcal{A}$' be a QPT adversary who aims to break the pseudorandomness property of $\mathcal{U}$ using $\mathcal{A}$, and works as follows:
$\mathcal{A}$' picks $|\psi\rangle$ as one of her chosen inputs in the learning phase of the pseudorandomness game. Then $\mathcal{A}$' also runs $\mathcal{A}$ internally on $|\psi\rangle$.
From the previous equation, we know that $\mathcal{A}$ can estimate the output of $U|\psi\rangle$ better than $U_\mu|\psi\rangle$ where U_μ is a Haar random unitary, by a non-negligible value. Also by definition, we know that the probability that any QPT algorithm estimates the output of any Haar randomly given unitary, is negligible, as the response maps to any random state in the Hilbert space $\mathcal{H}^d$ with exponential distribution [DCEL09, NC10]. Thus the equation implies that:

$$|\Pr_{U\leftarrow\mathcal{U}}[F(\mathcal{A}(|\psi\rangle), U|\psi\rangle) \geqslant \textit{non-negl}(\lambda)]| = \textit{non-negl}(\lambda). \tag{5.11}$$

Meaning that $\mathcal{A}$ can estimate the output with non-negligible fidelity if U had been picked from the family. Now $\mathcal{A}$' runs a quantum equality test on $U|\psi\rangle$ obtained in the learning phase and $\mathcal{A}(|\psi\rangle)$. In the case where U is picked from the PRU family, the estimated output and the real output have non-negligible fidelity, and the test returns equality with a non-negligible probability. Otherwise, the test shows they are not equal, and $\mathcal{A}$' can conclude that the unitary has been picked from Haar unitaries. Thus for $\mathcal{A}$', we have:

$$\Pr_{U\leftarrow\mathcal{U}}[\mathcal{A}'^{U}(1^\lambda) = 1] - \Pr_{U_\mu\leftarrow\mu}[\mathcal{A}'^{U_\mu}(1^\lambda) = 1] = \textit{non-negl}(\lambda) \tag{5.12}$$

Therefore we conclude the contradiction. □

We have shown that PRUs imply unknown unitaries, and combined with the results from the previous chapter, we conclude that PRUs make a set of universally unforgeable unitaries. Now we show that PRU can also be considered a qPUF

family. To do this, we need to show that the PUF requirements given in Definition 39 are satisfied. Since the δ_r-Robustness and δ_c-Collision Resistance are trivially satisfied by the unitarity, we only need to argue about the δ_u-Uniqueness requirement.

Theorem 32. *Let $\mathcal{U} = \{U_k\}_{k\in\mathcal{K}}$ be a family of PRUs, where each U_i is a unitary matrix over a d-dimensional Hilbert space and is universally-unforgeable. Then there exist a $\delta_u = non\text{-}negl(\lambda) = non\text{-}negl(polylog(d))$ such that $\mathcal{U}$ satisfies δ_u-uniqueness.*

Proof. We prove by contraposition and we assume that there exists no non-negligible δ_u to satisfy δ_u-uniqueness. This means that for any two unitary U_i and U_j picked uniformly at random from $\mathcal{U}$, the two unitary are ζ-close in the diamond norm with a high probability. Otherwise if there exist a minimum $\zeta_{min} = non\text{-}negl(\lambda)$ distance in diamond norm between any two unitaries we have already shown the δ_u exists. Hence we assume that we have the following condition:

$$Pr[\| (U_i - U_j)_{i\neq j} \|_\diamond \leqslant \zeta] \geqslant 1 - \varepsilon(\lambda) \tag{5.13}$$

where both ζ and $\varepsilon(\lambda)$ are negligible functions in the security parameter. Now we assume an adversary $\mathcal{A}$ wants to distinguish between $\mathcal{U}$ and the set of Haar-random unitaries. By assumption, we have that all the unitaries in $\mathcal{U}$ are universally unforgeable. So now we let $\mathcal{A}$ play the PRU game[2] while running the universal unforgeability game as a distinguishing subroutine. Let $\mathcal{C}$ be the honest party picking at random a bit $b \in \{0,1\}$ where if $b = 0$, a unitary U is picked at random from $\mathcal{U}$ and we are in the PRU world and otherwise U is picked from μ that denotes the set of Haar-random unitary matrices. Then $\mathcal{A}$ gets polynomial oracle access to the U and after the interaction, needs to guess b. Now, since there exists an efficient public generation algorithm Q for the PRU set, we let the adversary sample another unitary U' from Q locally and uniformly at random. According to the contraposition assumption give in Eq. (5.13), if $b = 0$, with high probability these two unitaries are ζ-close in the diamond norm, *i.e.* $\| (U - U') \|_\diamond \leqslant \zeta$. Given this promise, the adversary performs the following strategy: $\mathcal{A}$ locally plays the universal unforgeability game on U, by picking a state $|\psi\rangle$ uniformly at random from Haar measure and querying it to $\mathcal{C}$ as part of the polynomial oracle interaction with U. $\mathcal{A}$ will receive $U|\psi\rangle$ and can ask for multiple copies of it so long as the total number of queries to the oracle remains polynomial. Now we also rely on the fact that since PRU has the efficient computation property, meaning that $\mathcal{A}$ can locally compute $U'|\psi\rangle$ to get multiple copies. Now $\mathcal{A}$'s strategy to win the unforgeability game is to output $U'|\psi\rangle$ as the forgery for $|\psi\rangle$.

Again in the case of $b = 0$, since the two unitaries are negligibly close in the diamond norm with a high probability we have the following:

$$Pr[\| (U - U') \|_\diamond \leqslant \zeta] \geqslant 1 - \varepsilon \Rightarrow Pr[F(U|\psi\rangle, U'|\psi\rangle) \geqslant 1 - \zeta] \geqslant 1 - \varepsilon \tag{5.14}$$

[2] Which is the indistinguishability game version of the pseudorandomness property, similar to Game 4

This holds since the diamond norm is defined as a maximum over all density matrices. Therefore, if the two unitaries are very close in the diamond norm, their output over a random state is also very close on average. Thus, the adversary can run a local efficient verification test (for instance, a GSWAP test) between $U'|\psi\rangle$ and $U|\psi\rangle$ and use the output of the test as a distinguisher between pseudorandom and Haar-random world. If $b=0$, we have:

$$Pr[F(U|\psi\rangle, U'|\psi\rangle) \geqslant 1-\zeta] \geqslant 1-\varepsilon \Rightarrow Pr[1 \leftarrow \mathcal{G}^{U'}_{\mathsf{qUni}}(\lambda, \mathcal{A})] = non\text{-}negl(\lambda) \tag{5.15}$$

Hence $\mathcal{A}$ will win the game with a high probability. However, in the case of $b=1$ where U is a Haar-random unitary, we can use a lemma in [Kre21], that states for a fixed state $|\phi\rangle \in \mathcal{H}^d$ and a Haar-random state $|\psi\rangle \leftarrow \mu$, and any $\varepsilon > 0$ we have:

$$\Pr_{|\psi\rangle\leftarrow\mu}[|\langle\phi|\psi\rangle|^2 \geqslant \varepsilon] \leqslant e^{-\varepsilon d} \tag{5.16}$$

This implies we can take $U'|\psi\rangle = |\phi\rangle$ to be the fixed state. Since U is a Haar-random unitary then $U|\psi\rangle$ is also a Haar-random state and hence the probability that the fidelity $F(U|\psi\rangle, U'|\psi\rangle)$ is a non-negligible value (with respect to $polylog(d)$) such as $1-\zeta$, is exponentially low. Hence in case $b=1$, the probability that the adversary's state passes the verification is exponentially low. Hence using this strategy, there will always be a distinguisher that can distinguish between $\mathcal{U}$ and Haar-random unitaries *i.e.*:

$$\Pr_{U\leftarrow\mathcal{U}}[\mathcal{A}'^{U}(1^\lambda)=1] - \Pr_{U_\mu\leftarrow\mu}[\mathcal{A}'^{U_\mu}(1^\lambda)=1] = non\text{-}negl(\lambda) \tag{5.17}$$

But this is in contrast with the assumption that $\mathcal{U}$ is a PRU. Hence we have reached a contradiction, and the proof is complete. □

5.4 Pseudorandom unitaries and states from hardware assumptions

As discussed earlier pseudorandom quantum states can be constructed under the assumption of qPRF or quantum one-way functions. Given the relationship that we have explored in the previous section between the unforgeability of qPUF and quantum pseudorandomness, here we ask whether it is possible to construct pseudorandom quantum states under a different set of assumptions? In this section, we discuss how one can achieve PRUs and PRSs under hardware assumptions on a family of unitary transformations. These hardware assumptions are generally discussed in the context of quantum PUFs, nevertheless, our results can be in general applied to any sets of unitaries with the given properties.

Let $\mathcal{U} = \{U_i\}_{i=1}^{\mathcal{K}}$ be a family of unitaries, where each U_i is a unitary matrix over a d-dimensional Hilbert space. Let us bring a specific assumption offered by the physical nature of such unitaries. We want to use the above family as a PRU family or generators for PRS. As shown in [JLS18], if $\mathcal{U}$ is a PRU then it is also a

generators for PRS states *i.e.* $G(k) = U_k|0\rangle = |\phi_k\rangle$. To this end, we investigate the properties of a qPUF family that can be used to achieve pseudorandomness. In the last section, we have shown that PRU implies the notion of unknown unitary assumption, or in other words, single-shot unknownness. Now we explore the relation of PRUs and another notion of unknownness called *practical unknownness* by Kumar et al. [KMK21]. This definition is better suited for t-design unitary sets constructions and is defined as follows:

Definition 44 ($\varepsilon, t, d-$ Practical unkownnness [KMK21])**.** We say a unitary transformations U, from a set $\mathcal{U} \subseteq U(d)$[a] is (ε, t, d)- practically unknown if provided a bounded number $t \leqslant poly(\log_2 d)$ of queries $U\rho U^\dagger$, for any $\rho \in \mathcal{H}^d$, the probability that any $poly(\log_2 d)$-time adversary can perfectly distinguish U from a Haar distributed unitary is upper bounded by $1/2(1 + 0.5\varepsilon)$. Here $0 < \varepsilon < 1$, t are functions of $log_2 d$, and $\lim_{\log_2(d)\to\infty} \varepsilon = 0$.

[a] where $U(d)$ denotes the set of all unitary matrices over d-dimensional Hilbert space

For the sake of our proof, we need a variation of this definition which is for any polynomial number of queries in the security parameter. Hence, we define the following:

Definition 45 ($\varepsilon, d-$ Practical unkownnness)**.** We say a unitary transformations U, from a set $\mathcal{U} \subseteq U(d)$ is (ε, d)- practically unknown if it is (ε, t, d)-practically unknown for any $t = poly(\lambda) = poly(\log d)$.

Now we first show that the assumption of $\varepsilon, d-$ Practical unkownness implies PRU.

Theorem 33. *A family of (ε, d)- practically unknown unitaries where $\varepsilon = negl(\lambda)$ is a PRU family.*

Proof. We prove this by contraposition. Let $\mathcal{U} = \{U_k\}_{i=1}^{\mathcal{K}} \subseteq U(d)$ be a (ε, d)-practically unknown family, that is not a PRU. This means that there exists a QPT adversary $\mathcal{A}$ for which we have the following after some $q = poly(\lambda) = poly(\log(d))$ queries to the unitary oracle:

$$|\Pr_{k\leftarrow\mathcal{K}}[\mathcal{A}^{U_k}(1^\lambda) = 1] - \Pr_{U\leftarrow\mu}[\mathcal{A}^{U}(1^\lambda) = 1]| = \delta = non\text{-}negl(\lambda). \tag{5.18}$$

Equivalently, we can say that if a unitary is randomly picked from either of the set $\mathcal{U}$ or a set of Haar-random distributed unitaries with a random bit b, the advantage of the adversary in guessing bit b is a non-negligible function δ greater than $\frac{1}{2}$. If such an adversary exists, there also exists an adversary $\mathcal{A}'$ that querying the same q states, can distinguish the $U_k \in \mathcal{U}$ from a Haar-random unitary with the following probability:

$$Pr[\text{distinguish } U_k] \geqslant \frac{1}{2} + \delta \tag{5.19}$$

On the other hand, if $\mathcal{U}$ is (ε, d)-practically unknown this probability is equal to $\frac{1}{2}(1+0.5\varepsilon)$ where $\frac{\varepsilon}{4}$ is a negligible function while as δ is non-negligible. Hence we reach a contradiction and the proof is complete. □

We have shown that given the hardware assumption of practical unknownness, over a set of unitary transformations such as unitary qPUFs, one can get PRU and as a result generate PRSs by applying random elements of the set on the computational basis state. However, practical unknownnes is a stronger assumption than UU, and it is not surprising that it will lead to PRU. Now, we look at other properties of a qPUF family and see whether there exists a more interesting assumption under which pseudorandomness can be achieved.

One of the main requirements on a qPUF family is the *uniqueness* property (Requirement 2, Definition 39) that ensures any two qPUFs in the family are sufficiently distinguishable in the diamond norm. In what follows we show a family of unknown and (almost) maximally distinguishable unitary matrices, such as unitary qPUFs, also form a family of PRUs and are a generator for PRSs.

Theorem 34. *Let $\mathcal{U}_\mathcal{K} = \{U_k\}_{k=1}^{\mathcal{K}} \subseteq U(d)$ be a family of unitary transformation selected at random from a distribution $\chi_\mathcal{U}$ such that they satisfy almost maximal uniqueness* i.e. *for any randomly picked pairs of unitary matrices from $\mathcal{U}_\mathcal{K}$, we have $\| (U_i - U_j)_{i\neq j} \|_\diamond = 2-\varepsilon$ where $\varepsilon = negl(\lambda)$, then for a sufficiently large $\mathcal{K}$ and d, the $\mathcal{U}_\mathcal{K}$ is also a PRU.*

Proof. We first show that if the maximum uniqueness is on average satisfied for any pairs of unitary matrices of $\mathcal{U}_\mathcal{K}$, then the distribution $\chi_\mathcal{U}$ converges to Haar measure in the limits of large d. We attempt to prove this convergence for a specific degree of uniqueness which is $2-\varepsilon$ where the maximum of the diamond norm is 2. The general proof idea is to show that the distribution of the eigenvalues of $2-\varepsilon$-distinguishable unitary matrices looks like the eigenvalue distribution of a Haar-random matrix. We use the toolkit from the random matrix theory introduced in Chapter 2 (Section 2.4) to show this statement. First, note that we have,

$$\| (U_i - U_j)_{i\neq j} \|_\diamond = 2-\varepsilon = 2\sqrt{1-\delta(U_i^\dagger U_j)^2} \tag{5.20}$$

Where the $\delta(M) = \min_{|\phi\rangle} |\langle\phi| M |\phi\rangle|$ is the minimum of absolute value over the numerical range of the operator M. From the above equation we have:

$$\delta(U_i^\dagger U_j)^2 = \varepsilon - \frac{\varepsilon^2}{4} \approx 0 \tag{5.21}$$

Since the diamond norm is unitary invariant, we can multiply all the unitaries of the family by a fixed unitary matrix which results in the set including the identity matrix $\mathcal{I}$, hence the above equation can be rewritten as:

$$\delta(U_k')^2 = \varepsilon - \frac{\varepsilon^2}{4} \tag{5.22}$$

where the set of unitary matrices U' is equivalent to the initial set up to a unitary transformation. Now let $\{e^{i\theta_1}, \ldots, e^{i\theta_d}\}$ be the eigenvalues of U'_k. The eigenvalues of a unitary matrix lie on a unit circle $\mathbb{S}^1 \subset \mathcal{C}$. As shown in [KMK21], the following relation exists between the distribution of the eigenvalues of a general unitary matrix in an arc of size θ, and the function $\delta(U)$:

$$\delta(U'_k)^2 = \frac{1}{2} + \frac{1}{2}\cos\theta \tag{5.23}$$

Where $\theta = \theta_j - \theta_k$ for pairs of eigenvalues $\{e^{i\theta_j}, e^{i\theta_k}\}$. From the above equation we have:

$$\theta = \theta_j - \theta_k = \arccos(-1 + 2\varepsilon - \frac{\varepsilon^2}{2}) \approx \pi - \sqrt{\varepsilon} + \ldots \tag{5.24}$$

Now we can use Theorem 9. Let N_θ be a random variable that represents the number of eigenvalues in an arc of size θ. Then we have the expectation value of this random variable for the given distribution where the $\theta = \pi - \varepsilon'$, and $\varepsilon' = negl(\lambda)$, to be

$$\mathbb{E}_d[N_\theta] = \frac{d \times \theta}{2\pi} = \frac{d}{2} - \frac{\varepsilon' d}{2\pi} \tag{5.25}$$

which is close to half of the total number of eigenvalues since the second term is always smaller than 1. This means that in the limit of large d, every diameter of the unit circle divide the circle into two areas that each on average includes half of the eigenvalues. Also the variance of the random variable N_θ will be:

$$Var(N_\theta) = \frac{1}{\pi^2}(\log(d) + 1 + \gamma + \log|2\sin(\frac{\pi - \varepsilon'}{2})|) + o(1) \approx \frac{\log(d)}{\pi^2} + c' + o(1) \tag{5.26}$$

where $\gamma \approx 0.577$ and $c' < 1$. Next, we calculate the probability that for our given distribution, there are more than half of the eigenvalues in each half of the circle denoted by an arc or size $\pi - \varepsilon'$. Using the Chernoff bound we have:

$$Pr[N_{\pi-\varepsilon'} - \mathbb{E}_d[N_{\pi-\varepsilon'}]| > x\mathbb{E}_d[N_{\pi-\varepsilon'}]] \leqslant e^{-\frac{x^2}{2+x}\mathbb{E}_d[N_{\pi-\varepsilon'}]} \tag{5.27}$$

Here we want the $x\mathbb{E}_d[N_{\pi-\varepsilon'}]$ to be equal to $\frac{d}{2}$, so we have $x = \frac{d/2}{d/2 - d\varepsilon'/2\pi} = \frac{1}{1-\varepsilon'/\pi}$ and since the x is a small value the above inequality can be used. Substituting this into the above equation we will have:

$$Pr[N_{\pi-\varepsilon'} - \mathbb{E}_d[N_{\pi-\varepsilon'}]| > \frac{d}{2}] \leqslant e^{-\frac{(\frac{1}{1-\varepsilon'/\pi})^2}{2+\frac{1}{1-\varepsilon'/\pi}} \times (d/2 - \varepsilon' d/2\pi)} \approx e^{-d/6} \tag{5.28}$$

since ε' is negligible. This shows that with a very high probability, on every half of the unit circle, there exist half of the eigenvalues of the random matrix from our specified distribution. We conclude eigenvalues of a random unitary from the distribution $\chi_{\mathcal{U}}$ are uniformly distributed on the unit circle. Let us denote this uniform distribution on $\mathbb{S}^1$ by ν. In order to compare the distribution of $\chi_{\mathcal{U}}$ with the Haar measure, we use the empirical spectral measure introduces in Section 2.4. We denote the empirical spectral distance of $\chi_{\mathcal{U}}$ as $\tilde{\mu}_\chi$ and for

Haar measure we denote it as $\tilde{\mu}_H$. Since we have shown that the eigenvalues of matrices from $\chi_{\mathcal{U}}$ are distributed uniformly on $\mathbb{S}^1$, it is easy to see that $\mathbb{E}(\tilde{\mu}_\chi) = \nu$ and in the limit of large d we have the convergence in probability $\tilde{\mu} \xrightarrow{d\to\infty} \nu$. Now we use the Theorem 8 that implies the convergence of the empirical spectral measure of the set of unitaries picked from Haar measure to ν, in the limit of large d. Having the these two convergence and the properties of the limit we can conclude that the empirical spectral measure for $\chi_{\mathcal{U}}$ converges to the one for Haar measure. Then we look at Kolmogorov distance of the eigenvalues of these two distributions. We rely on the result given in [Mec19] that shows the Kolmogorov distance between the distributions of eigenvalues of random unitary matrices is given by $d_K(\mu,\nu) = \sup_{0\leqslant\theta<2\pi} |\frac{N_\theta}{d} - \frac{\theta}{2\pi}|$ and specifically for Haar measure it is bounded by

$$d_K(\mu_H,\nu) \leqslant c\frac{\log(d)}{d} \tag{5.29}$$

Where $c > 0$ is a universal constant. Given the fact that for the specific value of θ for the distribution of $\chi_{\mathcal{U}}$ the Kolmogorov distance $d_K(\mu_\chi,\nu)$ is of the order $\frac{1}{d}$ which is negligible and using the triangle inequality for the Kolmogorov distance we have

$$\begin{aligned} d_K(\mu_H,\mu_\chi) &\leqslant d_K(\mu_H,\nu) + d_K(\nu,\mu_\chi) \\ &\leqslant c\frac{\log(d)}{d} + negl(\lambda) \\ &\leqslant negl(\lambda) \end{aligned} \tag{5.30}$$

Thus the distribution of the eigenvalues of the random matrices of $\chi_{\mathcal{U}}$ is negligibly close to the Haar measure. Also for any randomly picked matrix from each of these distributions, the eigenvalues are fixed. As a result, the convergence between the distribution of the eigenvalues of matrices leads to the fact that in the limit of large d, $\chi_{\mathcal{U}}$ converges to the Haar measure on the unitary set.

Finally, we show that a polynomial time quantum adversary given a polynomial query to each unknown unitary U_k cannot distinguish any member of this family from Haar measure. This is straightforward since the two distributions are asymptotically close. Thus we have:

$$|\Pr_{k\leftarrow\mathcal{K}}[\mathcal{A}^{U_k}(1^\lambda) = 1] - \Pr_{U\leftarrow\mu}[\mathcal{A}^{U}(1^\lambda) = 1]| = negl(\lambda). \tag{5.31}$$

And we have shown that the set $\mathcal{U}_\mathcal{K}$ is a PRU. □

5.5 Discussion and conclusions

We have explored in this chapter, the connection between quantum pseudorandomness and quantum hardware assumptions such as quantum physical unclonability. As one of the main cryptographic properties of quantum physical unclonable functions is the notion of universal unforgeability, we have inspected whether quantum pseudorandomness would be enough as a challenge sampling requirement, to

achieve this level of unforgeability. We have formally proved that the answer to this question is positive. This result improves the practicality of qPUF-based constructions and protocols since it replaces the requirement of Haar-randomness on the challenge states, which is resourceful and experimentally challenging. We will articulate this improvement in the next chapter.

We have also established the link between the notions of *unknownness* of unitary families and PRUs. We proved that any family of PRUs is also a family of unknown unitaries and, hence they could be a potential candidate for the construction of qPUF devices. This result complements the result of [KMK21] where they show t-designs can also satisfy a similar notion, namely practical unknownness, which leads to an efficient proposal for constructing quantum PUFs.

Then we also looked at the problem of generating pseudorandom quantum states from hardware assumptions. Our results show that different physical assumptions proposed in the context of PUFs, such as uniqueness or practical unknownnes, can also imply quantum pseudorandomness. This result is of theoretical interest as it shows an alternative way of achieving quantum pseudorandomness which is different from current approaches based on post-quantum and computational assumptions. Apart from the cryptography perspective, having a different set of assumptions for PRSs and PRUs can find potential applications in physics where PRSs have been shown recently to be relevant in the AdS/CFT correspondence for the study of quantum gravity [BFV19]. Nonetheless, due to arguments given in [BFV19], the proposed PRS constructions in [JLS18], are not directly applicable within this framework. Given our results in this chapter, a potential follow-up question would be whether a PRS state derived from physical unclonability assumptions can be used as an alternative solution. Another interesting future direction would be to further explore the relationship between unclonability and quantum pseudorandomness that has initially been discussed in [JLS18], relying upon our new results.

The final open problem that we would like to bring forward to conclude the chapter, is that of establishing concrete bounds on the randomness and pseudorandomness of unitary families, given different degrees of uniqueness or distinguishability (not negligibly close to perfect distinguishability). We believe this question has an interesting and non-trivial relationship to the study of t-design unitaries. A curious inquiry is whether the random matrix theory toolkit and the potential extension of our last result in the current chapter, regarding the relationship between distinguishability and pseudorandomness can also lead to novel constructions for t-designs.

6

Applications of Quantum Physical Unclonable Functions

"If you wish to make an apple pie from scratch, you must first invent the universe."

– Carl Sagan

6.1 Introduction

In the last two chapters, we have studied quantum physical unclonable functions, both as theoretical objects and provably unforgeable hardware tokens, while also exploring the connection between physical unclonability and quantum pseudorandomness. Moving from foundations to applications, it is now time to introduce applications of qPUFs in quantum communication and quantum cryptography. This chapter is dedicated to the design and security analysis of protocols based on quantum PUFs. In the course of the chapter, we also attempt to move towards more efficient variants of the proposed protocols and make them more accessible for implementation.

The recent advances in developing the quantum internet have enabled a broad range of applications from simple secure communications all the way to delegated quantum computation, with often no counterparts in classical networks [BS16, WEH18, Fit17, Ver19, PPA+20, Dia19, DWT+19, KDW20, CCB18, CCT+20, Unr13]. For most of such applications, a key security feature is the ability of secure authentication which plays a central role in performing secure communications over untrusted channels [AM17, DGJ+20, BZ13a]. The general term of *authentication* encloses different definitions and levels depending on the strength of the security requirement and the nature of the subject of authentication, for instance, whether it is a message or an entity. Amongst different types of required security features, including confidentiality and authentication of data, mutual entity authentication is a crucial, yet sometimes neglected, aspect [KHH+18, Gol96]. Entity authentication also referred to as *Identification*, is a

method to prove the identity of one party called *prover* to another party called *verifier*.

The focus of this chapter is on secure identification as it is a central application of quantum communication, as well as a building block for many other applications of quantum networks. We aim to propose resource-efficient solutions for mutual entity authentication between two parties who can also be two nodes of a quantum network. We explore the advantages of quantum communication in achieving protocols with fewer assumptions or stronger security guarantees compared to their classical counterparts or existing solutions.

We consider both complementary scenarios where either the trusted verifier or a potentially malicious prover has limited resources. To better motivate the two scenarios, consider the quantum cloud service platforms that are commercially available today [AAB+19, Cro18, Rig, BIS+20, BWM21]. In the first setting, a client with a low quantum resource (such as the one defined in [BFK09]) wishes to identify a high-resource quantum centre that they perhaps have had a previous contract with, before proceeding to access their platform and load its sensitive data. In the complimentary setting, the quantum cloud provider wishes to verify the identity of its customer possessing low quantum resources before providing them with access. This asymmetry between the verifier and the prover calls for 'party resource-specific' identification protocols which exploit this asymmetry to enhance the efficiency. Another potential approach is include the mutual identification within one protocol which requires symmetrizing the parties as much as possible. We will explore both of these approaches via our proposals in this chapter.

Most of the typical classical solutions for authentication and identification rely on computational assumptions or a perfectly random key being securely shared between the two parties, or in some cases, both. Throughout this chapter, we replace these computational assumptions, or secure classical key sharing, with the hardware assumption of physical unclonability. The protocols have the structure of a symmetric-key protocol, although the *key* here is some unclonable hardware. Another prominent aspect of our proposals is the employment of quantum communication. Similar to most functionalities and protocols, if one wishes for the security in the quantum era, there are usually two options available: either to go for the post-quantum alternatives and use assumptions that are believed to be hard for quantum computers, or to take advantage of the power of quantum information and quantum communication to attain quantum security. We focus on the second option here, for achieving provable security against quantum adversaries under minimal assumptions. We note that each of these approaches has its pros and cons, and the comparison between them is not the purpose of this chapter. We remark that the spirit of the works presented in this thesis is closer to exploiting the physical properties and fundamental limitations of quantum mechanics in the design of protocols.

First, we propose two entity authentication protocols based on the quantum PUFs that we have defined in Chapter 4. Our first proposal is a secure qPUF-based device identification protocol which requires the prover to only have access

to the valid qPUF device without requiring any quantum memory or quantum computational resource, while the verifier is required to possess a local quantum database and the ability to perform quantum operations. This covers the scenario presented before where a quantum cloud provider wants to identify its customer.

Our second proposal is a qPUF based protocol where the prover has a high computational resource, while, the verifier runs a purely classical algorithm, hence does not require performing quantum operations. This protocol can enable an *almost classical* client, to identify a quantum server in a quantum network. Construction of this protocol has taken inspiration from the ideas of blind quantum computing [BFK09] to introduce the idea of randomly placing trap quantum states in-between the valid states. This, coupled with the unknownness property of the qPUF device provides provable security against any QPT adversary. We also provide a comparison between the two protocols on different aspects and resources to get a better picture of their use-case in different scenarios.

Next, we exploit the result we have established in Chapter 5, to improve the efficiency of our protocols and make them more amenable to implementation, while formally proving that this step, on the way to practicality can be made with no compromise in the security guarantee.

Finally, in attempting to propose a yet more practical solution, we explore a different construction for PUFs, which although weaker than full qPUFs, can still enable secure quantum entity authentication while also being implementable with the technology and infrastructures that are available today. This new construction, called *Hybrid PUF*, combines classical PUFs with quantum encoding and using some additional techniques from the world of classical hardware security, can lead to quantum-secure mutual identification that does not require quantum database or preparation of resourcefully complicated quantum states. The latest protocol we present has some further properties, such as the re-usability of challenge states during the protocol, which we will investigate in detail.

We believe that all these proposed protocols are just the start of the road for applications that utilize physical unclonability and quantum information since identification is essential yet quite a simple functionality. Our studies presented in this chapter show that there are still many applications to come using this newly introduced assumption.

6.1.1 Structure of the chapter

In Section 6.2 we present our client-server qPUF-based identification protocols. The protocol with high-resource verifier has been introduced in Subsection 6.2.2, and the low-resource verifier protocol in Subsection 6.2.3. A generalisation of the second protocol is also discussed in Subsection 6.2.4 and the comprehensive comparison between them is given in Subsection 6.2.5.

In Section 6.3 we use pseudorandom quantum states to reduce the assumptions and requirements for our proposed protocol and introduce a more efficient version.

Finally, Section 6.4 focuses on presenting the Hybrid PUF construction as well as the identification protocol that is based on it and its security analysis. In Sub-

section 6.4.3 the construction is given. In Subsection 6.4.3 an enhanced version this construction called *Hybrid Locked PUF* has been introduced which later is used within the identification protocol presented in Subsection 6.4.4. Subsection 6.4.5 discussed the security analysis of HPUF, HLPUF and related protocol and Subsection 6.4.6 investigates the challenge re-usability property.

6.1.2 Related works

The idea of taking advantage of quantum communication between the verifier and the prover in PUF-based identification protocols was first introduced by Skoric in [Sko10] with the concept of *quantum read-out of PUF (QR-PUF)*. The identification protocols based on this construction have been proposed in [Sko10, Sko12, GKB20, Nik21]. The security of the majority of these protocols has been proved against limited types of attacks including intercept-resend [Sko10, Sko12], and Quantum Cloning [YGLZ16] attacks. The practical realization of this protocol was shown by Goorden et al. [GHM+14, Nik21]. In another work (also mentioned in Chapter 4), Nikolopoulos and Diamanti introduce a different setup for QR-PUF-based identification protocols in which classical data is encoded to the continuous quadrature components of the quantized electromagnetic field of the probe [ND17]. The security of this scheme has also been proved in [Nik18, FNAF19] against a bounded adversary who can only prepare and measure the quantum states. The common feature of the mentioned protocols is full or partial knowledge of the unitary modelling of the QR-PUF. However, as thoroughly discussed in 4.4.4, this extra information usually compromises the security and as a result, such protocols can only be proven secure against specific types of adversarial attacks. The main advantage of our qPUF-based proposals over the previous ones is their provable security against the most general form of attacks considering a QPT adversary.

Related to our Hybrid construction, first we mention some classical constructions for classical PUFs such as [GCvDD02, GKST07, KL18]. The literature of classical PUFs, specifically regarding the implementation, is very vast and covering the full references to them is outside the scope of this chapter, but we refer the reader to [Mae13] for a detailed review of the constructions of classical PUFs. We have also mentioned that most such classical PUFs are vulnerable to machine learning attacks. Some of the attacks on classical PUFs have been performed and studied in [Bec15a, Bec15b, Del19, RSS+10]. Furthermore, we also borrow an idea from classical hardware security literature, known as the lockdown technique (or as we call them, locking mechanism), that has been introduced by Yu et al.[YHD+16] as a proposal to prevent such machine learning attacks on classical PUFs.

6.2 Quantum-secure identification protocols using quantum PUF

In this section, we aim to provide protocols for the task of identification, using quantum PUF and quantum communication as our main ingredients. Intending to perform low-cost secure identification of the prover by the verifier using qPUF, we categorise the resources into three major segments. First is the 'memory resource' which quantifies the type and amount of storage resources that a party requires. It can either be a classical memory that we label as low cost, or a quantum memory which is high cost since such a memory tends to be highly fragile and dissipative to the environment [LST09]. Second is the 'computing ability' resource which indicates the kind of operations a given party can perform. We denote a party with high computing ability as the one that can perform any bounded polynomial quantum circuit operations [Wat03], and a low ability party as the one that is restricted to generation and measurement of quantum states on a certain basis. And the third resource is the type and number of 'communication rounds' required between the parties to establish identification. Often it is not possible to devise an identification scheme that minimises all the three types of resources for both the involved parties without compromising the underlying security. Hence, in this work, we propose two qPUF based identification schemes that achieve similar security guarantees but are vastly different in terms of the resource requirement for the involved parties. This allows the flexibility to deploy either of these schemes specific to each application.

The first protocol allows a low-resource party who has only access to the qPUF, to prove its identity to a high resource party with more quantum capabilities such as quantum computing capability and quantum memory. In the second protocol, we explore the other direction and try to minimize the resources on the verifier's side as much as possible. This leads to a novel qPUF-based protocol, which is different from the usual PUF-based protocols known in the literature. We give complete formal security proofs for each of the protocols. Then, we also provide a comprehensive comparison between the two proposed protocols in terms of our categorised resources.

But before introducing the protocols, let us give a general description of an identification protocol to provide a better intuition of the functionality we are trying to achieve.

6.2.1 General description of device-based identification protocol

An identification protocol, also called a device-authentication protocol, is run between a verifier and a prover. A verifier's task is to check the identity of the prover by identifying whether the prover is the correct owner of a valid device. Our setting assumes that the verifier and the prover having a valid device behave honestly. The security is provided against an adversary with limited access to

the valid device[1]. The objective of the adversary is to successfully impersonate themselves as the valid owner of the device. Prior to providing the details of the construction of device identification protocols using qPUF, we describe a common structure in these protocols. Any such protocol consists of three sequential phases: *setup phase* (or enrollment phase), *identification phase* and *verification phase* [ND17, PRTG02, GKB20].

1. *Setup phase*: A setup phase is the beginning phase of the protocol. Here the verifier has the valid device (in this case a PUF/qPUF) and locally prepares a database consisting of multiple challenge and response pairs of this device. The challenges and responses, namely Challenge-Response pairs (CRPs) are stored in the verifier's local database. We assume that the verifier's quantum capabilities are restricted to quantum polynomial time, and a polynomial-size database. Once the local database is generated, the device is physically transferred to the prover over a public channel.

2. *Identification phase*: The setup phase is followed by the identification phase where the verifier sends one or multiple challenges, usually chosen at random, to the prover from the CRP database. The challenge(s) is sent over a public (quantum) channel to the prover. The prover who has the valid device obtains the responses to the received challenges by querying the device and obtaining the response. Then the prover sends either the response directly, or sends some classical or quantum information related to the response to the verifier. We note that qPUF-based identification protocols would mostly differ in this phase by varying the number of challenges sent to the prover and the type of information received by the verifier.

3. *Verification phase*: In the verification phase, the verifier runs a quantum or classical verification algorithm on the information received from the prover. We denote that the verifier correctly identifies the prover if the verification algorithm outputs 1. Otherwise, it aborts.

The **Correctness** or **Completeness** of an identification protocol is defined as the success probability of an honest prover over multiple rounds of identification, in the absence of any adversary or noise, should be one. The **Soundness** of an identification protocol ensures that the success probability of any adversary (depending on the adversarial model) in passing the verification phase over the multiple rounds of identification, should be negligible in the security parameter.

6.2.2 Quantum identification protocol with high-resource verifier

The identification protocol runs between a verifier and a prover. The verifier is tasked with correctly identifying the prover who owns the device. Our setting

[1]This the same QPT adversarial model we have described in the universal unforgeability game in Chapter 4. This is a primitive-level access, while as in the level of our protocols we (usually) do not assume any bound on the adversary.

assumes that the verifier and device owner behave honestly. The security has been shown against an adversary (computationally bounded in the learning phase) willing to be identified as the valid device owner. We propose the construction of two identification protocols using qPUFs which provide exponential security against any QPT adversary. The qPUF used in this protocol is a UqPUF as defined in Definition 40.

The first qPUF-based device identification protocol we propose is the quantum analogue of the standard PUF-based identification scheme between the verifier (Alice) and the prover (Bob) as shown in Fig. 6.1. Before detailing the protocol, we list its salient features,

- The prover is not required to have quantum memory, and computing ability resources[2], whereas the verifier is required to have high quantum memory and high computing ability resources (restricted to polynomial-size memory and QPT computation).
- The protocol requires a 2-way quantum communication link between the prover and verifier.
- The protocol has a quantum verification phase *i.e.* the prover sends information in quantum states to the verifier who then performs a verification test to certify if the device is valid.
- The protocol provides perfect completeness and an exponentially-high security guarantee against any adversary with QPT resources.

6.2.2.1 Protocol description

This protocol, referred as hrv-id, is run between the verifier, and the prover and it is divided into three sequential phases,

Protocol 1 (hrv-id(K,N,M,D))**.** qPUF-based Identification protocol with high-resource verifier:[3]

---

1. *Setup phase:*
   (a) Verifier has the qPUF device.
   (b) Verifier randomly picks $K \in \mathcal{O}(\text{poly}\log D)$ classical strings $\phi_i \in \{0,1\}^{\log D}$.
   (c) Verifier selects and applies a Haar-random state generator operation denoted by the channel $\mathcal{E}_{prep}$ to locally create the corresponding quantum states in $\mathcal{H}^D$: $\phi_i \overset{\mathcal{E}_{prep}}{\rightarrow} |\phi_i^c\rangle$, $\forall i \in [K]$.

[2]Here we note that the prover applies the qPUF transformation (the unitary) on the challenge states. Nevertheless, we do not consider this as computing ability of the prover, thus by no computing ability we refer to the fact that the prover does not need to run any extra quantum computations other than the physical interaction with the qPUF hardware.

[3]We drop the parameters (K, N, M, D) from now on for simplicity whenever we refer to this protocol.

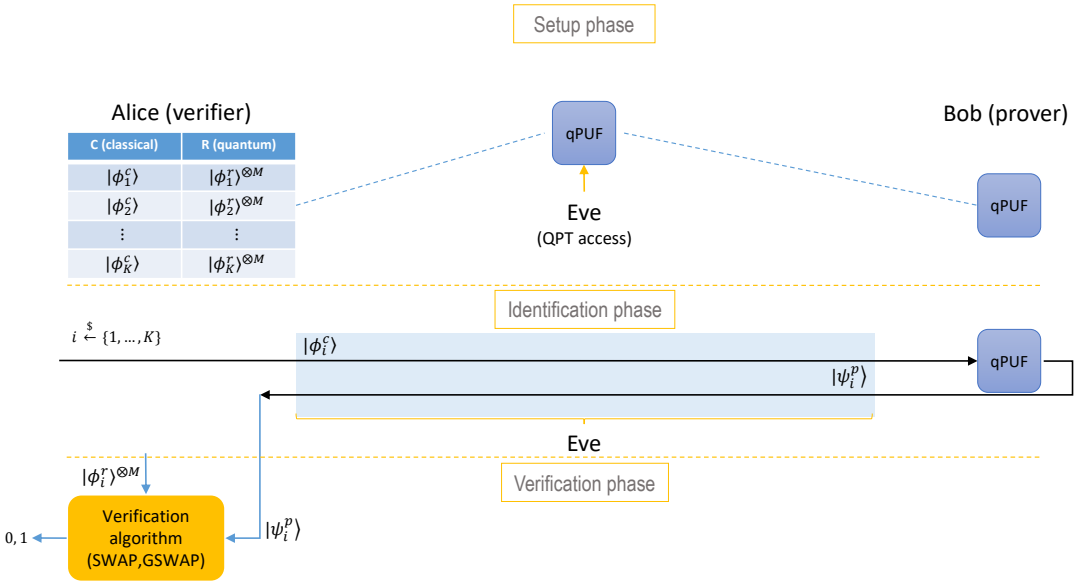

Figure 6.1: qPUF-based identification protocol with high-resource verification between Alice (verifier) and Bob (prover) (hrv-id). The protocol is divided into three sequential phases, *setup phase*, *identification phase*, and *verification phase*. The protocol is analysed in the presence of a QPT adversary (Eve) which can gain information about the device during the *setup phase*. In the last phase, verifier runs a quantum verification algorithm and outputs a classical bit '1' if prover's device is correctly identified. Otherwise, verifier outputs '0'.

(d) Verifier queries the qPUF individually with each challenge $|\phi_i^c\rangle$ a total of M number of times to obtain M copies of the response state $|\phi_i^r\rangle$ and stores them in their local database $S \equiv \{|\phi_i^c\rangle, |\phi_i^r\rangle^{\otimes M}\}_{i=1}^{K}$.

(e) Verifier publicly transfers the qPUF to prover.

To be able to investigate the security in a strong and general setting, we do not assume the qPUF's transition to be done securely, in the sense that any QPT adversary (Eve) is allowed to query the qPUF during transition an $\mathcal{O}(\text{poly}\log D)$ number of times and thus build its local database. Due to the conditions on universal unforgeability of the qPUF, it is important that verifier picks the challenges $|\phi_i^c\rangle \in S$ at random from a uniform distribution over the Hilbert space $\mathcal{H}^D$. This, in turn, implies that the encoding unitary operation $\mathcal{E}_{prep}$ is a Haar random unitary. We relax this condition in the upcoming section of this chapter.

2. *Identification phase*:

(a) Verifier uniformly selects a challenge labelled ($i \xleftarrow{\$} [K]$), and sends the state $|\phi_i^c\rangle$ over a public quantum channel to prover.

(b) Prover generates the output $|\phi_i^p\rangle$ by querying to the qPUF device, the challenge received from the verifier.

(c) The output state $|\phi_i^p\rangle$ is sent to verifier over a public quantum channel.

(d) This procedure is repeated with the same or different states a total of $R \leqslant K$ times[4].

3. *Verification phase*:

(a) Verifier runs a quantum equality test algorithm on the received response from the prover and the M copies of the correct response that exists in the database. This algorithm is run for all the R total number of CRP pairs.

(b) Verifier outputs '1' implying successful identification if the test algorithm returns '1' on all CRPs. Otherwise, outputs '0'.

Sections 6.2.2.2 and 6.2.2.3 describe the quantum verification algorithm run by the verifier.

For this protocol, we define the security in terms of completeness and soundness properties. Completeness of hrv-id protocol is the probability that verifier outputs '1' in the verification phase in absence of an adversary Eve. This implies that the verification algorithm must output '1' for all the R rounds of the protocol with a probability that differs negligibly in the security parameter from 1,

$$\Pr[\text{Ver accept}_{\text{H}}] = \Pr\Big[\prod_{i=1}^{R} (\texttt{qVer}(|\phi_i^p\rangle, |\phi_i^r\rangle) = 1)\Big] = 1 - negl(\lambda) \tag{6.1}$$

where the subscript H denotes the honest device holder.

Soundness of the protocol is defined as the probability that a QPT adversary (Eve) passes the verification test. We say the hrv-id is sound (or secure) if this probability is negligible in the security parameter:

$$\Pr[\text{Ver accept}_{\text{Eve}}] = \Pr\Big[\prod_{i=1}^{R} (\texttt{qVer}(\rho_i, |\phi_i^r\rangle) = 1)\Big] = negl(\lambda) \tag{6.2}$$

where ρ_i is the state sent by adversary in the i-th round.

Since our protocol is based on UqPUF, the verifier has no knowledge about the unitary of qPUF except the database S which can be obtained by querying. Consequently, the responses stored in S are unknown quantum states. This calls for quantum equality test verification algorithms to enable the verifier to validate the received states. We investigate the optimal one-sided error test, the SWAP test [BCWdW01] and the GSWAP [CDM+18] as described in Chapter 2, Section 2.2.1.

6.2.2.2 Verification with SWAP test

The first proposal for verifier's qVer algorithm is the SWAP test and the identification protocol using this test is called hrv-id-swap. Its single run inputs one copy

[4] $R = M \times N$

of each received state and verifier's response state and produces a binary outcome to determine the equality between two states. A single run, however, does not provide a low enough test error rate. To obtain an exponentially low rate, the test is repeated M times for the same challenge state where M is proportional to the inverse-log of the desired error probability. The error can be further decreased by choosing $N \leqslant K$ distinct challenge states such that the test is run for $R = N \times M$ number of times and the prover is successfully identified, only if he passes all the runs. In the next two theorems, we show that the SWAP-based test algorithm provides us with the desired completeness and soundness properties required in the protocol.

Theorem 35 (**Completeness of hrv-id with SWAP**). *In absence of an adversary Eve, the probability that the response state of an honest prover* $|\phi_i^p\rangle = U_{qPUF}\,|\phi_i^c\rangle$*, generated from the valid qPUF, passes all the* R SWAP *test runs is,*

$$\Pr[\text{Ver accept}_{\text{H}}] = \Pr\Big[\prod_{i=1}^{R}(\text{SWAP}(|\phi_i^p\rangle, |\phi_i^r\rangle) = 1)\Big] = 1 \tag{6.3}$$

Proof. When verifier receives prover's response $|\phi_i^p\rangle$ which is generated from the valid qPUF device for all the $i \in [R]$ copies of the challenge state, then $|\phi_i^p\rangle = |\phi_i^r\rangle$. This implies that $F(|\phi_i^p\rangle, |\phi_i^r\rangle) = 1$ for all $i \in [R]$. From Eq. (2.65), we see that,

$$\Pr\big[(\text{SWAP}(|\phi_i^p\rangle, |\phi_i^r\rangle) = 1] = \frac{1}{2} + \frac{1}{2}F(|\phi_i^p\rangle, |\phi_i^r\rangle) = 1, \;\; \forall i \in [R] \tag{6.4}$$

Since in the honest setting, the states received from prover over R rounds are all valid qPUF pure states which are unentangled to each other, hence the SWAP tests for all the R rounds are independent tests. This implies that,

$$\begin{aligned}\Pr[\text{Ver accept}_{\text{H}}] &= \Pr\Big[\prod_{i=1}^{R}(\text{SWAP}(|\phi_i^p\rangle, |\phi_i^r\rangle) = 1)\Big] \\ &= \prod_{i=1}^{R}\Pr\big[\text{SWAP}(|\phi_i^p\rangle, |\phi_i^r\rangle) = 1\big] \\ &= 1\end{aligned} \tag{6.5}$$

This completes the proof. □

To characterise the soundness, we bound Eve's success probability in passing the verification test *i.e.* the probability that the state ρ^R she sends to verifier passes all the R runs of the SWAP test. Even though the test runs are independent, if a generalised entangled state ρ^R is sent by Eve, her success probability across the runs may no longer be the product of success probability of individual test runs. This implies that Eve's strategy might result in a higher success probability in some rounds based on the results of previous rounds. However, we show that since the N distinct challenges being picked by verifier are all uniformly

random, Eve does not gain anything by entangling the states across rounds corresponding to different challenges. To this end, we assume Eve can achieve optimal success probability by sending the state $\bigotimes_{i=1}^{N}\rho_i^M$, where ρ_i^M is a generalised state sent to M runs of the SWAP test corresponding to the same challenge $|\phi_i^c\rangle$. Across these $j \in [M]$ runs corresponding to $|\phi_i^c\rangle$, the state received by verifier is $\rho_{i,j} = \mathrm{Tr}_{\{1\cdots M/j\}}(\rho_i^M)$, where $\rho_{i,j}$ is obtained by tracing out the M-1 instances $\{1,\cdots M/j\}$. Let $\rho_i^{\max}$ be Eve's response state corresponding to challenge $|\phi_i^c\rangle$, with the highest fidelity with the correct response, *i.e.*

$$\forall j \in M \quad F(\rho_i^{\max}, |\phi_i^r\rangle) = \langle\phi_i^r|\,\rho_i^{\max}\,|\phi_i^r\rangle \geqslant \langle\phi_i^r|\,\rho_{i,j}\,|\phi_i^r\rangle \tag{6.6}$$

Since the SWAP test success probability is directly proportional to the fidelity between the two input states, this implies that Eve can maximise her success probability by sending M unentangled states $\rho_i^{\max}$ to verifier instead of the generalised state ρ_i^M. The above equation Eq. (6.6) can be used to bound Eve's success probability in passing verifier's verification test,

$$\begin{aligned}
\Pr[\text{Ver accept}_{\text{Eve}}] &= \Pr\Big[\prod_{i=1}^{R}(\text{SWAP}(\rho_i, |\phi_i^r\rangle) = 1)\Big] \\
&= \prod_{i=1}^{N}\Pr\Big[\prod_{j=1}^{M}(\text{SWAP}(\rho_{i,j}, |\phi_i^r\rangle) = 1)\Big] \\
&\leqslant \prod_{i=1}^{N}\prod_{j=1}^{M}\Pr\big[\text{SWAP}(\rho_i^{\max}, |\phi_i^r\rangle) = 1\big] \\
&\leqslant \prod_{i=1}^{N}\Big(\frac{1}{2} + \frac{1}{2}F_i\Big)^M = \varepsilon
\end{aligned} \tag{6.7}$$

where $\rho_i = \mathrm{Tr}_{\{1\cdots R/i\}}(\rho^R)$, and $F_i = F(\rho_i^{\max}, |\phi_i^r\rangle)$.

Now using the fact that the qPUF device exhibits universal unforgeability against any QPT adversary (Theorem 29), we bound the success probability of Eve using the following theorem.

Theorem 36 (**Soundness of hrv-id with SWAP**)**.** *Let qPUF be a universally unforgeable UqPUF over $\mathcal{H}^D$. The success probability of any QPT adversary Eve, to pass the* SWAP*-test based verification of the hrv-id-swap protocol is at most ε, given that there are N different CRPs, each with M copies. The ε is bounded as follows:*

$$\Pr[\text{Ver accept}_{\text{Eve}}] \leqslant \varepsilon \approx \mathcal{O}(\frac{1}{2^{NM}}) \tag{6.8}$$

Proof. From Eq. (6.7), we see that the optimal strategy of Eve is to produce the response states $\rho_i^{\max}$ which maximises the fidelity F_i for each CRP $(|\phi_i^c\rangle, |\phi_i^r\rangle^{\otimes M})$. We provided an upper bound on the fidelity when Eve has polynomial access to

the qPUF in Theorem 28 stating that the fidelity F_i is bounded as,

$$\Pr[F_i \geqslant \delta] \leqslant \frac{d+1}{D} \tag{6.9}$$

for any $\delta > 0$. Here $d = poly(\lambda) = poly\log(D)$ is the dimension of subspace that Eve has learnt from $\mathcal{H}^D$. For $D = 2^d$, this implies that the maximum fidelity state that Eve can create on average is non-orthogonal to the valid response state $|\phi_i^r\rangle$ with a negligible probability $\approx \mathcal{O}(2^{-d})$. Hence $F_i = \delta \to 0$ with overwhelming probability. This bound holds true for all distinct CRPs labelled by $i \in [N]$.

Thus from Eq. (6.7) and Eq. (6.9), the probability that Eve passes verifier's SWAP based verification test is,

$$\begin{aligned}\Pr[\text{Ver accept}_{\text{Eve}}] &\leqslant \prod_{i=1}^{N} \left(\frac{1}{2} + \frac{1}{2}F_i\right)^M \\ &\leqslant \prod_{i=1}^{N} \left(\frac{1}{2} + \frac{1}{2}\delta\right)^M \\ &\approx \mathcal{O}(\frac{1}{2^{NM}}) = negl(\lambda)\end{aligned} \tag{6.10}$$

Note that here we also take into account the adaptive strategy of the adversary. That is even by assuming the previous rounds are added as extra states to Eve's learning phase, the dimension of the subspace d will remain polynomial in λ. This completes the proof. □

The bound indicated above shows that one can achieve an exponentially secure qPUF-based identification using SWAP test based verification protocol with just a single challenge state *i.e.* $N = 1$ and repeated for M instances. However, non-ideal cases would make identification with different challenge states necessary. Hence we provide a general recipe involving multiple distinct challenges each running for multiple instances. Our protocol requires $R = N \times M$ number of rounds and uses $T = 2R$ number of communicated states.

6.2.2.3 Verification with GSWAP test

The second proposal for verifier's `qVer` algorithm is the GSWAP test and the identification protocol using this test is called hrv-id-gswap. Its single run requires one copy of the received state and M copies of verifier's response state and produces a binary outcome to determine the equality between two states with a polynomial one-sided error *i.e.* $\propto 1/M$. To boost the security to exponentially low error with a polynomial number of copies, the verifier first runs the challenge phase with $R = N \subset K$ distinct challenge states, then uses the GSAWP test as `qVer` algorithm to test the equality. To this end, she consumes N received response states and $N \times M$ numbers of valid response states in her database. In the next two theorems, we show that GSWAP based test algorithm provides us with the desired completeness and soundness properties required in the protocol.

Theorem 37 (**Completeness of hrv-id with GSWAP**). *In absence of an adversary Eve, the probability that the response state of an honest prover, $|\phi_i^p\rangle = U_{qPUF}\,|\phi_i^c\rangle$ generated from the valid UqPUF passes all the $R = N$ test runs is,*

$$\Pr[\text{Ver accept}_{\text{H}}] = \Pr\Big[\prod_{i=1}^{N}(\text{GSWAP}(|\phi_i^p\rangle\,,|\phi_i^r\rangle^{\otimes M}) = 1)\Big] = 1 \tag{6.11}$$

Proof. When verifier receives prover's response $|\phi_i^p\rangle$ which is generated from the valid qPUF device for all the $i \in [R]$ copies of the challenge state, then $|\phi_i^p\rangle = |\phi_i^r\rangle$. This implies that $F(|\phi_i^p\rangle\,,|\phi_i^r\rangle) = 1$ for all $i \in [R]$. From Eq 2.67, we see that,

$$\Pr\big[(\text{GSWAP}(|\phi_i^p\rangle\,,|\phi_i^r\rangle^{\otimes M}) = 1] = \frac{1}{M+1} + \frac{M}{M+1}F(|\phi_i^p\rangle\,,|\phi_i^r\rangle) = 1,\ \ \forall i \in [N] \tag{6.12}$$

Since in the honest setting, the states received from prover over R rounds are all valid qPUF pure states which are unentangled to each other, hence the GSWAP tests for all the R rounds are independent tests. This implies that,

$$\begin{aligned}\Pr[\text{Ver accept}_{\text{H}}] &= \Pr\Big[\prod_{i=1}^{N}(\text{GSWAP}(|\phi_i^p\rangle\,,|\phi_i^r\rangle) = 1)\Big] \\ &= \prod_{i=1}^{N}\Pr\big[\text{GSWAP}(|\phi_i^p\rangle\,,|\phi_i^r\rangle) = 1\big] \\ &= 1\end{aligned} \tag{6.13}$$

This completes the proof. □

To characterise the soundness, we bound Eve's success probability in simultaneously passing the N runs of GSWAP test when she sends the generalised entangled state ρ^N to verifier. Similar to the argument provided for SWAP test soundness, Eve does not gain anything by entangling the states across different test runs. Thus Eve's probability in passing the verification test by sending the state $\bigotimes_{i=1}^{N}\rho_i$ is the same as that for a generalised state ρ^N, where ρ_i is the state sent to the instance of GSWAP test corresponding to the same challenge $|\phi_i^c\rangle$. As a result, Eve's optimal success probability can be expressed as a product of individual GSWAP instance success probability,

$$\begin{aligned}\Pr[\text{Ver accept}_{\text{Eve}}] &= \Pr\Big[\prod_{i=1}^{N}(\text{GSWAP}(\rho_i, |\phi_i^r\rangle^{\otimes M}) = 1)\Big] \\ &= \prod_{i=1}^{N}\Pr\big[\text{GSWAP}(\rho_i, |\phi_i^r\rangle^{\otimes M}) = 1\big] \\ &\leqslant \prod_{i=1}^{N}\Big(\frac{1}{M+1} + \frac{M}{M+1}F_i\Big) = \varepsilon\end{aligned} \tag{6.14}$$

where $F_i = F(\rho_i, |\phi_i^r\rangle)$ is the fidelity between Eve's state and the valid qPUF response state for the i-th round.

Theorem 38 (**Soundness of hrv-id with GSWAP**). *Let qPUF be a universally unforgeable UqPUF over $\mathcal{H}^D$. The success probability of any QPT adversary Eve, to pass the* GSWAP*-test based verification of the hrv-id-gswap protocol is at most ε, given that there are N different CRPs, each with M copies. The ε is bounded as follows:*

$$\Pr[\text{Ver accept}_{\text{Eve}}] \leqslant \varepsilon \approx \mathcal{O}\left(\frac{1}{(M+1)^N}\right) \tag{6.15}$$

Proof. From Eq. (6.14), we see that the optimal strategy of Eve is to produce the response states ρ_i which maximises the fidelity F_i for each CRP $(|\phi_i^c\rangle, |\phi_i^r\rangle^{\otimes M})$. We utilise the same universal unforgeability result to bound the fidelity F_i with which Eve can produce the states ρ_i,

$$\Pr[F_i \geqslant \delta] \leqslant \frac{d+1}{D} \tag{6.16}$$

for any $\delta > 0$. Here $d = poly(\lambda) = poly\log(D)$ is the dimension of subspace that Eve has learnt from $\mathcal{H}^D$. For $D = 2^d$, this implies that the maximum fidelity state that Eve can create on average is non-orthogonal to the valid response state $|\phi_i^r\rangle$ with a negligible probability $\approx \mathcal{O}(2^{-d})$. Hence $F_i = \delta \to 0$ with overwhelming probability. This bound holds true for all distinct CRPs labelled by $i \in [N]$.

Thus from Eq 6.14 and 6.16, the probability that Eve passes verifier's GSWAP based verification test is,

$$\begin{aligned}\Pr[\text{Ver accept}_{\text{Eve}}] &\leqslant \prod_{i=1}^{N}\left(\frac{1}{M+1} + \frac{M}{M+1}F_i\right) \\ &\leqslant \prod_{i=1}^{N}\left(\frac{1}{M+1} + \frac{M}{M+1}\delta\right) \approx \mathcal{O}\left(\frac{1}{(M+1)^N}\right) = negl(\lambda)\end{aligned} \tag{6.17}$$

We have also taken into account the adaptive strategy of Eve since our security is analysed for the most general attack strategy. This completes the proof. □

The last equation shows that to achieve an exponentially secure qPUF based identification using the GSWAP based verification protocol with only a polynomial sized register S, the protocol needs to be repeated for multiple N instances. Our protocol requires $R = N$ number of communication rounds and uses $T = 2R$ number of communicated states.

6.2.3 Quantum identification protocol with low-resource verifier

Our second protocol enables a weak verifier to identify a quantum server prover in the network. We achieve this by delegating the equality testing to the prover

thus effectively removing the quantum computational requirement on the verifier. While this might look like it could facilitate a malicious Eve to fool the weak verifier easily, we demonstrate due to the unforgeability of qPUF that the security is not affected. Before describing the details, we list the salient features of our protocol,

- The protocol requires the prover to hold quantum computing capability, whereas the verifier is just required to have quantum memory and no quantum computing resources during the identification and verification phase[5] (restricted memory and computation).
- The protocol requires a one-way quantum communication link directed from the verifier to the prover. The prover to the verifier directed link is a classical channel.
- The protocol has a classical verification phase *i.e.* the prover locally performs the verification test and sends the classical information to the verifier.
- The protocol provides perfect completeness and an exponentially-high security guarantee against any adversary with QPT resources.

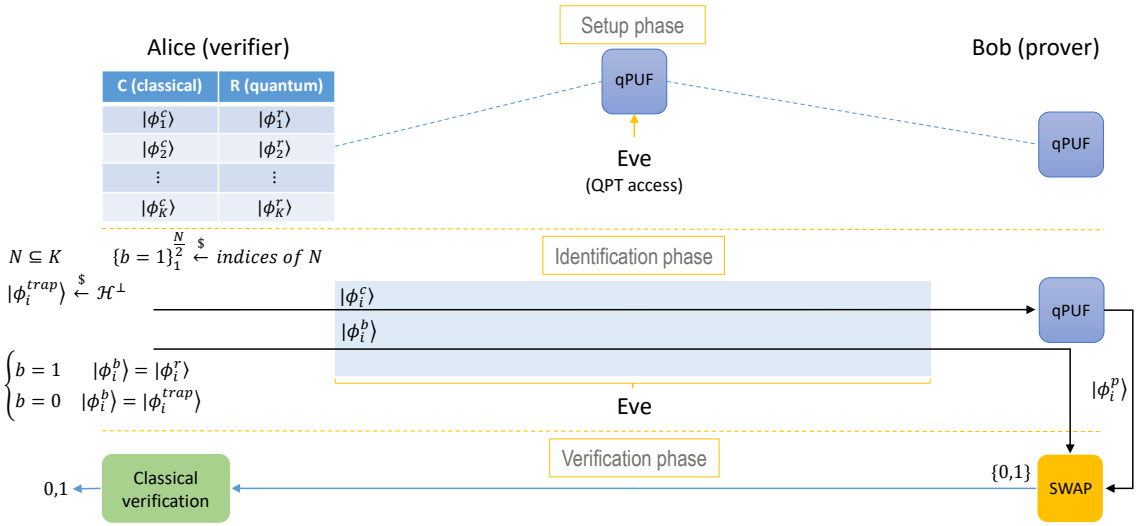

Figure 6.2: qPUF-based identification protocol with low-resource verification between Alice (verifier) and Bob (prover) (Irv-id). The protocol is divided into three sequential phases, *setup phase*, *identification phase* and *verification phase*. In the identification phase, Alice randomly picks a subset $N \subseteq K$ of challenges which are sent to Bob. She also employs a trap based scheme where she sends either the correct response state of the challenges or the trap states which are states orthogonal to the valid response states. Bob performs the SWAP-test verification and sends the classical bits back to Alice. Alice finally performs a classical verification to check.

[5]The state preparation phase happens in the setup phase of the protocol and it is a common property of all qPUF-based protocols. Here, we are mainly interested in the computing ability in the verification phase, which is the major difference between such protocols since verifying quantum states is a challenging task.

6.2.3.1 Protocol description

This protocol is run between a verifier, a prover in three sequential phases,

Protocol 2 (lrv-id(K,N,D))**.** qPUF-based identification protocol with low resource verifier and classical verification algorithm:[6]

---

1. *Setup phase*:

   (a) Verifier has the qPUF device.

   (b) Verifier randomly picks $K \in \mathcal{O}(\text{poly}\log D)$ classical strings $\phi_i \in \{0,1\}^{\log D}$.

   (c) Verifier selects and applies a Haar-random state generator operation denoted by the channel $\mathcal{E}$ to locally create the corresponding quantum states in $\mathcal{H}^D$: $\phi_i \xrightarrow{\mathcal{E}} |\phi_i^c\rangle$, $\forall i \in [K]$.

   (d) Verifier queries the qPUF individually with each quantum challenge $|\phi_i^c\rangle$ to obtain the response state $|\phi_i^r\rangle$.

   (e) Verifier creates states $|\phi_i^{\perp}\rangle$ orthogonal to $|\phi_i^c\rangle$ and queries the qPUF device with them to obtain the trap states labelled as $|\phi_i^{\text{trap}}\rangle$. The unitary property of qPUF device ensures that $\langle\phi_i^{\text{trap}}|\phi_i^r\rangle = 0$.

   (f) Verifier creates a local database $S \equiv \{|\phi_i^c\rangle, \{|\phi_i^r\rangle, |\phi_i^{\text{trap}}\rangle\}\}$ for all $i \in [K]$. Thus the S registers stores the challenge state $|\phi_i^p\rangle$ and the corresponding valid response state and the trap state which is orthogonal to the response state.

   (g) Verifier publicly transfers the qPUF to prover.

   The transition is non-secure and Eve is allowed $\mathcal{O}(\text{poly}\log D)$ query access to the qPUF to build her own local database.

2. *Identification phase*:

   (a) Verifier randomly selects a subset $N \subseteq K$ different challenges $|\phi_i^c\rangle$ and sends them over a public channel to prover.

   (b) Verifier randomly selects $N/2$ positions, marks them $b = 1$ and sends the valid response states $|\phi_i^1\rangle = |\phi_i^r\rangle$ to prover. On the remaining $N/2$ positions, marked as $b = 0$, the verifier sends the trap states $|\phi_i^0\rangle = |\phi_i^{\text{trap}}\rangle$.

3. *Verification phase*:

   (a) Prover queries the qPUF device with the challenge states received from verifier to generate the response states $|\phi_i^p\rangle$ for all $i \in [N]$.

---

[6] We drop the parameters (K, N, D) from now on for simplicity whenever we refer to this protocol.

(b) Prover performs a quantum equality test algorithm by performing a SWAP test between $|\phi_i^p\rangle$ and the response state $|\phi_i^b\rangle$ received from the verifier. This algorithm is repeated for all the N distinct challenges.

(c) Prover labels the outcome of N instances of the SWAP test algorithm by $s_i \in \{0,1\}$ and sends them over a classical channel to verifier.

(d) Verifier runs a classical verification algorithm $\texttt{cVer}(s_1,...,s_N)$ and outputs '1' implying that prover's qPUF device has been successfully identified, and outputs '0' otherwise.

Fig. 6.2 shows the qPUF based identification protocol with low-resource verification denoted as lrv-id. For the lrv-id protocol, completeness is the probability that Verifier's verification algorithm cVer returns an outcome '1' in absence of Eve. Ideally we require completeness to differ negligibly from 1,

$$\Pr[\text{Ver accept}_{\text{H}}] = \Pr\big[\texttt{cVer}(S_N) = 1\big] = 1 - negl(\lambda) \tag{6.18}$$

where λ is the security parameter.

Soundness of the protocol is the probability that cVer returns an outcome '1' in presence of Eve. For security, we require the soundness to be negligible in λ,

$$\Pr[\text{Ver accept}_{\text{Eve}}] = \Pr\big[\texttt{cVer}(S_N) = 1\big] = negl(\lambda) \tag{6.19}$$

We investigate the security of our protocol when the prover uses the SWAP test and the verifier uses the classical verification algorithm cVer. We remark that the prover can alternatively use GSWAP testing to generate the outcomes, however, this would require the verifier to send multiple copies of the same challenge state to the prover, thus incurring higher resources on the verifier's side.

6.2.3.2 cVer algorithm

The main ingredient of verification is the cVer classical test algorithm employed by the verifier to certify whether the prover's device has been identified. As described in Algorithm 2, cVer receives an N-bit binary string S_N as input. The algorithm is divided into two tests. test1 first checks whether in the $N/2$ positions marked as $b = 1$, *i.e.* the positions where the verifier had sent a valid qPUF response state to the prover if the corresponding bits in S_N are all 0.

If this test succeeds, then the algorithm proceeds to test2 which is a test on the positions where the verifier had sent the trap states to the prover. If on these positions, the expected number of bits in S_N which are 0 lie between $\{\kappa\frac{N}{2} - \delta_{er}, \kappa\frac{N}{2} + \delta_{er}\}$, then cVer algorithm outputs '1' indicating that the device has been identified. Here $\kappa\frac{N}{2}$ is the expected number of bits in $b = 1$ positions with outcome '0' that prover would obtain after the equality test algorithm measurement, in absence of any adversary Eve. In our case when the prover uses the SWAP test, $\kappa = 0.5$. Here, δ_{er} accounts for the statistical error in the measurement.

Algorithm 2 `cVer` algorithm

Description: Let $S_N = \{0,1\}^N$ be the input N-bit string. Let $P = \{i_k\}_{k=1}^{N/2}$ be the set of indices showing the rounds of the protocol where $b = 1$. Algorithm consists of two tests, `test1` and `test2` as follows:

`test1:`
forall *i in P* **do**
  if $s_i = 0$ **then**
    $count \leftarrow count + 1$
  end
end
if $count = \frac{N}{2}$ **then**
  return 1
else
  return 0
end

`test2:`
if *test1 = 0* **then**
  return 0
else
  forall *i not in P* **do**
    if $s_i = 1$ **then**
      $count \leftarrow count + 1$
    end
  end
  if $|count - \delta\frac{N}{2}| \leqslant \delta_{er}$ **then**
    return 1
  else
    return 0
  end
end

6.2.3.3 Verification using SWAP test and `cVer` algorithm

Here we explicitly describe and calculate the completeness and soundness probabilities of the lrv-id protocol which employs the verification algorithm involving the prover's SWAP test, followed by the verifier's `cVer` algorithm. This allows the verifier to efficiently identify the valid qPUF device even though the SWAP test algorithm has been delegated to the prover. A single instance of the prover's SWAP test requires a single copy of the response state received from the verifier (either the valid qPUF response state or the trap state) and the response state that the prover generates by querying the verifier's challenge state in his qPUF device. To obtain a desired low enough error rate in the verification algorithm, the SWAP test is performed on N distinct instances of the received response state and response state generated by prover by querying distinct challenges states. The

responses of the SWAP test instances are classical bits. Thus the N bit binary classical outcome string is sent to the verifier who employs the algorithm `cVer` described in Algorithm 2. An identification protocol performed using N distinct challenge states consumes a combined total of $2N$ copies of the received state and the response state generated by the verifier. In the next two sections, we show that SWAP based test algorithm provides us with the desired completeness and soundness properties required in the protocol.

Theorem 39 (`cVer` **Completeness**). *In absence of an adversary Eve, the probability that the N-bit string $S_N = \{s_1, ..., s_N\}$ sent by prover, passes the* `cVer`(S_N) *algorithm is,*

$$\Pr[\text{Ver accept}_{\text{H}}] = \Pr\big[\texttt{cVer}(S_N) = 1\big] = 1 - 2e^{-N/4} \tag{6.20}$$

Proof. To prove this theorem, we separately analyse the $N/2$ positions where verifer sends the valid qPUF response state to prover (marked as $b = 1$), and the remaining positions where she sends the trap state (marked as $b = 0$),

1. $b = 1$ positions: When prover prepares the response state $|\phi_i^p\rangle$ by querying her qPUF device with verifer's challenge state $|\phi_i^c\rangle$, then prover's generated response state is equal to verifer's response state sent to prover, *i.e.* $|\phi_i^r\rangle = |\phi_i^p\rangle$. This implies that $F(|\phi_i^p\rangle, |\phi_i^r\rangle) = 1$ for all $i \in [N]$ marked $b = 1$. From Eq. (2.65), we see that,

$$\Pr\big[\text{SWAP}(|\phi_i^p\rangle, |\phi_i^r\rangle) = 1\big] = \frac{1}{2} + \frac{1}{2}F(|\phi_i^p\rangle, |\phi_i^r\rangle) = 1, \tag{6.21}$$

   Note that $[\text{SWAP}(|\phi_i^p\rangle, |\phi_i^r\rangle) = 1]$ corresponds to the classical outcome 0. This implies that $s_i = 0$ for all $i \in [N]$ marked $b = 1$ with certainty. Thus when verifer employs the `cVer` algorithm, prover always achieves a $count = N/2$ in the `test1` and thus passes it with certainty,

$$\Pr[\texttt{test1 pass}] = 1 \tag{6.22}$$

2. $b = 0$ positions: These positions correspond to verifer sending the trap states $|\phi_i^{\text{trap}}\rangle$ to prover such that prover's generated response state $|\phi_i^p\rangle$ is orthogonal to the trap state. In other words, $F(|\phi_i^p\rangle, |\phi_i^{\text{trap}}\rangle) = 0$ for all $i \in [N]$ marked $b = 0$. This implies that,

$$\Pr\big[\text{SWAP}(|\phi_i^p\rangle, |\phi_i^{\text{trap}}\rangle) = 1\big] = \frac{1}{2} + \frac{1}{2}F(|\phi_i^p\rangle, |\phi_i^{\text{trap}}\rangle) = \frac{1}{2}, \tag{6.23}$$

   Thus, half of the $N/2$ positions would produce the classical outcome 1 on average. When verifer employs `test2` of the `cVer` algorithm, $\mathbb{E}[count] = N/4$. Using the Chernoff-Hoeffding inequality [MU17], for any constant $\delta_{er} > 0$,

$$\Pr[\texttt{test2 pass}] = \Pr\left[\left|count - \frac{N}{4}\right| \leqslant \delta_{er}\right] \geqslant 1 - 2e^{-N\delta_{er}^2} \tag{6.24}$$

From the above results and using the fact that $\delta_{er} = 0.5$ for SWAP test based algorithm,

$$\begin{aligned} \Pr[\text{Ver accept}_{\mathsf{H}}] &= \Pr\big[\texttt{cVer}(s_1, \cdots, s_N) = 1\big] \\ &= \Pr\big[\texttt{test1 pass} \wedge \texttt{test2 pass}\big] \\ &= \Pr[\texttt{test1 pass}] \cdot \Pr[\texttt{test2 pass}] \\ &\geqslant 1 - 2e^{-N/4} \end{aligned} \tag{6.25}$$

This completes the proof. □

The next section details the soundness proof of the lrv-id protocol.

6.2.3.4 Soundness of lrv-id protocol

To characterise the soundness, we bound Eve's success probability in passing the cVer test. Since the verification test is reduced to a classical test, we consider the soundness in the presence of two types of Eve. The first is a *classical Eve* who does not process any quantum resources. The second is a *quantum Eve*, which possesses QPT memory and computing capability. We separately analyse the security against both types of Eve and prove that *quantum Eve* gains only an exponentially-small advantage compared to the *classical Eve*, thus reducing the security to analysing only the classical adversary. We show that since the verification test is classical, the only way for a *quantum Eve* to succeed better than a *classical Eve* is to succeed at guessing the trap positions better than a random guess of *classical Eve*. We utilise the unforgeability property of qPUF to prove that a *quantum Eve* can have only negligible advantage in guessing the trap positions compared to a *classical Eve*, thus enabling the reduction.

(I) Security against classical adversary

Theorem 40 (**Soundness against classical Eve**). *The probability that any classical PPT adversary (Eve) produces an N-bit string $S_N = \{s_1, ..., s_N\}$ which passes the* cVer *algorithm is bounded as,*

$$\Pr[\text{Ver accept}_{\text{Eve}}] = \Pr\big[\texttt{cVer}(S_N) = 1\big] \leqslant \mathcal{O}(2^{-N}) \tag{6.26}$$

Proof. First, we remark that any classical Eve's strategy to produce a valid N-bit string S_N can be divided into two categories,

1. **Independent guessing strategy:** Under this strategy, Eve tries to independently guess each bit of the string S_N that would pass the verifier's cVer algorithm. This also relates to the strategy of independently finding valid responses and trap positions.

2. **Global strategy:** Here, Eve's strategy is to output a string S_N using the global properties of the cVer, such that it passes the verification test with

maximum probability. In contrast to the previous strategy, the probability of outputting each bit s_i is not necessarily independent of the global strategy.

We calculate the optimal success probability of Eve in both cases and show that by optimizing for both the strategies, we obtain a higher success probability for Eve in the optimal global strategy scenario. Although, the two strategies converge in the limit of large N. Hence we bound Eve's success probability by the optimal global strategy.

1. Independent guessing strategy: Under this strategy, Eve independently guesses each bit with the probability,

$$Pr[s_i = 0] = \alpha, \quad Pr[s_i = 1] = 1 - \alpha \tag{6.27}$$

where $\alpha \in [0, 1]$.

We denote the resulting string generated by Eve's strategy as $S_{id} = \{s_1, \cdots, s_N\}$. In order for S_{id} to pass the `cVer` verification algorithm, it must simultaneously pass the `test1` and `test2`. Since Eve's strategy is guessing each bit independently, hence the probability for her to pass the `test1` and `test2` are independent. Let us look at the probability of passing the `test1` (which corresponds to checking the $N/2$ positions marked $b = 1$,

$$\Pr[\texttt{test1 pass}] = \Pr[s_{p_1} = 0] \times \cdots \times \Pr[s_{p_{\frac{N}{2}}} = 0] = \alpha^{\frac{N}{2}} \tag{6.28}$$

where p_i correspond to the $b = 1$ marked positions.

If Eve's generated string passes `test1`, then verifer runs the `test2` to check if *count*, which is the number of bits that are 1 in the remaining $N/2$ bits marked with $b = 0$, lies within the interval $\left|count - \frac{N}{4}\right| \leqslant \delta_{er}$. Eve succeeds in passing this test with the probability,

$$\begin{aligned} \Pr[\texttt{test2 pass}] &= \sum_{x=N/4-\delta_{er}}^{N/4+\delta_{er}} (1-\alpha)^x \alpha^{\frac{N}{2}-x} \times \binom{N/2}{x} \\ &\approx (2\delta_{er}+1)(1-\alpha)^{\frac{N}{4}} \alpha^{\frac{N}{4}} \times \binom{N/2}{N/4} \end{aligned} \tag{6.29}$$

where the approximation holds since we assume that $\delta_{er} \ll N$. From the above results, we see that the probability that Eve's string S_{id} passes the `cVer` verification algorithm is,

$$\begin{aligned} \Pr[\text{Ver Accept}_{\text{Eve},\alpha}] &= \Pr[\texttt{test1 pass}_\alpha] \cdot \Pr[\texttt{test2 pass}_\alpha] \\ &\approx (2\delta_{er}+1)\alpha^{\frac{3N}{4}}(1-\alpha)^{\frac{N}{4}} \times \binom{N/2}{N/4} \end{aligned} \tag{6.30}$$

This is Eve's acceptance probability for a given α. An optimal strategy for Eve is to find the optimal value of α that maximises the acceptance probability. This corresponds to,

$$\frac{\partial}{\partial\alpha}\Pr[\text{Ver Accept}_{\text{Eve},\alpha}] \Rightarrow \frac{\partial}{\partial\alpha}(\alpha^{\frac{3N}{4}}(1-\alpha)^{\frac{N}{4}}) = 0 \Rightarrow \alpha = \frac{3}{4} \tag{6.31}$$

Thus the maximum acceptance probability of Eve using an independent guessing strategy is:

$$\Pr[\text{Ver Accept}_{\text{Eve}}] = (2\delta_{er}+1)\frac{3^{\frac{3N}{4}}}{2^{2N}} \times \binom{N/2}{N/4} \approx \mathcal{O}(2^{-N}) \tag{6.32}$$

2. Global strategy: The second category of Eve's strategy is to guess the N bit string which passes the `cVer` test algorithm with maximum probability. Here, Eve is not restricted to choosing each bit independently. To find the optimal global strategy we look at the `test1` and `test2` algorithms and extract essential properties that can be leveraged by Eve to pass the verification test. We note that

- Since the good and trap response positions corresponding to $b = 0$ and 1 are chosen uniformly randomly by verifer, hence verifer does not have any information on the index set P corresponding to $b = 1$ (thus no information on $b = 0$ positions too).
- Eve knows the statistics of 0's and 1's in the desired string to pass the `cVer`. For example, a string must have a minimum of $\approx 3N/4$ bits which are 0, otherwise, the string necessarily fails the `test1` or `test2` or both.

Based on the above facts, any global strategy for Eve should consist of optimizing the number of 0's and 1's to pass both verification tests.

Before considering the optimal global attack strategy, we give an example of a specific (non-optimal) attack strategy to provide intuition on the kind of strategies that Eve can adopt here.

Example of a global strategy: The first global strategy that one might think of is to try to guess P, since passing the `test1` reduces to finding the strings that have bits '0' is all the p_i positions *i.e.* positions marked $b = 1$. If Eve successfully manages to guess the $b = 1$ positions, then she has a deterministic strategy of winning the `test2`, since she also knows the $b = 0$ trap positions. Across these positions she can deterministically assign the bits such that the *count* of the number of 1 bits lie within the interval $\left|count - \frac{N}{4}\right| \leqslant \delta_{er}$.

We denotes Eve's generated string with this strategy to be S_g. Hence the probability of S_g passing `test1` is equal to correctly guessing the $\frac{N}{2}$ positions marked $b = 1$,

$$\Pr[\texttt{test1 pass}_{S_g}] = \Pr[\text{guess } b = 1 \text{ positions}] = \binom{N}{N/2}^{-1} \tag{6.33}$$

Once this test passes, then test2 passes with certainty. Now the probability of passing the `cVer` verification algorithm is,

$$\begin{aligned}\Pr[\text{Ver accept}_{\text{Eve},S_g}] &= \Pr[\texttt{test1 pass}_{S_g} \wedge \texttt{test2 pass}_{S_g}] \\ &= \Pr[\texttt{test1 pass}_{S_g}] \cdot \Pr[\texttt{test2 pass}_{S_g} | \texttt{test1 pass}_{S_g}] \\ &= \binom{N}{N/2}^{-1} \cdot 1 \\ &\leqslant N^{-\frac{N}{2}}\end{aligned} \tag{6.34}$$

We show that this global strategy is not optimal and Eve can design an optimal global strategy by properly utilising the part the second part of the information.

First, we argue that maximising the number of 0's will necessarily increase the success probability of passing `test1`. Let us assume that Eve sends an all '0' string S_g to the verifier. Since `test1` checks only if in the $b = 1$ marked positions are 0, so S_g will always pass the first test. However, this string necessarily fails the `test2` since the *count* for this test is $N/2$ which is much higher than the tolerated limit.

Thus there always exists a global strategy with an optimal number of bits (number of 1) in S_g in the case of $\delta_{er} = 0$, or more precisely a strategy that allows the flexibility of having a set of values for the number of '1' bits that the `test2` tolerates in case of $\delta_{er} \neq 0$.

Optimal global strategy: We say that an optimal global strategy $\mathcal{E}_{gop}$ is the one that outputs a string S_{gop} with c_1 number of 1 bits, where $c_1 \in m_{valid} = \{\frac{N}{4} - \delta_{er}, \ldots, \frac{N}{4} + \delta_{er}\}$.

Optimality argument: We prove the optimality of our test by the contradiction argument. Let us assume that there is a strategy $\mathcal{E}_g$ different from above which produces a string S_g that succeeds with the verification acceptance probability higher than S_{gop}. Now, either all the strings that $\mathcal{E}_g$ outputs have c_1 number of 1 bits, where c_1 lies within the optimal boundary m_{valid}. In this case $\mathcal{E}_g$ falls within the $\mathcal{E}_{gop}$ strategy set. Or, there is at least one string that $\mathcal{E}_g$ outputs with c_1 number of 1 bits such that $c_1 \notin m_{valid}$. In this case, that string will necessarily fail `test2`, even if it passes `test1`. This is because for the strategy $\mathcal{E}_g \notin \mathcal{E}_{gop}$ to pass, the bits in S_g which are 1 must necessarily appear in the positions marked $b = 0$ (trap positions). And since the number of 1 bits $c_1 \notin m_{valid}$, this implies it will fail the `test2`. Thus, $\Pr[\text{Ver Accept}_{\text{Eve},\mathcal{E}_g \notin \mathcal{E}_{gop}}] = 0$.

Note that the condition of $c_1 \in m_{valid}$ is necessary but not a sufficient condition for passing the verification algorithm `cVer` *i.e.* any string with with $c_1 \notin m_{valid}$, will always fail but not all strings with $c_1 \in m_{valid}$ will always pass the verification. Thus we can define the largest possible set of potentially valid strings which Eve needs to choose from to maximise her acceptance probability. As a result, we can define the optimal strategy $\mathcal{E}_{gop}$'s event space to be $\binom{N}{c_1}$. This is the set of all

strings with the number of bits $c_1 \in m_{valid}$. We can now find the optimal global probability which is the probability that both the tests of `cVer` pass,

$$\begin{aligned}\Pr[\text{Ver accept}_{\text{Eve},S_{gop}}] &= \Pr[\texttt{test1 pass}_{S_{gop}} \wedge \texttt{test2 pass}_{S_{gop}}] \\ &= \Pr[\texttt{test1 pass}_{S_{gop}}] \cdot \Pr[\texttt{test2 pass}_{S_{gop}} | \texttt{test1 pass}_{S_{gop}}]\end{aligned} \tag{6.35}$$

To calculate $\Pr[\texttt{test1 pass}_{S_{gop}}]$, we need to find the number of strings S_{gop} from the whole set of strings $\{0,1\}^N$ with $c_1 \in m_{valid}$ bits and which passes the first test. In other words, the string S_g must have bits 0 in all the $b=1$ marked positions and the bits 1 in the $b=0$ marked positions.

Thus there are $N/2$ positions out N where the bits 1 can be placed without the `test1` getting rejected.

For a specific c_1, the total number of such strings is equal to the possible ways of distributing c_1 objects (1's) in $N/2$ positions:

$$\#(\text{correct strings}) = \binom{N/2}{c_1} \tag{6.36}$$

If one of these 'correct strings' is picked, it will necessarily also satisfy the condition of the second test. Hence the conditional probability is

$$\Pr[\texttt{test2 pass}_{S_{gop}} | \texttt{test1 pass}_{S_{gop}}] = 1$$

. And the probability of passing the first test is,

$$\Pr[\texttt{test1 pass}_{S_{gop}}] = \binom{N/2}{c_1} \Big/ \binom{N}{c_1} \tag{6.37}$$

The above `test1` passing probability is for a single $c_1 \in m_{valid}$. Summing over the probabilities of all the accepted c_1 ,

$$\Pr[\texttt{test1 pass}_{S_{gop}}] = \sum_{c_1 \in m_{valid}} \frac{\binom{\frac{N}{2}}{c_1}}{\binom{N}{c_1}} = \sum_{k=-\delta_{er}}^{\delta_{er}} \frac{\binom{\frac{N}{2}}{\frac{N}{4}+k}}{\binom{N}{\frac{N}{4}+k}} = \frac{(\frac{N}{2})!}{N!} \sum_{k=-\delta_{er}}^{\delta_{er}} \frac{(\frac{3N}{4}-k)!}{(\frac{N}{4}-k)!} \tag{6.38}$$

In the limit $\delta_{er} \ll N$, the sum will converge,

$$\Pr[\texttt{test1 pass}_{S_{gop}}] = (2\delta_{er}+1) \cdot \frac{(\frac{N}{2})!(\frac{3N}{4})!}{N!(\frac{N}{4})!} \tag{6.39}$$

From the above equations, the probability that Eve passes the `cVer` algorithm using the global strategy,

$$\begin{aligned}\Pr[\text{Ver accept}_{\text{Eve},S_{gop}}] &= \Pr[\texttt{test1 pass}_{S_{gop}}] \cdot \Pr[\texttt{test2 pass}_{S_{gop}} | \texttt{test1 pass}_{S_{gop}}] \\ &= (2\delta_{er}+1) \times \frac{(\frac{N}{2})!(\frac{3N}{4})!}{N!(\frac{N}{4})!} \cdot 1 \\ &\leqslant \mathcal{O}(N^{-N/2})\end{aligned} \tag{6.40}$$

3. Probability comparison of Independent guessing strategy and Global strategy: To find the optimal classical attack, we compare the two categories of the attack strategies of Eve.

We fix the accepted tolerance value $\delta_{er} = 1$ for the comparison. The same result holds for other fixed δ_{er} values. Fig. 6.3 shows the behaviour of the acceptance probabilities of Eve in the independent guessing strategy and global strategy as an increasing function of the string length N.

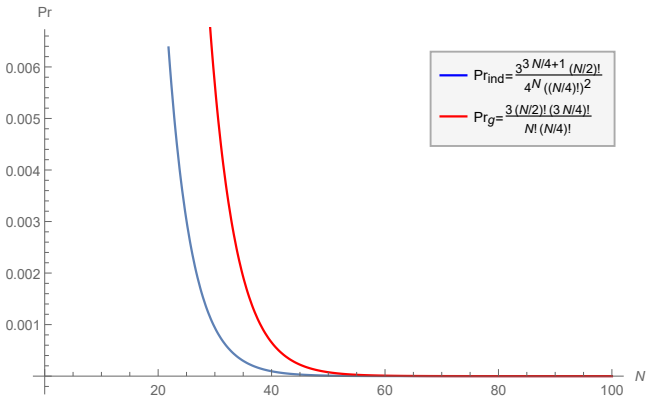

Figure 6.3: Comparison of the acceptance probabilities of a classical adversary (Eve) in the independent guessing strategy (in blue) and global strategy (in red) as a decreasing function of the string length N for the tolerance value $\delta_{er} = 1$

From the simulation, we infer that the two strategies have an inverse exponential form as expected. Also, they both converge for large enough N values. This also confirms the fact that the optimal strategy lies in finding the correct number of 1's in the string and the difference comes from our approximation in using the frequency interpretation of the probabilities in the smaller N. Using Stirling's approximation $n! \approx \sqrt{2n\pi}(\frac{n}{e})^n$ one can check that $\frac{1}{\binom{N}{\frac{N}{4}}} \approx (\frac{4}{3^{3/4}})^{-N}$ which gives exactly the same bound as the independent guessing strategy. Although, in small N the global strategy is slightly better. Finally, we use Stirling's approximation $\binom{2n}{n} \approx \frac{2^{2n}}{\sqrt{\pi n}}$ to obtain the common factor of both probabilities we can bound the adversary's optimal success probability as,

$$\Pr[\text{Ver Accept}_{\text{Eve}}] \approx \frac{3^{3N/4}}{2^{2N}} \times \frac{2^{N/2}}{\sqrt{\frac{\pi N}{4}}} = \frac{2}{\sqrt{N\pi}}(\frac{2^6}{3^3})^{-N/4} \approx \mathcal{O}(2^{-N}) \quad \text{for large enough N} \tag{6.41}$$

This completes the proof. □

(II) Security against quantum adversary

We now investigate the soundness property of the protocol against QPT Eve by modelling Eve's strategy with a completely positive trace preserving (CPTP) map that takes as input the target challenge $|\phi_i^c\rangle$, the unknown state $|\phi_i^b\rangle$, and ancilla qubits and outputs the classical bits which are sent to verifier for verification. This

map utilises the database information created by Eve during the qPUF transition. A QPT Eve's strategy can be divided into two categories,

1. **Collective attack strategy**: Eve applies an independent CPTP map on each of the N rounds.

2. **Coherent attack strategy**: Eve applies a CPTP map on the combined N distinct challenge and their corresponding response states that the verifier sends to the prover.

A collective strategy is a special case of Eve's coherent strategy. However, we show that independence in choosing the trap states by verifier reduces the coherent strategy to the collective strategy by Eve. We analyse the collective security first and then give a reduction of the coherent strategy to the collective strategy.

1. Collective strategy: Under this strategy, Eve optimises over all the CPTP maps that input verifier's states $|\phi_i^c\rangle$ and $|\phi_i^b\rangle$ and outputs a single bit s_i to maximise the acceptance probability. Fig. 6.4 shows Eve performing a general collective strategy.

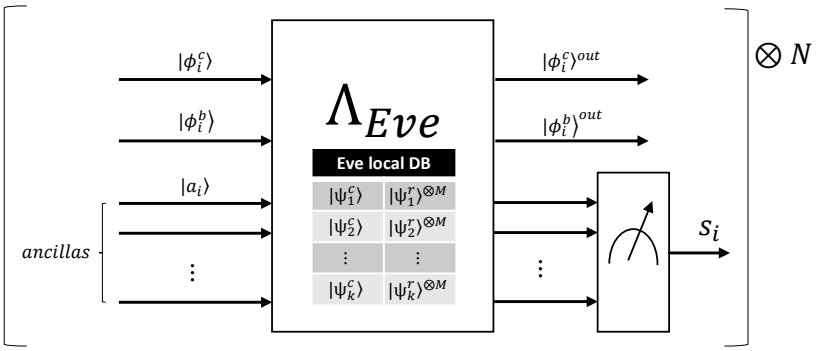

Figure 6.4: Quantum collective attack strategy performed by Eve on Irv-id protocol by applying the same local-database-dependent CPTP map on each round of the challenge and response state $|\phi_i^c\rangle$ and $|\phi_i^b\rangle$ respectively. The output of the single instance of the map is a bit S_i.

We denote Eve's quantum map to be,

$$\Lambda_{Eve} \equiv \bigotimes_{i=1}^{N} \Lambda_i \quad . \tag{6.42}$$

Contrary to the classical Eve who is unable to figure out the trap positions in any round with a probability higher than half, a QPT Eve, by leveraging her local database information, could be expected to do better than a random guess. More formally, we say that the Irv-id protocol is secure against any QPT Eve that performs a CPTP map Λ_i on the states $|\phi_i^c\rangle, |\phi_i^b\rangle$ for each $i \in [N]$ if the resulting success probability of correctly guessing the bit b for each position differs negligibly in the security parameter from half.

We need a small toolkit, which is the abstraction of an ideal test in a single instance case (when one is provided with a single copy of one quantum state

and multiple copies of the other state), in terms of fidelity. This definition is very similar to Definition 42 in Chapter 4 where we remove the quantifier δ for simplicity:

Definition 46 (Single Instance Ideal Test Algorithm)**.** We call a test algorithm according to Definition 12, a $\mathcal{T}_{ideal}$ test algorithm when one is provided a single copy of the state ρ and multiple copies of the state $|\psi\rangle$ (or vice-versa) with fidelity $F(\rho, |\psi\rangle\langle\psi|)$ the test responds as follows:

$$\mathcal{T}_{ideal} := \Pr[1 \leftarrow \mathcal{T}_{ideal}(\rho, |\psi\rangle\langle\psi|)] = F(\rho, |\psi\rangle\langle\psi|) \tag{6.43}$$

Now we can show the security of Irv-id against QPT adversaries with collective attack strategies with the following theorem:

Theorem 41 (**Security against collective attack**)**.** *The success probability of any QPT adversary Eve, in correctly guessing whether $|\phi_i^b\rangle = |\phi_i^r\rangle$ for each $i \in [N]$ differs negligibly from half,*

$$Pr[b \leftarrow \Lambda_i(|\phi_i^c\rangle, |\phi_i^b\rangle)] \leqslant \frac{1}{2} + \mathcal{O}(2^{-d}) \;\; \forall i \in [N] \tag{6.44}$$

where $d = \mathcal{O}(\text{poly}\log D)$ is the size of Eve's database and qPUF is in $\mathcal{H}^D$.

Proof. First, we use the symmetry of the problem to restrict ourselves to cases where $b = 1$. We prove the theorem by contradiction *i.e.*, suppose there exists an algorithm W that wins the quantum security game for each index $i \in [N]$ with a probability non-negligibly better than a random guess. In other words, $W = 1$ if the index b is correctly guessed, and $W = 0$ otherwise. Let $f(\lambda) \geqslant 0$ be a non-negligible function of the security parameter. The joint probabilities for all collective possible values of b and W can be written as,

$$\begin{aligned} \Pr[W=1, b=1] &= \frac{1}{4} + f(\lambda) \quad \Pr[W=1, b=0] = \frac{1}{4} - f(\lambda) \\ \Pr[W=0, b=0] &= \frac{1}{4} + f(\lambda) \quad \Pr[W=0, b=1] = \frac{1}{4} - f(\lambda) \end{aligned} \tag{6.45}$$

where the joint probabilities are higher when W correctly guesses b, and is lower otherwise. From this, we can define the following conditional probability of winning for cases where $b = 1$ as follows:

$$Pr[W = 1 | b = 1] = \frac{Pr[W = 1, b = 1]}{Pr[b = 1]} = \frac{1}{2} + f'(\lambda)$$

Where $f' = 2f$ is again a non-negligible function in the security parameter λ. This is the same probability of winning when $b = 0$ *i.e.* $\Pr[W = 0 | b = 0]$.

Now we show that the success probability of Eve in successfully guessing whether $|\phi_i^b\rangle = |\phi_i^r\rangle$ reduces to finding a CPTP map Λ_i which performs an optimal

quantum test to distinguish the response state $|\phi_i^b\rangle$ with the reference state $|\phi_i^{\mathcal{E}}\rangle$. As Eve has no access to the actual response $|\phi_i^r\rangle$, the reference state $|\phi_i^{\mathcal{E}}\rangle$ should be generated within the Λ_i itself. Thus without loss of generality, any attack map Λ_i, consists of two parts. The first part uses a generator algorithm $\mathtt{gen}$ to generate a reference state $|\phi_i^{\mathcal{E}}\rangle$, or more generally a mixed state ρ_e by using the local database and the input challenge state $|\phi_i^c\rangle$, and the second part performs a test algorithm $\mathcal{T}$ on $|\phi_i^b\rangle$ and ρ_e,

$$\Lambda_i \equiv \mathcal{T}(|\phi_i^b\rangle, \rho_e \leftarrow \mathtt{gen}(DB, |\phi_i^c\rangle)) \tag{6.46}$$

where DB is the local database of Eve generated in the *setup phase*. To further provide the capability to Eve, we assume that her test $\mathcal{T}$ is an optimal test equality test algorithm also referred to as the ideal test algorithm in Definition 46, *i.e.* $\mathcal{T} = \mathcal{T}_{ideal}$. Note that $\mathcal{T}_{ideal}$ is the optimal test allowed by quantum mechanics where the probability of succeeding in the equality test is proportional to the square of the fidelity distance of the two states. Now we state the following contraposition: Let us assume that there exists a winning algorithm W running $\Lambda = \mathcal{T}_{ideal}(|\phi_i^b\rangle, \rho_e)$ such that,

$$\Pr[1 \leftarrow \Lambda(|\phi_i^c\rangle, |\phi_i^b\rangle) | b = 1] \leqslant \frac{1}{2} + non\text{-}negl(\lambda) \tag{6.47}$$

From Definition 46, we see that $\mathcal{T}_{ideal}$ outputs 1 with probability $p = F(|\phi_i^b\rangle, \rho_e)$. In other words,

$$\Pr[1 \leftarrow \Lambda(|\phi_i^c\rangle, |\phi_i^b\rangle) | b = 1] = \Pr[1 \leftarrow \mathcal{T}_{ideal}] = F(|\phi_i^b\rangle, \rho_e) \leqslant \frac{1}{2} + non\text{-}negl(\lambda) \tag{6.48}$$

This implies that if an algorithm W exists for Eve, then she is able to generate the state ρ_e with non-negligible fidelity with the valid qPUF response (for b=1), and similarly with trap states (for b = 0). And this would hold for all $i \in [N]$. But this contrasts with the universal unforgeability of the qPUF which states that the success probability of any QPT adversary having polynomial-size access to the qPUF is bounded as $\frac{d_e+1}{D}$ (Theorem 28) where $d_e = poly(\lambda) = \text{polylog}(D)$ is the dimension of subspace that Eve has learnt from $\mathcal{H}^D$. Thus such Λ cannot exist even with the most efficient test $\mathcal{T}_{ideal}$. This concludes the proof. □

2. Coherent Strategy: The collective strategy is restricted to Eve applying individual unentangled maps in each round. A more generalised strategy, the coherent strategy, involves applying a CPTP map collectively on all the rounds thus potentially leveraging entanglement capabilities across rounds. Such a strategy takes as input the N challenge states $\otimes_{i=1}^N |\phi_i^c\rangle$, the N response state $\otimes_{i=1}^N |\phi_i^b\rangle$ and the ancilla qubits, and outputs a N bit string S_N which is sent to verifier for verification. Fig. 6.5 depicts this strategy. Eve's objective is to produce the S_N which maximises the $\mathtt{cVer}$ passing probability. We denote Eve's quantum map to be,

$$\Lambda_{Eve} \equiv \Lambda^N \tag{6.49}$$

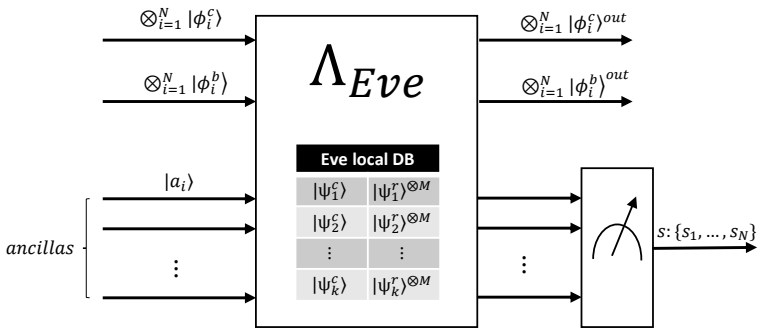

Figure 6.5: Quantum coherent attack strategy performed by Eve on Irv-id protocol by applying the general local-database-dependent, CPTP map on the combined N challenges and response states $|\phi_i^c\rangle$ and $|\phi_i^b\rangle$ respectively. The output is the N bit string $s : \{s_1, \cdots, s_N\}$.

We say that the Irv-id protocol is secure against any QPT Eve who performs the map Λ^N if the resulting success probability of correctly guessing the b value for all the N positions is negligibly small in the security parameter.

Theorem 42 (**Security against coherent attack**). *The success probability of any QPT adversary Eve, in correctly guessing the b values for all the N positions, denoted by* $[b_1, \cdots, b_N]$ *is,*

$$\Pr[\mathbf{b} \leftarrow \Lambda^N(|\phi^{\mathbf{c}}\rangle, |\phi^{\mathbf{b}}\rangle)] \leqslant \left(\frac{1}{2} + \mathcal{O}(2^{-d})\right)^N \tag{6.50}$$

where $\mathbf{b} : [b_1, \cdots, b_N]$ *are the bits corresponding to correct b values,* $|\phi^{\mathbf{c}}\rangle = \otimes_{i=1}^{N} |\phi_i^c\rangle$, $|\phi^{\mathbf{b}}\rangle = \otimes_{i=1}^{N} |\phi_i^b\rangle$ *and* $d = \mathcal{O}(\text{poly}\log D)$ is the size of Eve's database.

Proof. To prove this theorem, we notice that Eve applies a generalised map Λ^N on the challenge and the response states of verifer to be able to correctly distinguish whether the response states are $|\phi_i^b\rangle = |\phi_i^r\rangle$ for all $i \in [N]$. Thus the probability to correctly guess $\mathbf{b}$ reduces to Eve applying a CPTP map Λ^N to perform an optimal test to distinguish the response state $|\phi^{\mathbf{b}}\rangle$ with her reference state ρ_e^N, where ρ_e^N is the generalised entangled state. Thus without loss of generality, any attack map Λ^N, consists of two parts. The first part uses a generator algorithm gen_N to generate a reference state ρ_e^N by using the local database and the input challenge state $|\phi^{\mathbf{c}}\rangle$, and the second part performs a test algorithm $\mathcal{T}$ on $|\phi^{\mathbf{b}}\rangle$ and ρ_e^N,

$$\Lambda^N = \mathcal{T}(|\phi^{\mathbf{b}}\rangle, \rho_e^N \leftarrow gen(DB, |\phi^{\mathbf{c}}\rangle)) \tag{6.51}$$

where DB is the local database of Eve generated in the *setup phase*. Similar to the collective strategy proof, we assume Eve's testing algorithm $\mathcal{T}$ is the optimal test equality test algorithm, also referred as ideal test algorithm in Definition 46, *i.e.* $\mathcal{T} = \mathcal{T}_{ideal}$. Here $\mathcal{T}_{ideal}$ again relates to the fidelity distance between the two states,

$$\Pr[1 \leftarrow \Lambda^N(|\phi^{\mathbf{c}}\rangle, |\phi^{\mathbf{b}}\rangle)] = \Pr[1 \leftarrow \mathcal{T}_{ideal}] = F(|\phi^{\mathbf{b}}\rangle, \rho_e^N) \tag{6.52}$$

Since each b across the N positions are chosen independently and randomly, this implies at entangling the map across different rounds does not help Eve in any way. Thus to correctly guess the b values for all the N positions, the optimal attack strategy of Eve is to generate the reference state $\rho_{max}^{\otimes N}$, such that,

$$\forall i \in [N] \quad F(\rho^{\max}, |\phi_i^r\rangle) = \langle\phi_i^b|\,\rho^{\max}\,|\phi_i^b\rangle \geqslant \langle\phi_i^b|\,\rho_i\,|\phi_i^b\rangle \tag{6.53}$$

where $\rho_i = \mathrm{Tr}_{\{1\cdots N/i\}}(\rho_e^N)$, *i.e.* ρ_i is obtained by tracing out the N-1 instances $\{1, \cdots N/i\}$.

This further implies that attack map Λ^N is reduced to $\Lambda_{ind}^{\otimes N}$, where the map $\Lambda_{ind}^{\otimes N}$ involves a generator algorithm that produces the state ρ^{max} which maximises the average fidelity with verifier's response state across all the N rounds. This implies that,

$$\begin{aligned}\Pr[\{b_1, \cdots, b_N\} \leftarrow \Lambda^N(|\phi^{\mathbf{c}}\rangle, |\phi^{\mathbf{b}}\rangle)] &= \prod_{i=1}^{N} \Pr[b_i \leftarrow \Lambda_{ind}(|\phi_i^c\rangle, |\phi_i^b\rangle)] \\ &\leqslant \left(\frac{1}{2} + negl(\lambda)\right)^N\end{aligned} \tag{6.54}$$

where we used the result of Theorem 41 after the reduction from coherent to the collective attack. This completes the proof. □

(III) Comparing Classical and Quantum Strategies: Using the above Theorem 41 and Theorem 42 we show that a QPT Eve does not have any non-negligible advantage in passing the `cVer` verification test compared to the purely classical Eve. Thus, we can bound the success probability of a general QPT Eve which the success probability of the classical Eve from the Theorem 40,

$$\Pr[\text{Ver accept}_{\text{QPT Eve}}] \leqslant \Pr[\text{Ver accept}_{\text{Classical Eve}}] + \mathcal{O}(2^{-N}) \approx \mathcal{O}(2^{-N}) \tag{6.55}$$

6.2.4 Generalisation of low-resource protocol to arbitrary distribution of traps

In the original Irv-id protocol, verifier randomly picks half of the $N/2$ positions, and marks them $b = 1$. The rest is marked $b = 0$. Here, even though an adversary Eve does not know the locations of valid qPUF response states and the trap states, she knows that half of the positions are traps. In this section, we generalise the Irv-id protocol, to further hide the number of traps from Eve. This is done with the hope that hiding the number of trap and good response states could further decrease the probability of Eve passing the `cVer` test, especially against a fully classical Eve who only uses the statistics information to attack the protocol. Here verifier chooses an arbitrary number of trap positions. In other words, she randomly pics a value $p \in [0, 1]$, then randomly picks pN locations out of N and marks them $b = 1$ (valid response states). The rest of $(1-p)N$ positions are assigned $b = 0$ (trap positions). One can observe that the protocol on the prover's side does not depend on this value p hence verifier is not required to make the p value public.

We note that $b=1$ positions must all have bits valued 0, and $b=0$ positions must have half bits valued 0 and the rest are valued 1 (assuming $\delta_{er}=0$ for simplicity) if the N bits have to pass the classical verification algorithm `cVer`. Now, upon running the Irv-id protocol, there are in total $N(1+p)/2$ number of 0 bits and $N(1-p)/2$ number of '1' bits in the desired bit-string S_N which can pass the verification. Changing the tolerance value δ_{er} will not affect the result as we have seen in the previous section that by having a δ_{er} much smaller than N the probability only multiplies to a constant factor. We follow the same argument as in the proof of Theorem 40, for finding the optimal success probability of Eve generating successful bit-strings for the new classical verification. We say that the optimal strategies are the ones where their string space consists of exactly c_1 bits that are 1, where here $c_1=N(1-p)/2$. For the specific case of $p=0.5$, we have proven the optimality of such strategies. Hence in this specific case, we can refer to the same proof. In the generalised setting, the p value is unknown, and as a result, c_1 is unknown to Eve as well. Therefore the overall winning probability of Eve will depend on first guessing the correct values of c_1 and then the probability of such strings passing both tests. Also, we know that the probability of any strings with incorrect c_1 is necessarily 0, hence we can write the probability that Eve passes the verification test as follows,

$$\begin{aligned}\Pr[\text{Ver accept}_{\text{Eve}}] &= \Pr[\text{guess } c_1]\times\Pr[\text{Ver accept}_{\text{Eve},S_{gop}}|c_1=\frac{N(1-p)}{2}]\\ &= \Pr[\text{guess } c_1]\times\frac{\binom{N-Np}{\frac{N-Np}{2}}}{\binom{N}{\frac{N-Np}{2}}}\end{aligned} \tag{6.56}$$

Let us assume that verifer, in order to maximize the randomness over the correct choice of c_1, picks p completely uniformly from $[0,1]$. In this case, the number of trap responses can be any number between 0 (for $p=1$) and N (for $p=0$). Consequently, $c_1\in\{0,1,\ldots,\frac{N}{2}\}$ and if any of these values occur with equal probability, then Eve can guess c_1 with the following probability:

$$\Pr[\text{guess } c_1]=\frac{1}{\frac{N}{2}+1}$$

Now one can calculate the average wining probability of Eve over p:

$$\Pr[\text{Ver } \underset{p}{\text{accept}}_{\text{Eve}}]=\int_0^1\frac{2}{N+2}\frac{(N-Np)!(\frac{N+Np}{2})!}{N!(\frac{N-Np}{2})!}dp \tag{6.57}$$

Now, we approximate the above integral as follows:

Theorem 43 (Average success probability convergence with arbitrary distribution of traps)**.** *Let p be the probability of choosing correct responses in lrv-id protocol. Then the average wining probability of Eve over p, approximately converges as follows:*

$$Pr[Ver \underset{p}{accept}_{Eve}] \approx \overline{Pr_{win}} = \frac{2}{N+2}\sum_{k=0}^{N}\frac{(N-k)!(\frac{N+k}{2})!}{N!(\frac{N-k}{2})!} \approx \frac{6}{N(N+2)} = \mathcal{O}(\frac{1}{N^2}) \tag{6.58}$$

Proof. We approximate the following integral for the average probability that Eve wins the classical verification by performing the optimal classical strategy when p is chosen to be a uniform distribution.

$$\Pr[\text{Ver } \underset{p}{\text{accept}}_{\text{Eve}}] = \int_0^1 \frac{2}{N+2}\frac{(N-Np)!(\frac{N+Np}{2})!}{N!(\frac{N-Np}{2})!}dp$$

We choose $NP = k$ thus we have $Ndp = dk$ and we can rewrite the integral as:

$$\Pr[\text{Ver } \underset{p}{\text{accept}}_{\text{Eve}}] = \frac{2}{N(N+2)}\int_0^N \frac{(N-k)!(\frac{N+k}{2})!}{N!(\frac{N-k}{2})!}dk$$

Now we can approximate the integral for discrete $k \in \{0, 1, \ldots, N\}$. Hence we have:

$$\Pr[\text{Ver } \underset{p}{\text{accept}}_{\text{Eve}}] \approx \overline{\Pr[\text{Ver accept}_{\text{Eve}}]} = \frac{2}{N(N+2)}\sum_{k=0}^{N}\frac{(N-k)!(\frac{N+k}{2})!}{N!(\frac{N-k}{2})!}$$

The above series can be opened further as:

$$\begin{aligned}
\sum_{k=0}^{N}\frac{(N-k)!(\frac{N+k}{2})!}{N!(\frac{N-k}{2})!} &= 1 + \frac{(N-1)!}{N!}\times\frac{(\frac{N}{2}+\frac{1}{2})!}{(\frac{N}{2}-\frac{1}{2})!} + \frac{(N-2)!}{N!}\times\frac{(\frac{N}{2}+1)!}{(\frac{N}{2}-1)!} + \cdots + 1 \\
&= 1 + \frac{1}{N}\times\frac{(\frac{N}{2}+\frac{1}{2})\cancel{(\frac{N}{2}-\frac{1}{2})!}}{\cancel{(\frac{N}{2}-\frac{1}{2})!}} + \frac{1}{N(N-1)}\times\frac{(\frac{N}{2}+1)(\frac{N}{2})\cancel{(\frac{N}{2}-1)!}}{\cancel{(\frac{N}{2}-1)!}} + \cdots + 1 \\
&\underset{N\gg 1}{\approx} 2 + \frac{\frac{N}{2}}{N} + \frac{(\frac{N}{2})^2}{N^2} + \frac{(\frac{N}{2})^3}{N^3} + \ldots \\
&= 2 + \sum_{i=1}^{N-1}(\frac{1}{2})^i \approx 2 + (1 - 2^{1-N}) \approx 3
\end{aligned} \tag{6.59}$$

where the sum has been approximated for large N. Thus we can write the average probability in the limit of large N as follows,

$$\Pr[\text{Ver } \underset{p}{\text{accept}}_{\text{Eve}}] \approx \overline{\Pr[\text{Ver accept}_{\text{Eve}}]} = \frac{6}{N(N+2)} \tag{6.60}$$

This concludes the proof. □

This means that by choosing p form a uniform distribution, the average success probability of the adversary becomes polynomially small in N which reduces the security of the protocol to polynomial. This may seem a surprising result although the reason is that the probability function for $p=0$ and $p=1$ is 1. On the other hand, from the security result for $p=\frac{1}{2}$, we know that the probability function's behaviour can be inverse exponential. This gives rise to the interesting question of whether one can find a boundary for p in which $\Pr[\text{Ver accept}_{\text{Eve}}]$ is negligible. Before addressing this problem, it is worth mentioning that by hiding p, one can hope the protocol's security to be boosted by at most a polynomial factor $(\frac{1}{\mathcal{O}(N)})$ as Eve's probability of guessing the correct c_1 depends only on the different number of 1's in the string that results from different choices of p. Even though for large N this polynomial factor can be ignored, assuming that the verifier has a good choice of p which leads to exponential security, in relatively smaller N the hiding can practically boost the security of the identification.

Now to be able to analyse the $\Pr[\text{Ver accept}_{\text{Eve}}]$, we rewrite the factorials with Gamma function and we define $z=\frac{N-Np}{2}$ where $z\in\{0,1,\ldots,\frac{N}{2}\}$. Considering that $\Gamma(z+1)=z\Gamma(z)$, the probability is,

$$\begin{aligned}\Pr[\text{Ver accept}_{\text{Eve}}] &= \frac{(N-Np)!(\frac{N+Np}{2})!}{N!(\frac{N-Np}{2})!} = \frac{\Gamma(2z+1)\Gamma(N-z+1)}{N!\Gamma(z+1)} \\ &= \frac{2}{N!}\times\frac{\Gamma(2z)\Gamma(N-z+1)}{\Gamma(z)}\end{aligned} \tag{6.61}$$

Using properties of Gamma functions we have that $\frac{\Gamma(2z)}{\Gamma(z)}=\frac{2^{2z-1}}{\sqrt{\pi}}\Gamma(z+\frac{1}{2})$. Thus we can simplify the function to be:

$$\begin{aligned}\Pr[\text{Ver accept}_{\text{Eve}}] &= \frac{2}{\sqrt{\pi}}\times\frac{2^{2z-1}}{N!}\Gamma(z+\frac{1}{2})\Gamma(N-z+1) \\ &\approx \frac{2^{2z-1}}{N!}\Gamma(z+\frac{1}{2})\Gamma(N-z+1)\end{aligned} \tag{6.62}$$

For a large enough fixed N, the factor $\frac{2^{2z-1}}{N!}\ll 1$. However it is an increasing function in z and $\Gamma(z+\frac{1}{2})\Gamma(N-z+1)$ is a large factor which quickly decreases with z. Also at the beginning and the end of the period where $z=0, z=\frac{N}{2}$, the probability is 1, and it reduces to a small value for certain z. Thus we deduce that the function will necessarily have a minimum for any N. The Fig. 6.6, different $\Pr[\text{Ver accept}_{\text{Eve}}]$ for different N has been shown. We have renormalised the probabilities as a function of p to be able to compare them. As we can see, the function for all the different values of N falls exponentially in a minimum region where there are the desirable values of p. As N grows, the range of desirable p expands, which can be found in the top right plot where we compare the probability for $N=16, N=32$ and $N=64$. Also by comparing the probability range for $N=10, N=100, N=150$ one can see how the exponential security is achieved for a p which has been chosen in the *good* region. This specification of the success probability would be useful for the verifier to be able to optimise

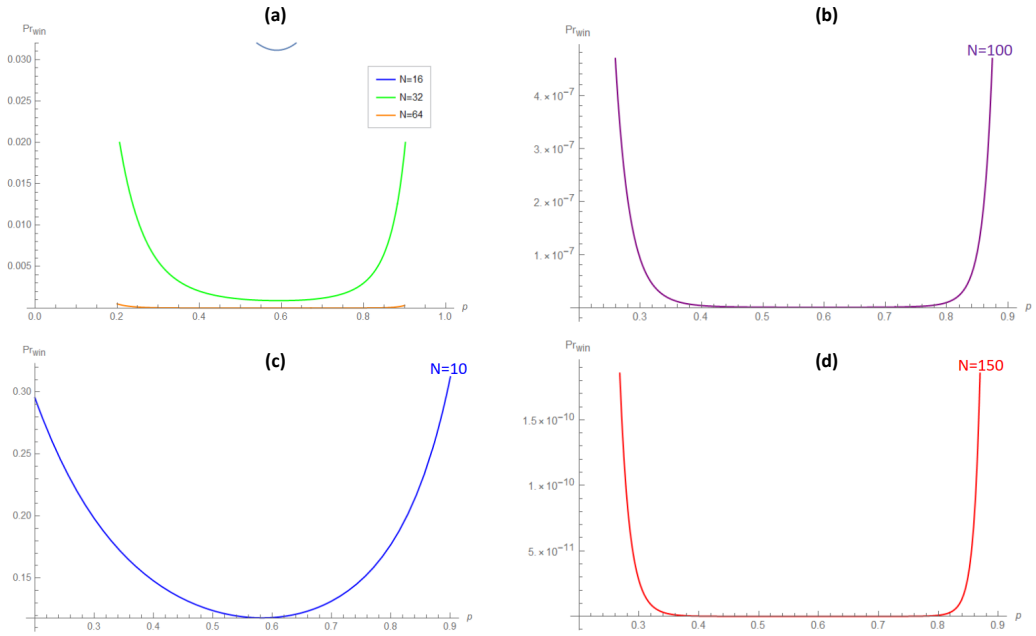

Figure 6.6: Behaviour of Eve's success probability Pr[Ver accept$_{\text{Eve}}$] as a function of p (corresponding to number of valid qPUF responses), for different values of N.

the protocol based on her resources. Moreover, the freedom of choosing traps according to desired distribution, conditioning that it bounds the value of p to the minimum region, enables the protocol to be useful in other scenarios.

6.2.5 Resource comparison of protocols

The two proposed qPUF-based identification protocols differ a great deal in terms of the type and amount of resources available to the concerned parties. We divide the resources into three categories: quantum memory, quantum computing ability, and the number of communication rounds required to achieve identification. Here, quantum memory is quantified by the number of quantum states stored in a register, and the computing ability resource is quantified in terms of the number of quantum gates required to implement a specific quantum circuit.

Table 6.1 compares the resources of the two protocols that we have introduced. For a fair comparison between the above protocols, we fix the maximum acceptance probability for any QPT adversary, Pr[Ver accept$_{\text{Eve}}$], to be ε, and compute the number of resources required to achieve that desired acceptance probability. In all the protocols, we assume that during one identification, N copies of different states, each with M identical copies are used. For the specific case of lrv-id protocol, $M = 1$. For the hrv-id-swap protocol, where the quantum verification is via the SWAP test circuit, the adversary's acceptance probability is $\varepsilon = \mathcal{O}(2^{-MN})$. In this protocol, the verifier requires $MN = \mathcal{O}(\log 1/\varepsilon)$ size quantum memory and computing ability of $\mathcal{O}(poly \log D)$ quantum gates, where D is the size of qPUF. The prover, on the other hand, requires no quantum memory and computing ability. The number of communication rounds required to achieve the desired

Protocol	Security		Quantum Memory		Verification computing ability		Communication round	
			Verifier	Prover	Verifier	Prover	Quantum	Classical
hrv-id-swap	ε	$= 2^{-MN}$	$\log 1/\varepsilon$	0	$poly \log D$	0	$\log 1/\varepsilon$	0
hrv-id-gswap		$= (M+1)^{-N}$	$\frac{M}{\log M+1} \log 1/\varepsilon$	0	$poly \log MD$	0	$\frac{1}{\log M+1} \log 1/\varepsilon$	0
lrv-id		$= 2^{-N}$	$\log 1/\varepsilon$	0	0	$poly \log D$	$\log 1/\varepsilon$	1

Table 6.1: Comparison of different qPUF-based identification protocols in terms of security ($\Pr[\text{Ver accept}_{\text{Eve}}] = \varepsilon$) against any QPT adversary and the three resource categories of the verifier and the prover: quantum memory, computing ability and number of communication rounds. Here all the resources are in $\mathcal{O}(.)$. All our proposed protocols exhibit ε exponential security with polynomial sized resource $\mathcal{O}(\log 1/\varepsilon)$ memory/communication and $\mathcal{O}(poly \log D)$ computing ability in both the parties. Here D is the size of qPUF.

security is $MN = \mathcal{O}(\log 1/\varepsilon)$. The protocol hrv-id-gswap, where the quantum verification is via the GSWAP test circuit, the adversary's acceptance probability is $\varepsilon = \mathcal{O}((M+1)^{-N})$. In this protocol, the verifier requires $MN = \mathcal{O}(\frac{M}{\log M+1} \log 1/\varepsilon)$ size quantum memory and a computing ability of $\mathcal{O}(poly \log MD)$ quantum gates. Similar to hrv-id-swap, the prover requires no quantum memory and computing ability. The number of communication rounds required to achieve the desired security is $N = \mathcal{O}(\frac{1}{\log M+1} \log 1/\varepsilon)$. Thus for large M values, the verifier's quantum memory requirement is less while using SWAP compared to GSWAP, but the number of communication rounds is higher using the SWAP test.

Now for the lrv-id protocol, the protocol with the low-resource verifier, the adversary's acceptance probability is $\varepsilon = \mathcal{O}(2^{-N})$. In this protocol, the verifier requires $N = \mathcal{O}(\log 1/\varepsilon)$ size quantum memory. Since the verifier performs classical verification, hence she does not require a quantum computing ability. The prover here requires no quantum memory but since he performs the SWAP test circuit, his computing ability is required to be $\mathcal{O}(poly \log D)$. The number of quantum communication rounds required to achieve the desired security is $N = \mathcal{O}(\log 1/\varepsilon)$. This protocol also requires a single round of classical communication transmitting N bits.

Fig. 6.7 demonstrates the graphical comparison of different resources among the three qPUF-based identification protocols. The plots show a tradeoff in resources between different protocols to achieve the desired success probability of ε. We choose the ε to range from 10^{-6} to 10^{-1}. Since the computing ability resource depends on the qPUF size D, we choose $D = 1/\varepsilon$ for comparison.

We identify that the difference in resources primarily comes about due to the different requirements of SWAP and GSWAP tests. To illustrate this graphically, we provide density plots in Fig. 6.8 to showcase the trade-off between the success probability ε and the memory and communication round resources required for different M ad N's for protocols based on SWAP vs GSWAP tests.

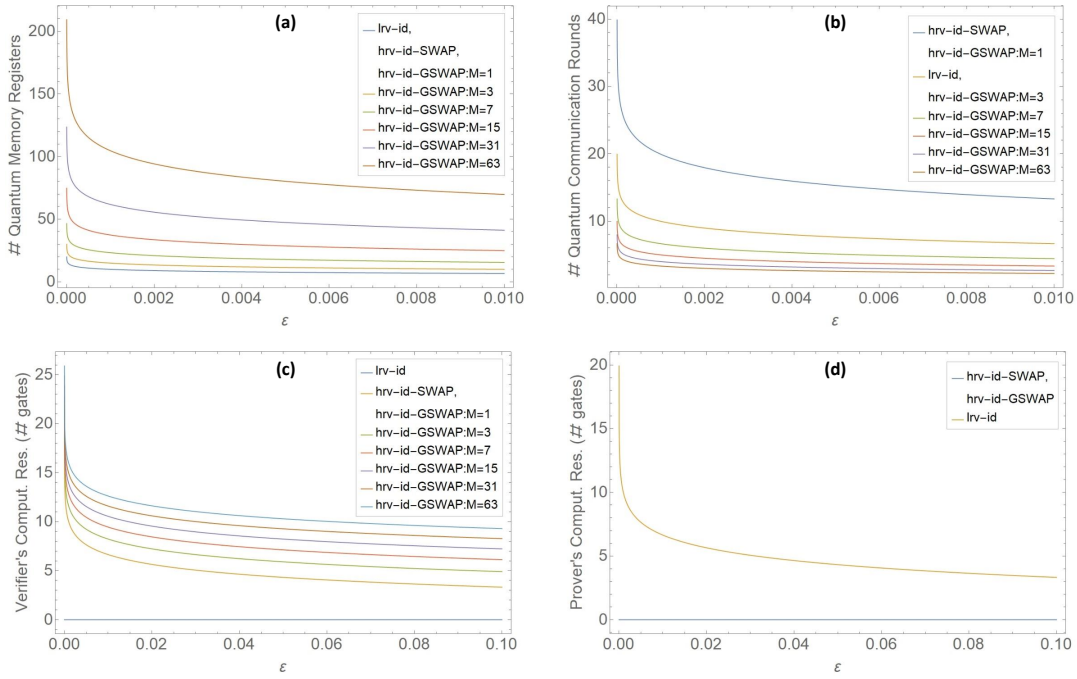

Figure 6.7: Comparison of the resources required by the prover and verifier in the three qPUF-based identification protocols (hrv-id-swap, hrv-id-gswap, and lrv-id) for varying security values ε. We choose the ε to range from 10^{-6} and 10^{-2} for the top row and between 10^{-6} and 10^{-1} for the bottom row. Plot top left compares the verifier's quantum memory resource vs ε for the three protocols. The plot shows that the least memory requirement is minimum in hrv-id-swap and lrv-id protocols while it increases by increasing the number of local copies M required in the GSWAP test for hrv-id-gswap protocol. We note that the prover's memory requirement is 0 in all three protocols. Plot top right similarly compares the number of quantum communication rounds in the three protocols. The requirement is minimum in the lrv-id while it increases with M in the hrv-id-gswap. The communication round in hrv-id is double compared to the lrv-id requirement to indicate the two-way quantum communication instead of one way in the latter. Plots bottom left and bottom right compares the computational resource vs ε for the verifier and prover respectively. Here we have taken $D = 1/\varepsilon$ for comparison.

6.3 Towards more efficient qPUF-based identification protocols

In this section, we show that using our results form Chapter 5, we can make the protocols we presented in the previous section yet more efficient. We have seen how these identification protocols exploit the unforgeability of qPUFs, to achieve exponential security against QPT adversaries in a polynomial number of rounds. We have already discussed that these protocols are resource-efficient in many aspects, but one of the main practical challenges in implementing them is to sample the challenge states at random from the Haar measure. We have seen that this requirement is crucial for achieving unforgeability for qPUFs. Nevertheless, Theorem 30 showed that universal unforgeability can still be achieved with the same security guarantee if the states are sampled from a PRS family instead of Haar-random states. Here, we show that the qPUF-based identifica-

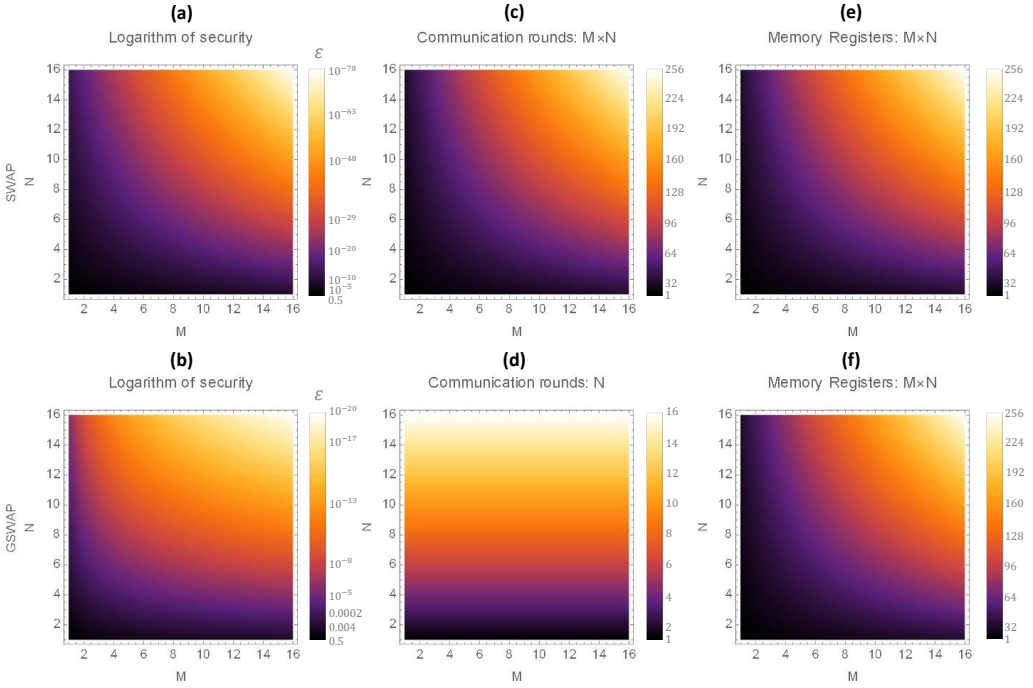

Figure 6.8: Comparison of verification based on SWAP and GSWAP for identification protocols. The top row is associated with SWAP and the button row with GSWAP. The x-axis of the plots are all M (the number of local copies) and the y-axes are all N (the number of different states) and the security, quantum memory and quantum communications have been shown with a colour spectrum. The left column shows the security ε where we have $\varepsilon = 2^{-MN}$ for SWAP and $\varepsilon = (M+1)^{-N}$ for GSWAP, in a logarithmic scale for more visibility. The middle column shows the required communication where we see that for GSWAP the communication rounds are independent of M and only linearly growing with N while for SWAP the communication rounds grow also linear by increasing the number of local copies. The right column shows the memory which has been fixed for both SWAP and GSWAP to $M \times N$. The comparison between security and communication plots shows a trade-off between SWAP and GSWAP as the quantum verification algorithm.

tion protocols can also achieve exponential security using PRS. This transition to a more efficient sampling of challenge states brings us one step closer to the practical implementations of quantum identification protocols with exponential security against powerful quantum adversaries and leads to promising solutions to the problem of untrusted manufacturers.

We start with hrv-id, and we introduce a computationally efficient variation of this protocol which we call *Efficient-hrv-id* protocol, by replacing the Haar-random challenges with pseudorandom quantum states in the setup phase as follows:

Protocol 3 (Efficient-hrv-id)**.** Efficient version of hrv-id protocol using pseudorandom challenges:

---

1. *Setup Phase*:
   (a) Verifier has the qPUF device with unitary evaluation U.
   (b) Verifier has also access to a family of PRS $\{|\phi_k\rangle \in S(\mathcal{H}^d)\}_{k\in\mathcal{K}}$ and randomly picks $Q \in \mathcal{O}(\text{poly}\log d)$ of them as the challenge states.

(c) Verifier queries the U individually with each challenge $|\phi_k\rangle$ a total of M number of times to obtain M copies of the response state $|\phi_k^r\rangle$ and stores them in their local database S.

(d) The verifier transfers the U to Prover.

2. *Identification phase*: same as hrv-id.

3. *Verification phase*: same as hrv-id.

The following statement which is a corollary of the previous results shows that the *Efficient-hrv-id* protocol is also exponentially secure against QPT adversaries with the same security bounds.

Corollary 6. *Let U be a UqPUF over $\mathcal{H}^D$. The success probability of any QPT adversary to pass the* SWAP*-test or* GSWAP*-test verification of the* Efficient-hrv-id *is at most ε, given that there are N different CRPs, each with M copies. The ε is bounded as follows for each verification:*

$$Pr[Ver\ accept_{\mathcal{A}}] \leqslant \varepsilon \qquad \varepsilon_{SWAP} \approx \mathcal{O}(\frac{1}{2^{NM}}) \qquad \varepsilon_{GSWAP} \approx \mathcal{O}(\frac{1}{(M+1)^N}) \tag{6.63}$$

Proof. First, Theorem 30 that states any unknown unitary satisfies efficient universal unforgeability where the challenge states are selected from a PRS family. Then we can directly use the result states in Theorem 36 and Theorem 38 using the SWAP and GSWAP test which shows the same security bound in the number of rounds and copies of challenge-response pairs. □

Similarly, we can improve the efficiency of the second protocol, namely lrv-id by substituting the Haar random states with PRS as follows:

Protocol 4 (Efficient-lrv-id)**.** Efficient version of lrv-id protocol using pseudorandom challenges

---

1. *Efficient-lrv-id Setup Phase*:

(a) Verifier has the qPUF device with unitary evaluation U.

(b) Verifier has also access to a family of PRS $\{|\phi_k\rangle \in S(\mathcal{H}^d)\}_{k\in\mathcal{K}}$ and randomly picks $Q \in \mathcal{O}(\text{poly}\log d)$ of them as the challenge states.

(c) Verifier queries the U individually with each challenge $|\phi_k\rangle$ a total of M number of times to obtain M copies of the response state $|\phi_k^r\rangle$ and stores them in their local database S.

(d) Verifier selects states $|\phi^{\perp}\rangle$ orthogonal to the selected challenge's subspace and queries the U with them to obtain the trap states labelled as $|\phi^{\text{trap}}\rangle$. The unitary property ensures that $\langle\phi^{\text{trap}}|\phi_k^r\rangle = 0$.

(e) The verifier transfers the U to Prover.

2. *Identification phase*: same as lrv-id.

3. *Verification phase*: same as lrv-id.

Again using the proof techniques presented in the previous sections, and our results from Chapter 5, we show that the *Efficient lrv-id protocol* satisfies exponential security against QPT adversary both under the coherent and collective attack models.

Corollary 7. *Let U be a UqPUF over $\mathcal{H}^D$. The success probability of a QPT adversary $\mathcal{A}$ to pass the verification of the* Efficient lrv-id protocol *is at most ε, in N rounds. The ε is bounded as follows:*

$$Pr[Ver\ accept_{\mathcal{A}}] \leqslant \varepsilon \qquad \varepsilon \approx \mathcal{O}(\frac{1}{2^N}) \tag{6.64}$$

Proof. First, we specify that we can directly use the result of Theorem 40 which bounds the success probability of a classical adversary in passing the classical verification algorithm. Then the success probability against a quantum adversary with the collective and coherent attack is defined as the advantage of the quantum adversary over that classical adversary in guessing the trap states, using all the side information obtained from the U in the learning phase. We use Theorem 30 that states any unknown unitary satisfy efficient universal unforgeability with PRS challenge states. Next, the conditions of Theorem 41 and Theorem 42 are satisfied and we can directly use those results which gives the following bounds.

$$\Pr[b \leftarrow \Lambda_{\mathcal{A}}] \leqslant \frac{1}{2} + \mathcal{O}(2^{-N}) \tag{6.65}$$

Where $\Lambda_{\mathcal{A}}$ denotes any map that $\mathcal{A}$ uses to distinguish the traps states. Finally, putting all the above results together we have

$$\Pr[\text{Ver accept}_{\mathcal{A}}] \leqslant \varepsilon = \Pr[\text{Ver accept}_{\text{Classical Adv}}] + \mathcal{O}(2^{-N}) \approx \mathcal{O}(2^{-N}) \tag{6.66}$$

This concludes the soundness proof of *Efficient lrv-id protocol*. □

6.4 Hybrid PUF: A practical solution

As our last contribution towards the practical realisation of qPUF-based applications, we propose a more implementation-friendly construction called Hybrid PUF (HPUF). This new type of PUF, as opposed to quantum PUFs which exploit quantum randomness, uses a classical PUF as a weak source of randomness and enhances it using quantum communication, hence the name *Hybrid PUF*.

Let us recall the main implementation challenges of qPUF-based identification protocols and see how our proposal for a hybrid construction can overcome these challenges. The first challenge is the implementation of the UqPUF itself, which requires either sampling unitaries from Haar-measure or equivalently a family of PRU or UU. As mentioned in Chapter 4, one of the promising candidates for qPUFs is optical devices. Nevertheless, in using them as qPUFs the main challenge will be to certify the unknownness property and dimension of the unitary since these parameters are directly related to the security of qPUFs. On the other hand, the literature on classical PUFs is rich, and there is a multitude of constructions available based on several different hardware technologies [GCvDD02, GKST07, KL18, Mae13]. Although all of those constructions can be manufactured quite easily and they provide unique and inexpensive hardware fingerprints, they all suffer from the lack of enough randomness and as a result, do not provide satisfactory unpredictability. Thus, most of the existing CPUF constructions are vulnerable against the machine learning modeling-based attacks [Bec15a, Bec15b, Del19, RSS+10]. In these types of attacks, the attacker first collects a sufficient number of CRPs by adaptively querying the PUF. Then the collected data is used to derive a numerical model that mimics the behaviour of the PUF, using the tools and techniques from machine learning. The central idea behind the Hybrid PUF is to use a classical PUF as an embedded hardware module that is easy to implement but does not offer suitable security, to construct a secure hardware token that uses commercially available quantum communication tools and provides sophisticated security guarantees. At the same time, we aim for a technologically available construction to overcome the manufacturing obstacle of a secure PUF that uses quantum CRPs.

The second major challenge regarding qPUF-based identification protocols becoming widely available today or in the near term is the fact that the verifier needs to store the CRP database on a quantum memory. Although there has been significant progress in the implementation of quantum memories in the recent years [LST09, WLZ+19, BRA+19, GI20, LRGR+21, BBFO+19, WMH+20, DKLP02], storing large quantum states for a considerable time is still infeasible given today's technology. The hybrid construction can solve this problem by fully removing the quantum memory requirement. Since an HPUF encodes classical responses in separable single-qubit states, verifying the response states can be done more easily using the underlying classical information from the CPUF, and as we will show, having a classical database will suffice to verify an HPUF.

Finally, through studying this construction, we will also address a long-standing open problem in the field of PUF-based identification, which is the re-usability of challenge-response pairs stored in the database. One significant drawback of the PUF-based authentication protocols is that the server cannot use the same challenge multiple times to authenticate a client due to man-in-the-middle attacks. There is no way to avoid this limitation for classical PUFs. However, in this section, we show that due to the entropic uncertainty principle of quantum information theory, with our given construction, the server can reuse a challenge as long as they have been successfully authenticated by the client using that challenge in

the previous rounds, overcoming this problem and proving for the first time, the challenge re-usability of PUF-based application.

In this section, we give a construction for HPUF based on *conjugate coding*, which enhances the security of classical PUFs against a weaker class of quantum adversaries (as opposed to adaptive QPT adversaries that we have been considering so far). Then by exploiting a technique known as *Locking mechanism*, we boost the security of this construction to our usual adaptive QPT adversaries, therefore presenting the concept of *Hybrid Locked PUF* (HLPUF) as our candidate for a secure and efficiently implementable quantum hardware token. We then show a secure HLPUF-based identification protocol. Finally, we formally prove the challenge re-usability property of this protocol.

But first, we introduce the theoretical model that we adopt for the CPUF which is going to be used as the building block of our construction.

6.4.1 CPUF model

We have introduced classical PUFs in Chapter 4. Here we provide some additional technical tools and definitions that we need to introduce our Hybrid construction. Classical PUFs are usually defined with probabilistic functions, due to their inherent physical randomness. Here we also consider them as probabilistic functions.

A classical PUF can be modelled as a probabilistic function $f : \mathcal{R} \times \mathcal{X} \to \mathcal{Y}$ where $\mathcal{X}$ is the input space, $\mathcal{Y}$ is the output space of f and $\mathcal{R}$ is the identifier. As defined in 4.2, the creation of a classical PUF is formally expressed by invoking a manufacturing process $f \leftarrow \mathrm{Gen}_C(\lambda)$, where λ is the security parameter. Here f is the evaluation algorithm of CPUF which needs to satisfy the requirements of robustness, collision-resistance and uniqueness as defined before [AMSY16]. For a fixed input $x \in \mathcal{X}$, and a random coin (or key) $R \leftarrow \mathcal{R}$, we denote the probability distribution of the output random variable $f(x) := f(R, x)$ over all $y \in \mathcal{Y}$ as,

$$p_x^f(y) := Pr[f(x) = y|x] = \sum_{r:f(r,x)=y} Pr[R = r]. \tag{6.67}$$

Now, let us define a *parameterised randomness* definition for the classical PUF f as follows:

Definition 47 (p-Randomness)**.** We define the p-randomness of a classical PUF $f : \mathcal{R} \times \mathcal{X} \to \mathcal{Y}$ as

$$p := \max_{\substack{x \in \mathcal{X} \\ y \in \mathcal{Y}}} p_x^f(y) = \max_{\substack{x \in \mathcal{X} \\ y \in \mathcal{Y}}} \Pr_R[f(R, x) = y]. \tag{6.68}$$

We use this definition to characterize the quality of a CPUF with a quantitative measure of its randomness. This parameter has some relation to a property of min-entropy for classical PUFs but is formally defined differently. Nevertheless, for technical purposes, we choose to use this definition.

6.4.2 Construction for Hybrid PUF

For our construction, we start with a classical PUF with a certain amount of randomness characterized by the p-randomness value we defined in Definition 47. We construct a new PUF that is the combination of a classical PUF, and a *quantum encoder*, which encodes the output of the CPUF into non-orthogonal quantum states. The output qubits are the response of the PUF and will be sent through the quantum communication channel. We refer to the entire system, *i.e.* the combination of CPUF and the quantum encoding, as *Hybrid PUF (HPUF)*. A HPUF receives a *classical* challenge and produces a *quantum* response. In Construction 3 we give a simple design of a HPUF based on conjugate coding [Wie83].

Construction 3 (Hybrid PUF)**.** Suppose $f : \{0,1\}^n \to \{0,1\}^{4m}$ be a classical PUF, that maps an n-bit string $x_i \in \{0,1\}^n$ to an $4m$-bit string output $y_i \in \{0,1\}^{4m}$. We denote the j-th bit of y_i as $y_{i,j} \in \{0,1\}$. From the $4m$-bit string, we prepare the set of $2m$-tuples $\{(y_{i,(2j-1)}, y_{i,2j})\}_{1 \leqslant j \leqslant 2m}$. The hybrid PUF encodes each of the tuples $(y_{i,(2j-1)}, y_{i,2j})$ into a single qubit *BB84 states*, $|\psi^{i,j}\rangle$. The exact expression of the encoding is defined in the following way,

$$|\psi^{i,j}_{out}\rangle\langle\psi^{i,j}_{out}| := \begin{cases} |0\rangle\langle 0| & (y_{i,(2j-1)}, y_{i,2j}) = (0,0) \\ |1\rangle\langle 1| & (y_{i,(2j-1)}, y_{i,2j}) = (1,0) \\ |+\rangle\langle +| & (y_{i,(2j-1)}, y_{i,2j}) = (0,1) \\ |-\rangle\langle -| & (y_{i,(2j-1)}, y_{i,2j}) = (1,1) \end{cases} \tag{6.69}$$

For any $x_i \in \{0,1\}^n$, the mapping of the HPUF $\mathcal{E}_f : \{0,1\}^n \to (\mathcal{H}^2)^{\otimes 2m}$ is defined as follows.

$$x_i \to |\psi^i_{out}\rangle\langle\psi^i_{out}| \quad (\text{or } |\psi_{f(x_i)}\rangle\langle\psi_{f(x_i)}|) \tag{6.70}$$

where $|\psi^i_{out}\rangle\langle\psi^i_{out}| = \bigotimes_{j=1}^{2m} |\psi^{i,j}_{out}\rangle\langle\psi^{i,j}_{out}|$.

Intuitively, if the adversary wants to extract the information about the $i,2j$-th bit out of the classical PUF corresponding to a challenge x_i, they need to guess whether the state is prepared in $Z = \{|0\rangle, |1\rangle\}$ basis or in $X = \{|+\rangle, |-\rangle\}$ basis, then knowing the encoded bit. In Section 6.4.5, we estimate the success probability of a (weak) adversary in winning the universal unforgeability game for the HPUF as a function of the number of required queries.

Another remark here is that HPUF can be considered and studied within the quantum PUF framework that we have defined in Chapter 4. However, one should consider it as a non-unitary qPUF since it includes a classical pre-processing and state preparation which can be described by CPTP maps but is not necessarily a unitary. Nevertheless, we treat the HPUF as a new type of PUF and prove its unforgeability in a stand-alone manner. Moreover, we require that the CPUF inside

the construction satisfies the robustness and collision resistance requirements, as have also been defined for the qPUFs. If these requirements are satisfied by the CPUF, the HPUF will also deliver them (to the same degree) since the encoding part of the construction is fully deterministic. Finally, we do not investigate the uniqueness property of the HPUF here as we use it as a single device within the other construction and protocol that we will present, which only assume the robustness, collision resistance and p-randomness from the classical PUF.

6.4.3 Hybrid Locked PUF

As discussed before, most classical PUFs are vulnerable to machine learning attacks. To perform such attacks, the adversary needs to get access to many CRPs of the PUF to use this data for training and obtaining a model for CPUF. One common scenario that makes such attacks viable is when an adversary intercepts the communication channel between the verifier and prover during an identification protocol and pretends to be the verifier. Then the adversary can send their favourite queries as challenges to the prover, who will provide the adversary with the correct response. In this way, an adversary can build a local database, even during the identification phase. To address this issue, a technique has been introduced in the literature of classical PUFs known as *Lockdown technique* (or locking mechanism) [YHD+16] that upper-bounds adversary's capability in querying CRPs by converting the adaptive adversary into a weak one. We recall from Section 3.5.3 that a *weak adversary*, in contrast to an *adaptive adversary* who queries the oracle or device with their chosen and potentially adaptive queries, has only access to a random set of challenges and responses (or input-output queries) that are selected at random by an honest party. we recall that this is equivalent to the random-message attack model as we have also discussed in Section 2.5.2.

The central idea of the locking mechanism is that the prover (client) can also identify the verifier (server) during the identification. This mutual identification prevents an adversary from querying the PUF arbitrarily. One method is that the verifier sends part of the response along with the challenge so that the prover having access to the PUF device can check if the challenge has really come from the server or from an adversary who is trying to increase their information on the PUF device. We adopt the idea of the locking mechanism and we apply it on a HPUF that leads to our next construction, namely Construction 4. We refer to this construction as Hybrid Locked PUFs (HLPUFs). First, we devide the output of the HPUF $\mathcal{E}_f : \{0,1\}^n \to (\mathcal{H}^2)^{\otimes 2m}$ corresponding to a classical PUF $f : \{0,1\}^n \to \{0,1\}^{4m}$ into two separate parts. The first part contains the first m qubits, and the second half contains the last m qubits of the outcome of HPUF. Note that, the first m qubits of the HPUF's outcome is generated from the first $2m$ bits outcome of the corresponding classical PUF f. For any challenge $x \in \{0,1\}^n$ we can write the outcome of the classical PUF as $f(x) = f_1(x)||f_2(x)$, where the mapping $f_1 : \{0,1\}^n \to \{0,1\}^{2m}$ denotes the first $2m$ bits of f and $f_2 : \{0,1\}^n \to \{0,1\}^{2m}$ denotes the last $2m$ bits of f. Similarly, we can rewrite the HPUF $\mathcal{E}_f$ as a tensor product of two mappings $\mathcal{E}_{f_1} : \{0,1\}^n \to (\mathcal{H}^2)^{\otimes m}$, and

$\mathcal{E}_{f_2} : \{0,1\}^n \to (\mathcal{H}^2)^{\otimes m}$, where for any challenge $x \in \{0,1\}^n$, $\mathcal{E}_{f_1}(x)$ denotes the first m qubits of $\mathcal{E}_f(x)$, and $\mathcal{E}_{f_2}(x)$ denotes the last m qubits of $\mathcal{E}_f(x)$.

The hybrid locked PUF, takes the classical input x_i and a quantum state $\tilde{\rho}_1$ and produces the second half of the response of the hybrid PUF, $|\psi_{f_2(x_i)}\rangle\langle\psi_{f_2(x_i)}|$, as an output if $\tilde{\rho}_1$ is equal to the first half of the output of the hybrid PUF $|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}|$. The construction is shown in Fig. 6.9. We formalise it as follows.

Construction 4 (HLPUF)**.** Suppose we have a hybrid PUF $\mathcal{E}_f$ where $f : \{0,1\}^n \to \{0,1\}^{4m}$ is a CPUF. The mapping of the HLPUF $\mathcal{E}_f^L : d_{in} \times \mathcal{H}^{d_{out_1}} \to \mathcal{H}^{d_{out_2}} \otimes \mathcal{H}^{\perp}$ corresponding to a hybrid PUF $\mathcal{E}$ is defined as follows:

$$(x_i, \tilde{\rho}_1) \to \begin{cases} |\psi_{f_2(x_i)}\rangle\langle\psi_{f_2(x_i)}| & \text{if } \texttt{Ver}(|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}|, \tilde{\rho}_1) = 1 \\ \perp & \text{otherwise.} \end{cases} \tag{6.71}$$

where $\texttt{Ver}(.,.)$ is verification algorithm that checks the equality of the first half of the response based on the classical response y_i^1.

More precisely, $\texttt{Ver}(.,.)$ is specified by measuring each qubit of the incoming quantum state with corresponding basis according to $\{y_{i,2j}\}_{1\leqslant j\leqslant 2m}$ of response y_i and check the equality $\texttt{Equal}(y_{i,2j}, \tilde{y}_{i,2j})_{1\leqslant j\leqslant 2m}$ in our construction.

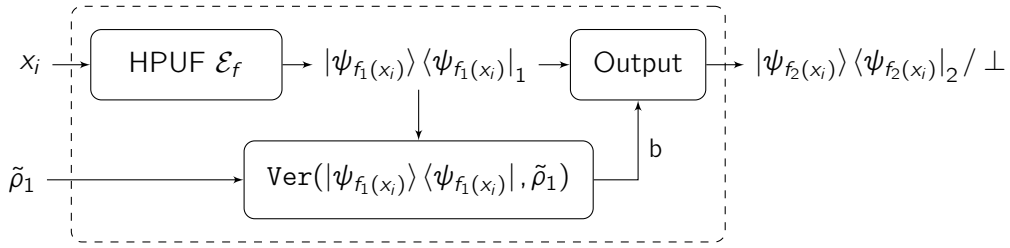

Figure 6.9: Hybrid Locked PUF (HLPUF) $\mathcal{E}_f^L$ with Construction 4

In [CDM+21] the possibility of exploiting the lockdown technique for quantum PUFs has also been investigated where we have developed the mathematical model for it. First of all, it is interesting to see whether the lockdown technique can enable to reduce the adversarial power in the quantum case as well. Moreover, this applicated is well-motivated, especially for the weaker types of qPUF that we have mentioned before, such as QR-PUFs, since arbitrarily querying the PUF with multiple copies allows for quantum emulation attacks (discussed in Chapter 4). One possible way to protect the PUF from such sophisticated attacks is to use the locking technique. However, similar to the HPUF setting, a central component of the locking mechanism is the verification subroutine. In the quantum case, this verification consists of testing the equality of unknown quantum states. We show a no-go result stating that due to the entanglement that can be generated by the unknown unitary over the subsystems of the response, the verification of such subsystems in a way that can be used for a locking mechanism is impossible,

unless in very limited cases. Nevertheless, we avoid presenting the details of this result here, and we conclude with this brief mention not to make the section unnecessarily long. We refer the reader to the main paper for more information.

6.4.4 Quantum identification protocol using Hybrid Locked PUF

After introducing the HLPUF construction, we describe an identification protocol based on it. The description of the protocol is given in Protocol 5. Note that we use $\tilde{\rho}_1$ and $\tilde{\rho}_2$ to denote the quantum state received by the prover/verifier respectively.

Protocol 5. [HLPUF-based Authentication] An authentication protocol based on HLPUF construction.

---

1. **Set-up:**
   (a) The Prover $\mathcal{P}$ equips a Hybrid Locked PUF: $\mathcal{E}_f^L$ with HPUF $\mathcal{E}_f : \{0,1\}^n \to (\mathcal{H}^2)^{\otimes 2m}$ constructed upon a classical PUF $f : \mathcal{X} \to \mathcal{Y}$. Here, the classical PUF f maps an n-bit string $x_i \in \{0,1\}^n$ to an $4m$-bit string output $y_i \in \{0,1\}^{4m}$.
   (b) The Verifier $\mathcal{V}$ has a classical database $D := \{(x_i, y_i)\}_{i=1}^d$ with all d CRPs of f, as well as the necessary quantum devices for preparing and measuring quantum states.
2. **Authentication:**
   (a) $\mathcal{V}$ randomly chooses a CRP (x_i, y_i) and splits the response equally into two partitions $y_i = f_1(x_i)||f_2(x_i) = y_i^1||y_i^2$ with length $2m$.
   (b) $\mathcal{V}$ then encodes the first partition of response into $|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}| := \bigotimes_{j=1}^m |\psi^{i,j}_{f_1(x_i)}\rangle\langle\psi^{i,j}_{f_1(x_i)}|$ and issues the joint state $(x_i, |\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}|)$ to the client.
   (c) $\mathcal{P}$ receives the joint state $(x_i, \tilde{\rho}_1)$ and queries Hybrid Locked PUF $\mathcal{E}_f^L$. If the verification algorithm $\texttt{Ver}(|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}|, \tilde{\rho}_1) \geqslant 1-\varepsilon(\lambda)$ with negligible $\varepsilon(\lambda)$, $\mathcal{P}$ obtains $|\psi_{f_2(x_i)}\rangle\langle\psi_{f_2(x_i)}| := \bigotimes_{j=1}^m |\psi^{i,j}_{f_2(x_i)}\rangle\langle\psi^{i,j}_{f_2(x_i)}|$ from $\mathcal{E}_f^L$ and sends back to $\mathcal{V}$. Otherwise, the authentication aborts.
   (d) $\mathcal{V}$ receives the quantum state $\tilde{\rho}_2$ and performs the the verification algorithm $\texttt{Ver}(.,.)$. If the verification $\texttt{Ver}(|\psi_{f_2(x_i)}\rangle\langle\psi_{f_2(x_i)}|, \tilde{\rho}_2) \geqslant 1-\varepsilon(\lambda)$ with negligible $\varepsilon(\lambda)$, the authentication passes. Otherwise it aborts.

We note that this protocol requires only a classical database, but two-way quantum communication. Nevertheless, the quantum states used in this protocol are easy to prepare and measure given the infrastructures that already exist for QKD and the current stage of a quantum internet [WEH18].

6.4.5 Security analysis

Now, we give a comprehensive security analysis of the proposed protocol. We will prove the security step-by-step. First, we show that using hybrid construction will exponentially improve security compared to CPUF. More precisely, it will exponentially decrease the success probability of a *weak* quantum adversary in the universal unforgeability game, compared to a classical PUF with the same number queries in the learning phase. For this part, we will use the weak quantum adversarial model and the respective variant of the universal unforgeability game as defined in Section 3.5.3. This result shows how much quantum communication can improve the security of a weaker classical PUF against quantum adversaries. Using this improvement, we propose an efficient and secure construction using existing classical PUFs. Then, we analyse the completeness and security of the HLPUF-based device authentication protocol and show that given that the intrinsic classical PUF is not fully broken against a weak quantum adversary, the HLPUF-based protocol will be secure against a QPT adaptive adversary. The HLPUF-based protocol, in fact, provides mutual authentication of both prover/client and verifier/server due to the specific construction of the quantum lock. However, in our security analysis, we only formally prove the authentication of the prover to the verifier, and the lock has only been shown to reduce the adversary's capability. Nevertheless, since the verification mechanism is similar on both sides, the alternative side of authentication can be proven similarly. Let us first formalise our assumption on the underlying CPUF:

Assumptions on the CPUFs
For the security analysis of our constructions we consider the following assumptions of the CPUFs $f : \{0,1\}^n \to \{0,1\}^{4m}$.

1. For any input $x \in \{0,1\}^n$ the probability distributions of the $4m$ output bits $f(x)_1, \ldots, f(x)_{4m}$ are independent and identically distributed (i.i.d)[7].

2. The output distributions $\{p_x^f(y)\}_{y\in\{0,1\}^{4m}}$ for all the inputs x are independent and identically distributed (i.i.d).

6.4.5.1 Universal Unforgeability of HPUF

Intuitively the security of our HPUF comes from the indistinguishability property of the non-orthogonal quantum states. First we show that the HPUFs are at least as secure as the underlying CPUFs.

[7] This assumption is not strictly required in practice, for the HLPUF construction to be secure, as our simulation results show in [CDM+21]. It is mostly required for our theoretical bounds and even so, we parameterised the deviation from perfect randomness or identical distribution with the randomness parameter of CPUF, although we require that the encoded qubits are independent.

Theorem 44. *Let $f : \{0,1\}^n \to \{0,1\}^{2m}$ be a classical PUF. If there is no QPT weak adversary who can win the universal unforgeability game for CPUF with a non-negligible probability in the security parameter, then the HPUF constructed from f according to Construction 4, is also universally unforgeable.*

Proof. We show the contrapositive statement that if you can break HPUF you can also break underlying CPUF. Here we give the proof for $m = 1$, and it can easily be generalised for any arbitrary integer $m > 0$. Suppose for the HPUF, a q-query weak-adversary win the unforgeability game with a non-negligible probability $P(m = 1, p, q)$. This implies, given a database of q random challenge response from the HPUF, the adversary can produce $|\psi_{f(x^*)}\rangle$ corresponding to a random challenge $x^* \in \{0,1\}^n$ with a non-negligible probability $P(m = 1, p, q)$. Note that, for the deterministic adversarial strategy, the adversary can produce multiple copies of the forged state $|\psi_{\tilde{f}(x^*)}\rangle$ for a random challenge x^*. For the random adversaries we can produce the multiple copies of the same forged state $|\psi_{\tilde{f}(x^*)}\rangle$ just by fixing the internal randomness parameter of the adversarial strategy. Hence, both the random and deterministic adversary can produce multiple copies of the forged state $|\psi_{\tilde{f}(x^*)}\rangle$ for a random challenge x^*. From the multiple (say K) such copies of $|\psi_{\tilde{f}(x^*)}\rangle$, the adversary will extract $\tilde{f}(x^*)$ using the following strategy.

Algorithm 3 Algorithm to Forge CPUF from HPUF

Require: $K \geqslant 2$-copies of the forged state $|\psi_{\tilde{f}(x^*)}\rangle$

- Measure the 1-st copy of the state $|\psi_{\tilde{f}(x^*)}\rangle$ in $\{|0\rangle, |1\rangle\}$-basis.
- Let $z_1 \in \{0,1\}$ be the measurement outcome.

for $i = 2$; $i \leqslant (K-1)$; $i{+}{+}$ **do**
- Measure the i-th copy of the state $|\psi_{\tilde{f}(x^*)}\rangle$ in $\{|0\rangle, |1\rangle\}$-basis.
- Let $z_i \in \{0,1\}$ be the measurement outcome.
- **if** $z_i \neq z_{i-1}$ **then**
  - **break** ▷ Implies $|\psi_{\tilde{f}(x^*)}\rangle \in \{|+\rangle, |-\rangle\}$.

if $i = K$ **then**
- **Return** $\tilde{f}(x^*) = (0, z_i)$

else
- Measure the $i+1$-th copy in $\{|+\rangle, |-\rangle\}$-basis.
- Let z_{i+1} be the measurement outcome.
- **Return** $\tilde{f}(x^*) = (1, z_{i+1})$.

If $|\psi_{f(x^*)}\rangle = |\psi_{\tilde{f}(x*)}\rangle \in \{|0\rangle, |1\rangle\}$ then in Algorithm 3 all the measurement outcomes z_i (for $1 \leqslant i \leqslant K$) would be the same, and $\tilde{f}(x^*) = f(x^*)$. However, if $|\psi_{f(x^*)}\rangle = |\psi_{\tilde{f}(x*)}\rangle \in \{|+\rangle, |-\rangle\}$ then we $\tilde{f}(x^*) \neq f(x^*)$ if and only if all the measurement outcomes z_i are equal ($1 \leqslant i \leqslant K$). This happens with probability $\frac{1}{2^K}$. Therefore, we get

$$\Pr_{x^*}[\tilde{f}(x^*) = f(x^*) || \psi_{f(x^*)}\rangle = |\psi_{\tilde{f}(x*)}\rangle] \geqslant 1 - \frac{1}{2^K}. \tag{6.72}$$

If the adversary successfully forges the HPUF with a non-negligible probability $P(m = 1, p, q)$ then from Eq. (6.72) we get that the adversary manages the CPUF with probability at least $P(m = 1, p, q) = 1 - \frac{1}{2^K}$, that is an overwhelming probability. Therefore, if an adversary successfully wins the unforgeability game for the HPUF with a non-negligible probability, then using the same forging strategy it can also win the unforgeability game for the corresponding CPUF with a non-negligible probability. This implies, that if no QPT weak adversary can win the universal unforgeability game with a non-negligible probability for the CPUF then no QPT adversary can win the universal unforgeability game with a non-negligible probability for the corresponding HPUF. This concludes the proof. □

The above theorem is an intuitive result that shows HPUF is stronger or at least as strong as the underlying CPUF. Although we want to prove a more powerful and explicit statement regarding HPUFs by quantifying how much the hybrid construction will boost the security. In fact, we want to show that one can construct a secure unforgeable HPUF against a quantum adversary even if the underlying CPUF is breakable (with a certain probability) against the classical forger. To this end, we compare the success probability of a QPT adversary in breaking the HPUF in the universal unforgeability game, with the success probability of the adversary who breaks the CPUF with a certain non-negligible probability, in a fixed query setting. This, allows us to show that some of the weak and considerably broken CPUFs can still be used to construct an asymptotically secure HPUF against stronger quantum adversaries since the quantum encoding drastically decreases the success probability. Before giving our main theorem, we need to prove two lemmas. In the first one, we give an upper bound on the adversary's guessing probability of the response $f(x_i)$ corresponding to a challenge x_i and a single copy of the quantum response state $|\psi_{f(x_i)}\rangle$.

Lemma 4. *Let $f : \{0,1\}^n \to \{0,1\}^{4m}$ be a CPUF with the following property,*

$$\forall \;\; x_i \in \{0,1\}^n, \forall \;\; 1 \leqslant j \leqslant 4m, \;\; p^f_{x_i}(y_{i,j} = 0) = \frac{1}{2} + \delta_r, \tag{6.73}$$

with a biased distribution $p = \frac{1}{2} + \delta_r$ where $0 \leqslant \delta_r \leqslant \frac{1}{2}$, and $\mathcal{E}_f$ be a HPUF corresponding to f that we construct using Construction 3. Let a quantum adversary $\mathcal{A}$ extract the value $y_{i,(2j-1)}$ out of $(y_{i,(2j-1)}, y_{i,2j})$ from quantum state $|\psi^{i,j}_{out}\rangle\langle\psi^{i,j}_{out}|$ corresponding to a random challenge x_i. If all the output bits of the CPUF are independent and identically distributed, then for any quantum adversary $\mathcal{A}$, and $\forall\; x_i \in \{0,1\}^n$ then,

$$\begin{aligned} p_{guess} &:= \Pr[\mathcal{A}(x_i, |\psi^{i,j}_{out}\rangle\langle\psi^{i,j}_{out}|) = y_{i,(2j-1)}] \\ &\leqslant p(1 + \sqrt{p^2 + (1-p)^2}) \\ &\leqslant p(1 + \sqrt{2}p) \end{aligned} \tag{6.74}$$

Proof. According to Construction 4, for a given x_i, we use the $2j$-th bit $y_{i,2j} \in \{0,1\}$ of the outcome of the CPUF to choose the basis (either $\{|0\rangle, |1\rangle\}$-basis or $\{|+\rangle, |-\rangle\}$-basis) of the j-th qubit output of the HPUF. Further we use the $y_{i,(2j-1)} \in \{0,1\}$ to choose a state from the chosen basis. Here, if $y_{i,(2j-1)} = 0$ then from an adversarial point of view, the output state is $\rho_0 = (\frac{1}{2} + \delta_r)|0\rangle\langle 0| + (\frac{1}{2} - \delta_r)|+\rangle\langle +|$. Similarly, if $y_{i,(2j-1)} = 1$ then from an adversarial point of view, the output state is $\rho_1 = (\frac{1}{2} + \delta_r)|1\rangle\langle 1| + (\frac{1}{2} - \delta_r)|-\rangle\langle -|$. For the adversary, the probability of correctly guessing $y_{i,(2j-1)}$ is the same as distinguishing the two states ρ_0, ρ_1. Here $Pr[\mathcal{A}(x_i, |\psi^{i,j}_{out}\rangle\langle\psi^{i,j}_{out}|) = y_{i,(2j-1)}]$ denotes the optimal probability of guessing the basis correctly. From the Holevo-Helstorm bound [Hol73] (see Section 2.2) we get,

$$\begin{aligned} \Pr[\mathcal{A}(x_i, |\psi^{i,j}_{out}\rangle\langle\psi^{i,j}_{out}|) = y_{i,(2j-1)}] &\leqslant p[1 + \max_E \mathsf{Tr}[E(\rho_0 - \rho_1)]] \\ &= p[1 + \frac{1}{2}\|\rho_0 - \rho_1\|_1]. \\ &= p(1 + \sqrt{p^2 + (1-p)^2}) \\ &\leqslant p(1 + \sqrt{2}p) \end{aligned} \tag{6.75}$$

This concludes the proof. □

The next lemma shows that the adversary needs to extract the classical information $f(x)$ that is encoded in the quantum state $|\psi_{f(x)}\rangle$ for the forgery of the HPUFs. This will be a key step in our proof, since using this lemma we can put bounds on the maximum amount of information the adversary can extract from the overall response state using quantum information tools we have presented in the preliminaries, in Section 2.1.5.

Lemma 5. *Let* $|D_q\rangle = \bigotimes_{i=1}^{q} \left(|x_i\rangle_C \otimes |\psi_{f(x_i)}\rangle_R \right)$ *denotes the adversary's database of q random CRPs that are generated from a HPUF* $\mathcal{E}_f : \{0,1\}^n \to (\mathcal{H}^2)^{\otimes m}$. *Let* $E(D_q)$ *denote the optimal measurement strategy for forging the HPUF with probability* p_{forge} *using the database* D_q, *then the following measure-then-forge strategy can optimally forge the HPUF with the same probability* p_{forge}.

- *Adversary extracts the classical encoding* $\{f(x_i)\}_{1\leqslant i\leqslant q}$ *from* $|D_q\rangle$. *Let* $\{\tilde{f}(x_i)\}_{1\leqslant i\leqslant q}$ *denotes the extracted classical string.*
- *The QPT adversary applies a forging strategy using the extracted data set* $\{\tilde{f}(x_i)\}_{1\leqslant i\leqslant q}$.

Proof. For a successful forgery, the adversary needs to win the universal unforgeability game that we define in Game 2 in Chapter 3. This implies, using the measurement strategy $E(D_q)$ the adversary needs to produce a quantum state $|\psi_{f(x^*)}\rangle$ corresponding to a challenge $x^* \in_R \{0,1\}^n$ that is chosen uniformly at random. Without loss of generality we can write the measurement strategy as a POVM with two outcomes $E(D_q) = \{E_{\text{forge}}(D_q, x^*), E_{\text{fail}}(D_q, x^*)\}$, where $E_{\text{forge}}(D_q, x^*), E_{\text{fail}}(D_q, x^*)$ denote the measurement operators corresponding to the successful forgery and the failure forgery respectively. Therefore, we can write the successful forging probability p_{forge} as follows.

$$p_{\text{forge}} = \text{Tr}[E_{\text{forge}}(D_q, x^*)\rho_{D_q}^{x^*}], \tag{6.76}$$

where $\rho_{D_q}^{x^*} := |D_q\rangle\langle D_q| \otimes |x^*\rangle\langle x^*| \otimes |0^m\rangle_{out}\langle 0^m|$. Here the out register would contain the forged state. If we write $E_{\text{forge}}(D_q, x^*) = M_{\text{forge}}^{\dagger}(D_q, x^*)M_{\text{forge}}(D_q, x^*)$, then we can rewrite the post-measurement state corresponding to the successful forgery as follows:

$$\frac{M_{\text{forge}}(D_q, x^*)|D_q\rangle \otimes |x^*\rangle \otimes |0^{m'}\rangle_{out}}{\sqrt{p_{\text{forge}}}} = \frac{|\tilde{D}_q\rangle_R \otimes |x^*\rangle \otimes |\psi_{f(x^*)}\rangle_{out} \otimes |\tilde{a}\rangle_{out}}{\sqrt{p_{\text{forge}}}}, \tag{6.77}$$

where $|\tilde{D}_q\rangle_R$ denotes the post-measurement database state, and $|\tilde{a}\rangle_{out}$ is the post-measurement state of the ancillary system which is a $(m' - m)$ dimensional state while as $|\psi_{f(x^*)}\rangle_{out}$ is m dimensional. As $\bigotimes_{i=1}^{q} |x_i\rangle_C$ is a classical state, in the rest of the proof we don't write them in the expressions.

Using the *Neimark's theorem*[8] we can replace the POVM measurement strategy $E(D_q)$ with the combination of a unitary acting on an extended system including an ancilla $|anc\rangle_A$, followed by a projective measurement. Let us denote the unitary as $U_{D_q}^{x^*}$ which couples the input state $|D_q\rangle \otimes |0^{m'}\rangle_{out}$ with the ancillary system $|anc\rangle_A$, and let $\{|v\rangle\}$ be the basis on which the projective measurement is

[8] The version of Neimark's theorem, is similar but more genral than the one we introduced in Section 2.2.

applied to the ancilla. We first rewrite the impact of the unitary $U_{D_q}^{x^*}$ on the input state:

$$U_{D_q}^{x^*}\left(\bigotimes_{i=1}^{q}|\psi_{f(x_i)}\rangle_R\otimes|0\rangle_{out}\otimes|anc\rangle_A\right)=U_{D_q}^{x^*}\left(|\Psi_f^q\rangle_R\otimes|0\rangle_{out}\otimes|anc\rangle_A\right)$$
$$=\sum_v\sqrt{p_v}\,|\Psi_v^q\rangle_R\otimes|\tilde{\psi}_v\rangle_{out}\otimes|v\rangle_A. \quad (6.78)$$

where in the second line we have rewritten everything after applying the unitary in the $\{|v\rangle\}$-basis. Now, the adversary performs a projective measurement on the state Eq. (6.78) in this basis. Suppose for the correct forgery, the ancilla is projected into the $|v_{\text{forge}}\rangle_A$ state. Therefore we can rewrite the expression of p_{forge} as follows:

$$p_{\text{forge}}=\sum_{v:v=v_{\text{forge}}}p_v|\langle v_{\text{forge}}|v\rangle|^2. \quad (6.79)$$

Overall, following this strategy, the purification of the adversary's post-measurement state with an optimal POVM measurement, can be written as the following:

$$\frac{|\tilde{D}_q\rangle_R\otimes|x^*\rangle\otimes|\psi_{f(x^*)}\rangle_{out}\otimes|v_{\text{forge}}\rangle_A}{\sqrt{p_{\text{forge}}}}, \quad (6.80)$$

where $|\tilde{D}_q\rangle$ denotes the post-measurement database state. Note that, due to Neimark's theorem the post-measurement database states in Equation Eq. (6.77), and Eq. (6.80) are the same, if the same ancillary systems has been assumed after the purification and POVM, *i.e.* if $|v_{\text{forge}}\rangle_A=|\tilde{a}\rangle_{out}$.

Now, let us use the unitary $U_{D_q}^{x^*}$ and the measurement basis $\{|v\rangle\}$ to construct a *measure-then-forge* strategy. As the unitary $U_{D_q}^{x^*}$ only depends on the input x^* and D_q, we can rewrite it in the basis that is diagonalised with respect to the states $\{|\Psi_v^q,v\rangle\}_v$. For the post-measurement state $|v_{\text{forge}}\rangle$, of the ancilla, the adversary applies $U_{D_q,\Psi_{\text{forge}}^q,v_{\text{forge}}}^{x,x^*}$ on the $|0\rangle_{out}$ register. Note that, the adversary doesn't have any information about the $\{f(x_i)\}_{1\leqslant i\leqslant q}$ before measuring the ancillary sub-system in the $\{|v\rangle\}$-basis. Hence, the measurement basis $\{|v\rangle\}$ choice only depends on the classical challenges x_i's and x^*. Therefore, the adversary can use the same information to find the $\{|v\rangle\}$-basis, and first performs the measurement on the RA register in $\{|\Psi_v^q,v\rangle\}$-basis, and obtains the state $|\Psi_{\text{forge}}^q,v_{\text{forge}}\rangle$ with the same probability p_{forge}. After the measurement, the adversary applies the unitary $U_{D_q,\Psi_{\text{forge}}^q,v_{\text{forge}}}^{x^*}$ on $|0\rangle_{out}$, and get the forged state $|\psi_{f(x^*)}\rangle$. Therefore, with this strategy the adversary also win the unforgeability game with the probability p_{forge}.

Note that, there always exists a unitary U such that $U(\bigotimes_{i=1}^q|\tilde{f}(x_i)\rangle)\otimes|anc\rangle=|\Psi_{\text{forge}}^q,v_{\text{forge}}\rangle$, where $\tilde{f}(x_i)$ denotes the extracted information about $f(x_i)$'s from the encoded database $|D_q\rangle$. Therefore, from any generalised measurement strategy $E(D_q)$ we can construct a strategy for the measure-then-forge protocol that can win the universal unforgeability game with the same probability p_{forge}. This concludes the proof. □

Lemma 5 suggests that an optimal strategy of the adversary including general POVM strategies, is equivalent to optimally extracting the classical information from the database (state $|D_q\rangle$) and then performing the most optimal forgery strategy on the measurement results. In general, if the extracted classical information $\{\tilde{f}(x_i)\}_{1\leqslant i\leqslant q}$ from the database state $|D_q\rangle$ is very far from the original encoded string $\{f(x_i)\}_{1\leqslant i\leqslant q}$ then the forgery will not perform well and the overall forgery attack will have a low probability. Based on this, we now bound the overall amount of information that can be extracted from the outputs of HPUF, using the information quantities such as min-entropy and we will show how much the quantum encoding will contribute to reducing the success probability.

Theorem 45. *Let $f : \{0,1\}^n \to \{0,1\}^{4m}$ be a classical PUF with p-randomness, where $p = \frac{1}{2} + \delta_r$. Let $p_{forge}^{classic}(m,p,q)$ denote the optimal success probability of any q-query weak quantum adversary to wins the universal unforgeability game for the CPUF f. Then a q-query weak quantum adversary can win the universal unforgeability game for the HPUF $\mathcal{E}_f$ at most the following probability*

$$p_{forge}^{quant}(m,p,q) = p_{forge}^{classic}(m,p,q) \times (p(1+\sqrt{2}p))^{2mq} \tag{6.81}$$

Proof. We want to quantify the success probability of the QPT adversary in attacking HPUF, in comparison with the QPT adversary who attacks the classical PUF with a fixed number of queries. Let $\mathcal{A}_c$ be the QPT adversary attempting to forge CPUF, where they produce a classical forgery $f(x^*)$ for a randomly selected challenge x^*, from a classical database consisting of q pairs of $\{(x_i, f(x_i))\}_{i=1}^q$ input-outputs of CPUF with probability $p_{\text{forge}}^{\text{classic}}(m,p,q)$. Note that in general the success probability, is a function of the CPUF's randomness parameter p, the output size m and the number of queries q. Let $\mathcal{A}_h$ be a quantum adversary who plays the unforgeability game against the HPUF. $\mathcal{A}_h$ has access to q queries of HPUF included in the database state $|D_q\rangle$. The goal of $\mathcal{A}_h$ is to produce a valid forgery $|\psi_{f(x^*)}\rangle$ for a random challenge x^*. Let $p_{\text{forge}}^{\text{quant}}$ be the optimal success probability of any adversary in successfully doing so.

Now, for the purpose of the proof, we introduce another intermediate quantum adversary $\mathcal{B}$ who plays the same version of the unforgeability game as $\mathcal{A}_h$, although has access to a combined database of $\mathcal{A}_h$ and $\mathcal{A}_c$ *i.e.* the triplet $\{(x_i, f(x_i), |\psi_{f(x_i)}\rangle)\}_{i=1}^q$, and will output a valid forgery of the form $|\psi_{f(x^*)}\rangle$ for a random challenge x^*. We show that the adversaries $\mathcal{A}_c$ and $\mathcal{B}$ are equivalent in the success probability up to a negligible factor. First, note that $\mathcal{B}$ is obviously at least as strong as $\mathcal{A}_c$ since has an extended database, and can simply ignore the quantum encoded detest and run $\mathcal{A}_c$ as a subroutine so we have $p_{\mathcal{B}}(m,p,q) \geqslant p_{\text{forge}}^{\text{classic}}(m,p,q)$. But on the other hand, $\mathcal{A}_c$ can also locally construct the third column of the database $|\psi_{f(x_i)}\rangle$ easily and run $\mathcal{B}$ as a subroutine. Then $\mathcal{A}_c$ can use this to produce polynomial many copies of $|\psi_{f(x^*)}\rangle$ and measure them with the optimal POVM measurement and get $f(x^*)$ with very high probability (See the proof of Theorem 44). Also note that by definition of the

unforgeability game, $\mathcal{B}$ produces the forgery with non-negligible success probability, meaning that the $|\psi_{f(x^*)}\rangle$ is close to the actual BB84 encoding of $f(x^*)$ (in terms of fidelity). Thus $\mathcal{B}$ has at most a negligible advantage over $\mathcal{A}_c$ and we can conclude: $p_{\mathcal{B}}(m,p,q) \approx p^{\text{classic}}_{\text{forge}}(m,p,q)$.

Now we have two adversaries who produce quantum state as a forgery and we can compare the success probability of $\mathcal{A}_h$ with $\mathcal{B}$, which are both QPT adversaries producing the same quantum forgery while having access to different input databases since $\mathcal{B}$ has the underlying classical information $f(x_i)$ for each encoded quantum query, while $\mathcal{A}_h$ has only access to the encoded states in the form of $|\psi_{f(x_i)}\rangle$. We compare the success probability of these two adversaries by comparing the accessible amount of information via entropy inequalities. First we note that according to Lemma 5, the optimal forgery for a quantum adversary includes optimally extracting the classical information then applying the forgery (which is equivalent to the classical forgery on the classical database). This exactly quantifies the relation between the success probabilities of $\mathcal{A}_h$ with $\mathcal{B}$ which will give us the reduction to the problem of extracting information from the quantum-classical database of $\mathcal{A}_h$.

To do so, we reformulate the forging probability (the success probability of the adversary in the unforgeability game) in terms of quantum processing. We consider $\mathcal{B}$ as a CPTP map over a database of size N, denoted as $\rho^{\mathcal{B}^N}$, where each input $\rho_i^{\mathcal{B}} = |x_i\rangle\langle x_i| \otimes |f(x_i)\rangle\langle f(x_i)| \otimes |\psi_{f(x_i)}\rangle\langle\psi_{f(x_i)}|$ is a classical-quantum state. Adversary $\mathcal{A}_h$ can be defined directly from $\mathcal{B}$ through another CPTP map which we denote by Λ_h. For each record we have $\rho_i^{\mathcal{A}_h} = \Lambda_h(\rho_i^{\mathcal{B}})$. Assuming the queries to be i.i.d and the way we have defined these two adversaries, The overall action of the CPTP map $\mathcal{A}_h$ is captured by the density matrix $\rho^{\mathcal{A}_h^N} = \Lambda_h^{\otimes N}(\rho^{\mathcal{B}^N})$. Also, let F represent the random variable of getting the correct output of the PUF, over the uniform choice of the input x^*, by processing the given input database.

We now use the inequality for conditional min-entropy to relate the above success probability. Using Lemma 5 stating that the optimal strategy is equivalent to extracting the underlying classical information (optimal state discrimination) and then run an optimal forgery algorithm that depends on the extracted information. Thus we can write the success probability of $\mathcal{A}_h$ in terms of the extraction probability as follows:

$$p^{\text{quant}}_{\text{forge}}(m,p,q) = p_{\text{extract}} \times p_{\mathcal{B}}(m,p,q) = 2^{-H_{min}(F|C^N)} \tag{6.82}$$

where C^N denotes $\rho^{\mathcal{A}_h^N}$ as the full quantum system of $\mathcal{A}_h$ and also let C denotes $\rho^{\mathcal{A}_h}$ which is the single-qubit database of $\mathcal{A}_h$. We can rewrite the success probability of forgery for adversary $\mathcal{B}$, in terms of the min-entropy as follows:

$$p_{\mathcal{B}}(m,p,q) = p^{\mathcal{B}}_{\text{forge}} = 2^{-H_{min}(F|B^N)} \tag{6.83}$$

Where B^N represent the full system of $\mathcal{B}^N$ *i.e.* $\rho^{\mathcal{B}^N}$ for short. Now for the single-qubit database, the following relation holds:

$$H_{min}(F|C) \geqslant H_{min}(\tilde{Y}|C) + H_{min}(F|\tilde{Y}) \tag{6.84}$$

Where $\tilde{Y}$ denotes the random variable of the estimated output $Y = f(X)$ which is the 2-bit fraction of the output of CPUF. We also note that $H_{min}(F|\tilde{Y}) \geqslant H_{min}(F|B)$, and the equality is when the $\tilde{Y}$ is arbitrarily close to Y. As a result we have[9]:

$$H_{min}(F|C) \geqslant H_{min}(\tilde{Y}|C) + H_{min}(F|B) \tag{6.85}$$

Extending to N-fold database we have:

$$H_{min}(F|C^N) \geqslant H_{min}(\tilde{Y}^N|C^N) + H_{min}(F|B^N) \tag{6.86}$$

Next, we need to relate the first term of the right-hand side, which denoted the min-entropy of extracting information from all N given qubits to the min-entropy quantity $H_{min}(\tilde{Y}|C)$. For that, we use the quantum-classical AEP that we have introduced in Theorem 4 as follows:

$$H^{\varepsilon}_{min}(\tilde{Y}^N|C^N) \geqslant N(H(\rho_{\tilde{Y}C}) - H(\rho_C)) - N\eta \tag{6.87}$$

where $\eta := (2H_{max}(\rho_F)+3)\sqrt{\frac{\log(\frac{1}{\varepsilon})}{N}+1}$, is a function of the smoothing parameter ε and N. Here we select the smoothing parameter ε such that $N\eta$ becomes a negligible function in the security parameter. Given that $H(\tilde{Y}|C) = H(\rho_{\tilde{Y}C}) - H(\rho_C)$ and the fact that $H_{min}(\tilde{Y}|C) \leqslant H(\tilde{Y}|C)$, we have:

$$H^{\varepsilon}_{min}(\tilde{Y}|C^N) \geqslant NH_{min}(\tilde{Y}|C) \tag{6.88}$$

By substituting the above inequality back into 6.86, we can conclude the following:

$$H_{min}(F|C^N) \geqslant NH_{min}(\tilde{Y}|C) + H_{min}(F|B^N) \tag{6.89}$$

The final step is to determine N. We note that there exist q number of i.i.d queries in each database, but each query itself consists of a $2m$ number of qubits in tensor product form. Although the effective size or the amount of information included in these $2m$ qubits depends on the bias of the PUF. In general, using quantum data compression inequality, the effective size of such $2m$ tensor product states is given by $2mS(\rho_f)$, where $S(\rho_f)$ is the von-Neumann entropy of each encoded state ρ_f given as follows concerning the HPUF construction and the PUF bias:

$$\rho_f = (\frac{1}{2}+\delta_r)^2 |0\rangle\langle 0| + (\frac{1}{4}-\delta_r^2)(|1\rangle\langle 1| + |+\rangle\langle +|) + (\frac{1}{2}-\delta_r)^2 |-\rangle\langle -| \tag{6.90}$$

We can then calculate $S(\rho_f)$ which gives the following result while we are discarding $\mathcal{O}(\delta^3)$ and higher:

$$S(\rho_f) = 1 - (\frac{1}{2}-\delta_r)\log(1-2\delta_r) - (\frac{1}{2}+\delta_r)\log(1+2\delta_r) \tag{6.91}$$

[9]This inequality in fact gives and improvement to the data processing inequality (introduced in Section 2.1.5) for our specific case. Since according to the data-processing inequality we have that the entropy min-entropy increases via any CPTP channel acting on the joint state of the system, which results in $H_{min}(F|B) \leqslant H_{min}(F|C)$ which denotes that the success probability of adversary $\mathcal{B}$ who has access to additional classical output is higher than $\mathcal{A}_h$. The new inequality quantifies the bound on the difference.

Let us call $g(\delta_r) = (\frac{1}{2} - \delta_r)\log(1-2\delta_r) + (\frac{1}{2} + \delta_r)\log(1+2\delta_r)$ Thus we have $N \approx 2mq - 2mqg(\delta_r)$. Also, for small enough values of δ_r, we have $g(\delta_r) \approx 0$. Let us also denote $H_{min}(\tilde{Y}|C) = -\log(p^1_{\text{extract}})$. Where p^1_{extract} is the probability of extracting the classical information from a single qubit. We can then conclude the following relations between the success probabilities:

$$p^{\text{quant}}_{\text{forge}}(m,p,q) \leqslant p_{\mathcal{B}}(m,p,q) \times (p^1_{\text{extract}})^{2mq-2mqg(\delta_r)} \tag{6.92}$$

For our given construction $p^1_{\text{extract}} = p_{\text{guess}} = p(1+\sqrt{2}p)$ according to the optimal discrimination probability given in Lemma 4. Thus we can rewrite the above equation as:

$$\begin{aligned} p^{\text{quant}}_{\text{forge}}(m,p,q) &\leqslant p^{\text{classic}}_{\text{forge}}(m,p,q) \times (p(1+\sqrt{2}p))^{2mq-2mqg(\delta_r)} \\ &\approx p^{\text{classic}}_{\text{forge}}(m,p,q) \times (p(1+\sqrt{2}p))^{2mq} \end{aligned} \tag{6.93}$$

which is the bound we wanted to prove, and we have also used that $p^{\text{classic}}_{\text{forge}}(m,p,q) \approx p_{\mathcal{B}}(m,p,q)$. As a final remark, we note that the optimal probability is a function of the number of queries, thus we can show that the optimal overall probability is achieved given the adversary optimises on the number of queries used for extracting the information for the forgery. Analysing the upper bound of the probability as a function of q, one can see that the first term is a non-decreasing function of q while the second term is always strictly decreasing with q. As a result, the combined function has necessary an extremum over q, which we denote by q_{opt}. Assuming the two cases where the given number of queries is smaller or larger than q_{opt}, we have the following bounds:

$$\begin{aligned} p^{\text{quant}}_{\text{forge}} &\leqslant p^{\text{classic}}_{\text{forge}}(m,p,q) \times (p_{\text{guess}})^{2mq(1-g(\delta_r))} && q < q_{opt} \\ p^{\text{quant}}_{\text{forge}} &\leqslant p^{\text{classic}}_{\text{forge}}(m,p,q_{opt}) \times (p_{\text{guess}})^{2mq_{opt}(1-g(\delta_r))} && q \geqslant q_{opt} \end{aligned} \tag{6.94}$$

Summarizing the above cased and given that $g(\delta_r)$ is small we have:

$$p^{\text{quant}}_{\text{forge}} \leqslant \sup_q [p^{\text{classic}}_{\text{forge}}(m,p,q) \times (p_{\text{guess}})^{2mq}] \tag{6.95}$$

which concludes the proof.[10] □

Finally, let us present the following corollary that ensures the universal unforgeability of an HPUF constructed from a CPUF that does not provide suitable security, yet is not totally broken with overwhelming probability.

[10] In the latest version of the paper [CDM+21], we have given an alternative version of this proof which does not use the AEP, and instead relies on giving the bound on the forgery probability over a noisy database where the bound is slightly different although very similar. The exponential decay in the probability happens in both cases. In that proof, one no longer needs to give an optimality argument over the number of queries. However, since proving the result, in this way seemed more intuitive and used a rather nice quantum information tools, we have decided to present this proof in the thesis. The reader can refer to the paper for the other result.

Corollary 8. *Let the success probability of any QPT weak-adversary in the universal unforgeability game with a CPUF $f : \{0,1\}^n \to \{0,1\}^{4m}$ with p-randomness, be at most $p_{forge}^{classic}$, where $0 \leqslant p_{forge}^{classic} \leqslant 1 - non\text{-}negl(2m)$. Then, the success probability of any QPT adversary in the universal unforgeability game for the HPUF $\mathcal{E}_f$, is at most $\varepsilon(2m)$, which is a negligible function in the security parameter. Hence such HPUFs are universally unforgeable.*

Proof. This directly follows from Theorem 45 where $p_{\text{forge}}^{\text{classic}} = p_{\text{forge}}^{\text{classic}}(m, p, q)$ for any $q = poly(m)$ is a value between 0 and 1, and not negligibly close to 1. As shown in the proof of Theorem 45 the second term of the probability, becomes negligibly small (in $2m$) and hence the overall probability becomes a negligible function $\varepsilon(2m)$. □

6.4.5.2 Universal Unforgeability of HLPUF

So far, we have analysed the security of the HPUFs only against weak adversaries. In the next theorem, we show that if the HPUF is secure against weak adversaries, then using the locking mechanism, we can make the HLPUF secure against adaptive adversaries.

Theorem 46. *Let $\mathcal{E}_f : \{0,1\}^n \to (\mathcal{H}^2)^{\otimes m} \otimes (\mathcal{H}^2)^{\otimes m}$ be a hybrid PUF that we construct from a classical PUF $f : \{0,1\}^n \to \{0,1\}^{2m} \times \{0,1\}^{2m}$ and let $\mathcal{E}_f^L : \{0,1\}^n \times (\mathcal{H}^2)^{\otimes m} \to (\mathcal{H}^2)^{\otimes m}$ denotes the HLPUF that we construct from $\mathcal{E}_f$ using the Construction 4. If $\mathcal{E}_f = \mathcal{E}_{f_1} \otimes \mathcal{E}_{f_2}$ and if each of the mappings $\mathcal{E}_{f_1}, \mathcal{E}_{f_2}$ has (ε, m)-universal unforgeability against the q-query weak adversaries, then the corresponding HLPUF $\mathcal{E}_f^L$ is (ε, m)-secure against the q-query adaptive adversaries.*

Proof. At the i-th round, the HLPUF $\mathcal{E}_f^L$ receives the queries of the form $(x_i, \tilde{\rho}_1)$, where the classical string $x_i \in \{0,1\}^n$, and $\tilde{\rho}_1 \in (\mathcal{H}^2)^{\otimes m}$. The HLPUF returns $\mathcal{E}_{f_2}(x_i)$ if $\texttt{Ver}(\tilde{\rho}_i', \mathcal{E}_{f_1}(x_i)) = 1$, otherwise it returns an abort state $|\perp\rangle\langle\perp|$ corresponding to $\perp$. Hence, to avoid getting state $|\perp\rangle$ from the HLPUF, the adaptive adversaries $\mathcal{A}_{ad}$ need to produce a query of the form $(x_i, \mathcal{E}_{f_1}(x_i))$. As the adversary doesn't have any direct access to the mapping $\mathcal{E}_{f_1}$, the only way it can get any information about $\mathcal{E}_{f_1}(x_i)$ by intercepting the challenges that are sent by the server to the client. Suppose that the adaptive adversary has access to a set of q queries $X_{[q]} := \{X_i\}_{1 \leqslant i \leqslant q}$ and the corresponding responses $\Psi_{[q]} := \{\mathcal{E}_{f_1}(x_i)\}_{1 \leqslant i \leqslant q}$. Here each X_i follows a uniform distribution over the challenge set $\{0,1\}^n$. Hence, for the mapping $\mathcal{E}_{f_1}$ the power of the adaptive adversary reduces to the power of a weak adversary. As $\mathcal{E}_{f_1}$ has the universal unforgeability property against any q-query weak adversary, hence we get, for any random challenge $X \notin X_{[q]}$,

$$\Pr_{X,X_{[q]}}[1 \leftarrow \mathcal{G}^{\mathcal{E}_{f_1}}(\mathcal{A}_{ad}, m, X, X_{[q]})] = \Pr_{X,X_{[q]}}[1 \leftarrow \mathcal{G}^{\mathcal{E}_{f_1}}(\mathcal{A}_{weak}, m, X, X_{[q]})] \leqslant \varepsilon(m). \tag{6.96}$$

This implies, that using the set of challenges $X_{[q]}$ and responses $\Psi_{[q]}$ the adversary cannot produce the response corresponding to a random challenge $X \notin X_{[q]}$. Suppose from the query set $X_{[q]}$ and the responses, the adaptive adversary successfully generates a set $X'_{[q']}$ of q' adaptive queries, and corresponding responses $\Psi_{[q']}$ for the HLPUF $\mathcal{E}_f^L$. Without any loss of generality we assume that for all of the queries $X'_i \in X'_{[q']}$ the HLPUF returns a non-abort state. We assume that the adaptive adversary wins the universal unforgeability game using the query set $X_{\text{ad}} = X_{[q]} \cap X'_{[q']}$. This implies,

$$\Pr_{X,X^{\mathcal{E}_L}_{[q]_{\text{ad}}}}[1 \leftarrow \mathcal{G}^{\mathcal{E}_f^L}(\mathcal{A}_{ad}, m, X, X_{\text{ad}})] \geqslant \text{non-negl}(m). \tag{6.97}$$

From the HLPUF Construction 4, we get that winning the universal unforgeability game with the HLPUF $\mathcal{E}_f^L$ implies winning the universal unforgeability with $\mathcal{E}_{f_2}$. Hence, we can rewrite Eq. (6.97) in the following way,

$$\Pr_{X,X_{\text{ad}}}[1 \leftarrow \mathcal{G}^{\mathcal{E}_{f_2}}(\mathcal{A}_{ad}, m, X, X_{\text{ad}})] \geqslant \text{non-negl}(m). \tag{6.98}$$

Note that, if the adaptive adversary manages to get non-abort outcomes from the HLPUF corresponding to all $X'_i \in X_{\text{ad}}$ then from the Construction 4 we get, $1 \leftarrow \mathcal{G}^{\mathcal{E}_{f_1}}(\mathcal{A}_{ad}, m, X'_i, X_{\text{ad}})$. Due to the unforgeability assumption of Equation Eq. (6.96) we have,

$$\Pr_{X,X_{[q]}}[1 \leftarrow \mathcal{G}^{\mathcal{E}_{f_1}}(\mathcal{A}_{weak}, m, X, X_{[q]})] = \Pr_{X,X_{\text{ad}}}[1 \leftarrow \mathcal{G}^{\mathcal{E}_{f_1}}(\mathcal{A}_{ad}, m, X, X_{\text{ad}})] \leqslant \varepsilon(m). \tag{6.99}$$

Note that, the main difference between adaptive and weak adversaries lies in the choice of the query set. If we fix the query set X_{ad}, then both adaptive $\mathcal{A}_{ad}$ and a weak adversary can extract the same amount of information from the responses corresponding to the query set X_{ad}. Therefore, their winning probability of the universal unforgeability game becomes equivalent. This implies, we can rewrite Equation Eq. (6.99) in the following way,

$$\Pr_{X,X_{\text{ad}}}[1 \leftarrow \mathcal{G}^{\mathcal{E}_{f_1}}(\mathcal{A}_{ad}, m, X, X_{\text{ad}})] = \Pr_{X,X_{\text{ad}}}[1 \leftarrow \mathcal{G}^{\mathcal{E}_{f_1}}(\mathcal{A}_{weak}, m, X, X_{\text{ad}})] \leqslant \varepsilon(m). \tag{6.100}$$

By combining Equation Eq. (6.99) and Equation Eq. (6.100) we get, both the random variables $X_{[q]}$ and X_{ad} are equivalent. From the universal unforgeability property of the PUF $\mathcal{E}_{f_2}$ against any q-query weak adversary, we get

$$\Pr_{X,X_{[q]}}[1 \leftarrow \mathcal{G}^{\mathcal{E}_{f_2}}(\mathcal{A}_{weak}, m, X, X_{[q]})] \leqslant \varepsilon(m). \tag{6.101}$$

As both of the random variables $X_{[q]}$ and X_{ad} are equivalent, so we get,

$$\begin{aligned}
&\Pr_{X,X_{[q]}}[1 \leftarrow \mathcal{G}^{\mathcal{E}_{f_2}}(\mathcal{A}_{weak}, m, X, X_{[q]})] \\
&= \Pr_{X,X_{\text{ad}}}[1 \leftarrow \mathcal{G}^{\mathcal{E}_{f_2}}(\mathcal{A}_{weak}, m, X, X_{\text{ad}})] \\
&= \Pr_{X,X_{\text{ad}}}[1 \leftarrow \mathcal{G}^{\mathcal{E}_{f_2}}(\mathcal{A}_{ad}, m, X, X_{\text{ad}})] \leqslant \varepsilon(m).
\end{aligned} \tag{6.102}$$

The second equality follows from the fact that for a fixed query set X_{ad} the adaptive adversary $\mathcal{A}_{ad}$ and weak adversary $\mathcal{A}_{weak}$ becomes equivalent. Note that, only one of Eq. (6.98) and Eq. (6.102) is true. The Eq. (6.102) is true because of the unforgeability of $\mathcal{E}_{f_2}$. Hence, our assumption of Eq. (6.98) is wrong. Therefore, Eq. (6.97) is also not true. Hence, with conclude our proof by contradiction. □

Apart from the theoretical results provided in this section, we have also simulated the design of HPUF constructions with underlying silicon CPUFs instantiated by *pypuf* [Wis21] which is a python-based emulator that features different existing CPUFs. Furthermore, we simulate the situation where an adversary acquires classical challenges and quantum-encoded responses from HPUF and converts the response into classical bit string by measurement behaviour. The adversary then performs some machine learning-based attacks with CRPs to reproduce a model that accurately predicts enough the behaviours of underlying CPUF. Such an adversary possibly forges the HPUF given (exponentially in the security parameter) many CRPs. Our simulation results assist to demonstrate the exponential gap in the security between CPUF and HPUF in a regime outside the polynomial-size database and for classical PUFs that are commercially available. Since the simulations have not been done by the author, we have excluded them from this chapter. However, we refer the reader to [CDM+21] for the full work, including the simulation results.

6.4.5.3 Security of the HLPUF-based authentication protocol:

We now have all the elements to be able to prove the completeness and security (or soundness) of our HLPUF-based authentication protocol. Firstly, we define the completeness and security property for Protocol 5. Then, in Theorem 47 we will prove that they are satisfied. We start with the completeness:

Definition 48 (Completeness of HLPUF-based Authentication Protocol 5)**.** We say the HLPUF-based authentication Protocol 5 satisfies completeness if in the absence of any adversary, an honest verifier and prover generating $|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}|$ and $|\psi_{f_2(x_i)}\rangle\langle\psi_{f_2(x_i)}|$ with a valid HLPUF for any selected challenge x_i, can pass the verification algorithms with overwhelming probability:

$$Pr[\texttt{Ver}(|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}|, \tilde{\rho}_1) = \texttt{Ver}(|\psi_{f_2(x_i)}\rangle\langle\psi_{f_2(x_i)}|, \tilde{\rho}_2) = 1] \geqslant 1 - \varepsilon(\lambda) \tag{6.103}$$

We also define the security of the protocol, in relation to the universal unforgeability game as follows:

Definition 49 (Security of the HLPUF-based Authentication Protocol 5)**.** We say the HLPUF-based authentication Protocol 5 is secure if the success probability of any QPT adaptive adversary $\mathcal{A}_{ad}$ in winning the universal unforgeability game to forge an output of HLPUF according to Construction 4, for any randomly selected challenge of the form $\tilde{c} = (x, |\psi_{f_1(x)}\rangle\langle\psi_{f_1(x)}|)$ is at most negligible in the security parameter:

$$Pr[1 \leftarrow \mathcal{G}^{HLPUF}(\mathcal{A}_{ad}, \lambda)] \leqslant \varepsilon(\lambda) \tag{6.104}$$

where the verification algorithm of the universal unforgeability game checks the adversary's output σ_2, with the output of the HLPUF, $|\psi_{f_2(x)}\rangle\langle\psi_{f_2(x)}|$.

Through the following theorem, we can see that Protocol 5 satisfies both completeness and security according to the above definitions.

Theorem 47. *If the HLPUF $\mathcal{E}_f^L$ is constructed from a hybrid PUF $\mathcal{E}_f$ using Construction 4, then the HLPUF-based authentication Protocol 5 satisfies both the completeness and security conditions.*

Proof. In Protocol 5 with hybrid PUF $\mathcal{E}_f = \mathcal{E}_{f_1} \otimes \mathcal{E}_{f_2}$, the verifier(server) chooses the classical input $x_i \in \mathcal{X}$, encodes the quantum state corresponding to $2m$ bits of $f_1(x_i)$ and issues the joint state to the prover(client). If there is no adversary, the prover receives the joint state and queries $\mathcal{E}_f$ with x_i and $\tilde{\rho}_1$, where $\tilde{\rho}_1 = \mathcal{E}_{f_1}(x_i) = |\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}|$ for the first m qubits of $\mathcal{E}_f(x_i)$. Hence we have:

$$Pr\left[\texttt{Ver}(|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}|, \tilde{\rho}_1) = 1\right] = 1 \tag{6.105}$$

On the prover's side, since the verification algorithm of HLPUF $\mathcal{E}_f^L$ always passes with $\texttt{Ver}(|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}|, \tilde{\rho}_1) = 1$, and they return the quantum state $\mathcal{E}_{f_2}(x_i) = |\psi_{f_2(x_i)}\rangle\langle\psi_{f_2(x_i)}|$ corresponding to $2m$ bits of $f_2(x_i)$ to the verifier. Without the presence of adversary, the verifier always receives the state with $\tilde{\rho}_2 = |\psi_{f_2(x_i)}\rangle\langle\psi_{f_2(x_i)}|$, and we obtain the equation similarly to Eq. (6.105). Therefore,

we can say the HLPUF-based authentication protocol satisfies the completeness condition with

$$\Pr\left[\mathtt{Ver}(|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}|,\tilde{\rho}_1)=\mathtt{Ver}(|\psi_{f_2(x_i)}\rangle\langle\psi_{f_2(x_i)}|,\tilde{\rho}_2)=1\right]=1 \tag{6.106}$$

On the other hand, for the security property, we rely on Theorem 46 that the HLPUF $\mathcal{E}_f^L$ is (ε, m)-secure against any QPT adaptive adversary (a q-query adaptive adversary for any q polynomial in the security parameter). For both $\mathcal{E}_{f_1}$ and $\mathcal{E}_{f_2}$ of HPUF $\mathcal{E}_f$, we show that the power of an adaptive adversary can be reduced to the power of a weak adversary, due to the locking mechanism. Also since $\mathcal{E}_{f_1}$ has the universal unforgeability against a weak adversary by definition, for any adaptive query of the form (x_i, σ_1) that an adaptive adversary issues to the HLPUF, the following applies:

$$\Pr\left[\mathtt{Ver}(|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}|,\sigma_1)=1\right]\leqslant\varepsilon(m) \tag{6.107}$$

Where $|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}|$ is the correct response constructed from CPUF according to HPUF construction. Thus the power of the adaptive adversary reduces to the power of weak adversary and we have:

$$Pr[1\leftarrow\mathcal{G}^{HLPUF}(\mathcal{A}_{ad},m)]\approx Pr[1\leftarrow\mathcal{G}^{HLPUF}(\mathcal{A}_{weak},m)] \tag{6.108}$$

Now given the fact that the adaptive adversary cannot boost from the weak-learning phase to the HPUF, producing a forgery σ_2 for the HLPUF that passes the verification $\mathtt{Ver}(|\psi_{f_2(x_i)}\rangle\langle\psi_{f_2(x_i)}|,\sigma_2)$, reduces to forging the HPUF $\mathcal{E}_{f_2}$. Again by assumption, $\mathcal{E}_{f_2}$ has the universal unforgeability against weak adversary, hence we have:

$$Pr[1\leftarrow\mathcal{G}^{HLPUF}(\mathcal{A}_{ad},m)]=Pr[1\leftarrow\mathcal{G}^{HPUF}(\mathcal{A}_{weak},m)]\leqslant\varepsilon(m) \tag{6.109}$$

This concludes the proof. □

Therefore, we have shown that Protocol 5, under certain reasonable assumptions on the underlying classical PUF, is correct and achieves suitable security against QPT adversaries.

6.4.6 Challenge re-usability

In any PUF-based protocol relying on the classical communication of challenges and responses of the PUF, each challenge can only be used once as the adversary can simply copy and record the challenges and responses and have a perfect copy of the challenger's database which later they can use to falsely identify themselves. This is an important limitation of the classical PUFs [SD07, HYKD14]. Quantum communication can solve this issue due to the unclonability of quantum states. In this section, we discuss how our hybrid construction can allow for challenge states to be used several times during the authentication, under the circumstances of previous successful authentication rounds. This property will resolve an important

practical issue as the challenger can avoid storing a big database or renewing the database of challenge responses frequently.

First, we need to clarify the conditions under which the challenge can be reused. We assume the challenger's database to only include q number of challenge-response pairs such that q is polynomial in the security parameter. We also need to recall that in our hybrid construction, the challenges are still being sent as *classical* bit-strings over the public channel, hence the adversary, after polynomial rounds of communication, can have the same challenge set as the server's database. Due to this fact, we should emphasize that the adversary does not get any physical access to the internal classical PUF in the HLPUF construction during the authentication and no query can be directly issued to the CPUF by the adversary. This condition is satisfied using our locking mechanism. Thus, the adversary has access to the following information: a pre-learnt polynomial-size local database of challenge-responses of the CPUF, a set of classical challenges used during the protocol, and the set of quantum states that encode either the first or second half of the response, in the BB84 states.

It is a straightforward observation that the challenges for which the verification test has failed should never be used again. A trivial attack, in this case, would be that the adversary intercepts the communication and stores the response state, and later when the same challenge has been queried again, will re-send the stored correct response state to pass the verification. As a result, all the challenges in the failed rounds should be discarded.

Nonetheless, we argue that in the events of successful authentication, the challenges can be re-used. Here, by successful identification, we mean that the received response state passes the verification on the client and server sides and the prover identifies an honest party. Even though the events of false identification of an adversary, is still possible (for example, if the challenge is the same as one of the challenges that previously exists in the adversary's local database), the unforgeability of PUF and our security proof for the hybrid construction, ensures that these events occur only with negligible probability.

We are thus interested in the eavesdropping attacks by the adversary on the first and second half of the response states that are of the form $|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}| = \bigotimes_{j=1}^{m} |\psi^{i,j}_{f_1(x_i)}\rangle\langle\psi^{i,j}_{f_1(x_i)}|$ and $|\psi_{f_2(x_i)}\rangle\langle\psi_{f_2(x_i)}| = \bigotimes_{j=1}^{m} |\psi^{i,j}_{f_2(x_i)}\rangle\langle\psi^{i,j}_{f_2(x_i)}|$ Note that eavesdropping on the states which encode the first part of the response will lead to breaking the locking mechanism while eavesdropping on the second half will lead to an attack on the identification. Without loss of generality, we only consider one of the cases where the adversary wants to eavesdrop on the first (or second) half to break the protocol in the upcoming rounds where the challenge is re-used. The arguments will hold equivalently for both cases since the states and verification are symmetric.

Given all these considerations, the challenge re-usability problem will reduce to the optimal probability of the eavesdropping attack on state $|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}|$ which is in fact m qubit states encoded in conjugate basis same as BB84 states. In the most general case, the adversary can perform any arbitrary quantum operation on

the state $\bigotimes_{j=1}^{m} |\psi^{i,j}_{f_1(x_i)}\rangle\langle\psi^{i,j}_{f_1(x_i)}|$ or separately on each qubit state $|\psi^{i,j}_{f_1(x_i)}\rangle$, together with a local ancillary system and sends a partial state of this larger state to the verifier to pass the verification test, and keep the local state to extract the encoded response bits. Let ρ_{SEC} be the joint state of the server, the eavesdropper and the client. Since the states used in the protocol are from Mutually Unbiased Basis (MUB) states *i.e.* from either $Z = \{|0\rangle, |1\rangle\}$ or $X = \{|+\rangle, |-\rangle\}$, in order to show the optimal attack, we can rely on the entropy uncertainty relations that have been used for the security proof of QKD. The measurements for verification are also performed in the $\{Z, X\}$ bases accordingly. We use the entropy uncertainty relations from [CBTW17] where the security criteria for QKD have been given in terms of the conditional entropy for MUBs measurements. Using these results we show that the entropy of Eve in guessing the correct classical bits for the response is very high if the state sent to the verification algorithm passes the verification with a high probability. Intuitively this is due to the uncertainty that exists related to the commutation relation between X and Z operators in quantum mechanics. Hence we conclude that the success probability of Eve in extracting information from the encoded halves of the response is relatively low. Also, we show that this uncertainty increases linearly with m similar to the number of rounds for QKD. This argument results in the following theorem. In proving this theorem, we have used the entropic uncertainty relation introduced in Chapter 2, Section 2.1.5.

Theorem 48. *In Protocol 5, let x be a challenge and $(y_1, \ldots, y_{2m})$ be the response of a classical PUF used within the HPUF construction, with randomness bias $p = (\frac{1}{2} + \delta_r)^{2m}$ over the classical responses. If the verification algorithm for a state $\tilde{\rho}$ passes with probability $1 - \varepsilon(m)$, then Eve's conditional min-entropy H_{min}^{Eve} in terms of von Neumann entropy over the verifier/prover's (server/client) classical response, satisfies the following inequality:*

$$H_{min}^{Eve} = H_{min}(S^m|ER^m) \geqslant m - \varepsilon(m) \tag{6.110}$$

Proof. We prove this theorem based on the first half of the state used in Protocol 5, *i.e.* the state $|\psi_{f_1(x_i)}\rangle\langle\psi_{f_1(x_i)}| = \bigotimes_{j=1}^{m} |\psi^{i,j}_{f_1(x_i)}\rangle\langle\psi^{i,j}_{f_1(x_i)}|$ that is being sent by the verifier/server denoted by (S) and received and measured by the prover/client denoted by (C). Nevertheless, we note that the same proof applies for the second state due to the symmetry of the states and the protocol.

Let $R^m = (R_1, \ldots, R_m)$ be the randomness bitstring showing the choice of the basis encoding of the response, $S^m = (S_1, \ldots, S_m)$ be the server's bit encoded in the R^m bases. Note that both R^m and S^m are produced according to the bitstring $(y_1, \ldots, y_{2m})$ which is the first half of the response of CPUF to a given challenge x. Also, let $C^m = (C_1, \ldots, C_m)$ be the client's correct bit string. We denote the arbitrary joint state of three systems by $\rho_{S^m E C^m}$ where E denotes any arbitrary quantum system held by the eavesdropper. Now, let the the Client's measurement outcomes, after the verification be $\tilde{Y}^m = (\tilde{Y}_1, \ldots, \tilde{Y}_m)$ which shows the estimated bits by the Client. Now we can write the tripartite uncertainty principle, in terms

of the von Neumann entropy, for MUB measurements and MUB states as follows:

$$H(X_1X_2Z_3X_4\dots X_{m-1}Z_m|E)+H(Z_1Z_2X_3Z_4\dots Z_{m-1}X_m|C)\geqslant \log_2(\frac{1}{c})^m \tag{6.111}$$

where $c=\max_{x,z}c_{xz}$ and $c_{xz}=\| \sqrt{M^x}\sqrt{N^z} \|^2$ for an arbitrary POVM sets $M=\{M^x\}_x$ and $N=\{N^z\}_z$. We note that if the CPUF creates perfect random bit-string for R^m then states are perfect MUB states and $c=\frac{1}{2}$. Nonetheless we consider a weaker CPUF with a biased distribution of $p=(\frac{1}{2}+\delta_r)^{2m}$ in creating 0s and 1s in the response. Hence, we can translate this imperfectness into a disturbance in the measurement bases. Let $M^0=|0\rangle\langle 0|$ and $M^1=|1\rangle\langle 1|$ be the usual measurement in the computational basis but let the N measurements be a slightly shifted version of the measurements in the X basis. Consider the following states:

$$\begin{aligned}|\psi_N\rangle&=\sqrt{\frac{1}{2}+\delta_r}\,|0\rangle+\sqrt{\frac{1}{2}-\delta_r}\,|1\rangle\\ |\psi_N^\perp\rangle&=\sqrt{\frac{1}{2}-\delta_r}\,|0\rangle-\sqrt{\frac{1}{2}+\delta_r}\,|1\rangle\end{aligned} \tag{6.112}$$

We define the new N projective operators according to the following states as $N^0=|\psi_N\rangle\langle\psi_N|$ and $N^1=|\psi_N^\perp\rangle\langle\psi_N^\perp|$. Now we calculate the operator norm for all the pairs of measurements and we have:

$$\begin{aligned}\| \sqrt{M^0}\sqrt{N^0} \|^2&=\frac{1}{2}+\delta_r, \quad \| \sqrt{M^0}\sqrt{N^1} \|^2=\frac{1}{2}-\delta_r\\ \| \sqrt{M^1}\sqrt{N^0} \|^2&=\frac{1}{2}-\delta_r, \quad \| \sqrt{M^1}\sqrt{N^1} \|^2=\frac{1}{2}+\delta_r\end{aligned} \tag{6.113}$$

Thus we conclude that $c=\frac{1}{2}+\delta_r$ and the Equation Eq. (6.111) can be re-written as follows:

$$H(X_1X_2Z_3X_4\dots X_{m-1}Z_m|E)+H(Z_1Z_2X_3Z_4\dots Z_{m-1}X_m|C)\geqslant m-m\log_2(1+2\delta_r) \tag{6.114}$$

Now, we use the data processing inequality [CBTW17], we have got the following security criteria that show Eve's uncertainty (in terms of the von Neumann entropy) of the actual response bits S^m:

$$H(S^m|ER^m)+H(S^m|\tilde{Y}^m)\geqslant m-m\log_2(1+2\delta_r) \tag{6.115}$$

We can get the same inequality in terms of smooth min and max entropy [CBTW17, TR11] (see Section 2.1.5), which is more appropriate for ensuring the security in the finite size, for min and max entropy we equivalently have:

$$H^\varepsilon_{min}(S^m|ER^m)\geqslant m-H^\varepsilon_{max}(S^m|\tilde{Y}^m)-m\log_2(1+2\delta_r) \tag{6.116}$$

To calculate the above bound we need to find the bound on the second term of the right-hand side, *i.e.* $H^\varepsilon_{max}(S^m|\tilde{Y}^m)$. Here we use another result from [TR11]

where it states that for any bitstring X of n bit and the respective measurement outcome X', which at most a fraction ζ of them disagree according to the performed statistical test, then the smooth max entropy is bounded as follows:

$$H^{\varepsilon}_{max}(X|X') \leqslant nh(\zeta) \tag{6.117}$$

where $h(.)$ denotes the classical binary Shannon entropy. Now we can use this result and our assumption of successful verification together. Given the assumption that the verification is passed with a probability $1-\varepsilon(m)$, and the verification algorithm consists of measuring the states in the Z and X bases, we can conclude that the final bits differ in at most a fraction $\zeta = \varepsilon(m)$ where $\varepsilon(m)$ is a negligible function. As a result, we have:

$$H^{\varepsilon}_{max}(S^m|\tilde{Y}^m) \leqslant mh(\zeta) \approx m\varepsilon(m) \tag{6.118}$$

Putting Equations Eq. (6.116) and Eq. (6.118) together, we have:

$$H^{\varepsilon}_{min}(S^m|ER^m) \geqslant m - m\varepsilon(m) - m\log_2(1+2\delta_r) \tag{6.119}$$

On the right-hand side of the above inequality, the second term is still a negligible function and the third term depends on the CPUF bias probability distribution. We assume the CPUF satisfies p-randomness, as defined in the Definition 47, thus the δ_r is a small value and hence the term $(1+2\delta_r)$ is negligibly close to 1, which means that the third term, is negligibly close to 0 in the security parameter which is m. Finally, we conclude that:

$$H^{Eve}_{min} = H^{\varepsilon}_{min}(S^m|ER^m) \geqslant m - \varepsilon'(m) \tag{6.120}$$

where $\varepsilon'(m)$ is a negligible function and the proof is complete. □

Let us see how the abouve information-theoretic bound can be used to prove the challenge-reusability of the Protocol 5. First, define the re-usability in relation with the unforgeability game and then using Theorem 48, we prove the challenge re-usability of our protocol.

Definition 50 (Challenge (k-)re-usability in the universal unforgeability game)**.** Let $\mathcal{G}_{re}(\lambda, \mathcal{A}, x_{k+1})$ be a special instance of the universal unforgeability game, where a challenge x, picked uniformly at random by the challenger, has been previously used k times. We are interested in the events where the same challenge is used in the $(k+1)$-th round, which we denote by x_{k+1}. We say the challenge x is *(k-)re-usable* if the success probability of any QPT adversary in winning $\mathcal{G}_{re}(\lambda, \mathcal{A}, x_{k+1})$, i.e, in forging message x_{k+1}, is negligible in the security parameter:

$$p_{forge}(\mathcal{A}, x_{k+1}) = Pr[1 \leftarrow \mathcal{G}_{re}(\lambda, \mathcal{A}, x_{k+1})] \leqslant \varepsilon(\lambda) \tag{6.121}$$

Theorem 49 (Challenge re-usability of HLPUF-based Authentication Protocol 5)**.** *A challenge x can be reused k times during the Protocol 5 as long as the received respective response σ for each round passes the (client's or server's) verification with overwhelming probability. In other words, given the successful verification, the success probability of any quantum adversary in passing the $(k+1)$-th round with the same challenge x is bounded as follows:*

$$p_{forge}(\mathcal{A}, x_{k+1}) \leqslant k2^{-m} \approx \varepsilon(m). \tag{6.122}$$

Proof. To prove this theorem, we use Theorem 48 directly. First, we assume that x has been used one time before in a previous round. Given the assumption that the verification is passed with probability $1-\varepsilon(m)$, and this theorem, we conclude that the uncertainty of the adversary in guessing the encoded response of the HLPUF is larger than $m-\varepsilon(m)$. In our case, the joint quantum state between the server and the adversary is a classical-quantum state (server has the classical description of $f(x)$, and the adversary has the quantum state $|\psi_{f(x)}\rangle$). For such states, Eve's uncertainty, H_{min}^{Eve} is same as $-\log p_{guess}^{Eve}$, where p_{guess}^{Eve} is Eve's guessing probability of the classical information encoded in the quantum state [KRS09]. Therefore,

$$\begin{aligned} p_{guess}^{Eve} &= 2^{-H_{min}^{Eve}} \\ &\leqslant 2^{-m+\varepsilon(m)} \end{aligned} \tag{6.123}$$

This probability is negligible in the security parameter, which means that after performing any arbitrary quantum operations, the adversary's local state includes at most, a negligible amount of information on the response of x, each round that the state x is reused. Now, we can use the union bound (See Preliminaries, **??**) to show that this success probability only linearly scales with k:

$$p_{guess}^{Eve,k} = Pr(\bigcup_{i=1}^{k} E_{guess}^{i}) \leqslant \sum_{i=1}^{k} p(E_{guess}^{i}) \approx k2^{-m} \tag{6.124}$$

where E_{guess}^{i} are the events where Eve correctly guesses the response and where $p(E_{guess}^{i}) = (p_{guess}^{Eve})^{i}$ is the success probability of Eve in guessing in the i-th round. Finally, let the success probability of an adversary in the universal unforgeability game for the HLPUF be upper-bounded by $\varepsilon_1(m)$ which is a negligible function in the security parameter since we assume that the HLPUF satisfies the universal unforgeability. This is the same as the success probability of the adversary in passing the verification for a new challenge, chosen at random from the database. Now in the $(k+1)$-th round, where the same x is reused, the success probability is at most boosted by the guessing probability over the previous k-th rounds, hence we will have:

$$p_{forge}(\mathcal{A}, x_{k+1}) \leqslant \varepsilon_1(m) + k2^{-m} = \varepsilon(m) \tag{6.125}$$

As long as k is polynomial in the security parameter, the second term is also a negligible function and since the sum of two negligible probabilities will be also negligible. This concludes the proof. □

6.5 Discussion and conclusions

We have proposed three different identification protocols based on quantum PUFs and Hybrid classical-quantum PUFs which provide exponential (or relatively exponential) security against any QPT adversary by exploiting physical unclonability (both in the quantum and classical sense) as a hardware assumption instead of the usual cryptographic assumptions. The first two protocols use full quantum PUFs and the last one combines a classical PUF with quantum encoding to give rise to our hybrid construction. Our primary classification in the first two protocols has come about from the practical scenarios in a network, *i.e.* parties with varying capabilities should be able to run a secure identification protocol. The first protocol, hrv-id, is proposed to be suited more in the mobile-like device settings *i.e.* provers having low resources would want their device to be correctly identified by a high resource verifier. This protocol can be used as a subroutine for many other applications, specifically in a quantum internet or a quantum network with *star-like* architecture [CHZ+09, PB16, LMR+17], where a server (central node) can use this protocol to identify each of the clients. Also, since the identification protocol requires multi-round communication between the prover and the verifier, we have proposed efficient quantum equality-testing verification to reduce the communication overhead requirement.

Our second protocol, lrv-id, is suited to the quantum verification setting *i.e.* a low-resource and classical-like verifier would want to verify the identity of a high resource quantum device, like a quantum cloud server. The advantage of this protocol is that a purely classical verification algorithm is sufficient to verify the prover's device with provable security. lrv-id is based on the idea of trapification, where the verifier inserts random trap states in between the communication rounds which facilitates a secure delegation of the quantum testing to the prover. This allows the verifier to simply run a classical algorithm on the quantum test outcomes to perform successful identification. We have also shown an extension of the lrv-id protocol that generalises it to an arbitrary distribution of traps instead of randomly inserting them in half of the positions as proposed in the current version of the protocol. With this generalisation on hiding the trap distribution, one hopes for further enhancement in security against a QPT adversary. We draw non-trivial conclusions from this generalisation, including the worsening of security to polynomial in the number of communication rounds (instead of exponential as our current protocol) when the number of trap positions is chosen uniformly over the total positions. We also remark that some distributions provide a polynomial enhancement over the current exponential security bound, thus justifying the need for hiding the number of trap positions. We also believe that this generalisation provides a potential use case of a similar protocol for a certain degree of verification and certification of quantum devices using embedded quantum PUFs. We see this direction as an engaging future subject of study since qPUFs provide natural and physical randomness, which combined with more enhanced trapification techniques, can potentially lead to a new class quantum verification/certification protocols.

An important future direction regarding the practicality of the presented protocols is to study the effect of noise and robustness of the protocols under honest noise. So far, the protocols have been studied in the noiseless setting, and the requirements of verification are strictly rigid (for instance, all rounds of verification have to be successfully passed), however, both protocols can be made more robust by relaxing some of these requirements and allowing for a tolerance parameter while remaining secure. We leave the study of the robustness and security properties of the protocol in the noisy setting for future work.

Then, we have used our result regarding efficient universal unforgeability, to reduce the Haar-random sampling of our proposed protocols to pseudorandom quantum states that are efficient to generate, as a result, introducing a more implementation-friendly version of our qPUF-based protocols. Quantum pseudorandomness is yet a very young field of research, and we believe the advancement in this field can assist the qPUF constructions and the protocols based on them to become more and more practical.

Finally, in our latest proposal, we have exploited one of the main sources of security in the quantum information world, namely the concept of conjugate coding [BS16], to propose a new construction that uses an internal (and almost weak) classical PUF and enhances its unpredictability to a high degree using this quantum encoding. Nonetheless, this enhancement is only against non-adaptive adversaries. For security against our usual QPT adversary that is in general adaptive, we use an additional technique, namely the locking mechanism that provides us with our HLPUF construction used in our proposed authentication protocol. An important property of the new construction is the combination of classical challenges and quantum responses, which harnesses the power of quantum information over an untrusted quantum channel while the verifier does not need to store the responses quantumly, which fully removes the quantum-memory requirement. As a result, the implementation of hybrid PUF is practical nowadays with quantum communication technology. Another advantage of the HLPUF-based protocol is that each challenge-response pair used for a successful authentication round can be used several times for authentication due to the unclonability and other fundamental quantum mechanical properties of the response quantum states. Therefore, with our solution, a server can continue the client authentication protocol for a longer period without exhausting its CRP database. This result overcomes the fundamental drawbacks of the existing classical PUF-based authentication protocols and offers a novel use case, not only for our construction but also for quantum communication in general.

However, there are several thought-provoking questions, yet to be explored. In bounding the success probability of a QPT adversary against HPUF in Theorem 45, we have established a connection between unforgeability and the learnability of a classical function (in our case CPUF's evaluation function) from a quantumly encoded random set of data, using quantum informatics approaches. Despite the current proof being specific to our construction, we believe that most of the techniques we have used are fairly general and can be used to formally establish a link between cryptographic properties and learning problems, using

quantum information theory. Furthermore, if the same result can be generalised to bound the success probability of a QPT adversary respective to a classical PPT adversary, for learning a classical function from a general quantum encoding, then one can relate the problem to the open question of the advantage of *supervised learning* with quantum models [SSM21, SG04]. The use of quantum information theory in learning theory has recently led to influential results regarding the comparison between classical and quantum machine learning [HKP21]. Therefore, we conjecture that further expanding this connection to cryptography and quantum-information-based cryptography can provide new insights into these very challenging and exciting problems.

We finally discuss the experimental requirement of our HLPUF-based proposal. We note that by selecting the quantum encoding to be BB84 states, the protocol can be implemented with resources similar to quantum key distribution. QKD technology is one of the most mature quantum technologies. The long-distance QKD networks are already implemented and used in many different countries like the USA, UK, China, EU, Japan, [SFI+11, SLB+11, PPM08, WCY+14, Cou16] etc. Many commercially available QKD infrastructures provide almost 300kb/s secret key rate over optical fibre links of length 120km [FLD+17]. Moreover, the availability of the mature QKD on-chip technology [SEG+17, SSH+20, BLL+18] makes all the proposed constructions in this work implementable inside the IoT devices. Given all these available technologies, our proposal can solve almost all of the shortcomings of the device authentication problem. To further study the feasibility and practicality of hybrid PUF constructions, an important future direction would be toward the experimental implementation of our proposal and the HLPUF-based authentication protocol. Furthermore, we believe that due to the matching of required resources and the strong security guarantee of our protocol, it can be easily incorporated with the QKD itself as a promising solution for providing the authenticated channel that is required for QKD [SBPC+09]. Hence an immediate and important future research direction would be the usage of the protocol for the task of message authentication and the composition of this protocol with QKD. Furthermore, since QKD has been proven composable secure [BOHL+05, Ren08, TL17, Lev15], an important future direction would be to study the security of the proposed protocol within the existing composable frameworks such as universal composability framework [Can01] or abstract cryptography framework [Mau05, Mau12].

7

Variational Quantum Cloning: A New Cryptanalysis Toolkit

"Everything in this world is magic, except to the magician."
– Robert Ford, Westworld (S1.Ep2: Chestnut)

7.1 Introduction

We have set out on a long journey to understand new aspects of 'Unclonability', from a foundational point of view in discovering its relationship to randomness and learnability, all the way to introducing new applications in quantum cryptography. The three past chapters navigate around the concept of physical unclonability. In this chapter, we come back to the more familiar notion of quantum unclonability, that is, the no-cloning theorem and unclonability of quantum states.

In Chapter 2 (Section 2.3), we have covered the no-cloning theorem and the concept of approximate cloning. We have seen that it is both possible to create *imperfect* copies of unknown quantum states (approximate cloning) or to have a quantum operation that only *sometimes* gives you two perfectly similar copies of general quantum states (probabilistic cloning). Among these two categories, approximate cloning is particularly interesting for us since it somehow matches the idea that we have pursued in this thesis in understanding the fundamental relation between unclonability, learnability and the level of *unknownness*. Intuitively we focus on the amount of information that exists prior to performing the cloning mechanism, about the entity that one aims to clone. As we have seen in Section 3.2 and Section 3.3.1. This specific prior information leads to different classes of cloners with the ability to clone the particular family of states corresponding to that information, where the quality (or more technically, fidelity) of these clones is higher than the universal cloner. To roughly summarise this argument, the more you know (or learn) about the states you want to copy, the better you can copy them. Having this inherent relation in mind, now we move to the field of quantum cryptography, where no-cloning is at the heart of the security of many quantum

protocols (QKD, coin-flipping, verifiable blind quantum computing, etc.). Here, the following question arises:

'In what ways the ability to clone a specific family of states with partial prior information can affect the security of quantum protocols?'

Despite the fact that the study of approximate cloning was born many years ago with the remarkable discovery of *Buzeh and Hillery* [BBHB97], and despite the existence of a rich literature on the subject, there are still very limited classes of states that are known how to be cloned approximately with optimal fidelity [SIGA05, FWJ+14]. More importantly, even for some of these known classes of cloning, the unitary circuits (or circuit decompositions) of these cloners are not known. Having these circuits and unitaries explicitly is not only important for practical applications of quantum cloning machines, but also a very relevant problem to quantum compilation. Notably, in touching on quantum cryptanalysis, this 'prior information' becomes much more general and can include cases where our knowledge about the cloning machines (for the specific problem of interest) is narrow. Moreover, if cloning based on a specific family of the state is to be used as an attack model, the complexity of the circuit and technological feasibility of performing those cloners will be considerable factors, which once again calls for being able to have an explicit form of the cloning machines.

Quantum cloning has been previously considered as an attack model for some quantum protocols such as QKD. It turns out that in some cases, these types of attacks are in fact, optimal [SIGA05]. In cases where they are not, cloning provides a means to determine lower bounds on the strategies of an adversary [XSW+12]. However, *implementing* such cloning-based attacks might be non-trivial in practice due to the difficulties mentioned above. Also, the effect of decoherence and errors in NISQ devices makes the production of high-quality clones out of reach for the adversary, which limits the power of practical quantum cryptanalysis in the NISQ era. On the other hand, there has been much interest in implementing quantum cloning and cryptographic attacks on protocols via specific and tailored experiments (for example [LLSHB02, Fiu03, CZSD07, BL13, B+17]), but these may not be easily reconfigurable or generalizable to other scenarios. In summary, finding and constructing quantum cloning circuits for preparing high fidelity clones on NISQ hardware is challenging.

All the arguments given above, motivate us to seek a new approach to efficiently produce optimal cloning machines and their circuits for specific classes of states. We are in particular interested in the ones that are implementable efficiently on NISQ devices. We also note that existing approaches in the literature are by no means optimal if one wants to generalise this question in order to target applications, especially targeting applications in cryptanalysis given the everyday-expanding variety of quantum protocols. The known analytical approaches proceed as follows: Firstly, the most general unitary for the cloning machine has been considered. Secondly, by imposing the symmetries and conditions on the specified family of states, the output fidelity of the machine has been optimised, irrespective of the characterisation of the unitary. Thirdly, one should try to find a unitary

matrix that achieves that maximum fidelity, which is often a non-trivial and challenging task. Some inspiring alternative approaches have also been given for the case of equatorial states[1] [CIVA02], which exploit the symmetries in a much more intriguing way than the previous formalism for cloning. However, these approaches do not seem generalisable to other classes of cloning.

The approach we propose here takes a very different path. We start with the idea of *letting a quantum machine learn how to clone* a specific given family of quantum states. We give a novel algorithm: 'Variational Quantum Cloning' (VarQlone) which uses quantum machine learning (QML) [WHT15, BWP+17, Kop18, SP18a, SP18b] techniques to *learn* how to clone quantum states in an end-to-end manner. VarQlone is made possible by recent advances and techniques in the field of *variational* quantum algorithms (VQAs) [MRBAG16, Bia21, ECBY21, WHT15, CAB+21]. VQAs are intentionally tailored to be useful on NISQ devices, which are limited in scale and noisy to implement 'coherent' algorithms with speedups, such as factoring large prime numbers [Sho94]. However, such devices are capable of performing tasks which cannot be simulated by any classical device in reasonable time [AAB+19, ZWD+20, CKDK21]. This motivates the search for dedicated applications for a topic of likely practical relevance.

Variational quantum algorithms have been proposed and used for various applications, including quantum chemistry [PMS+14] and combinatorial optimization [FGG14]. The core quantum component is typically a parameterised quantum circuit (PQC) [DHLT20] (as mentioned in Section 2.6.4.1). When VQAs are applied to machine-learning problems, they have come to be seen as quantum neural networks (QNNs) [BLSF19, KBA+19]. This is because they can achieve many of the same tasks as classical neural networks, [MNKF18, GBC+18] and can outperform them in certain cases [WM20, CMDK20, CCL19]. Furthermore, machine learning techniques, both quantum [MTB18, KLP+19, JB22, HSNF18, BPLC+20, XSE+21, HBR21] and classical [KMF+16, MPNK+18, ONK19, NMR+19, WMDB20] have proven to be useful in *discovering* and providing insights into quantum algorithms and subroutines. This line of study even extends to the foundations of quantum mechanics. We refer the reader to this paper [ACS+19] about variational consistent histories.

VarQlone is different from other variational algorithms in that it can be viewed as the first step into a new area of applications, *variational quantum cryptanalysis*. Specifically, by using QML techniques to learn to clone quantum states, VarQlone can discover unique ways to attack quantum protocols, in particular those whose underlying security can be reduced to quantum cloning. Furthermore, in developing such techniques more generally, we can determine the relationship between classical machine learning and deep learning, with classical cryptography [AMS+15, MPP16, PMSW16, Ala19].

We believe this new approach for approximate cloning is reasonably general and can be used to deepen and widen our understanding of approximate cloning. However, as a concrete case study and proof of concept of our new approach to crypt-

[1] We recall that we have defined these states in Chapter 2, Section 2.3.1.2

analysis, we map two well-known families of approximate cloning machines, *i.e. phase-covariant cloning* [BCMDM00] and *fixed-overlap state-dependent cloning* [BDE+98, BM99] to two well-known families of quantum protocols, namely *quantum key distribution* and *quantum coin flipping* respectively [MSCK18, ATSVY00]. The latter is probably more exciting since, to the best of our knowledge, this is the first time such a connection has been revealed in the case of state-dependent cloning. One can also see this work as an attempt to uncover the core ingredient of security of different quantum protocols from the perspective of the family of states they rely on.

As part of our research in developing this algorithm, we define suitable cost functions which depend on the symmetries used in the cloning problem, and we use them to prove theoretical guarantees for them (including notions of faithfulness [KLP+19]). As we have discussed in Section 2.6.4.1 since VQAs are heuristic techniques, usually being able to provide such theoretical guarantees is rare in the field (also one of the reasons we have started the chapter with 'everything in this world is magic, except to the magician'!) and this work is one of the few exceptions.

Finally, to underline the practical potential of our approach we implemented it on the Rigetti `Aspen` quantum computer and show how VarQlone can learn to clone states with a higher fidelity on this device than previously known 'analytic' quantum circuits, highlighting the flexibility of our approach. Furthermore, the nature of VarQlone allows us to improve cloning fidelities generically, on quantum computers available through the cloud [LaR19], without requiring significant tweaking and custom-built experimental hardware.

7.1.1 Structure of the chapter

First, in Section 7.2 we introduce how different classes of cloning can be used as a cryptanalysis toolkit to attack protocols that benefit from using certain classes of states. For our case study we introduce cloning attacks on the BB84 protocol in 7.2.1 in relation to phase covariant cloning, and two different quantum coin-flipping protocols in Section 7.2.2 in connection to state-dependent cloning. We present our cloning-based attacks based on an optimal cloner which we then replace with the cloning machines learned through our variational algorithm. In Section 7.3 we introduce some theoretical aspects and specifications of VarQlone such as cost functions (Section 7.3.1), gradient of the proposed cost functions (Section 7.3.2) and more importantly, theoretical guarantees for our algorithm based on the proposed cost function (Section 7.3.3). Finally, in Section 7.4 we present both simulation and experimental results based on VarQlone, including the circuits found for our specified problems and the probability analysis of the attacks based on them.

7.2 Quantum cryptanalysis based on different classes of cloning

In the first part of this chapter, and before introducing our machine learning algorithm, we want to establish the theoretical ground for our cryptanalysis based on different types of cloning with *partial prior information* about the states. Here, we only look at cryptanalysis using approximate quantum cloning, and we do not study the use-cases and relevance of probabilistic cloners. Using cloning as an attack strategy has been considered previously only in limited settings, for instance, regarding the QKD protocol, where it has been shown that there exists an optimal cloning-based attack on the BB84 protocol [SIGA05]. Nevertheless, here we take a more methodologic approach. Keeping in mind that approximate cloning is categorised into different classes characterised by *specific* families of states, we look for quantum protocols that use the same state families for which we have a cloning machine.

Among different types of cloning, universal cloning does not seem very thrilling for this purpose, as it is state-agnostic. Moreover, as we have seen, restricting the class of states that one wants to clone,*i.e.* having prior information about the states leads to higher fidelity clones. In cryptanalysis, we ever optimise over all the adversarial strategies, which in this case includes using the best cloning strategy given the protocol's characteristics.

We choose two main families of approximate cloning for our purpose, namely *phase-covariant cloning* (see Section 2.3.1.2) and *fixed-overlap state-dependent cloning* (see Section 2.3.1.3). For the former, our target example protocol is QKD, where we will explain the optimal cloning-based attack. This analysis serves as an important tool as we use the calculations later for our variational cloning attacks. The latter class of approximate cloning, on the other hand, has never been used before for the purpose of cryptanalysis to the best of our knowledge. We notice that the specific family of states used in this type of cloning matches the class of states used in quantum coin-flipping. Hence, we introduce, for the first time, cloning-based attacks on two different quantum coin-flipping protocols.

Finally, we emphasise that the purpose of this study is mostly the illustrations of applications of variational cloning based techniques, for practical cryptanalysis.[2] Nonetheless, the attacks we present in this section either saturate the optimal bounds, where there exists a rigorous security proof or even lead to a complete novel security attack, in the absence of such strong proof.

[2]To clarify, let us take QKD as an example. This protocol has been proven information-theoretic secure, so we do not intend to break it using variational attacks as it seems pointless. But rather QKD will serve as an example of phase-covariant cloning attacks in general. Moreover, as we will see it is not the case for all the coin-flipping protocols we study in this chapter.

7.2.1 Cryptanalysis based on phase-covariant cloning

Let us begin with phase-covariant cloning and the quantum key distribution protocols by focusing on the BB84 protocol [BB84, BB14]. In this protocol, one honest party, Alice, sends single-qubit states in two orthogonal bases (for instance, the eigenstates of the Pauli X and Pauli Y matrices, $|\pm\rangle$ and $|\pm i\rangle$) to a second honest party, Bob, via a quantum channel that is susceptible to an eavesdropping adversary, Eve. Eve's goal is to extract the secret information exchanged between Alice and Bob, encoded in the states. It turns out that the optimal 'individual' (or incoherent) Eve's attack [SIGA05] on this protocol is given by cloning so-called *phase-covariant* [BCMDM00] states of the form:

$$|\psi_{xy}(\eta)\rangle = \frac{1}{\sqrt{2}}\Big(|0\rangle + e^{i\eta}|1\rangle\Big) \tag{7.1}$$

For these states, some *analytic* circuits are given in [BBHB97, FMWW01, FWJ+14]. For this family of states, Eve can construct a cloning machine with fidelity $F_{\mathrm{L,opt}}^{\mathrm{PC,E}} \approx 0.85$.

There are different families of attacks considered on such protocols. The simplest attack by Eve is a so-called 'incoherent' or individual attack, where Eve attacks each quantum state individually before the reconciliation phase of the protocol. The security of this protocol relies on the information-theoretic bounds on the information shared between Alice and Bob as compared to the information that Eve was able to extract from the key. In the incoherent attacks, the security condition states that a secret key can only be extracted as long as the amount of Eve's information is less than what Bob has received. Thus, one key parameter in the protocol is what is called the *critical error rate*, D_{crit}, which defines the threshold above which Alice and Bob abort the protocol and conclude that the channel is insecure.

For incoherent attacks, the optimal error rate for the ideal incoherent attack is $D_{\mathrm{crit}}^{\mathrm{incoh}} = 1 - F_{\mathrm{L,opt}}^{\mathrm{PC,E}} \approx 14.6\%$ [SIGA05].

However, as discussed in [SIGA05], this is not the best way of analysing the cloning-based attacks against this protocol, since it does not allow for a comparison between a cloning machine that uses the ancilla, and one that does not. The importance of this comparison is that the cloning machine with ancillary inputs may provide Eve with extra information about both parties. A more appropriate way of calculating the key rate which generalises the strategies is via the Holevo quantity, denoted as χ which is defined as follows from von Neumann entropy:

$$\chi(Q:E) := S(\rho_E) - \frac{1}{2}S(\rho_E^0) - \frac{1}{2}S(\rho_E^1) \tag{7.2}$$

In Eq. (7.2), ρ_E denotes the mixed state of Eve over all of the combinations of Alice's choice of input, and ρ_E^0 and ρ_E^1 denote the states of Eve for the random variables that encode 0 and 1 in the protocol respectively.

Combining mutual information with Holevo quantity gives a concrete and simple method for calculating the key rate in QKD protocols. The key rate for QKD

can be defined as follows:

$$R = I(A:B) - \min\{\chi(A:E_Q), \chi(B:E_Q)\} \tag{7.3}$$

where $I(A:B)$ denotes the mutual information between Alice and Bob and the index Q denotes that Eve may employ general quantum strategies. Intuitively, Eq. (7.3) states that no key can be extracted at $R = 0$ which is when Alice and Bob's mutual information is the minimum value between Alice and Eve and Bob and Eve. At any point after that, Eve has increased the correlation to the key, to the point that the key is compromised.

For calculating the quantity D_{crit}, one needs to calculate the Holevo quantity for Eve, set $R = 0$, compute the mutual information, $I(A:B) = 1 - H(D_{\text{crit}})$ and finally solve the resulting equation for D_{crit}. For the ideal incoherent attack, this value is again proven to be $D_{\text{crit}}^{\text{incoh}} \approx 14.6\%$. We go back to this calculation in Section 7.4.1 where we will show that the cloning transformations we learn using our variational cloning algorithm give an approximately close critical error rate while being experimentally superior to the ideal proposed circuits.

7.2.2 Cryptanalysis based on state-dependent cloning

Now, let us examine the class of states used in *state-dependent* cloning, which are states with fixed and known overlap. Non-orthogonal quantum states are among the elements that are often present in quantum cryptography since they can encode information that is not easily decodable for an adversary who does not know the basis. The most famous example is, of course, conjugate coding [Wie83]. But despite the generality of the state-dependent cloning framework, it is surprising that this type of cloning machine has not been studied as a concrete attack model, and to the best of our knowledge, this is the first time we use state-dependent cloning as a cryptanalysis tool. A concrete example of the protocols that exploit such states is the family of *quantum coin-flipping* protocols. Coin-flipping is a cryptographic task where two mutually distrustful parties, who are usually spatially separated and want to agree on a common random bit (see Section 2.5.7 for more details about coin-flipping). Classical and quantum coin-flipping have a vast literature, but here for our case study, we focus on two specific *strong quantum coin-flipping* protocols. The two protocols we consider are that of Mayers et al. [MSCK99], and that of Aharonov et al. [ATSVY00]. First, let us introduce the common aspect of these protocols as well as the states used in the protocols.

7.2.2.1 Quantum coin flipping states

Let us first introduce quantum coin flipping in more detail. The task of the quantum coin flipping is similar to the classical one, only the parties can have quantum capabilities. We say the coin is 'biased' when one outcome is more likely to occur than the other, for example, with the following probabilities:

$$\begin{aligned} \Pr(y=0) &= 1/2 + \varepsilon \\ \Pr(y=1) &= 1/2 - \varepsilon \end{aligned} \tag{7.4}$$

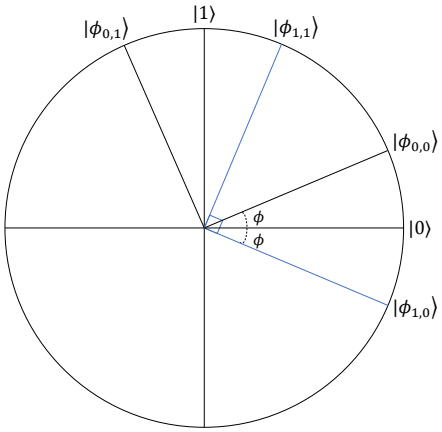

Figure 7.1: States used for quantum coin-flipping. The first bit represents the *basis*, while the other represents one of the two orthogonal states.

where y is the output bit. The above coin is an ε-biased coin with a bias towards the outcome 0. In contrast, a fair coin would correspond to $\varepsilon = 0$.

We recall from Section 2.5.7 that it is impossible in an information-theoretic manner, to achieve a perfectly secure (with zero bias $\varepsilon = 0$) strong coin-flipping protocol in both the classical and quantum setting [Blu83, LC98, MSCK99]. Several protocols have been proposed for ε-biased strong coin flipping [MSCK99, ATSVY00, BB14, BBBG09], the states used by these protocols share a common structure. Here we introduce a more general form of these states, which will be useful for our purpose. The following set of states (illustrated in Fig. 7.1) have been used in the protocols that we will investigate:

$$|\phi_{x,a}\rangle = \begin{cases} |\phi_{x,0}\rangle = \cos\phi\,|0\rangle + (-1)^x \sin\phi\,|1\rangle \\ |\phi_{x,1}\rangle = \sin\phi\,|0\rangle + (-1)^{x\oplus 1} \cos\phi\,|1\rangle \end{cases} \tag{7.5}$$

where $x \in \{0,1\}$, and the angle ϕ determines the overlap between the pairs of states.

Additionally, these protocols share the following shared structure: One of the parties, or sometimes both, will encode some random classical bits into the above states and then exchange some classical/quantum information as part of the protocol. The attack of the malicious party who is trying to bias the coin is then, reduced to the ability to learn the encoded classical bit, from the state (or in some cases, to prepare states deviating from the perfect ones). To see why this is the case, we need to look at the impossibility of the classical coin-flipping task [Blu83]. The intuitive reason behind this impossibility lies in the *order* or asymmetry between the two parties. In other words, since the outcome has to be determined after a certain number of communication rounds in the protocol consisting of sending some messages, one can always find a message such that, before the message is sent, the outcome is not yet determined, but once the message has been sent it becomes determined. This means that the party who

receives this 'extra' information first can always bias the protocol. This limitation cannot be overcome in the classical world. Therefore, there is no value of $\varepsilon < 0.5$ for which, the protocol can be secure. However, using non-orthogonal quantum states gives the parties the ability to 'hide' their choice of the random bit before the other party also flips a coin. Therefore the existence of quantum coin-flipping protocols with a certain degree of bias is closely related to the fact that an adversary, is fundamentally bounded by quantum mechanics, in performing the task of distinguishing non-orthogonal quantum states.

Now, going back to our selected protocols, we have the two following cases:

1. The protocol of Mayers *et. al.* [MSCK99] (denoted by $\mathcal{P}_1$) in which the states, $\{|\phi_{0,0}\rangle, |\phi_{1,0}\rangle\}$ are used (which have a fixed overlap $s = \cos(2\phi)$).

2. The protocol of Aharonov *et. al.* [ATSVY00] (denoted by $\mathcal{P}_2$), which uses the full set of states, *i.e.* $\{|\phi_{x,a}\rangle\}$.

This set of states is conveniently related through a reparameterisation of the angle ϕ [BM06], which makes them easier to deal with mathematically.

In general for the security analysis of strong quantum coin-flipping protocols, one considers both cases where Alice or Bob are being dishonest. Here, for simplicity of comparison and since our goal is to demonstrate cloning-based attacks, we only focus on a dishonest Bob who tries to bias the bit by cloning the non-orthogonal states sent by Alice.

In the following two subsections, the biases are computed assuming access to the *ideal* cloning machine (*i.e.* the one which clones the input states with the optimal, analytical fidelities). Later, we compare these ideal biases with those achievable using the quantum cloning machines learned by our variational quantum cloner.

7.2.2.2 Cloning attack on 2-state quantum coin flipping protocol

The Mayers' protocol was incidentally one of the first protocols proposed for strong quantum coin-flipping. Here, Alice utilizes the states[3] $|\phi_0\rangle := |\phi_{0,0}\rangle$ and $|\phi_1\rangle := |\phi_{1,0}\rangle$ such that the angle between them is $\phi := \frac{\pi}{18} \implies s := \cos(\frac{\pi}{9})$. In the following, we describe the general version of the protocol with k rounds (of quantum communication). We also discuss the proposed attack in more detail and prove the relevant theorems for fixed-overlap state-dependent cloning attacks.

Protocol 6 ($\mathcal{P}_1$ with k rounds)**.** During the protocol, each party sends multiple copies of either $|\phi_0\rangle \otimes |\phi_1\rangle$ or $|\phi_1\rangle \otimes |\phi_0\rangle$

---

1. Alice and Bob now choose k random bits, $\{a_1, \ldots, a_k\}$ and $\{b_1, \ldots, b_k\}$ respectively. The final bit is now equal to the XOR of input bits over all k

[3] Since the value of the overlap is the only relevant quantity, the different parameterisation of these states compared to the ones in Eq. (7.5) does not make a difference for our purposes. However, we note that explicit cloning unitary would be different in both cases.

rounds *i.e.*,

$$x = \bigoplus_j a_j \oplus \bigoplus_j b_j \tag{7.6}$$

2. Each round has n steps identified by i: In each round $j = 1, \ldots, k$ of the protocol, and for every step $i = 1, \ldots, n$ within each round, Alice uniformly picks a random bit $c_{i,j}$ and sends the state $|\phi_c^{i,j}\rangle := |\phi_{c_{i,j}}\rangle \otimes |\phi_{\overline{c_{i,j}}}\rangle$ to Bob (where $\overline{c_{i,j}}$ denotes the complement of $c_{i,j}$).

3. Bob uniformly picks a random bit $d_{i,j}$ and sends the state $|\phi_d^{i,j}\rangle := |\phi_{d_{i,j}}\rangle \otimes |\phi_{\overline{d_{i,j}}}\rangle$ to Alice.[4]

4. For each j and i, Alice announces the value $a_j \oplus c_{i,j}$.

5. If $a_j \oplus c_{i,j} = 0$, Bob returns the second state of the pair (i,j) back to Alice, and sends the first state otherwise.

6. Bob announces $b_j \oplus d_{i,j}$.

7. Alice returns one of the states back to Bob accordingly (similar to step 5).

8. a and b are announced by both sides.

9. Alice measures the remaining states with the projectors, $(E_b, E_b^\perp)$ and the returned states by Bob with $(E_{\overline{a}}, E_{\overline{a}}^\perp)$ (Eq. (7.7)). She aborts the protocol if the measurement result corresponds to $\perp$, and declares Bob as being dishonest[5].

Considering steps (5) to (7) of the protocol, we argue that it is sufficient to only consider a single round in the protocol from the point of view of a cloning attack. This is because a dishonest Bob can bias the protocol if he learns about Alice's bit a_j (for any choice of j), which he can do by guessing $c_{i,j}$ with probability better than $1/2$. With this knowledge, Bob only needs to announce a single false $b_j \oplus d_{i,j}$ to cheat, and so this strategy can be deferred to the final round [MSCK99]. Hence a single round of the protocol is sufficient for analysis, and we herein drop the j index.

In the last phase of the protocol, after a and b are announced by both sides (so x can be computed by both sides), Alice performs the measurements $(E_b, E_b^\perp)$ and $(E_{\overline{a}}, E_{\overline{a}}^\perp)$ on the remaining states. (as defined in Eq. (7.7)) for checking whether Bob has cheated or not. In this sense, the use of quantum states in this protocol is purely for cheat-detection.

$$E_l = |\phi_l\rangle\langle\phi_l|^{\otimes n} \tag{7.7}$$

$$E_l^\perp = \mathbb{1} - |\phi_l\rangle\langle\phi_l|^{\otimes n}, \qquad l \in \{0,1\} \tag{7.8}$$

[4] Note that if $c_{i,j}$ and $d_{i,j}$ are chosen independently of a_j and b_j, no information about the primary bits has been transferred.

[5] The similar verification measurement is performed by Bob to verify Alice, however here we skip that part since we are only interested in half of the protocol, *i.e.* dishonest Bob

A Cloning Attack on $\mathcal{P}_1$: Next, we present the explicit attack and calculation that can be implemented by Bob on $\mathcal{P}_1$. Without loss of generality, we assume that Bob wishes to bias the bit towards $x = 0$. For clarity, we give the attack for when Alice only sends one copy of the state ($n = 1$), but we discuss the general case later:

Attack 1 (Cloning Attack on $\mathcal{P}_1$ with $k = 1$)**.** The goal is to bias the bit towards 0, *i.e.* $p(x = 0) > 1/2$

Inputs. Random bit for Alice ($a \xleftarrow{\$} \{0,1\}$) and Bob ($b \xleftarrow{\$} \{0,1\}$). Bob receives a state $|\phi_c^i\rangle$ from Alice.
The attack:

1. for $i = 1, \ldots, n$:

   (a) **Step 1:** Alice announces $a \oplus c_i$. If $a \oplus c_i = 0$, Bob sends the second qubit of $|\phi_c^i\rangle$ to Alice, otherwise he sends the first qubit.

   (b) **Step 2:** Bob runs a $1 \to 2$ state-dependent cloner on the qubit he has to return to Alice, producing 2 approximate clones. He sends her one clone and keeps the other.

   (c) **Step 3:** Bob runs an optimal state discrimination on the remaining qubit (and any other auxiliary output of the cloner, if exists), and finds c_1 with a maximum success probability $P^{\text{opt}}_{\text{disc},\mathcal{P}_1}$. He then guesses a bit a' such that $P_{\text{succ},\mathcal{P}_1}(a' = a) := P^{\text{opt}}_{\text{disc},\mathcal{P}_1}$.

   (d) **Step 4:** If $a' \oplus b = 0$ he continues the protocol honestly and announces $b \oplus d_1$, otherwise he announces $a' \oplus d_1$. The remaining qubit on Alice's side is $|\phi_a^i\rangle$.

Now, we find the success probability of the above attack:

Theorem 50. *[Bias of ideal cloning attack on $\mathcal{P}_1$] Bob can achieve a bias of $\varepsilon \approx 0.27$ using an ideal state-dependent cloning attack on the protocol $\mathcal{P}_1$ using a single copy of Alice's state.*

Proof. As mentioned in the previous section, the final measurements performed by Alice on her remaining n states, plus the n states returned to her by Bob allow her to detect his nefarious behaviour. If he performed a cloning attack, the $\perp$ outcomes would be detected by Alice with some probability. We must compute both probabilities: the probability of guessing the value of Alice's bit a (by guessing the value of the bit c_1), and the probability of being detected by Alice. This would provide us with Bob's final success probability in cheating, hence the bias probability.

At the start of the attack, Bob has a product state of either $|\phi_0\rangle \otimes |\phi_1\rangle$ or $|\phi_1\rangle \otimes |\phi_0\rangle$ (but he does not know which). After the announcement stage, depending on Alice's announced bit, Bob proceeds to clone one of the qubits, sends

one copy to Alice and keeps the other to himself. Without loss of generality, we assume that Alice's announced bit is 0. In this case, at this point of the attack, he has one of the following pairs: $|\phi_0\rangle\langle\phi_0| \otimes \rho_c^1$ or $|\phi_1\rangle\langle\phi_1| \otimes \rho_c^0$, where ρ_c^1 and ρ_c^0 are leftover clones (the second state of the cloner together with any existing ancillary systems) for $|\phi_1\rangle$ and $|\phi_0\rangle$ respectively.

Bob must now discriminate between the following density matrices:

$$\rho_1 = |\phi_0\rangle\langle\phi_0| \otimes |\phi_1\rangle\langle\phi_1| \tag{7.9}$$

$$\text{and} \qquad \rho_2 = |\phi_1\rangle\langle\phi_1| \otimes \rho_c^0 \tag{7.10}$$

Alternatively, if Alice announced $a \oplus c_i = 1$, he would have:

$$\rho_1 = |\phi_1\rangle\langle\phi_1| \otimes |\phi_0\rangle\langle\phi_0|, \tag{7.11}$$

$$\text{and} \qquad \rho_2 = |\phi_0\rangle\langle\phi_0| \otimes \rho_c^1 \tag{7.12}$$

In either case, we have that the minimum discrimination error for two density matrices is given by the Holevo-Helstrom bound [Hol73, Hel69] (also see Section 2.2 for more information) bound as follows[6]:

$$P_{\text{disc}}^{\text{opt}} = \frac{1}{2} + \frac{1}{4}\|\rho_1 - \rho_2\|_{\text{Tr}} = \frac{1}{2} + \frac{1}{2}d_{\text{Tr}}(\rho_1, \rho_2) \tag{7.13}$$

The ideal symmetric cloning machine for these states will have an output of the form:

$$\rho_c = \alpha|\phi_0\rangle\langle\phi_0| + \beta|\phi_1\rangle\langle\phi_1| + \gamma(|\phi_0\rangle\langle\phi_1| + |\phi_1\rangle\langle\phi_0|) \tag{7.14}$$

where α, β and γ are functions of the overlap $s = \langle\phi_0|\phi_1\rangle = \cos\frac{\pi}{9}$. Now, using Eq. (7.9), ρ_2 can be written as follows:

$$\begin{aligned}\rho_2 =& \alpha|\phi_1\rangle\langle\phi_1| \otimes |\phi_0\rangle\langle\phi_0| + \beta|\phi_1\rangle\langle\phi_1| \otimes |\phi_1\rangle\langle\phi_1| \\ &+ \gamma(|\phi_1\rangle\langle\phi_1| \otimes |\phi_0\rangle\langle\phi_1| + |\phi_1\rangle\langle\phi_1| \otimes |\phi_1\rangle\langle\phi_0|)\end{aligned} \tag{7.15}$$

Finally, by plugging in the values of the coefficients in Eq. (7.14) for the optimal local cloning machine [BDE+98] and finding the eigenvalues of $\sigma := (\rho_1 - \rho_2)$, we can calculate the corresponding value for Eq. (7.13), and recover the following minimum error probability:

$$P_{\text{fail},\mathcal{P}_1} = P_{\text{disc},\mathcal{P}_1}^{\text{er}} = 1 - P_{\text{disc},\mathcal{P}_1}^{\text{opt}} \approx 0.214 \tag{7.16}$$

This means that Bob can successfully guess c_1 with $P_{\text{succ},\mathcal{P}_1}^1 = 78.5\%$ probability.

Now we look at the probability of a cheating Bob being detected by Alice. We note that whenever Bob guesses a successfully, the measurements $(E_b, E_b^\perp)$ will be passed with probability 1, hence we use $(E_{\bar{a}}, E_{\bar{a}}^\perp)$ where the states sent by Bob will be measured. Using Eq. (2.82) (in Section 2.3.1.3) with the value of overlap $s = \cos(\pi/9)$, the optimal fidelity is $F_{\text{L}} \approx 0.997$ and so the probability

[6] This also is because we assume a symmetric cloning machine for both $|\phi_0\rangle$ and $|\phi_1\rangle$. If this is not the case, the guessing probability is instead the average of the discrimination probabilities of both cases.

of Bob getting caught is at most 1%. Putting this together with Bob's guessing probability for a gives his overall success probability of 77.5%.

This implies that Bob is able to successfully create a bias of $\varepsilon \approx 0.775 - 0.5 = 0.275$. □

We also have the following corollary, for a general number n of exchanged states, which shows the protocol can be completely broken and Bob can enforce an arbitrary bias:

Corollary 9. *The probability of Bob successfully guessing Alice's bit a, over n rounds and from all n copies of the received stated, has the property:*

$$\lim_{n\to\infty} P^n_{\text{succ},\mathcal{P}_1} = 1 \tag{7.17}$$

Proof. If Bob repeats the above Attack 1 over all n copies, he will guess n different bits $\{a'_i\}_{i=1}^n$. He can then take a majority vote and announce b such that $a^* \oplus b = 0$, where we denote a^* as the bit he guesses in at least $\frac{n}{2}+1$ of the rounds.

If n is even, he may have guessed a' to be 0 and 1 an equal number of times. In this case, the attack becomes indecisive and Bob is forced to guess at random. Hence we separate the success probability for even and odd n as follows:

$$P^n_{\text{succ},\mathcal{P}_1} = \begin{cases} \sum_{k=\frac{n+1}{2}}^{n} \binom{n}{k}(1-P_{\text{fail}})^k P_{\text{fail}}^{n-k} & n \text{ odd}, \\ \sum_{k=\frac{n}{2}+1}^{n} \binom{n}{k}(1-P_{\text{fail}})^k P_{\text{fail}}^{n-k} + \frac{1}{2}\binom{n}{n/2}(1-P_{\text{fail}})^{\frac{n}{2}} P_{\text{fail}}^{\frac{n}{2}} & n \text{ even} \end{cases} \tag{7.18}$$

By substituting the value of P_{fail} one can see that the function is uniformly increasing with n so $\lim_{n\to\infty} P^n_{\text{succ},\mathcal{P}_1} = 1$[7]. This concludes the proof. □

7.2.2.3 Cloning attack on 4-state quantum coin flipping protocol

Another class of coin-flipping protocols are those which require all the four states in Eq. (7.5). One such protocol was proposed by Aharonov *et al.* [ATSVY00], where the optimal ϕ is set as $\frac{\pi}{8}$ *i.e.* resulting in the following states:

$$|\phi_{x,a}\rangle = \begin{cases} |\frac{\pi}{8}_{x,0}\rangle = \cos\left(\frac{\pi}{8}\right)|0\rangle + (-1)^x \sin\left(\frac{\pi}{8}\right)|1\rangle \\ |\frac{\pi}{8}_{x,1}\rangle = \sin\left(\frac{\pi}{8}\right)|0\rangle + (-1)^{x\oplus 1} \cos\left(\frac{\pi}{8}\right)|1\rangle \end{cases} \tag{7.19}$$

In protocols of this form, Alice encodes her bit as 'basis information' of the family of states. More specifically, her random bit is encoded in the state $|\phi_{x,a}\rangle$. For instance, we can take $\{|\phi_{0,0}\rangle, |\phi_{1,0}\rangle\}$ to encode the bit $a = 0$; and $\{|\phi_{0,1}\rangle, |\phi_{1,1}\rangle\}$ to encode $a = 1$. The goal again is to produce a final 'coin flip' $y = a \oplus b$, while

[7] Although, as Bob's success probability in guessing correctly increases with n, the probability of his cheating strategy getting detected by Alice will also increase, yet does not converge to 1 as fast. We also note that this strategy is independent of k, the number of different bits used during the protocol.

ensuring that no party has biased the bit y. A similar protocol has also been proposed using BB84 states [BB14] where $|\phi_{0,0}\rangle := |0\rangle, |\phi_{0,1}\rangle := |1\rangle, |\phi_{1,0}\rangle := |+\rangle$ and $|\phi_{1,1}\rangle := |-\rangle$. In this case, the states (as well as some protocol steps) are different but the angle between them is the same as with the states in $\mathcal{P}_2$. A fault-tolerant version of $\mathcal{P}_2$ has also been proposed in Ref. [BBBG09], which uses a generalized angle as in Eq. (7.5).

Protocol 7 ($\mathcal{P}_2$ (Aharonov's coin flipping))**.** The protocol uses all four possible states from Eq. (7.109).

---

1. Alice selects two random bits $a \xleftarrow{\$} \{0,1\}$ and $x \xleftarrow{\$} \{0,1\}$.
2. Alice sends one of the states, $|\phi_{x,a}\rangle$ to Bob.
3. Bob selects his random bit $b \xleftarrow{\$} \{0,1\}$ and sends to Alice.
4. One of two following things happens:
   (a) (either) Alice will send the bits x and a to Bob, who measures the qubit on a suitable basis to check if Alice was honest.
   (b) (or) Bob is asked to return the qubit $|\phi_{x,a}\rangle$ to Alice, who measures it and verifies if it is correct.
5. If no party declares cheating, the final output bit, will be $c = a \oplus b$.

Now, we can discuss the cheating strategies of each of the players. Examples of the cheating strategies for Alice include incorrect preparation of $|\phi_{x,a}\rangle$ and giving Bob the wrong information about (x, a), or Bob trying to determine the bits x, a from $|\phi_{x,a}\rangle$ before Alice has revealed them classically. We again focus only on Bob's strategies here to use cloning arguments. We note that the information-theoretic achievable bias of $\varepsilon = 0.42$ proven in Ref. [ATSVY00] applies only to Alice's strategy since she has greater control of the protocol (she prepares the original state). In general, with a cloning based attack strategy, Bob will be able to achieve a lower bias, as we show next. As mentioned above, Bob randomly selects his own bit b and sends it to Alice. He then builds a QCM to clone all 4 states in Eq. (7.109).

We next sketch the two cloning attacks on Bob's side of $\mathcal{P}_2$. Again, as with the protocol, $\mathcal{P}_1$, Bob can cheat using as much information as he can gain about a and again, once Bob has performed the cloning, his strategy boils down to the problem of state discrimination. In both attacks, Bob will use a state-dependent cloning machine.

In the first attack model (which we denote `I` - see Fig. 7.6(a) in Section 7.4.2.2 where we introduce the variational cloning version of the attack) Bob measures *all* the qubits outputted from the cloner to guess (x, a). As such, it is the *global* fidelity that will be the relevant quantity. This strategy would be useful in the

first possible challenge in the protocol, where Bob is not required to send anything back to Alice. We will discuss how using cloning in this type of attack can also reduce practical resources for Bob from a general POVM to projective measurements, which may be of independent interest. The main attack here boils down to Bob measuring the global output state from his QCM using the projectors $\{|v\rangle\langle v|, |v^{\perp}\rangle\langle v^{\perp}|\}$, and from this measurement, determines a. These projectors are constructed explicitly relative to the input states using the Neumark theorem [BK15] (see Section 2.2).

The second attack model (which we denote `II` - see Fig. 7.6(a) in Section 7.4.2.2) is instead a *local* attack and as such will depend on the optimal local fidelity. It may also be more relevant in the scenario where Bob is required to return a quantum state to Alice. We note that Bob could also apply a global attack in this scenario but we do not consider this possibility here to give more interesting and distinct examples. In what follows we explain the attacks in detail. For simplicity, we compute a bias assuming he does not return a state to Alice thus the bias will be equivalent to his discrimination probability. The analysis could be tweaked to take a detection probability for Alice into account as well. In this scenario, Bob again applies the QCM, but now he only uses one of the clones to perform state discrimination (given by the *Discriminator* in Fig. 7.6(a)).

Attack `I` on $\mathcal{P}_2$:
We note that attack `I`, is a 4 state *global* attack on $\mathcal{P}_2$ and that this attack model (*i.e.* based on cloning) can be considered a constructive way of implementing the optimal discrimination strategy of the states Alice is to send. To bias the bit, Bob needs to discriminate between the four pure states in Eq. (7.5) or equivalently between the ensembles of states encoding $a = \{0, 1\}$, where the optimal discrimination is done via a set of POVM measurements.

However, by implementing a cloning based attack, we can simplify the implementation of optimal discrimination strategies. This is because the symmetric state-dependent cloner (which is a unitary) has the interesting feature that for either case ($a = 0$ or $a = 1$), the cloner's output is a pure state in the 2-qubit Hilbert space. As such, the states (after going through the QCM) can be optimally discriminated via a set of projective measurements $\{P_v, P_{v^{\perp}}\}$, rather than general POVMs. This may not seem very important at this stage, but later we will see that it will relate a theoretical bound to an implementational attack strategy. Especially when we introduce the variational cloner that can learn to optimally clone these states efficiently, which, in turn, will assist the benchmarking of existing protocols. Let us now establish this bound and prove it.

Theorem 51. *[Ideal Cloning Attack (`I`) Bias on $\mathcal{P}_2$] Using a cloning attack on the protocol, $\mathcal{P}_2$, (in attack model I) Bob can achieve a bias:*

$$\varepsilon^{\mathsf{I}}_{\mathcal{P}_2,\text{ideal}} \approx 0.35 \tag{7.20}$$

Proof. The attack involves the global output state of the cloning machine. For

this attack we can use the fixed overlap $1 \to 2$ cloner with the global fidelity given by Eq. (2.80):

$$F_{\mathrm{G}}^{\mathrm{FO,opt}}(1,2) = \frac{1}{2}\left(1 + s^3 + \sqrt{1-s^2}\sqrt{1-s^4}\right) \approx 0.983 \tag{7.21}$$

where $s = \sin(2\phi) = \cos(\frac{\pi}{4})$ for $\mathcal{P}_2$. Also alternatively we can use the 4-state cloner which clones the two states with a fixed overlap plus their orthogonal set. For both of these cloners, we are interested in the global state of the cloner which we denote as $|\psi_{x,a}^{1\to 2}\rangle$ for an input state $|\phi_{x,a}\rangle$.

In order for Bob to guess a he must discriminate between $|\phi_{0,0}\rangle$ (encoding $a=0$) and $|\phi_{1,1}\rangle$ (encoding $a=1$) or alternatively the pair of states $\{|\phi_{0,1}\rangle, |\phi_{1,0}\rangle\}$. This is due to the pairs $\{|\phi_{0,0}\rangle, |\phi_{0,1}\rangle\}$ being orthogonal and $\{|\phi_{1,0}\rangle, |\phi_{1,1}\rangle\}$ both encode $a=0$, so the only choice is to discriminate between $|\phi_{0,0}\rangle$ and $|\phi_{1,1}\rangle$. Due to the symmetry and without an ancilla, the cloner preserves the overlap between each pairs *i.e.* $\langle\psi_{0,0}^{1\to 2}|\psi_{1,1}^{1\to 2}\rangle = \langle\phi_{0,0}|\phi_{1,1}\rangle = s$ (we also have $\langle\psi_{0,1}^{1\to 2}|\psi_{1,0}^{1\to 2}\rangle = s$).

Now we select the projective measurements $P_v = |v\rangle\langle v|$ and $P_{v^\perp} = |v^\perp\rangle\langle v^\perp|$ such that $\langle v|v^\perp\rangle = 0$. One can show that the discrimination probability is optimal when $|v\rangle$ and $|v^\perp\rangle$ are symmetric with respect to the target states according to the Neumark theorem. We have that $\langle v|v^\perp\rangle = 0$ so $2\theta + 2\phi = \frac{\pi}{2} \Rightarrow \theta = \frac{\pi}{4} - \phi$. Finally, writing the cloner's states for $\{|\psi_{0,0}^{1\to 2}\rangle, |\psi_{1,1}^{1\to 2}\rangle\}$ in the basis $\{|v\rangle, |v^\perp\rangle\}$ gives:

$$\begin{aligned} |\psi_{0,0}^{1\to 2}\rangle &= \cos\left(\frac{\pi}{4} - \phi\right)|v\rangle + \sin\left(\frac{\pi}{4} - \phi\right)|v^\perp\rangle, \\ |\psi_{1,1}^{1\to 2}\rangle &= \cos\left(\frac{\pi}{4} - \phi\right)|v\rangle - \sin\left(\frac{\pi}{4} - \phi\right)|v^\perp\rangle \end{aligned} \tag{7.22}$$

where it can be checked that $\langle\psi_{0,0}^{1\to 2}|\psi_{1,1}^{1\to 2}\rangle = \cos\left(\frac{\pi}{2} - 2\phi\right) = \sin(2\phi) = s$. Hence $|v\rangle$ and $|v^\perp\rangle$ can be explicitly derived. Note that these bases are also symmetric with respect to the other pair i.e $\{|\psi_{0,1}^{1\to 2}\rangle, |\psi_{1,0}^{1\to 2}\rangle\}$. Finally, the success probability of this measurement is then given by:

$$P_{\mathrm{disc},\mathcal{P}_2}^{\mathrm{opt,I}} = \frac{1}{2} + \frac{1}{2}\langle\psi_{0,0}^{1\to 2}|\psi_{1,1}^{1\to 2}\rangle = \frac{1}{2} + \frac{1}{2}\sin 2\phi = 0.853 \tag{7.23}$$

which is the maximum cheating probability for Bob. From this, we derive the bias as:

$$\varepsilon_{\mathcal{P}_2,\mathrm{ideal}}^{\mathrm{I}} = P_{\mathrm{disc},\mathcal{P}_2}^{\mathrm{opt,I}} - \frac{1}{2} = 0.353 \tag{7.24}$$

which completes the proof. □

Attack `II` on $\mathcal{P}_2$:

Finally, we consider a second attack model (attack `II`) on the protocol, $\mathcal{P}_2$, which is in the form of a 'local' attack. Here, we further consider two scenarios:

1. A cloning machine which is able to clone *all* 4 states $|\phi_{0,0}\rangle, |\phi_{1,1}\rangle$ *and* $|\phi_{0,1}\rangle, |\phi_{1,0}\rangle$,

2. A cloning machine tailored for only the two states, $|\phi_{0,0}\rangle$ and $|\phi_{1,1}\rangle$ (which Bob needs to discriminate between).

We focus on the former scenario since it connects in a more clear way to the VarQlone clone fidelities, while the second scenario enables a stronger attack (in the ideal scenario).

Scenario 1:

In this case, we can compute an exact discrimination probability, but it will result in a non-optimal attack (smaller success probability compared to the second one).

Theorem 52. *[Ideal Cloning Attack (II) Bias on $\mathcal{P}_2$ in Scenario 1.] Using a cloning attack on the protocol $\mathcal{P}_2$, (in attack model II with 4-states) Bob can achieve the following bias:*

$$\varepsilon^{\mathrm{II}}_{\mathcal{P}_2,\mathrm{ideal}} = 0.25 \tag{7.25}$$

Proof. Considering the 4 states to be in the $\mathsf{X}-\mathsf{Z}$ plane of the Bloch sphere, the density matrices of each state can be represented as:

$$\rho_{ij} = \frac{1}{2}(\mathbb{1} + m^x_{ij}\sigma_x + m^z_{ij}\sigma_z) \tag{7.26}$$

where σ_x and σ_z are Pauli matrices and m^x_{ij} and m^z_{ij} are 3 dimensional vectors given by:

$$\begin{aligned} m_{00} &:= [\sin(2\phi), 0, \cos(2\phi)] \\ m_{01} &:= [-\sin(2\phi), 0, -\cos(2\phi)] \\ m_{10} &:= [-\sin(2\phi), 0, \cos(2\phi)] \\ m_{11} &:= [\sin(2\phi), 0, -\cos(2\phi)] \end{aligned} \tag{7.27}$$

After the cloning (in the ideal case), the density matrix of each clone will become:

$$\rho^c_{ij} = \frac{1}{2}(\mathbb{1} + \eta_x m^x_{ij}\sigma_x + \eta_z m^z_{ij}\sigma_z) \tag{7.28}$$

where η_x and η_z are the shrinking factors in each direction given as follows:

$$\eta_x = \sin^2(2\phi)\sqrt{\frac{1}{\sin^4(2\phi)+\cos^4(2\phi)}}, \qquad \eta_z = \cos^2(2\phi)\sqrt{\frac{1}{\sin^4(2\phi)+\cos^4(2\phi)}} \tag{7.29}$$

For the states used in $\mathcal{P}_2$, we have $\phi = \frac{\pi}{8}$ and hence $\eta_x = \eta_z := \eta = \frac{1}{\sqrt{2}}$. Again, we can reduce the problem to the discrimination probability between the

two ensembles encoding $a=0$ and $a=1$ in Eq. (7.32). Let us define ρ^c to be the output clone that Bob chooses to use ($c \in \{1,2\}$). We have:

$$\begin{aligned} P^{\text{opt,II}}_{\text{disc},\mathcal{P}_2} &= \frac{1}{2}+\frac{1}{4}\left|\left|\rho_{(a=0)}-\rho_{(a=1)}\right|\right|_{\text{Tr}} \\ &= \frac{1}{2}+\frac{1}{4}\left|\left|\frac{1}{2}\left[(\rho^c_{00}-\rho^c_{11})+(\rho^c_{10}-\rho^c_{01}\right]\right|\right|_{\text{Tr}} \\ &= \frac{1}{2}+\frac{1}{4}\left|\left|\frac{\eta}{4}((m^x_{00}-m^x_{11}+m^x_{10}-m^x_{01})\sigma_x+(m^z_{00}-m^z_{11}+m^z_{10}-m^z_{01})\sigma_z\right|\right|_{\text{Tr}} \\ &= \frac{1}{2}+\frac{\eta\cos(2\phi)}{4}\left|\left|\sigma_z\right|\right|_{\text{Tr}} \\ &= \frac{1}{2}+\frac{\eta\cos(2\phi)}{2}=\frac{3}{4} \end{aligned} \tag{7.30}$$

Computing the bias in the same way as above completes the proof. □

Scenario 2:

Here, we give a bound on the success probabilities of Bob in terms of the local fidelities of the QCM where the cloning machine is only tailored to clone two fixed-overlap states. We rely on the fact that Bob can discriminate between the two ensembles of states (for $a=0$, $a=1$) with equal probabilities.

Theorem 53. *The optimal discrimination probability for a cloning attack on the protocol P_2, (in attack model II, with 2 states) is:*

$$0.619 \leqslant P^{\text{opt,II}}_{\text{disc},\mathcal{P}_2} \leqslant 0.823 \tag{7.31}$$

Proof. For each of the input states, $|\phi_{i,j}\rangle$ in Eq. (7.109), we denote ρ^c_{ij} to be a clone outputted from the QCM. Due to symmetry, we only need to consider one of the two output clones. We can now write the effective states for each encoding ($a=0, a=1$) as:

$$\rho_{(a=0)} := \frac{1}{2}(\rho^c_{00}+\rho^c_{10}), \qquad \rho_{(a=1)} := \frac{1}{2}(\rho^c_{01}+\rho^c_{11}) \tag{7.32}$$

Dealing with these two states is sufficient since it can be shown that discriminating between these two density matrices, is equivalent to discriminating between the entire set of 4 states in Eq. (7.5).

Again, we use the discrimination probability from the Holevo-Helstrom bound:

$$P^{\text{opt,II}}_{\text{disc},\mathcal{P}_2} := P^{\text{opt}}_{\text{disc}}(\rho_{(a=0)},\rho_{(a=1)}) := \frac{1}{2}+\frac{1}{2}d_{\text{Tr}}(\rho_{(a=0)},\rho_{(a=1)}) \tag{7.33}$$

Now, we have:

$$\begin{aligned} d_{\mathrm{Tr}}(\rho_{(a=0)},\rho_{(a=1)}) &= \frac{1}{2}\left|\left|\rho_{(a=0)}-\rho_{(a=1)}\right|\right|_{\mathrm{Tr}} \\ &= \frac{1}{2}\left|\left|\frac{1}{2}(\rho^c_{00}-\rho^c_{11})+\frac{1}{2}(\rho^c_{10}-\rho^c_{01})\right|\right|_{\mathrm{Tr}} \\ &\leqslant \frac{1}{4}||(\rho^c_{00}-\rho^c_{11})||_{\mathrm{Tr}}+||(\rho^c_{10}-\rho^c_{01})||_{\mathrm{Tr}} \\ &\leqslant \frac{1}{2}[d_{\mathrm{Tr}}(\rho^c_{00},\rho^c_{11})+d_{\mathrm{Tr}}(\rho^c_{01},\rho^c_{10})] \end{aligned} \tag{7.34}$$

$$\begin{aligned} \implies P^{\mathrm{opt}}_{\mathrm{disc}}(\rho_{(a=0)},\rho_{(a=1)}) &\leqslant \frac{1}{2}(P^{\mathrm{opt}}_{\mathrm{disc}}(\rho^c_{00},\rho^c_{11})+P^{\mathrm{opt}}_{\mathrm{disc}}(\rho^c_{01},\rho^c_{10})) \\ &= P^{\mathrm{opt}}_{\mathrm{disc}}(\rho^c_{00},\rho^c_{11}) \end{aligned}$$

The last equality follows since for both ensembles, $\{|\phi_{0,0}\rangle, |\phi_{1,1}\rangle\}$ and $\{|\phi_{0,1}\rangle, |\phi_{1,0}\rangle\}$, we have that their output clones having equal discrimination probability:

$$P^{\mathrm{opt}}_{\mathrm{disc}}(\rho^c_{00},\rho^c_{11}) = P^{\mathrm{opt}}_{\mathrm{disc}}(\rho^c_{01},\rho^c_{10}) \tag{7.35}$$

This is because the QCM is symmetric, and depends only on the overlap of the states (we have in both cases $\langle\phi_{00}|\phi_{11}\rangle = \langle\phi_{01}|\phi_{10}\rangle = \sin(2\phi)$).

Furthermore, since the cloning machine can only lower the discrimination probability between two states, we have:

$$P^{\mathrm{opt}}_{\mathrm{disc}}(\rho^c_{00},\rho^c_{11}) \leqslant P^{\mathrm{opt}}_{\mathrm{disc}}(\rho^c_{00},|\phi_{1,1}\rangle\langle\phi_{1,1}|) =: \overline{P^{\mathrm{opt}}_{\mathrm{disc}}} \tag{7.36}$$

Now, using the relationship between fidelity and the trace distance (Eq. (**??**)), we have the following bounds:

$$\frac{1}{2}+\frac{1}{2}\left(1-\sqrt{\langle\phi_{1,1}|\rho^c_{00}|\phi_{1,1}\rangle}\right) \leq \overline{P^{\mathrm{opt}}_{\mathrm{disc}}} \leq \frac{1}{2}+\frac{1}{2}\sqrt{1-\langle\phi_{1,1}|\rho^c_{00}|\phi_{1,1}\rangle} \tag{7.37}$$

By plugging this inequality in the observed density matrix for the output clone, we can find this discrimination probability.

As in the previous section, the output density matrix from the QCM for an output clone can be written as Eq. (7.14):

$$\rho^c_{00} = \alpha|\phi_{0,0}\rangle\langle\phi_{0,0}|+\beta|\phi_{1,1}\rangle\langle\phi_{1,1}|+\gamma(|\phi_{0,0}\rangle\langle\phi_{1,1}|+|\phi_{1,1}\rangle\langle\phi_{0,0}|) \tag{7.38}$$

Hence the output state has a local fidelity, $F_{\mathrm{L}} = \langle\phi_{0,0}|\rho^c_{00}|\phi_{0,0}\rangle = \alpha+s^2\beta+s\gamma$. On the other hand, we have $F(\rho^c_{00},|\phi_{1,1}\rangle\langle\phi_{1,1}|) = \langle\phi_{1,1}|\rho^c_{00}|\phi_{1,1}\rangle = s^2\alpha+\beta+s\gamma$. Combining these two, we then have:

$$F(\rho^c_{00},|\phi_{1,1}\rangle\langle\phi_{1,1}|) = F_{\mathrm{L}}+(s^2-1)(\alpha-\beta) \tag{7.39}$$

Plugging in F_{L} from Eq. (2.82), and $\alpha-\beta = \sqrt{\frac{1-s^2}{1-s^4}}$ (for an optimal state-dependent cloner), we get:

$$\frac{1}{2}+\frac{1}{2}\left[1-\sqrt{F_{\mathrm{L}}+(s^2-1)\sqrt{\frac{1-s^2}{1-s^4}}}\right] \leqslant P^{\mathrm{opt,II}}_{\mathrm{disc},\mathcal{P}_2} \leqslant \frac{1}{2}+\frac{1}{2}\sqrt{1-F_{\mathrm{L}}-(s^2-1)\sqrt{\frac{1-s^2}{1-s^4}}} \tag{7.40}$$

To complete the proof, we use $F_{\mathsf{L}} \approx 0.989$ and $s = 1/\sqrt{2}$ which gives the numerical discrimination probabilities above. □

7.3 Variational Quantum Cloner: specifications of the algorithm

Now we introduce our machine learning algorithm *Variational Quantum Cloner* or VarQlone that given a specific family of states, learns the circuit that optimally clones that family. We recall that our motivation is to find short-depth circuits to clone a given family of states, and also use this toolkit to investigate the family of states where the optimal figure of merit is unknown.

VarQlone is a variational quantum algorithm with similar core parts to other VQAs. We have given an overview of such techniques in the preliminaries (Section 2.6.4.1). In particular, this variational method uses a parameterised state, denoted by $\rho_{\boldsymbol{\theta}}$, typically prepared by some short-depth parameterised unitary on some initial state $\rho_{\boldsymbol{\theta}} := U(\boldsymbol{\theta}) |0\rangle \langle 0| U^{\dagger}(\boldsymbol{\theta})$. The parameters are then optimized by minimizing (or maximizing) a *cost function*, typically a function of k observable measurements on $\rho_{\boldsymbol{\theta}}$, O_k. This resembles a classical neural network, and indeed, techniques and ideas from classical machine learning can be borrowed and adapted to our setting. Nevertheless, we need to develop the core ingredients such as differentiable cost functions for gradient-descent based optimisation and theoretical guarantees on these cost functions specific to our problem. Additionally, for all the results we give here we use the gradient-descent-based optimizer (as discussed in Section 2.6.4.4) [KB17] with our cost functions.

We also note that other than the core theoretical subjects discussed in this section, the machine learning algorithm itself as well as the codes and simulations have not been developed and run by the author and therefore have been excluded from this thesis. In this section, we only present theoretical results where the author has contributed, such as defining suitable cost functions for the cloning problem and providing theoretical guarantees on them. For more details on the algorithm and related information, we refer the reader to the paper [CDKK22].

7.3.1 Cost functions

In this section, we propose several cost functions for our problems and discuss their advantages and differences. Primarily, we propose the so-called 'local' cost functions of the following functional form:

$$C_{\mathsf{loc}}^{M\to N}(\boldsymbol{\theta}) := \mathop{\mathbb{E}}_{|\psi\rangle\in S} f(\mathsf{O}_{\mathsf{L}}^{\psi}, \rho_{\boldsymbol{\theta}}, M, N) \tag{7.41}$$

Where $|\psi\rangle$ denotes target state of the cloning machine, S denotes the set of states to be cloned, and M and N are the number of copies for the input states and output clones respectively. Choosing f, or f_{sq} (denoting the squared cost

function) as follows:

$$f_{\text{sq}} := \sum_{i=1}^{N} (1 - F_{\text{L}}^{i}(\boldsymbol{\theta}))^2 + \sum_{i<j}^{N} (F_{\text{L}}^{i}(\boldsymbol{\theta}) - F_{\text{L}}^{j}(\boldsymbol{\theta}))^2 \tag{7.42}$$

results in what for brevity we refer to as the *squared* cost function, a generalization of the cost also proposed in [JJB+19]. Here $F_{\text{L}}^{j}(\boldsymbol{\theta}) := F_{\text{L}}(|\psi\rangle\langle\psi|, \rho_{\boldsymbol{\theta}}^{j})$ is the local fidelity of the parameterized state relative to output clone j. This is generated using the observable $\text{O}_{\text{sq}}^{\psi} = |\psi\rangle\langle\psi|$ for the specific instance of state to be cloned from the set, $|\psi\rangle \in S$. As such, we define $\text{C}_{\text{sq}}^{M\to N}(\boldsymbol{\theta}) := \mathbb{E}_{|\psi\rangle\in S}[f_{\text{sq}}]$.

Let us now elaborate a bit on the alternative choices of cost functions. The second local cost, which we call the *linear* local cost or 'local cost' again for brevity, is given by:

$$\text{C}_{\text{L}}^{M\to N}(\boldsymbol{\theta}) := \mathop{\mathbb{E}}_{|\psi\rangle\in S}\left[\text{C}_{\text{L}}^{\psi}(\boldsymbol{\theta})\right] := \mathop{\mathbb{E}}_{|\psi\rangle\in S}\left[\text{Tr}(\text{O}_{\text{L}}^{\psi}\rho_{\boldsymbol{\theta}})\right], \quad \text{O}_{\text{L}}^{\psi} := \mathbb{1} - \frac{1}{N}\sum_{j=1}^{N} |\psi\rangle\langle\psi|_j \otimes \mathbb{1}_{\bar{j}} \tag{7.43}$$

where $|\psi\rangle \in S$ is state to be cloned.
These first two cost functions, are related only in that they are both functions of *local* observables, or in other words, the local fidelities. The third cost, on the other hand, is fundamentally different compared to the other two proposals, since it captures the global fidelity *i.e.* uses global observables, and as such, we refer to it as the '*global cost*':

$$\text{C}_{\text{G}}^{M\to N}(\boldsymbol{\theta}) := \mathop{\mathbb{E}}_{|\psi\rangle\in S}\left[\text{Tr}(\text{O}_{\text{G}}^{\psi}\rho_{\boldsymbol{\theta}})\right], \qquad \text{O}_{\text{G}}^{\psi} := \mathbb{1} - |\psi\rangle\langle\psi|^{\otimes N} \tag{7.44}$$

The second local cost, and our global cost functions are adapted from the literature on variational algorithms [LTOJ+19, CSV+21, KLP+19, SKCC20]. For compactness, we will drop the superscript $M \to N$ when the meaning is clear from context.

Now, we motivate our choices for the above cost functions. For Eq. (7.41), if we restrict to the special case of $1 \to 2$ cloning (*i.e.* we have only two output parties, $j \in \{B, E\}$), and remove the expectation value over states, we recover the cost function used in Ref. [JJB+19]. A useful feature of this cost is that symmetry is explicitly enforced by the difference term $(F_i(\boldsymbol{\theta}) - F_j(\boldsymbol{\theta}))^2$.

In contrast, the local and global cost functions are inspired by other variational algorithm in the literature [LTOJ+19, CSV+21, KLP+19, SKCC20] where their properties have been extensively studied, particularly in relation to the phenomenon of 'barren plateaus' [MBS+18, CSV+21]. Since we have not covered the topic of barren plateaus in Chapter 2, we will give a brief description here. Barren plateaus is a phenomenon where the gredient-based optimisation in the quantum landscape ends up with no interesting search directions to go. It has been demonstrated that hardware efficient Ansätze are untrainable using a global cost function similar to the one given in Eq. (7.44), since they have exponentially vanishing gradients [SBG+19], often leading to a barren plateau . In contrast,

local cost functions (Eq. (7.43), Eq. (7.41)) are shown to be efficiently trainable with $\mathcal{O}(\log N)$ depth hardware efficient Ansätze [CSV+21] (see Section 2.6.4.3 for more details about different types of Ansätze).

We also remark that typically global cost functions are usually more favourable from the point of view of *operational meaning*. For example in variational compilation [KLP+19], this cost function compares the closeness of two global unitaries. In this respect, local cost functions are usually used as a proxy to optimize a global cost function.

In our case, the nature of quantum cloning allows VarQlone local cost functions to have immediate operational meaning, illustrated through the following example (using the local cost, Eq. (7.43)) for $1 \to 2$ cloning:

$$C_{\mathsf{L}}^{\psi}(\boldsymbol{\theta}) = \mathsf{Tr}\left[\left(\mathbb{1} - \frac{1}{2}\sum_{j=1}^{2} |\psi\rangle\langle\psi|_j \otimes \mathbb{1}_{\bar{j}}\right)\rho_{\boldsymbol{\theta}}\right]$$
$$\implies C_{\mathsf{L}}(\boldsymbol{\theta}) = 1 - \frac{1}{2}\mathbb{E}\left[F_{\mathsf{L}}\left(|\psi\rangle\langle\psi|, \rho_{\boldsymbol{\theta}}^1\right) + F_{\mathsf{L}}\left(|\psi\rangle\langle\psi|, \rho_{\boldsymbol{\theta}}^2\right)\right]$$

where $\mathbb{E}[F_{\mathsf{L}}]$ is the average fidelity [SIGA05] over the possible input states. The final expression of $C_{\mathsf{L}}(\boldsymbol{\theta})$ in the above equation follows from the expression of fidelity when one of the states is pure. Similarly, the global cost function relates to the global fidelity of the output state concerning the input state(s).

7.3.2 Cost function gradients

In the gradient-descent-based optimization approach which we use for developing our algorithm, we require efficient computation of the gradients. Here, we derive the analytic gradients for our cost functions. We use the local cost function, Eq. (7.41) as an explicit example and the derivations for the other cost functions follow straightforwardly. As a reminder, the squared cost is given by:

$$C_{\mathsf{sq}}^{M\to N}(\boldsymbol{\theta}) := \mathop{\mathbb{E}}_{|\psi\rangle\in S}\left[\sum_{i=1}^{N}(1 - F_{\mathsf{L}}^{i}(\boldsymbol{\theta}))^2 + \sum_{i<j}^{N}(F_{\mathsf{L}}^{i}(\boldsymbol{\theta}) - F_{\mathsf{L}}^{j}(\boldsymbol{\theta}))^2\right] \tag{7.45}$$

where the expectation is taken over the set of states with uniform distribution. For example, in the phase-covariant cloner of the states Eq. (7.1), the parameter η is sampled uniformly from the interval $[0, 2\pi)$.

Now, the derivative of Eq. (7.45), with respect to a single parameter, θ_l is given by:

$$\frac{\partial C_{\mathsf{sq}}(\boldsymbol{\theta})}{\partial\theta_l} = 2\mathop{\mathbb{E}}_{|\psi\rangle\in S}\left[\sum_{i=1}^{N}(1 - F_{\mathsf{L}}^{i}(\boldsymbol{\theta}))\left[-\frac{\partial F_{\mathsf{L}}^{i}(\boldsymbol{\theta})}{\partial\theta_l}\right] + \sum_{i<j}^{N}(F_{\mathsf{L}}^{i}(\boldsymbol{\theta}) - F_{\mathsf{L}}^{j}(\boldsymbol{\theta}))\left[\frac{\partial F_{\mathsf{L}}^{i}(\boldsymbol{\theta})}{\partial\theta_l} - \frac{\partial F_{\mathsf{L}}^{j}(\boldsymbol{\theta})}{\partial\theta_l}\right]\right] \tag{7.46}$$

We can rewrite the expression for the fidelity of the j^{th} clone as:

$$F_{\mathsf{L}}^{j}(\boldsymbol{\theta}) = \langle\psi|\rho_j(\boldsymbol{\theta})|\psi\rangle = \mathsf{Tr}\left[|\psi\rangle\langle\psi|\rho_j\right] = \mathsf{Tr}\left[|\psi\rangle\langle\psi|\,\mathsf{Tr}_{\bar{j}}\left(U(\boldsymbol{\theta})\rho_{\mathsf{init}}U(\boldsymbol{\theta})^{\dagger}\right)\right] \tag{7.47}$$

Using the linearity of the trace, the derivative of the fidelities with respect to the parameters, θ_l, can be computed as:

$$\frac{\partial F_{\mathrm{L}}^{j}(\boldsymbol{\theta})}{\partial \theta_l} = \mathrm{Tr}\left[|\psi\rangle\langle\psi| \,\mathrm{Tr}_{\bar{j}}\left(\frac{\partial U(\boldsymbol{\theta})\rho_{\mathrm{init}}U(\boldsymbol{\theta})^{\dagger}}{\partial \theta_l}\right)\right] \tag{7.48}$$

Using the *parameter shift rule* (Theorem 12 from Section 2.6.4.4) technique the explicit expression of the cost function's gradients can be calculated. The calculation has been given in Appendix A.6.

7.3.3 Cost function guarantees

One of the interesting problems in the area of theoretical machine learning is showing theoretical guarantees for the cost function, *i.e.* achieving the cost minimum indicates a solution to the problem in question [KLP+19, BPLC+20]. This property is known as *faithfulness*.

For our approximate quantum cloning problem, due to the information-theoretic limits, the above costs cannot have a minimum at 0, but instead at some finite positive value (say $\mathrm{C}_{\mathrm{L}}^{\mathrm{opt}}$ for the local cost).

Despite this, we can still derive certain theoretical guarantees about them. Specifically, we consider notions of *strong* and *weak* faithfulness, relative to the learner's error in our solution. Our goal is to provide statements about the *generalization performance* of the cost functions, by considering how close are the states we output by our cloning machine, to those which would be outputted from the '*optimal*' cloner, relative to some metrics. In the following, we denote $\rho_{\mathrm{opt}}^{\psi,j}$ ($\rho_{\boldsymbol{\theta}}^{\psi,j}$) to be the optimal (VarQlone learned) reduced state for qubit j, for a particular input state $|\psi\rangle$. If the superscript j is not present, we mean the global state of all clones. Let us give the definitions of faithfulness.

Definition 51 (Strong Faithfulness)**.** A cloning cost function, C, is strongly faithful if for all $|\psi\rangle \in S$, optimising the closeness in cost function implies the the optimally close states *i.e.*:

$$\mathrm{C}(\boldsymbol{\theta}) = \mathrm{C}^{\mathrm{opt}} \implies \rho_{\boldsymbol{\theta}}^{\psi} = \rho_{\mathrm{opt}}^{\psi} \qquad \forall |\psi\rangle \in S \tag{7.49}$$

where $\mathrm{C}^{\mathrm{opt}}$ is the minimum value achievable (allowed by quantum mechanics) for the cost C, and S is the given set of states to be cloned.

Definition 52 (ε-Weak Local Faithfulness)**.** A local cloning cost function, $\mathrm{C_L}$, is ε-weakly faithful if for all $|\psi\rangle \in S$ and for all the *local* clones, the closeness of local cost function to its optimal value implies the closeness of local clone states *i.e.*:

$$|\,\mathrm{C_L}(\boldsymbol{\theta}) - \mathrm{C_L^{opt}}\,| \leqslant \varepsilon \implies D(\rho_{\boldsymbol{\theta}}^{\psi,j}, \rho_{\mathrm{opt}}^{\psi,j}) \leqslant f(\varepsilon), \qquad \forall |\psi\rangle \in S, \forall j \tag{7.50}$$

where $D(\cdot,\cdot)$ is a chosen metric in the Hilbert space between the two states and f is a polynomial function of ε.

Definition 53 (ε-Weak Global Faithfulness)**.** A global cloning cost function, $\mathrm{C_G}$, is ε-weakly faithful if for all $|\psi\rangle \in S$ the closeness of global cost function to its optimal value implies the closeness of global optimal state *i.e.*:

$$|\,\mathrm{C_G}(\boldsymbol{\theta}) - \mathrm{C_G^{opt}}\,| \leqslant \varepsilon \implies D(\rho_{\boldsymbol{\theta}}^{\psi}, \rho_{\mathrm{opt}}^{\psi}) \leqslant f(\varepsilon) \;\; \forall |\psi\rangle \in S \tag{7.51}$$

One could also define local and global versions of the strong faithfulness, but this is less attractive as it is included in the other case. Thus we do not focus on it here. Let us begin by examining the squared local cost function. For this case, we will provide the most extensive analysis, and faithfulness proofs for the other cost functions can be derived using similar methods.

Squared Cost Function

First, we start with the squared cost function which we rewrite as:

$$\mathrm{C_{sq}^{M\to N}}(\boldsymbol{\theta}) = \frac{1}{\mathcal{N}} \int_S \left[\sum_{j=1}^{N} (1 - F_i(\boldsymbol{\theta}))^2 + \sum_{i<j}^{N} (F_i(\boldsymbol{\theta}) - F_j(\boldsymbol{\theta}))^2 \right] d\psi \tag{7.52}$$

where the expectation of a fidelity F_i over the states in distribution S is defined as $\mathbb{E}[F_i] = \frac{1}{\mathcal{N}} \int_S F_i \cdot d\psi$, with the normalisation condition being $\mathcal{N} = \int_S d\psi$. For qubit states, if the normalisation is over the entire Bloch sphere in $SU(2)$, then $\mathcal{N} = 4\pi$. For notation simplicity, we herein denote the $\mathrm{C_{sq}^{M\to N}}(\boldsymbol{\theta})$ as $\mathrm{C_{sq}}(\boldsymbol{\theta})$. We begin with a proof of the fact that the cost function is *strongly* faithful.

Theorem 54. *[Strong faithfulness of the squared cost function] The squared local cost function is locally strongly faithful,* i.e.*:*

$$C_{sq}(\boldsymbol{\theta}) = C_{sq}^{\mathrm{opt}} \implies \rho_{\boldsymbol{\theta}}^{\psi,j} = \rho_{\mathrm{opt}}^{\psi,j} \qquad \forall |\psi\rangle \in S, \forall j \in [N] \tag{7.53}$$

Proof. The cost function $\mathrm{C_{sq}}(\boldsymbol{\theta})$ achieves a minimum at the joint maximum of $\mathbb{E}[F_i(\boldsymbol{\theta})]$ for all $i \in [N]$. In symmetric $M \to N$ cloning, the expectation value of all the N output fidelities peak at $F_i = F_{\mathrm{opt}}$ for all input states $|\psi\rangle$. This corresponds to a unique optimal joint state $\rho_{\mathrm{opt}}^{\psi,j} = U_{\mathrm{opt}} |\psi^{\otimes M}, 0^{\otimes N-M}\rangle \langle \psi^{\otimes M}, 0^{\otimes N-M}| U_{\mathrm{opt}}^{\dagger}$ for

each $|\psi\rangle \in S$ and for any $j \in [N]$, where U_{opt} is the unitary producing the the optimal state. Since the joint optimal state and the corresponding fidelities are unique for all input states in the distribution, we conclude that the cost function achieves a minimum under precisely the unique condition *i.e.* $\mathbb{E}[F_j(\boldsymbol{\theta})] = F_{\text{opt}}$ for all $j \in [N]$. This condition implies that,

$$\rho_{\boldsymbol{\theta}}^{\psi,j} = \rho_{\text{opt}}^{\psi,j}, \qquad \forall |\psi\rangle \in S, \forall j \in [N] \tag{7.54}$$

We note that since F_{opt} is the same for all the reduced states $j \in [N]$, this implies that the optimal reduced states are all the same for a given $|\psi\rangle \in S$. Thus Eq. (7.54) provides the necessary guarantee that minimizing the cost function over the parameter space, results in the corresponding circuit's output, being equal to the optimal cloned state for all the inputs. □

Now, we take the weaker notion of faithfulness into account. Computing the exact fidelities of the output states requires an infinite number of copies. In reality, we run the iteration only a finite number of times and thus, our cost function can only reach the optimal cost up to some precision. This is also relevant when running the circuit on devices in the NISQ era which would inherently introduce noise in the system. Thus, we can only hope to minimise the cost function up to some precision of the optimal cost. This statement can be formalised via the following lemma:

Lemma 6. *Suppose the cost function is ε-close to the optimal cost in symmetric cloning*

$$C_{\text{sq}}(\boldsymbol{\theta}) - C_{\text{sq}}^{\text{opt}} \leqslant \varepsilon \tag{7.55}$$

Then we have,

$$\text{Tr}\left[(\rho_{\text{opt}}^{\psi,j} - \rho_{\boldsymbol{\theta}}^{\psi,j})\,|\psi\rangle\langle\psi|\right] \leqslant \frac{\mathcal{N}\varepsilon}{2(1-F_{\text{opt}})}, \qquad \forall |\psi\rangle \in S, \forall j \in [N] \tag{7.56}$$

Proof. In $M \to N$ symmetric cloning, the optimal cost function value is achieved when each output clone achieves the fidelity F_{opt}. Thus, using Eq. (7.41) (or Eq. (7.45)), the optimal cost function value is given by,

$$\mathrm{C}_{\text{sq}}^{\text{opt}} = N \cdot (1 - F_{\text{opt}})^2 \tag{7.57}$$

The optimal cost function is achieved when all output clones have the same fidelity. Therefore, as we begin to minimize the cost $\mathrm{C}_{\text{sq}}(\boldsymbol{\theta})$, all the output clones start to produce states with approximately the same fidelity. This is explicitly enforced by taking the limit $\varepsilon \to 0$, in which case the difference terms of Eq. (7.45) vanish. Thus, the cost function explicitly enforces the symmetry property. Let us assume

$\varepsilon \to 0$, and consider the quantity $\mathsf{C}_{\mathsf{sq}}(\boldsymbol{\theta}) - \mathsf{C}_{\mathsf{sq}}^{\mathsf{opt}}$:

$$\begin{aligned}
\mathsf{C}_{\mathsf{sq}}(\boldsymbol{\theta}) - \mathsf{C}_{\mathsf{sq}}^{\mathsf{opt}} &= \frac{1}{\mathcal{N}} \int_S \left[\sum_i^N (1 - F_i(\boldsymbol{\theta}))^2 + \sum_{i<j}^N (F_i(\boldsymbol{\theta}) - F_j(\boldsymbol{\theta}))^2 \right] d\psi - N \cdot (1 - F_{\mathsf{opt}})^2 \\
&\underset{\varepsilon \to 0}{\approx} \frac{1}{\mathcal{N}} \int_S \left[\sum_j^N (1 - F_j(\boldsymbol{\theta}))^2 - N \cdot (1 - F_{\mathsf{opt}})^2 \right] d\psi \\
&\approx \frac{1}{\mathcal{N}} \int_S \left[\sum_j^N (F_{\mathsf{opt}} - F_j(\boldsymbol{\theta}))(2 - F_{\mathsf{opt}} - F_j(\boldsymbol{\theta})) \right] d\psi \\
&\geqslant \frac{2(1 - F_{\mathsf{opt}})}{\mathcal{N}} \int_S \left[\sum_j^N (F_{\mathsf{opt}} - F_j(\boldsymbol{\theta})) \right] d\psi \\
&= \frac{2(1 - F_{\mathsf{opt}})}{\mathcal{N}} \left[\sum_j^N \int_S \mathsf{Tr}[(\rho_{\mathsf{opt}}^{\psi,j} - \rho_{\boldsymbol{\theta}}^{\psi,j}) |\psi\rangle \langle\psi|] d\psi \right]
\end{aligned} \tag{7.58}$$

The second line follows since F_{opt} is the same for each input state $|\psi\rangle$. Utilizing the inequality in Eq. (7.55) and Eq. (7.58), we obtain,

$$\begin{aligned}
&\sum_j^N \int_S \mathsf{Tr}\left[(\rho_{\mathsf{opt}}^{\psi,j} - \rho_{\boldsymbol{\theta}}^{\psi,j}) |\psi\rangle \langle\psi| \right] d\psi \leqslant \frac{\mathcal{N}\varepsilon}{2(1 - F_{\mathsf{opt}})} \\
&\implies \mathsf{Tr}\left[(\rho_{\mathsf{opt}}^{\psi,j} - \rho_{\boldsymbol{\theta}}^{\psi,j}) |\psi\rangle \langle\psi| \right] \leqslant \frac{\mathcal{N}\varepsilon}{2(1 - F_{\mathsf{opt}})}, \quad \forall |\psi\rangle \in S, \forall j \in [N]
\end{aligned} \tag{7.59}$$

This concludes the proof. □

The above inequality allows us to quantify the closeness of the state produced by VarQlone and the unique optimal clone for any $|\psi\rangle \in S$. We quantify this closeness of the states in a popular distance measure in quantum information, namely the Fubini-Study (or Bures angle) distance between two quantum states (introduced in Section 2.1.4). Using the above lemma, we can prove the following two theorems for the squared local cost function:

Theorem 55. *[Weak faithfulness of the local squared const function] The squared cost function as defined in Eq. (7.41), is ε-weakly faithful with respect to the Bures angle Θ_{BA} (or alternatively Fubini-distance measure d_{FS}). In other words, if the squared cost function is ε-close to its minimum,* i.e.*:*

$$C_{\text{sq}}(\boldsymbol{\theta}) - C_{\text{sq}}^{\text{opt}} \leqslant \varepsilon \tag{7.60}$$

where $C_{\text{sq}}^{\text{opt}} := \min_{\boldsymbol{\theta}} \sum_{i}^{N}(1-F_i(\boldsymbol{\theta}))^2 + \sum_{i<j}^{N}(F_i(\boldsymbol{\theta})-F_j(\boldsymbol{\theta}))^2 = N(1-F_{\text{opt}})^2$ is the optimal theoretical cost using fidelities produced by the ideal symmetric *cloning machine, then the following holds:*

$$\Theta_{BA}(\rho_{\boldsymbol{\theta}}^{\psi,j}, \rho_{\text{opt}}^{\psi,j}) \leqslant \frac{\mathcal{N}}{2(1-F_{\text{opt}})\sin(F_{\text{opt}})} \cdot \varepsilon := f_1(\varepsilon), \qquad \forall |\psi\rangle \in S, \forall j \in [N] \tag{7.61}$$

Proof. To prove this theorem, we revisit and rewrite the Bures angle from Eq. (2.23) (see Section 2.1.4):

$$\Theta_{\text{BA}}(\rho,\sigma) = \arccos\sqrt{F(\rho,\sigma)} = \arccos\langle\phi|\tau\rangle \tag{7.62}$$

where $|\phi\rangle$ and $|\tau\rangle$ are the purifications of ρ and σ respectively which maximize the overlap. We note that $\Theta_{\text{BA}}(\rho,\sigma)$ lies in the interval $[0,\pi/2]$, with the value $\pi/2$ corresponding to the unique solution $\rho=\sigma$. Since this distance is a metric, it obeys the triangle's inequality, *i.e.*, for any three states ρ,σ and δ,

$$\Theta_{\text{BA}}(\rho,\sigma) \leqslant \Theta_{\text{BA}}(\rho,\delta) + \Theta_{\text{BA}}(\sigma,\delta) \tag{7.63}$$

Rewriting the result of Lemma 6 in terms of fidelity for each $|\psi\rangle \in S$ and correspondingly in terms of Bures distance using Eq. (7.62) is,

$$\begin{aligned} &F(\rho_{\text{opt}}^{\psi,j}, |\psi\rangle) - F(\rho_{\boldsymbol{\theta}}^{\psi,j}, |\psi\rangle) \leqslant \varepsilon' \\ &\implies \cos^2(\Theta_{\text{BA}}(\rho_{\text{opt}}^{\psi,j}, |\psi\rangle)) - \cos^2(\Theta_{\text{BA}}(\rho_{\boldsymbol{\theta}}^{\psi,j}, |\psi\rangle)) \leqslant \varepsilon' \end{aligned} \tag{7.64}$$

where $\varepsilon' = \mathcal{N}\varepsilon/2(1-F_{\text{opt}})$. Let us denote $D_{\pm}^{\psi} = \Theta_{\text{BA}}(\rho_{\text{opt}}^{\psi,j}, |\psi\rangle) \pm \Theta_{\text{BA}}(\rho_{\boldsymbol{\theta}}^{\psi,j}, |\psi\rangle)$ This inequality in Eq. (7.64) can be further rewritten as,

$$\begin{aligned} \cos(\Theta_{\text{BA}}(\rho_{\text{opt}}^{\psi,j}, |\psi\rangle)) - \cos(\Theta_{\text{BA}}(\rho_{\boldsymbol{\theta}}^{\psi,j}, |\psi\rangle)) &\leqslant \frac{\varepsilon'}{\cos(\Theta_{\text{BA}}(\rho_{\text{opt}}^{\psi,j}, |\psi\rangle)) + \cos(\Theta_{\text{BA}}(\rho_{\boldsymbol{\theta}}^{\psi,j}, |\psi\rangle))} \\ \cos(\Theta_{\text{BA}}(\rho_{\text{opt}}^{\psi,j}, |\psi\rangle)) - \cos(\Theta_{\text{BA}}(\rho_{\boldsymbol{\theta}}^{\psi,j}, |\psi\rangle)) &\lessapprox \frac{\varepsilon'}{2\cos(\Theta_{\text{BA}}(\rho_{\text{opt}}^{\psi,j}, |\psi\rangle))} \\ 2\sin\left(\frac{D_{+}^{\psi}}{2}\right)\sin\left(\frac{\mathrm{D}_{-}^{\psi}}{2}\right) &\leqslant \frac{\varepsilon'}{2\cos(\Theta_{\text{BA}}(\rho_{\text{opt}}^{\psi,j}, |\psi\rangle))} \\ \implies \mathrm{D}_{-}^{\psi} &\leqslant \frac{\varepsilon'}{\sin(\Theta_{\text{BA}}(\rho_{\text{opt}}^{\psi,j}, |\psi\rangle))} = \frac{\mathcal{N}\varepsilon}{2(1-F_{\text{opt}})\sin(F_{\text{opt}})} \end{aligned} \tag{7.65}$$

where we have used the approximations that in the limit $\varepsilon \to 0$, $\Theta_{\text{BA}}(\rho_{\text{opt}}^{\psi,j}, |\psi\rangle) \approx \Theta_{\text{BA}}(\rho_{\boldsymbol{\theta}}^{\psi,j}, |\psi\rangle)$ and the trigonometric identities $\cos(x-y) = 2\sin\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right)$, and $\sin 2x = 2\sin x \cos x$.

Further, using the Fubini-Study metric triangle's inequality on the set of states $\{\rho_{\text{opt}}^{\psi,j}, \rho_{\boldsymbol{\theta}}^{\psi,j}, |\psi\rangle\}$ results in,

$$\Theta_{\text{BA}}(\rho_{\boldsymbol{\theta}}^{\psi,j}, |\psi\rangle) \leqslant \Theta_{\text{BA}}(\rho_{\text{opt}}^{\psi,j}, |\psi\rangle) + \Theta_{\text{BA}}(\rho_{\boldsymbol{\theta}}^{\psi,j}, \rho_{\text{opt}}^{\psi,j}) \tag{7.66}$$

Combining the above inequality and Eq. (7.65) results in,

$$\Theta_{\text{BA}}(\rho_{\boldsymbol{\theta}}^{\psi,j}, \rho_{\text{opt}}^{\psi,j}) \leqslant \frac{\mathcal{N}}{2(1-F_{\text{opt}})\sin(F_{\text{opt}})} \cdot \varepsilon, \quad \forall |\psi\rangle \in S \tag{7.67}$$

This bounds the closeness of the trained output state and the optimal output state as a function of ε. □

A similar result can be derived relative to trace distance instead of the Bures/Fubini-Study distance. However, we avoid presenting the result here, since it is very similar in nature. Instead, we refer the reader to [CDKK22] for the faithfulness result using the trace distance.

Local Cost Function

Next, we prove analogous results for the local cost function, defined for $M \to N$ cloning. We rewrite the cost function with an average integral form over the set S:

$$\mathrm{C}_{\mathrm{L}}(\boldsymbol{\theta}) := \mathbb{E}\left[1 - \frac{1}{N}\left(\sum_{j=1}^{N} F_j(\boldsymbol{\theta})\right)\right] = 1 - \frac{1}{N\mathcal{N}}\int_S \sum_{j=1}^{N} F_j(\boldsymbol{\theta}) d\psi \tag{7.68}$$

where $\mathcal{N} = \int_S d\psi$ is the normalisation condition. As above, we can show this cost function also exhibits strong faithfulness:

Theorem 56 (Strong faithfulness of the local cost function)**.** *The local squared cost function is locally strongly faithful:*

$$C_L(\boldsymbol{\theta}) = C_L^{\text{opt}} \implies \rho_{\boldsymbol{\theta}}^{\psi,j} = \rho_{\text{opt}}^{\psi,j} \qquad \forall |\psi\rangle \in S, \forall j \in [N] \tag{7.69}$$

Proof. Similar to the faithfulness arguments of the squared cost function, one can immediately see that the cost function $\mathrm{C}_{\mathrm{L}}(\boldsymbol{\theta})$ achieves a unique minimum at the joint maximum of $\mathbb{E}[F_j(\boldsymbol{\theta})]$ for all $j \in [N]$. Thus, the minimum of $\mathrm{C}_{\mathrm{L}}(\boldsymbol{\theta})$ corresponds to the unique optimal joint state with its unique local reduced states $\rho_{\text{opt}}^{\psi,j}$ for each $j \in [N]$, and for each input state $|\psi\rangle \in S$. Thus the cost function achieves a minimum under precisely the unique condition *i.e.* the output state is equal to the optimal clone state. □

Now, we can also prove analogous versions of weak faithfulness. Many of the steps in the proof follow similarly to the squared cost derivations above, so we omit them for brevity where possible. As above, we first have the following lemma:

Lemma 7. *Suppose the cost function is ε-close to the optimal cost in symmetric cloning*

$$C_L(\boldsymbol{\theta}) - C_L^{\text{opt}} \leqslant \varepsilon \tag{7.70}$$

where we assume $\lim_{\varepsilon\to 0} |\mathbb{E}[F_i(\boldsymbol{\theta})] - \mathbb{E}[F_j(\boldsymbol{\theta})]| \to 0, \forall i,j$, *and therefore* $C_{\text{opt}} := 1 - F_{\text{opt}}$. *Then,*

$$\text{Tr}[(\rho_{\text{opt}}^{\psi,j} - \rho_{\boldsymbol{\theta}}^{\psi,j})\,|\psi\rangle\langle\psi|] \leqslant \mathcal{N}\varepsilon, \qquad \forall\,|\psi\rangle \in S, \forall j \in [N] \tag{7.71}$$

The proof of Lemma 7 follows almost identically to Lemma 6, but with the exception that we can write $\mathrm{C_L}(\boldsymbol{\theta}) - \mathrm{C_L^{opt}} = \mathbb{E}\,(F_{\text{opt}} - F(\boldsymbol{\theta}))$ in the symmetric case, assuming $F_i(\boldsymbol{\theta}) \approx F_j(\boldsymbol{\theta})$, $\forall i \neq j \in [N]$. Thus we skip the proof and we show the weak faithfulness in the following theorem:

Theorem 57. *The local cost function, Eq. (7.43), is ε-weakly faithful with respect to Θ_{BA}*

$$C_L(\boldsymbol{\theta}) - C_L^{\text{opt}} \leqslant \varepsilon \tag{7.72}$$

Then the following holds:

$$\Theta_{BA}(\rho_{\boldsymbol{\theta}}^{\psi,j}, \rho_{\text{opt}}^{\psi,j}) \leqslant \frac{\mathcal{N}\varepsilon}{\sin(F_{\text{opt}})} =: f_2(\varepsilon), \quad \forall\,|\psi\rangle \in S, \forall j \in [N] \tag{7.73}$$

where $C_L^{\text{opt}} := 1 - F_{\text{opt}}$

Proof. We rewrite the Eq. (7.71) in terms of the Bures angle,

$$\begin{aligned} &F(\rho_{\text{opt}}^{\psi,j}, |\psi\rangle) - F(\rho_{\boldsymbol{\theta}}^{\psi,j}, |\psi\rangle) \leqslant \mathcal{N}\varepsilon \\ &\implies \cos^2(\Theta_{\text{BA}}(\rho_{\text{opt}}^{\psi,j}, |\psi\rangle)) - \cos^2(\Theta_{\text{BA}}(\rho_{\boldsymbol{\theta}}^{\psi,j}, |\psi\rangle)) \leqslant \mathcal{N}\varepsilon \end{aligned} \tag{7.74}$$

Following the derivation in the squared cost function section, we obtain the Bures angle/Fubini-Study closeness as,

$$\Theta_{\text{BA}}(\rho_{\boldsymbol{\theta}}^{\psi,j}, \rho_{\text{opt}}^{\psi,j}) \leqslant \frac{\mathcal{N}\varepsilon}{\sin(F_{\text{opt}})}, \qquad \forall\,|\psi\rangle \in S, \forall j \in [N] \tag{7.75}$$

This concludes the proof. □

Global Cost Function

Finally, we show in the next theorems that the global cost function exhibits similar notions of faithfulness:

Theorem 58. *[Strong faithfulness of the global cost function] The global cost function is globally strongly faithful, meaning the following implication holds for all the states $|\psi\rangle \in S$:*

$$C_G(\boldsymbol{\theta}) = C_G^{\text{opt}} \implies \rho_{\boldsymbol{\theta}}^{\psi} = \rho_{\text{opt}}^{\psi} \qquad \forall |\psi\rangle \in S \tag{7.76}$$

Proof. The global cost function $C_G(\boldsymbol{\theta})$ achieves the minimum value C_G^{opt} at a unique point corresponding to $\mathbb{E}[F_G(\boldsymbol{\theta})] = F_G^{\text{opt}}$, where F_G^{opt} corresponds to the fidelity term for C_G^{opt}. This corresponds to the unique global clone state ρ_{opt}^{ψ}. Thus the cost function, achieves a unique minimum under precisely the unique condition *i.e.* the output global state is equal to the optimal clone state for all inputs in the distribution. □

Now, we provide a statement of weak faithfulness that is much more relevant in the practical implementation of the cloning scheme using global optimization.

Lemma 8. *Suppose the cost function is ε-close to the optimal cost in symmetric cloning*

$$C_G(\boldsymbol{\theta}) - C_G^{\text{opt}} \leqslant \varepsilon \tag{7.77}$$

where $C_G^{\text{opt}} := 1 - F_G^{\text{opt}}$. Then,

$$\text{Tr}\left[(\rho_{\text{opt}}^{\psi} - \rho_{\boldsymbol{\theta}}^{\psi}) |\psi\rangle^{\otimes 2} \langle\psi|^{\otimes 2}\right] \leqslant \mathcal{N}\varepsilon, \qquad \forall |\psi\rangle \in S \tag{7.78}$$

Proof. The proof follows identically to Lemma 7 but with the exception that $C_G(\boldsymbol{\theta}) - C_G^{\text{opt}} = \mathbb{E}[F_G^{\text{opt}} - F_G(\boldsymbol{\theta})]$. □

Finally, we have the following theorem regarding the weak faithfulness of the global cost function:

Theorem 59. *[Weak faithfulness of the global cost function] Suppose the cost function is ε-close to the optimal cost in symmetric cloning*

$$C_G(\boldsymbol{\theta}) - C_G^{\text{opt}} \leqslant \varepsilon \tag{7.79}$$

where $C_G^{\text{opt}} := 1 - F_G^{\text{opt}}$. Then,

$$\Theta_{BA}(\rho_{\boldsymbol{\theta}}^{\psi}, \rho_{\text{opt}}^{\psi}) \leqslant \frac{\mathcal{N}\varepsilon}{\sin(F_G^{\text{opt}})} =: f_3(\varepsilon), \qquad \forall |\psi\rangle \in S \tag{7.80}$$

Proof. The proof follows along the same lines as the proof of closeness of the Bures angl/Fubini-Study distance for local cost function provided in Theorem 57. □

7.3.3.1 Global versus Local Faithfulness

This section explores the relationship between local and global cost function optimization for different cloners (universal, phase-covariant, etc.). In particular, we address the question of whether optimizing a cloner with a local or a global cost function also achieves an optimal solution relative to the other cost (operational meaning). If the answer is affirmative, we can use whichever cost exhibits the most desirable qualities and be confident they will achieve the same results. If not, we must be more careful as the choice may not lead to the optimal behaviour we desire and so will be application dependent.

We note that this relationship only manifests in *symmetric* cloning since there is no possibility to enforce asymmetry in the global cost function. The tradeoff between local and global faithfulness turns out to be subtle when dealing with cloning problems and is in contrast to similar studies in analogous variational algorithm literature [KKR06, CSV+21]. To begin, we have the following theorem:

Theorem 60. *For the general case of $M \to N$ cloning, the global cost function $C_G(\boldsymbol{\theta})$ and the local cost function $C_L(\boldsymbol{\theta})$ satisfy the inequality,*

$$C_L(\boldsymbol{\theta}) \leqslant C_G(\boldsymbol{\theta}) \leqslant N \cdot C_L(\boldsymbol{\theta}) \tag{7.81}$$

Proof. We first prove the first part of the inequality,

$$\begin{aligned} \mathrm{C_G}(\boldsymbol{\theta}) - \mathrm{C_L}(\boldsymbol{\theta}) &= \frac{1}{\mathcal{N}} \int_S \mathrm{Tr}((\mathrm{O_G^\psi} - \mathrm{O_L^\psi})\rho_{\boldsymbol{\theta}}^{\psi}) d\psi \\ &= \frac{1}{\mathcal{N}N} \int_S \mathrm{Tr}\left(\left(\sum_{j=1}^{N} \left(|\psi\rangle\langle\psi|_j \otimes \mathbb{1}_{\bar{j}} - |\psi\rangle\langle\psi|_1 \otimes \cdots |\psi\rangle\langle\psi|_N \right) \right) \rho_{\boldsymbol{\theta}}^{\psi} \right) \geqslant 0 \\ &\implies \mathrm{C_G}(\boldsymbol{\theta}) \geqslant \mathrm{C_L}(\boldsymbol{\theta}) \end{aligned} \tag{7.82}$$

where $\mathrm{O_L^\psi}$ is defined in Eq. (7.43), and the inequality in the second line holds due to the following

$$\sum_{j=1}^{N} \left(|\psi\rangle\langle\psi|_j \otimes \mathbb{1}_{\bar{j}} - |\psi\rangle\langle\psi|_1 \otimes \cdots |\psi\rangle\langle\psi|_N \right) = \sum_{j=1}^{N} |\psi\rangle\langle\psi|_j \otimes (\mathbb{1}_{\bar{j}} - |\psi\rangle\langle\psi|_{\bar{j}}) \geqslant 0, \quad \forall |\psi\rangle \in S \tag{7.83}$$

For the second part of the inequality, we consider the operator $NO_L^\psi - O_G^\psi$,

$$\begin{aligned}
NO_L^\psi - O_G^\psi &= (N-1)\mathbb{1} - \sum_{j=1}^{N}\left(|\psi\rangle\langle\psi|_j \otimes \mathbb{1}_{\bar{j}}\right) + |\psi\rangle\langle\psi|_1 \otimes \cdots |\psi\rangle\langle\psi|_N \\
&= \sum_{j=1}^{N-1}\left(\mathbb{1}_j \otimes \mathbb{1}_{\bar{j}} - |\psi\rangle\langle\psi|_j \otimes \mathbb{1}_{\bar{j}}\right) - |\psi\rangle\langle\psi|_N \otimes \mathbb{1}_{\bar{N}} + |\psi\rangle\langle\psi|_1 \otimes \cdots |\psi\rangle\langle\psi|_N \\
&= \sum_{j=1}^{N-1}\left((\mathbb{1} - |\psi\rangle\langle\psi|)_j \otimes \mathbb{1}_{\bar{j}}\right) - \bigotimes_{j=1}^{N-1}(\mathbb{1} - |\psi\rangle\langle\psi|)_j \otimes |\psi\rangle\langle\psi|_N \\
&= (\mathbb{1} - |\psi\rangle\langle\psi|)_1 \otimes \left(\mathbb{1}_{\bar{1}} - \bigotimes_{j=2}^{N-1}(\mathbb{1} - |\psi\rangle\langle\psi|_j) \otimes |\psi\rangle\langle\psi|_N\right) \\
&+ \sum_{j=2}^{N-1}\left((\mathbb{1} - |\psi\rangle\langle\psi|)_j \otimes \mathbb{1}_{\bar{j}}\right) \\
&\geqslant 0
\end{aligned} \tag{7.84}$$

where the second last line is positive because each individual operator is positive for all $|\psi\rangle \in S$. □

A similar inequality was proven in the work of [BPLC+20]. But interestingly, the inequality proven in Theorem 60 (unlike in [BPLC+20]) does not allow us make statements about the similarity of individual clones from the closeness of the global cost function and vice versa. This can be seen as follows:

$$\begin{aligned}
C_G(\boldsymbol{\theta}) - C_G^{\text{opt}} \leqslant \varepsilon &\implies C_L(\boldsymbol{\theta}) - C_L^{\text{opt}} \leqslant \varepsilon - (C_G(\boldsymbol{\theta}) - C_L(\boldsymbol{\theta})) + (C_L^{\text{opt}} - C_G^{\text{opt}}) \\
&\implies C_L(\boldsymbol{\theta}) - C_L^{\text{opt}} \leqslant \varepsilon + (C_L^{\text{opt}} - C_G^{\text{opt}}) \\
&\not\Rightarrow C_L(\boldsymbol{\theta}) - C_L^{\text{opt}} \leqslant \varepsilon
\end{aligned} \tag{7.85}$$

Here we have used the result of Theorem 60 that $C_G(\boldsymbol{\theta}) \geqslant C_L(\boldsymbol{\theta})$ and we note that $C_L^{\text{opt}} - C_G^{\text{opt}} \neq 0$ for all the $M \to N$ cloning. In particular, for $1 \to 2$ cloning, $C_L^{\text{opt}} = 5/6$, while $C_G^{\text{opt}} = 2/3$. This is due to the non-vanishing property of these cost functions, and highlights the subtlety of the case in hand.

While we are unable to leverage generic inequalities for our purpose, based on the cost functions, we can make statements in *specific* cases. In other words, by restricting the cloning problem to a specific input set of states, we can guarantee that optimizing *globally* will be sufficient to also optimize *local* figures of merit.

In particular, in the following, we establish these strong and weak faithfulness guarantees for the *special cases* of universal and phase-covariant cloning by analyzing problem-specific features.

Theorem 61. *The* global *cost function is* locally *strongly faithful for a universal symmetric cloner, i.e,:*

$$C_G(\boldsymbol{\theta}) = C_G^{\text{opt}} \iff \rho_{\boldsymbol{\theta}}^{\psi,j} = \rho_{\text{opt}}^{\psi,j} \qquad \forall |\psi\rangle \in \mathcal{H}, \forall j \in \{1,\ldots,N\} \tag{7.86}$$

Proof. In the symmetric universal case, C_L^{opt} has a unique minimum when, each local fidelity saturates:

$$F_L^{\text{opt}} = \frac{M(N+2)+N-M}{N(M+2)} \tag{7.87}$$

achieved by local reduced states $\{\rho_{\text{opt}}^{\psi,j}\}_{j=1}^{N}$. Now, it has been shown that the optimal global fidelity F_G that can be reached [BBHB97, SIGA05] is,

$$F_G^{\text{opt}} = \frac{N!(M+1)!}{M!(N+1)!} \tag{7.88}$$

which also is the corresponding unique minimum value for C_G^{opt}, achieved by some global state ρ_{opt}^{ψ}.

Finally, it was proven in [Wer98, KW99] that the cloner which achieves one of these bounds is unique and also saturates the other bound, and therefore must also achieve the unique minimum of both global and local cost functions, C_G^{opt} and C_L^{opt}. Hence, the local states which optimize C_L^{opt} must be the reduced density matrices of the global state which optimizes C_G^{opt} and so:

$$\rho_{\text{opt}}^{\psi,j} := \text{Tr}_{\bar{j}}(\rho_{\text{opt}}^{\psi}), \quad \forall j \tag{7.89}$$

Thus for a universal cloner, the cost function with respect to both local and global fidelities will converge to the same minimum. □

Now, before proving an analogous statement in the case of phase-covariant cloning, we first need the following lemma (we return to the notation of B, E and E^* for clarity):

Lemma 9. *For any* $1 \to 2$ *phase-covariant cloning machine which takes states* $|0\rangle_B \otimes |\psi\rangle_E$ *and an ancillary qubit* $|A\rangle_{E^*}$ *as input, where* $|\psi\rangle := \frac{1}{\sqrt{2}}(|0\rangle + e^{i\theta}|1\rangle)$*, and outputs a 3-qubit state* $|\Psi_{BEE^*}\rangle$ *in the following form:*

$$\begin{aligned} |\Psi_{BEE^*}\rangle = &\frac{1}{2}[|0,0\rangle + e^{i\phi}(\sin\eta\,|0,1\rangle + \cos\eta\,|1,0\rangle))\,|0\rangle_{E^*} \\ &+ e^{i\phi}|1,1\rangle + (\cos\eta\,|0,1\rangle + \sin\eta\,|1,0\rangle)\,|1\rangle_{E^*}] \end{aligned} \tag{7.90}$$

the global and local fidelities are simultaneously maximized at $\eta = \frac{\pi}{4}$ *where* $0 \leqslant \eta \leqslant \frac{\pi}{2}$ *is the 'shrinking factor'.*

Proof. To prove this, we follow the formalism that was adopted by Cerf *et. al.*[CIVA02]. This uses the fact that a symmetric phase-covariant cloner induces a mapping of the following form [SIGA05]:

$$\begin{aligned}
&|0\rangle|0\rangle|0\rangle \to |0\rangle|0\rangle|0\rangle \\
&|1\rangle|0\rangle|0\rangle \to (\sin\eta|0\rangle|1\rangle + \cos\eta|1\rangle|0\rangle)|0\rangle \\
&|0\rangle|1\rangle|1\rangle \to (\cos\eta|0\rangle|1\rangle + \sin\eta|1\rangle|0\rangle)|1\rangle \\
&|1\rangle|1\rangle|1\rangle \to |1\rangle|1\rangle|1\rangle
\end{aligned} \tag{7.91}$$

Next, we calculate the global state by tracing out the ancillary state to get $\rho_{\mathrm{G}}^{\mathrm{opt}}$:

$$\rho_{\mathrm{G}}^{\mathrm{opt}} = \mathrm{Tr}_{E^*}(|\Psi_{BEE^*}\rangle\langle\Psi_{BEE^*}|) = |\Phi_1\rangle\langle\Phi_1| + |\Phi_2\rangle\langle\Phi_2| \tag{7.92}$$

where,

$$|\Phi_1\rangle := \frac{1}{2}\left[|0,0\rangle + e^{i\phi}(\sin\eta|0,1\rangle + \cos\eta|1,0\rangle)\right], \tag{7.93}$$

$$|\Phi_2\rangle := \frac{1}{2}\left[e^{i\phi}|1,1\rangle + (\cos\eta|0,1\rangle + \sin\eta|1,0\rangle)\right] \tag{7.94}$$

Hence the global fidelity can be computed as:

$$F_{\mathrm{G}}^{\mathrm{opt}} = \mathrm{Tr}(|\psi\rangle\langle\psi|^{\otimes 2}\rho_{\mathrm{G}}^{\mathrm{opt}}) = |\langle\psi^{\otimes 2}|\Phi_1\rangle|^2 + |\langle\psi^{\otimes 2}|\Phi_2\rangle|^2 = \frac{1}{8}(1+\sin\eta+\cos\eta)^2 \tag{7.95}$$

Now, optimising $F_{\mathrm{G}}^{\mathrm{opt}}$ with respect to η, we see that $F_{\mathrm{G}}^{\mathrm{opt}}$ has only one extremum value between $[0, \frac{\pi}{2}]$ specifically at $\eta = \frac{\pi}{4}$. We can also see that the local fidelity is also achieved for the same η and is equal to:

$$F_{\mathrm{L}}^{\mathrm{opt}} = \frac{1}{2}\left(1 + \frac{\sqrt{2}}{2}\right) \tag{7.96}$$

which is the upper bound for local fidelity of the phase-covariant cloner. □

With Lemma 9 established, we can next prove:

Theorem 62. *The global cost function is locally strongly faithful for phase-covariant symmetric cloner,* i.e.*:*

$$C_G(\boldsymbol{\theta}) = C_G^{\mathrm{opt}} \iff \rho_{\boldsymbol{\theta}}^{\psi,j} = \rho_{\mathrm{opt}}^{\psi,j} \qquad \forall|\psi\rangle \in S, \forall j \in \{B,E\} \tag{7.97}$$

where S is the distribution corresponding to phase-covariant cloning.

Proof. We have shown in Lemma 9 that the global and local fidelities of a phase-covariant cloner are both achieved with a cloning transformation of the form in Eq. (7.91). Applying this transformation unitary to $|\psi\rangle|\Phi^+\rangle_{BE}$ (where $|\Phi^+\rangle_{BE}$ is a Bell state) leads to Cerf's formalism for cloning. Furthermore, we can observe

that due to the symmetry of the problem, this transformation is unique (up to global phases) and so any optimal cloner must achieve it.

Furthermore, one can check that the ideal circuit in Fig. 7.3(b) does indeed produce an output in the form of Eq. (7.90) once the preparation angles have been set for phase-covariant cloning. By a similar argument to the above, we can see that a variational cloning machine which achieves an optimal cost function value, *i.e.* $C_G(\boldsymbol{\theta}) = C_G^{\text{opt}}$ much also saturate the optimal cloning fidelities. Furthermore, by the uniqueness of the above transformation (Eq. (7.91)) we also have that the local states of VarQlone are the same as the optimal transformation, which completes the proof. □

7.3.4 Summary of other specifications

In this section for completeness , we give an overview of some of the other specifications of the algorithm. We will not go in-depth to prove them as they are neither the author's contribution nor directly relevant to the main topic of this thesis. Nevertheless, it will contribute to understanding the algorithm for potential future applications.

First, we start with the choice of Ansatz (see Section 2.6.4.3). A key element in variational algorithms is the choice of Ansatz that is used in parameterized quantum circuits. The primary Ansatz we choose is one with a *variable* structure. This allows us to learn cloning circuits in an end-to-end manner. The idea is to optimize over both the continuous parameters of a quantum circuit, but also over the gates within the circuit itself, which come from a discrete set. The goal is to solve the following optimization problem [LFC+20]:

$$(\boldsymbol{\theta}^*, \boldsymbol{g}^*) = \arg\min_{\boldsymbol{\theta}, \boldsymbol{g} \in \mathcal{G}} C(\boldsymbol{\theta}, \boldsymbol{g}) \tag{7.98}$$

where $\mathcal{G}$ denotes the gate set. Such variable-structure Ansätze approaches can be broadly dubbed as the *Quantum Architecture Search* (QAS) [ZHZY21] similar to *Neural Architecture Search* (NAS) in classical ML [YWC+19, LSY18]. Approaches to QAS have appeared in many forms [CSSC18, GEBM19, OGB21, CSU+20, LFC+20, PT21]. In this work, $\mathcal{G}$ is a gateset *pool*, from which a particular sequence $\boldsymbol{g}$ is chosen. As a summary, to solve this problem, we iterate over $\boldsymbol{g}$, swap out gates, and re-optimize the parameters $\boldsymbol{\theta}$, until a minimum of the cost, $C(\boldsymbol{\theta}^*, \boldsymbol{g}^*)$ is found. This is a combination of a discrete and continuous optimization problem, where the discrete parameters are the indices of the gates in $\boldsymbol{g}$ (*i.e.*, the circuit structure), and the continuous parameters are represented by $\boldsymbol{\theta}$. Each time the circuit structure is changed (a subset of gates are altered), the continuous parameters are re-optimized, as in [CSSC18]. Variations of this approach have been proposed in [DHY+20, LFC+20] which could be easily incorporated, and we leave such investigation to future work. For the results shown for $1 \to 2$ cloning phase-covariant states, we use the following three qubit gate pool:

$$\mathcal{G}_{\text{PC}} := \{\, \mathsf{R}_z^2(\theta), \mathsf{R}_z^3(\theta), \mathsf{R}_z^4(\theta), \mathsf{R}_x^2(\theta), \mathsf{R}_x^3(\theta), \mathsf{R}_x^4(\theta), \\ \mathsf{R}_y^2(\theta), \mathsf{R}_y^3(\theta), \mathsf{R}_y^4(\theta), \mathsf{CZ}_{2,3}, \mathsf{CZ}_{3,4}, \mathsf{CZ}_{2,4} \,\} \quad (7.99)$$

In order to attack protocol $\mathcal{P}_1$ using $1 \to 2$ state dependent cloning, we use the following pool:

$$\mathcal{G}_{\mathcal{P}_1^{1\to 2}} := \left\{ \mathsf{R}_j^i(\theta), \mathsf{CZ}_{2,3}, \mathsf{CZ}_{3,4} \right\} \quad \forall i \in \{2,3,4\}, \forall j \in \{x,y,z\} \qquad (7.100)$$

where R_j^i indicates the j^{th} Pauli rotation with angle θ acting on the i^{th} qubit and CZ is the controlled-Z gate. In both cases, we use the qubits indexed 2, 3 and 4 in an `Aspen-8` sublattice. Note that in the latter case, we allow only a linear, nearest-neighbour (NN) connectivity, which removes the need for inserting SWAP gates by the quantum compiler. For more detailed specifications about the algorithm via supplementary numerical results, we refer to [CDKK22].

Next, we talk about the sample complexity of VarQlone. We discussed that VarQlone requires classical minimisation of one of the cost functions $\mathsf{C}(\boldsymbol{\theta}) := \{\mathsf{C}_{\text{sq}}(\boldsymbol{\theta}), \mathsf{C}_{\text{L}}(\boldsymbol{\theta}), \mathsf{C}_{\text{G}}(\boldsymbol{\theta})\}$ to achieve the optimal cost value. To do so, we must be able to efficiently evaluate the cost function of choice. In our case, this can be achieved via a method that allows the computation of the fidelity between quantum states. This estimation can be done either via SWAP test, which is a powerful toolkit that we have used many times so far in this thesis; or via computing an estimator for the true cost $\mathsf{C}(\boldsymbol{\theta}) = \mathbb{E}_{|\psi\rangle \in S}[\mathsf{C}^{\psi}(\boldsymbol{\theta})]$ using K different states sampled from S. Estimating the overlap in the mentioned way is sufficient for our purposes since this coincides with the fidelity when at least one of the states is a pure state:

$$F(|\psi\rangle\langle\psi|, \rho) = \langle\psi|\rho|\psi\rangle = \mathsf{Tr}(|\psi\rangle\langle\psi|\rho) \qquad (7.101)$$

Since VQAs are heuristic algorithms, there are no guarantees on the number of training iterations over $\boldsymbol{\theta}$ to converge to C^{opt}. However, one can at least provide guarantees on the number of samples required to estimate the cost, for a particular instance of the parameters. Since this is a necessary subroutine in the algorithm, it must be efficient. It can be shown that the number of samples $L \times K$, where K is the number of distinct states $|\psi\rangle$ sampled uniformly at random from the distribution S, and L is the number of copies of each input state, required to estimate the cost function $\mathsf{C}(\boldsymbol{\theta})$ up to ε'-additive error with a success probability δ is,

$$L \times K = \mathcal{O}\left(\frac{1}{\varepsilon'^2} \log \frac{2}{\delta}\right) \qquad (7.102)$$

We refer to [CDKK22] for the proof.

Finally, we can examine VarQlone for the existence of barren plateaus. We prove in [CDKK22], that the local cost function that we have presented does not exhibit barren plateaus for a sufficiently shallow alternating layered Ansatz, *i.e.* $U(\boldsymbol{\theta})$ contains blocks W, acting on alternating pairs of qubits [CSV+21].

7.4 Practical cryptanalysis based on the numerical results of VarQlone

At last, we present the numerical and experimental results of VarQlone for the cryptographic problems that we have introduced in Section 7.2. Let us start with the phase covariant case study.

7.4.1 Variational phase-covariant cloning

We start with the attack we have presented in Section 7.2.1 to attack the BB84 protocol, while here we replace the theoretical cloner with VarQlone. Fig. 7.2 demonstrates at a high level, how VarQlone is inserted into an attack (on QKD or similar protocols).

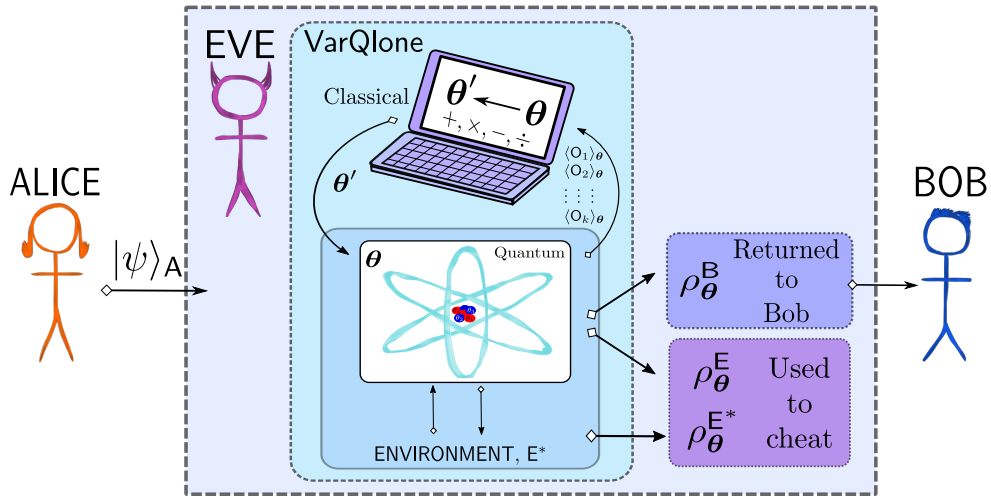

Figure 7.2: Cartoon overview of VarQlone in a cryptographic attack. Here an adversary Eve, E, implements a $1 \to 2$ cloning attack on states used in a quantum protocol (for example QKD) between Alice and Bob. Eve intercepts the states sent by Alice $|\psi\rangle_A$ and may interact with an ancillary 'environment', E^*. This interaction is trained (an optimal parameter setting $\boldsymbol{\theta}$ is found) by Eve to optimally produce clones, $\rho_{\boldsymbol{\theta}}^B, \rho_{\boldsymbol{\theta}}^E$. In order to attack the protocol, Eve will return $\rho_{\boldsymbol{\theta}}^B$ to Bob and use the rest (her clone, $\rho_{\boldsymbol{\theta}}^E$ plus the remaining environment state, $\rho_{\boldsymbol{\theta}}^{E^*}$) to cheat. The training procedure consists of using a classical computer to optimize the quantum parameters, via a cost function. The cost is a function of k observables, O_k, measured from the output states, which are designed to extract fidelities of the states to compare against the ideal state.

Here VarQlone has K layers in the ansatz, in each layer there is a fixed structure. For simplicity, we choose each layer to have parameterised single-qubit rotations, $\mathsf{R}_y(\theta)$, and nearest neighbour CZ gates. Our primary target is $1 \leftarrow 2$ cloning, so we use 3 qubits and therefore we have 2 CZ gates per layer. Not surprisingly, in the experiment, we observe convergence to the minimum as the number of layers increases, saturating at $K = 3$.

Now we show a proof of principle implementation of our methods for phase-covariant cloner. The results of this can be seen in Fig. 7.3.

Let us begin by describing some problem parameters. Firstly, we allow 3 qubits (2 output clones plus 1 ancilla) in the circuit. Next, we give the VarQlone the fully

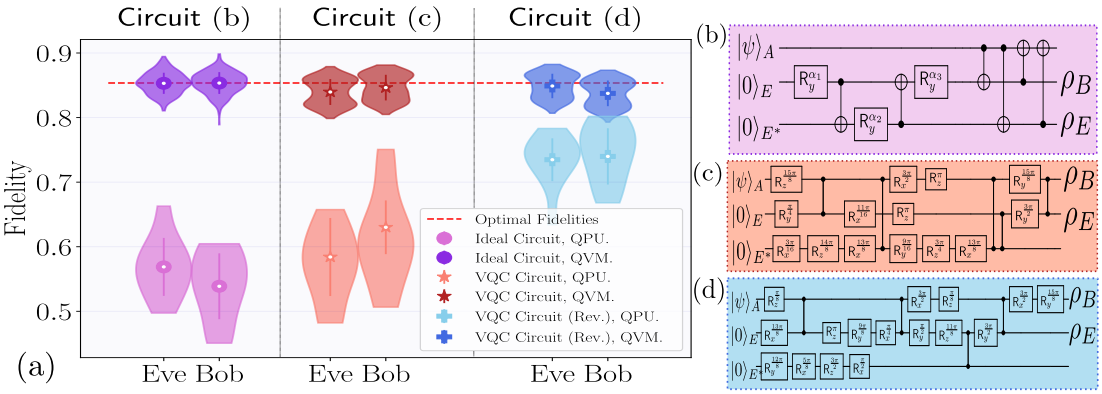

Figure 7.3: Variational Quantum Cloning implemented on phase-covariant states using three qubits of the Rigetti `Aspen-8` chip (QPU), plus simulated results (QVM). Violin plots in (a) show the cloning fidelities, for Bob and Eve, found using each of the circuits shown in (b)–(d) respectively. Shown in red is the maximal possible fidelity for this problem. (b) is the ideal circuit with clones appearing in registers 2 and 3. (c) shows the structure-learned circuit for the same scenario, using one less entangling gate. (d) demonstrates the effect of allowing clones to appear in registers 1 and 2. In the latter case, only four (nearest-neighbour) entangling gates are used, demonstrating a significant boost in performance on the QPU.

connected (FC) gateset pools introduced in Eq. (7.99). Let us now analyse the two candidate resulting circuits of VarQlone in Fig. 7.3(c,d) in comparison with the optimal 'analytic' circuit Fig. 7.3(b) introduced in [BBHB97, FWJ+14].

Firstly, we note that all three circuits approximately saturate the optimal bound for phase-covariant cloning ($F_L = 0.85$) when simulated (*i.e.* without quantum noise). But we notice that the ideal circuit in Fig. 7.3(b) suffers degradation in performance when implemented on the QPU since it requires 6 entangling gates as it is attempting to transfer the information across the circuit. Furthermore, since the `Aspen-8` chip does not have any 3 qubit loops in its topology, it is necessary for the compiler to insert SWAP gates.

Next, we compare the ideal circuit to two examples learned by VarQlone. Firstly, we force the qubit clones to appear in registers 2 and 3, (demonstrated in Fig. 7.3(c)) exactly as in Fig. 7.3(b). Secondly, we allow the clones to appear instead in registers 1 and 2 (demonstrated in Fig. 7.3(d) - The circuit labeled 'Rev.' ('Reverse').) The ability to make such a subtle change demonstrates clearly the advantage of our flexible approach. We notice that the restriction imposed in Fig. 7.3(c) results in only slightly improved performance over the ideal. However, by allowing the clones to appear in registers 1 and 2, VarQlone can find much more conservative circuits, having fewer entangling gates, and are directly implementable on a linear topology. This gives a significant improvement in the cloning fidelities, of about 15% when the circuit is run on the QPU, as observed in Fig. 7.3(a). For all results shown using a variable structure ansatz, we use the `forest-benchmarking` library [CGH+19] to reconstruct the output density matrix in order to mitigate the effect of quantum noise.

Finally, we can calculate the success probability of a real implementable attack on BB84 using today's quantum hardware as a result of VarQlone. We specifically

analyse the performance of one of these VarQlone-learned circuits (Circuit(c)) in such an attack. The reason for this choice is that while the circuit in Fig. 7.3(d) achieves higher fidelities on the Aspen hardware, it does not actually make use of the ancillary qubit (one can observe that the sequence of gates acting on it, is approximately an identity gate). We will do so by computing the corresponding critical error rate, D_{crit}, using Eq. (**??**) as discussed in Section 7.2.1. for the BB84 protocol run in $X-Y$ Pauli basis. First, we compute the resulting mixed states outputted over all input states of the cloning machine, for each basis state: $\{|+\rangle, |-\rangle, |+i\rangle, |-i\rangle\}$ so ρ_E is given by:

$$\rho_E := \frac{1}{4}(\rho_E^{+} + \rho_E^{-} + \rho_E^{+i} + \rho_E^{-i}) \tag{7.103}$$

Similarly, ρ_E^0, ρ_E^1 in Eq. (**??**) are the mixed states encoding the random bit 0 (corresponding to $\{|+\rangle, |+i\rangle\}$) and bit 1 (corresponding to $\{|-\rangle, |-i\rangle\}$), so are given by:

$$\rho_E^0 := \frac{1}{2}(\rho_E^{+} + \rho_E^{+i}), \qquad \rho_E^1 := \frac{1}{2}(\rho_E^{-} + \rho_E^{-i}) \tag{7.104}$$

Calculating the minimum Holevo quantity χ_{min} for the above density matrices outputted by the circuit in Fig. 7.3(c) gives the following:

$$\begin{aligned} &1 - H(D_{\text{crit}}) - \chi_{min} = 0 \\ &\implies 1 - \chi_{min} + (D_{\text{crit}} \log_2(D_{\text{crit}}) + (1 - D_{\text{crit}}) \log_2(1 - D_{\text{crit}}))) = 0 \\ &\implies D_{\text{crit}} = 15.8\%. \end{aligned} \tag{7.105}$$

Recalling the optimal bound for the individual attack, one can see that the D_{crit} obtained by the result of VarQlone is very close to that bound. Nevertheless, as pointed out in [SIGA05, FL12], the same bound can be reached by a collective attack (where Eve defers all the measurements until the end of the reconciliation phase and applies a general strategy to all collected states) so long as the individual quantum operations are still given by the optimal phase-covariant cloner. Thus, the VarQlone learned circuits can be used to perform collective attacks and almost saturate the optimal collective bound.

Finally, one may observe that Circuit (d) in Fig. 7.3 achieves an even higher fidelity on the actual hardware, but it does so without using the ancilla to reduce the circuit depth. Therefore, it is a more suited and non-trivial circuit for purely performing phase covariant cloning.

7.4.2 Variational state-dependent cloning

In this section, we present the results of VarQlone when learning to clone the states used in the two coin-flipping protocols described in Section 7.2.2.2 and Section 7.2.2.3. Firstly, we focus on the states used in the original protocol, $\mathcal{P}_1$ for $1 \to 2$ cloning, and then move to the 4 state protocol, $\mathcal{P}_2$. In the latter we also extend from $1 \to 2$ cloning to $1 \to 3$ and $2 \to 4$. These extensions will allow us

to probe certain features of VarQlone, in particular explicit symmetry in the cost functions. In all cases, we use the variable structure Ansatz, and once a suitable candidate has been found, the solution is manually further optimised. The learned circuits that are used to produce the figures and results in this section, are given in Fig. 7.5 and Fig. 7.8.

7.4.2.1 Variational cloning attack on 2-state quantum coin-flipping

As a reminder, the two states used in this protocol are:

$$|\phi_0\rangle := |\phi_{0,0}\rangle = \cos\left(\frac{\pi}{18}\right)|0\rangle + \sin\left(\frac{\pi}{18}\right)|1\rangle \tag{7.106}$$

$$|\phi_1\rangle := |\phi_{0,1}\rangle = \cos\left(\frac{\pi}{18}\right)|0\rangle - \sin\left(\frac{\pi}{18}\right)|1\rangle \tag{7.107}$$

The cloning-based attacks and the obtained fidelities achieved by the VarQlone learned circuit can be seen in Fig. 7.4 where we use the gate pool Eq. (7.100) which allows a linear entangling connectivity.

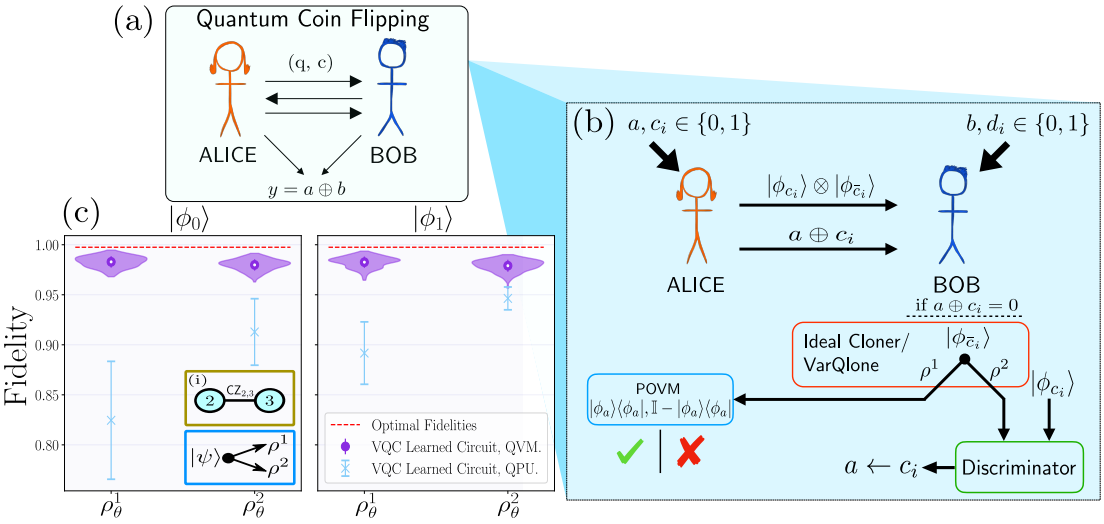

Figure 7.4: Overview of cloning-based attack on the protocol of Mayers *et. al.* [MSCK99], plus corresponding numerical results for VarQlone. (a) Cartoon of coin flipping protocols, Alice and Bob send quantum (q) and/or classical (c) information to agree on a final 'coin flip' bit, y. (b) The relevant part of the protocol of Mayers *et. al.*, $\mathcal{P}_1$, plus a cloning based attack on Bob's side. He builds a cloning machine using VarQlone to produce two clones of Alice's sent states, one of which he returns, and the other is used to guess Alice's input bit, a. (c) Fidelities of each output clone, ρ^j_θ achieved using VarQlone when $(1 \to 2)$ cloning the family of states used in, $\mathcal{P}_1$. In the left (right) panel, $|\phi_0\rangle$ ($|\phi_1\rangle$) is. Figure shows both simulated (QVM - purple circles) and on Rigetti hardware (QPU - blue crosses). For the QVM (QPU) results, 256 (5) samples of each state are used to generate statistics. Violin plots show complete distribution of outcomes and error bars show the means and standard deviations. Inset (i) shows the two qubits of the `Aspen-8` chip which were used, with the allowed connectivity of a CZ between them. Note an ancilla was also allowed, but VarQlone chose not to use it in this example.

In the above figure, a deviation from the optimal fidelity can be seen, even in the simulated case. We believe this is mostly due to tomographic errors in reconstructing the cloned states. Before analysing the results for our given attack,

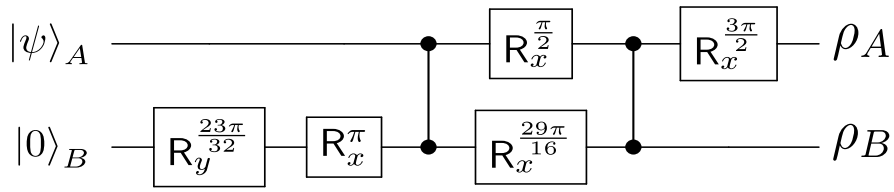

Figure 7.5: Circuit learned by VarQlone in to clone states, $|\phi_0\rangle, |\phi_1\rangle$, with an overlap $s = \cos(\pi/9)$ in the protocol $\mathcal{P}_1$. For example, ρ_A is the clone sent back to Alice, while ρ_B is kept by Bob.

let us also have a look at the circuits learned by VarQlone for the task of state-dependent cloning with fixed-overlap.

Fig. 7.5 shows the circuit used to achieve the fidelities in the attack on $\mathcal{P}^1$. In training, we still allowed an ancilla to aid the cloning, but the example in Fig. 7.5 did not make use of it (in other words, VarQlone only applied gates which were equivalent to applying identity on the ancilla), so we remove it to improve hardware performance. This repeats the behaviour seen for the circuits learned in phase-covariant cloning. We mention again, that some of the learned circuits did make use of the ancilla with similar performance. This mimics the behaviour seen in the previous example of phase-covariant cloning. As such, we only use the two qubits shown in the inset (i) of the figure when running on the QPU to improve performance.

Now, we proceed with calculating the success probability of the attack on $\mathcal{P}_1$ given the above experimental results. For illustration, let us return to the example in Eq. (7.9), where instead the cloned state is now produced from our VarQlone circuit, $\rho_c^0 \to \rho_{\text{VarQlone}}^0$.

Theorem 63. *[VarQlone Attack Bias on $\mathcal{P}_1$] Bob can achieve a bias of $\varepsilon \approx 0.29$ using a state-dependent VarQlone attack on the protocol $\mathcal{P}_1$, with a single copy of Alice's state.*

Proof. For the proof, we compute the success probability in the same way as in Theorem 50, as follows:

$$P_{\text{succ},\mathcal{P}_1}^{\text{VarQlone}} = \frac{1}{2} + \frac{1}{4}\text{Tr}|\rho_1 - |\phi_1\rangle\langle\phi_1| \otimes \rho_{\text{VarQlone}}^0| \approx 0.804 \tag{7.108}$$

Here $\rho_1 = |\phi_0\rangle\langle\phi_0| \otimes |\phi_1\rangle\langle\phi_1|$ (similar to the case in Eq. (7.9)). Here, we have a higher probability for Bob to correctly guess Alice's bit a, but correspondingly, the detection (by Alice) probability is higher than in the ideal case, due to a lower local fidelity of $F_{\text{L}}^{\text{VarQlone}} = 0.985$. □

7.4.2.2 Variational cloning attack on 4-state quantum coin-flipping

For the attacks on $\mathcal{P}_2$ using VarQlone, again we recall the family of states:

$$|\phi_{x,a}\rangle = \begin{cases} |\frac{\pi}{8}_{x,0}\rangle = \cos\left(\frac{\pi}{8}\right)|0\rangle + (-1)^x \sin\left(\frac{\pi}{8}\right)|1\rangle \\ |\frac{\pi}{8}_{x,1}\rangle = \sin\left(\frac{\pi}{8}\right)|0\rangle + (-1)^{x\oplus 1} \cos\left(\frac{\pi}{8}\right)|1\rangle \end{cases} \tag{7.109}$$

Here first we mention the attack that uses $1 \to 2$ cloning similar to the ones discussed in Section 7.2.2.3. But then we generalise the result to $1 \to 3$ and $2 \to 4$ cloning as well. One interesting aspect of the result given here is that there is no explicit analytical circuit known so far prior to our results, for these types of state-dependent cloning.

$1 \to 2$ **Cloning.**

Firstly, we use the same gate set and subset of the `Aspen-8` lattice ($\mathcal{G}_{\mathcal{P}_2^{1\to2}} = \mathcal{G}_{\mathcal{P}_1^{1\to2}}$). We use the local cost, Eq. (7.43), to train the model, with a sequence length of 35 gates. The attack model and the numerical results are shown in Fig. 7.6. Specifically, in part (b) the results for both QVM (simulation) and QPU (experiment) are given. We note that the solution exhibits some small degree of asymmetry in the output states, due to the form of the local cost function. This asymmetry is particularly pronounced as we scale the problem size and aim to produce N output clones, which we further discuss in the next section.

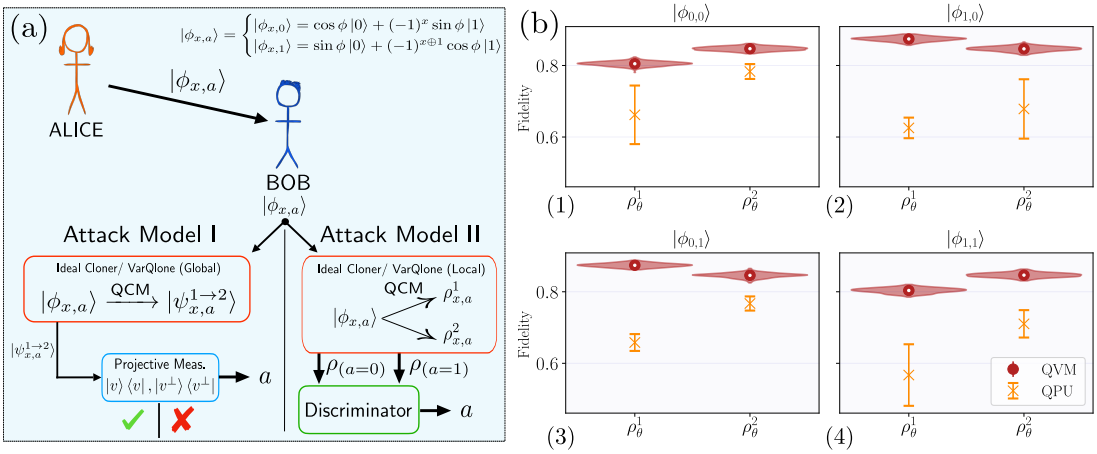

Figure 7.6: Cloning attacks and numerical results for the protocol, $\mathcal{P}_2$.(a) The two cloning based attacks we consider. In attack model `I` (left), Bob measures both output states with a set of fixed projective measurements, defined relative to the cloner output states, $|\psi^{1\to2}\rangle_{a,x}$ and guesses Alice's bit, a. In attack model `II`, Bob keeps one clone for either testing Alice later or to send back the deposit qubit requested by Alice. He uses then the other local clone to discriminate and guess a. (b) The fidelities achieved cloning the each state, $\{|\phi_{x,a}\rangle\}$ used in $\mathcal{P}_2$ with VarQlone. These numerics relate to scenario 1 from attack model `II`. Each panel (1-4) shows both simulated (QVM - red circles) and on Rigetti hardware (QPU - orange crosses). We indicate the fidelities of each clone received by Alice and Bob. For the QVM (QPU) results, 256 (3) samples of each state are used to generate statistics. Violin plots show the complete distribution of outcomes and error bars show the means and standard deviations. Inset (i) shows the connectivity we allow in VarQlone for this example.

Now, we can relate the performance of the VarQlone cloner to the attacks discussed in Section 7.2.2.3. We do this by explicitly analyzing the output states produced in the circuits used to achieve fidelities shown in Fig. 7.6(b) and following the derivation in Section 7.2.2.3, we show in Theorem 64 and Theorem 65:

Theorem 64. *[VarQlone Cloning Attack (I) Bias on $\mathcal{P}_2$] Using a cloning attack on the protocol $\mathcal{P}_2$, (in attack model I) Bob can achieve a bias:*

$$\varepsilon^{\mathrm{I}}_{\mathcal{P}_2,\mathit{VarQlone}} \approx 0.345 \tag{7.110}$$

Similarly, we have the bias which can be achieved with attack II:

Theorem 65. *[VarQlone Cloning Attack (II) Bias on $\mathcal{P}_2$] Using a cloning attack on the protocol $\mathcal{P}_2$, (in attack model II) Bob can achieve a bias:*

$$\varepsilon^{\mathrm{II}}_{\mathcal{P}_2,\mathit{VarQlone}} = 0.241 \tag{7.111}$$

The small variation between these results and the ideal biases proved in Theorem 51 and Theorem 52 is primarily due to the small degree of asymmetry induced by the heuristics of VarQlone. However, we emphasize that these biases can now be achieved via constructive attacks on the hardware.

$1 \to 3$ and $2 \to 4$ Cloning.

Finally, we extend the above analysis to the more general scenario of $M \to N$ cloning, taking $M = 1, 2$ and $N = 3, 4$. The result for $1 \to 3$ and $2 \to 4$ are illustrated in Fig. 7.7.

These examples are illustrative since they demonstrate the strengths of the squared local cost function in Eq. (7.41) over the local cost function in Eq. (7.43). In particular, we find that the local cost function does not enforce symmetry strongly enough in the output clones and using only the local cost function, suboptimal solutions are found. We particularly observed this in the example of $2 \to 4$ cloning, where VarQlone tended to take a shortcut by allowing one of the input states to fly through the circuit (resulting in nearly 100% fidelity for that clone). It then attempts to perform $1 \to 3$ cloning with the remaining input state. By strongly enforcing symmetry in the output clones using the squared cost, this can be avoided.

We also test two connectivities in these examples, a fully connected (FC) and the nearest neighbour (NN) architecture as allowed by the following gate sets:

$$\mathcal{G}^{\mathrm{NN}}_{\mathcal{P}_2^{1\to3}} = \{\mathrm{R}^i_z(\theta), \mathrm{R}^i_x(\theta), \mathrm{R}^i_y(\theta), \mathrm{CZ}_{2,3}, \mathrm{CZ}_{3,4}, \mathrm{CZ}_{4,5}\} \quad \forall i \in \{2,3,4,5\} \tag{7.112}$$

and

$$\mathcal{G}^{\mathrm{FC}}_{\mathcal{P}_2^{1\to3}} = \{\mathrm{R}^i_z(\theta), \mathrm{R}^i_x(\theta), \mathrm{R}^i_y(\theta), \mathrm{CZ}_{2,3}, \mathrm{CZ}_{2,4}, \mathrm{CZ}_{2,5}, \mathrm{CZ}_{3,4}, \mathrm{CZ}_{3,5}, \mathrm{CZ}_{4,5}\} \quad \forall i \in \{2,3,4,5\} \tag{7.113}$$

Note, that for $1 \to 3$ ($2 \to 4$) cloning, we actually use 4 (5) qubits, with one being an ancilla. The results of these experiments are also given in Fig. 7.7.

Finally, we give the the circuits learned by VarQlone and approximately clone all four states in Eq. (7.109) in the protocol, $\mathcal{P}_2$, for $1 \to 2, 1 \to 3$ and $2 \to 4$

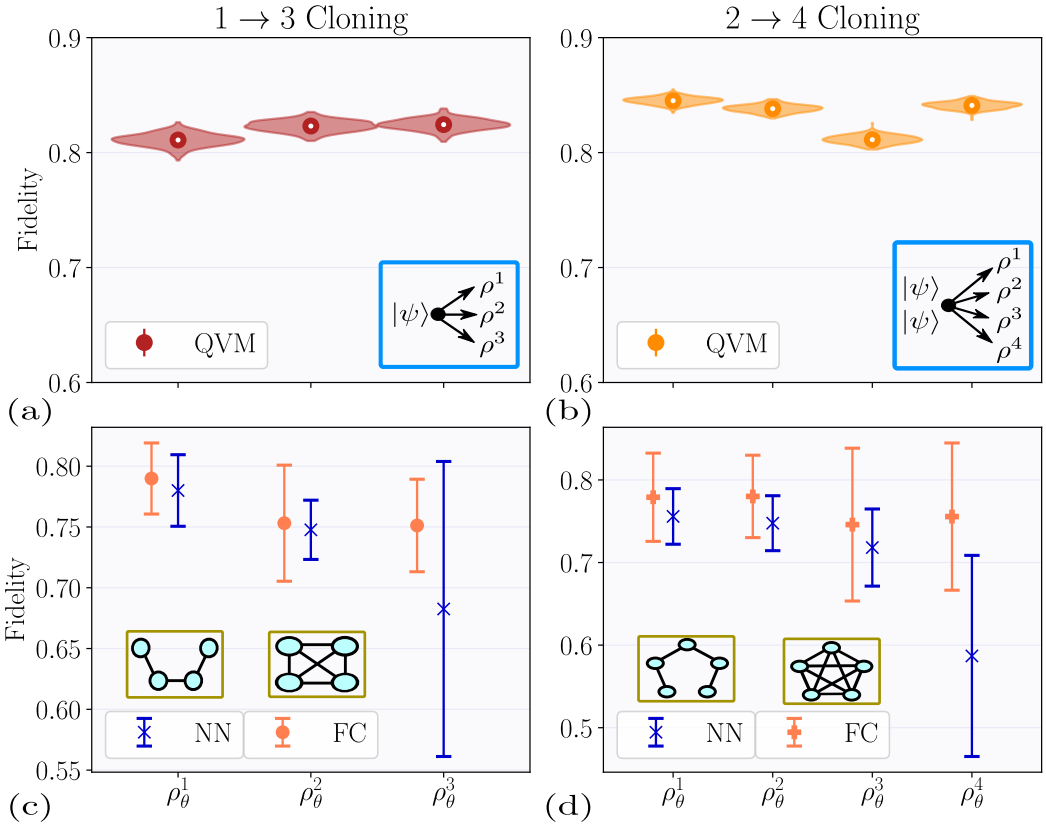

Figure 7.7: Clone fidelities for optimal circuits learned by VarQlone for (a) $1 \to 3$ and (b) $2 \to 4$ cloning of the states used in the coin-flipping protocol of [ATSVY00] *et. al.*. Mean and standard deviations of 256 samples are shown (violin plots show the full distribution of fidelities), where the fidelities are computed using tomography only on the Rigetti QVM. In both cases, VarQlone is able to achieve average fidelities $> 80\%$. (c-d) shows the mean and standard deviation of the optimal fidelities found by VarQlone over 15 independent runs (15 random initial structures, $\boldsymbol{g}$) for the nearest neighbour (NN - purple) versus (d) fully connected (FC - pink) entanglement connectivity allowed in the variable structure Ansatz for $1 \to 3$ and $2 \to 4$ cloning of $\mathcal{P}_2$ states. Insets of (c-d) shown corresponding allowed CZ gates in each example.

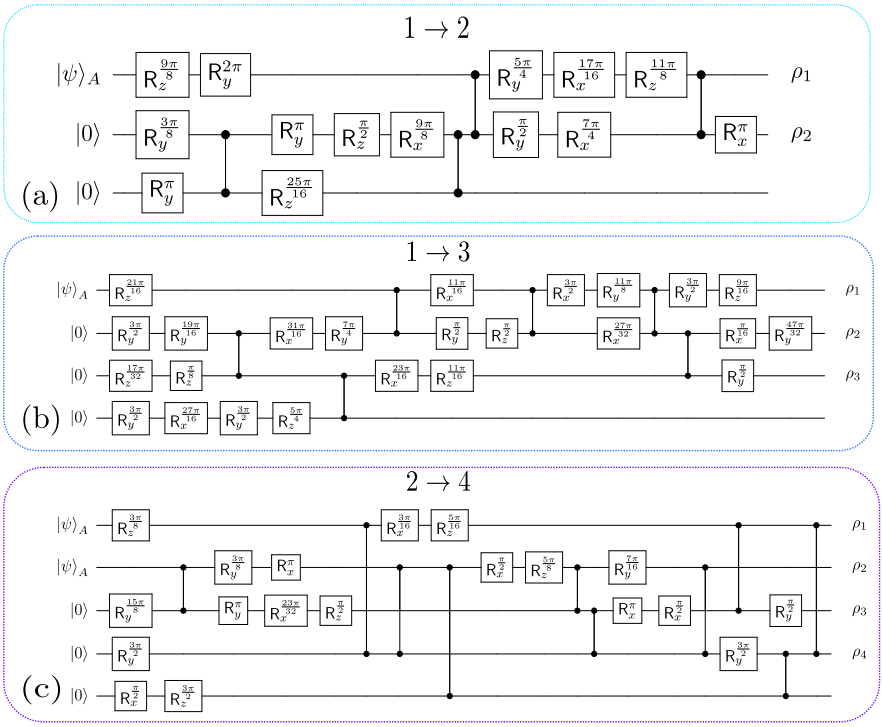

Figure 7.8: Circuits learned by VarQlone to clone states from the protocol, $\mathcal{P}_2$ for (a) $1 \to 2$, (b) $1 \to 3$ and (c) $2 \to 4$ cloning. These specific circuits produce the fidelities in Fig. 7.6(b) for $1 \to 2$, (using the local cost function), and in Fig. 7.7 for $1 \to 3$ and $2 \to 4$ (using the squared cost function). We allow an ancilla for all circuits, and ρ_k indicates the qubit which will be the k^{th} output clone.

cloning in Fig. 7.8. These are the specific circuits used to produce the fidelities in Fig. 7.6(b) and Fig. 7.7.

7.5 Discussion and conclusions

We have shown, throughout this chapter, yet another face of *unclonability*, not only as a core ingredient for quantum cryptanalysis but also with roots in foundational questions of quantum mechanics. Our attempts in this chapter have given partial answers to the following fundamental question: 'How do we construct efficient, flexible, and noise-tolerant circuits to perform approximate cloning? and 'How this ability will impact the security of real-life quantum protocols?' This latter question is especially pertinent in the current NISQ era, where the search for beneficial applications on small-scale noisy quantum devices remains at the forefront. On the other hand, this is an important and relevant question from a quantum communication perspective, given the existing gaps between the real implementations of quantum protocols and the proven theoretical results. In this work, exploring the exciting era between cryptanalysis and quantum machine learning, we have proposed our variational quantum cloner (VarQlone), a cloning device that utilizes the capability of short-depth quantum circuits and the power of classical computation to learn the ability to clone specific set of states. This brings

into view a whole new domain of performing realistic implementation of attacks on quantum cryptographic systems. We note, however, that in order to fully implement realistic and practical attacks, one must consider all aspects of the protocol environment, including, for example, the input and output mechanisms to the quantum cloner. Incorporating VarQlone into the full analysis of the experimental implementation of quantum protocols, for example as in [BL13, B+17], is a fruitful avenue for future work.

We remark that our work opens new frontiers for analyzing quantum cryptographic schemes using quantum machine learning. In particular, this is applicable to secure quantum communication schemes which are becoming increasingly relevant in the quantum internet era.

We also note that one of the applications of our work is to find new cloning circuits that perform better on specific hardware. We specify that even though the states used in our examples and experiments are of low dimensions, in which case the emulator or the state vector machines will also provide the required result for cryptanalysis, using the VQA in VarQlone provides a circuit with better fidelity than the optimal theoretical (or emulated) circuits when run on real hardware, as we have seen in Fig. 7.3. This point establishes a unique use-case of quantum machine learning techniques for a problem which is quantum in nature.

We also believe that the tools we have developed in this study can be used to clone a new family of states with partial prior information, which leads to expanding our fundamental understanding of approximate cloning and unclonability in general. Therefore we conclude that finding new classes of cloners and their circuits is a potential use-case of our work and a new appealing future research direction.

8

Conclusion

"Imagination is the only weapon in the war with reality."
– Cheshire Cat - Alice in Wonderland

This thesis started with questions about unclonability in quantum mechanics and its role in quantum cryptanalysis. We were also seeking to grasp a deeper insight into the capacities of a quantum entity whose purpose is to attack the quantum and classical cryptosystems, specifically given the recent advancement in theoretical and experimental verges of the research on quantum technologies. In doing so, we have joined paths with various concepts such as unforgeability, unknownness, pseudorandomness, learnability, physical unclonability and variational algorithms, each of which has helped us to uncover a new connection to unclonability and let us to this point to conclude the thesis with a brief summary of this tortuous road. Hoping that the reader is not too weary by now, we will also discuss a general outlook and future direction.

In Chapter 3 we studied a more general notion of unclonability with a new perspective that related the unknownness of quantum states and processes to their unclonability. Following this direction, we correspondingly discussed the relation to learnability and eventually to a related cryptographic notion: unforgeability. We then devised a framework in which the quantum unforgeability of cryptographic primitives can be studied, regardless of being classical or quantum. This game-based security framework has been our cryptographic handbook in the majority of our security proofs throughout the thesis.

In Chapter 4, we met a new kind of unclonability, the physical unclonability, which we have formally defined as a mathematical concept in the quantum world. We managed to thoroughly study the *unpredictability* of some physical devices with the physical unclonability properties in terms of their unforgeability. Thus the notion of quantum physical unclonable functions has helped us to formalise some of our intuitions about unforgeability, unclonability and learnability, which we have put forward earlier. We have shown that the unforgeability of quantum PUFs is a provable consequence of their unknownness as hardware assumptions, which is unlike classical PUFs as they usually require assuming requirements that are morally equivalent to their unpredictability. Another point worth mentioning is the midst of proving the unforgeability of the unitary qPUF family, we appre-

hended the crucial role of quantum randomness. We have followed this thread in Chapter 5 where we studied the computational or cryptographic counterpart of quantum randomness, *i.e.*, quantum pseudorandomness. The investigation of quantum pseudorandomness concerning physical unclonability has uncovered interesting facts, some of which we conjecture and hope to be of interest in different areas of physics and cryptography.

As the thesis title includes "...: from foundations to applications", one would expect that this road reaches 'Applications' at some point! That point was Chapter 6, where we search for applications of quantum physical unclonable function. We provide several proposals that demonstrate the applicability and relevance of this notion in designing a new genre of quantum protocols: secure quantum protocols based on hardware assumptions. Even though our proposed protocols aim for rather non-complex functionality, they are significant as a building block for other more complicated protocols and functionalities. This factor is particularly relevant if one takes a modular and composable view over quantum protocols, which we believe should be the next era in quantum protocol design and security analysis.[1]

Finally, we aimed to utilise recent developments in quantum computing for the purpose of cryptanalysis. One of the most recent tools and topics of research in this area is quantum machine learning and variational quantum algorithms. As a result, in the last chapter (Chapter 7), we turned to another type of application: practical cryptanalysis using variational algorithms. However, since we could not resist a gaze into foundations, this chapter also includes questions and contributions regarding foundations and, more specifically, approximate quantum cloning. Our attempts led to the design of VarQlone, our variational quantum cloner, using which we have performed a cloning-based security analysis on different quantum protocols. Although we have demonstrated specific case studies, we argue that the foremost importance of this contribution is not the particular examples we have investigated, but rather the new method and mentality that it uses for cryptanalysis. We believe this method can be used in a handful of scenarios, and most importantly, for protocols and cryptosystems for which we do not have full security proof while having such a tool can provide valuable insight. Yet another significance of this work is its compatibility with NISQ devices since it would allow hardware-efficient and high-quality cloning circuits according to the available hardware.

At the end of each chapter, we have discussed the potential future direction, and remaining questions in each of the topics, which we do not intend to repeat

[1] Historically, quantum protocols have not been designed with this mindset. One reason perhaps, is that they have been developed separately and by very different communities (physics, cryptography, math). However, we have adopted this kind of modular view in the development of *quantum protocol zoo* [Ver19] where we have gathered and studied different quantum protocols and showed their composition into simpler subroutines and building blocks. This composition and modularity are beneficial in designing advanced functionalities, as well as security proofs (Especially in composable frameworks such as universal composability or abstract cryptography). We have excluded our contributions on this topic from the thesis to keep it more coherent, though we believed it would be worth a short remark, in the conclusion.

here. Instead, we discuss a broader outlook and prospective research direction in this field.

I believe one of the most exciting realms to step into, as also probably mentioned several times in the thesis, is the relationship between cryptography and learning theory in the quantum world. One evident reason is that learning theory offers powerful tools, both concretely theoretical and heuristic, which is maybe unconventional (compared to the approaches used in what I call 'hard-core cryptography') but exciting approaches towards cryptanalysis. Although the two fields have quite different *languages*, it appears to me that in many cases, they talk about closely related concepts, maybe from different perspectives. Thus, an idea that hopefully, this thesis has managed to convey to some extent is that there might be some level of correspondence between the two fields that could be capturable inside a new framework (maybe too naively and ambitiously). This sort of generalisation is, in my humble opinion, more feasible with quantum systems since; first, its mathematical framework is enough to include classical cases, and second, it inherently encompasses *physics* or the *actual systems* into the picture, while it is often neglected in cryptography or classical learning theory. Additionally, the recent works regarding learning different properties of the quantum states, or the relationship between quantum information and quantum machine learning, are an indication of this potential. An example of a concrete question I can propose here is proving tight bounds for quantum unforgeability using these tools, which would be in turn of interest in terms of learnability.

Next, let me go back to my other favourite research areas: physics and foundations. The history of quantum cryptography shows how physics can influence (and has influenced) cryptography. But can cryptography do the same? Can powerful mathematical techniques and frameworks of cryptography help us to better understand nature? Especially the cryptography that has been already armed with physical phenomena such as unclonability and entanglement. To this end, a deep understanding of concepts such as unclonability would become handy since it is both central to physics and cryptography. Yet another concept that we have discussed in this thesis and can be considered in this regime is quantum pseudorandomness. Computational randomness is a cryptographic concept, however, as we have discussed in Chapter 5, pseudorandom quantum states have already been of interest for fundamental physics and quantum gravity. One cannot help oneself to wonder if perhaps there is a deeper level to this "rabbit hole" (which would be a very convenient term if we were to study black holes, for instance). Maybe this is motivating enough for 'Alice's in the future to continue the study of quantum pseudorandomness and quantum physical unclonability in this context.

And finally, applications! Quantum hardware security (or quantum hardware cryptography) is a very young[2], yet promising field in terms of application. In this field, the ultimate goal is to exploit the unique properties of quantum hardware to reduce or remove computational assumptions or resource-intensive machinery

[2]I believe it did not really exist as a concrete field of research when I started my PhD, despite the fact that there have been several works in this area.

and achieve secure and efficient quantum protocols on this ground. This goal, however exciting, and despite our attempts in this area, is admittedly still not too close to reality. Regarding qPUFs, many open questions and potential extensions exist, among which I can mention the realisation of efficient and secure quantum PUFs, and certifying existing hardware to satisfy the criteria of qPUF as the most influential ones. Both are challenging and potentially intriguing problems that can help close the gap between the theoretical security analysis and the hardware implementation of a quantum PUF. Even considering our proposed hybrid PUF construction, which gives a significant practicality improvement, analysis of noise and several experimental aspects has remained untouched. A more general remark on the field is that PUF is not the only potential subject in hardware security, and the discovery and study of other existing hardware assumptions in the quantum setting can be remarkably fruitful.

I conclude this chapter and this thesis with a slightly less scientific and more personal point, as I let my doubtful inner scientist elaborate. Despite the considerable recent progress in building quantum computers, it is still a possibility that either the noisy nature of these systems or our technological limitations in other ways will defeat us in the conquest to achieve large scale and fault-tolerant quantum computers (in which case there will be no need to protect ourselves against them). Also, despite the strong complexity theory evidence, it is still a possibility for quantum computation to be proven to have no advantages over classical. Even more drastically, we might find ourselves in a situation where quantum mechanics turns out to be insufficiently correct or severely incomplete (which is a possibility every scientist concerning any scientific theory should be prepared for, even if one develops affection for the beauty of a theory like quantum mechanics). In the unlikely event of any of these happening in the future, I am aware that it will hugely affect the validity and relevance of this thesis. Yet this thesis and my works during the period of my PhD have still been a (hopefully meaningful) attempt toward understanding, and if any piece of this attempt will ever create any tiny bit of imagination, curiosity or excitement in anyone, I can hope that all this effort has not been in vain, as “Imagination is the only weapon in the war with reality.”

Appendix A

Additional proofs and derivations

A.1 Proof of Theorem 13 in Chapter 3

Here we give the full proof of Theorem 13 as follows:

Proof. We prove the theorem by induction. For the first block ($K=1$), according to Eq. (3.8) and letting $|\chi_0\rangle = |\psi\rangle$ we have:

$$|\chi_1\rangle = \frac{1}{2}[(\mathbb{I} - R(\phi_r))\,|\psi\rangle\,|0\rangle + R(\phi_i)(\mathbb{I} + R(\phi_r))\,|\psi\rangle\,|1\rangle] \tag{A.1}$$

where the term $\mathbb{I} - R(\phi_r) = 2\,|\phi_r\rangle\langle\phi_r|$ projects the previous state to $|\phi_r\rangle$ with the coefficient $\langle\phi_r|\psi\rangle$ and the term $R(\phi_i)(\mathbb{I} + R(\phi_r))$ is equal to:

$$R(\phi_i)(\mathbb{I} + R(\phi_r)) = 2[\mathbb{I} - |\phi_r\rangle\langle\phi_r| - 2\,|\phi_i\rangle\langle\phi_i| + 2\,\langle\phi_i|\phi_r\rangle\,|\phi_i\rangle\langle\phi_r|]. \tag{A.2}$$

Thus, the final relation between all the parameters in the first block is as follows.

$$\begin{aligned}|\chi_1\rangle = \langle\phi_r|\psi\rangle\,|\phi_r\rangle\,|0\rangle + |\psi\rangle\,|1\rangle - \langle\phi_r|\psi\rangle\,|\phi_r\rangle\,|1\rangle \\ - 2\,\langle\phi_1|\psi\rangle\,|\phi_1\rangle\,|1\rangle + 2\,\langle\phi_r|\psi\rangle\,\langle\phi_r|\phi_1\rangle\,|\phi_1\rangle\,|1\rangle\end{aligned} \tag{A.3}$$

As can be seen, it satisfies the form of Eq. (3.9) where the first sum is zero and in the second sum $g_{10} = -1, g_{11} = +1$, $l'_{10} = l'_{11} = 1$, $x'_{10} = z'_{10} = 0, y'_{10} = 1$, $x'_{11} = z'_{11} = 1$ and $y'_{11} = 0$.

Now we write $|\chi_K\rangle$ according to recursive relation of Eq. (3.8). We assume $|\chi_{K-1}\rangle$ is written in form of Eq. (3.9) and show $|\chi_K\rangle$ also satisfies this equation.

$$\begin{aligned}|\chi_K\rangle = \langle\phi_r|\chi_{K-1}\rangle\,|\phi_r\rangle\,|0\rangle + |\chi_{K-1}\rangle\,|1\rangle - \langle\phi_r|\chi_{K-1}\rangle\,|\phi_r\rangle\,|1\rangle - 2\,\langle\phi_K|\chi_{K-1}\rangle\,|\phi_K\rangle\,|1\rangle \\ + 2\,\langle\phi_r|\chi_{K-1}\rangle\,\langle\phi_r|\phi_K\rangle\,|\phi_K\rangle\,|1\rangle\end{aligned} \tag{A.4}$$

By substituting $|\chi_{K-1}\rangle$ with its equivalent based on Eq. (3.9), we calculate each term in the above formula. Note that the coefficient in the third term is the same as the first one with a minus sign, and the ancillary state for the first term is $|0\rangle$

while for the third term is $|1\rangle$. Thus, we only show the details of the calculation for the first term:

$$\begin{aligned}\langle\phi_r|\rangle\chi_{K-1}|\phi_r\rangle|0\rangle = \\ \langle\phi_r|\psi\rangle|\phi_r\rangle|0\rangle^{\otimes K} + \langle\phi_r|\psi\rangle|\phi_r\rangle|1\rangle^{\otimes K-1}|0\rangle - \langle\phi_r|\psi\rangle|\phi_r\rangle|1\rangle^{\otimes K-1}|0\rangle + \\ + \sum_{i=1}^{K-1}\sum_{j=0}^{i}[f_{ij}2^{l_{ij}}|\langle\phi_r|\psi\rangle|^{x_{ij}}|\langle\phi_i|\psi\rangle|^{y_{ij}}|\langle\phi_r|\phi_i\rangle|^{z_{ij}}]|\phi_r\rangle|q_{anc}(i,j)\rangle|0\rangle \\ + \sum_{i=1}^{K-1}\sum_{j=0}^{i}[g_{ij}2^{l'_{ij}}|\langle\phi_r|\psi\rangle|^{x'_{ij}}|\langle\phi_i|\psi\rangle|^{y'_{ij}}|\langle\phi_r|\phi_i\rangle|^{z'_{ij}+1}]|\phi_i\rangle|q'_{anc}(i,j)\rangle|0\rangle\,.\end{aligned} \tag{A.5}$$

The second term is calculated as follows:

$$\begin{aligned}|\chi_{K-1}\rangle|1\rangle = \langle\phi_r|\psi\rangle|0\rangle^{\otimes K-1}|1\rangle + |\psi\rangle|1\rangle^{\otimes K} - \langle\phi_r|\psi\rangle|\phi_r\rangle|1\rangle^{\otimes K} + \\ + \sum_{i=1}^{K-1}\sum_{j=0}^{i}[f_{ij}2^{l_{ij}}|\langle\phi_r|\psi\rangle|^{x_{ij}}|\langle\phi_i|\psi\rangle|^{y_{ij}}|\langle\phi_r|\phi_i\rangle|^{z_{ij}}]|\phi_r\rangle|q_{anc}(i,j)\rangle|1\rangle \\ + \sum_{i=1}^{K-1}\sum_{j=0}^{i}[g_{ij}2^{l'_{ij}}|\langle\phi_r|\psi\rangle|^{x'_{ij}}|\langle\phi_i|\psi\rangle|^{y'_{ij}}|\langle\phi_r|\phi_i\rangle|^{z'_{ij}}]|\phi_i\rangle|q'_{anc}(i,j)\rangle|1\rangle\,.\end{aligned} \tag{A.6}$$

The forth term $-2\langle\phi_K|\chi_{K-1}\rangle|\phi_K\rangle|1\rangle$ has the coefficient $-2\langle\phi_K|\chi_{K-1}\rangle$, which produces the same sigma terms while only $l'_{i,j}, x'_{i,j}, y'_{i,j}$ and $z'_{i,j}$ are increased by one. The fifth term $2\langle\phi_r|\chi_{K-1}\rangle\langle\phi_r|\phi_K\rangle|\phi_K\rangle|1\rangle$ has the coefficient $2\langle\phi_r|\chi_{K-1}\rangle\langle\phi_r|\phi_K\rangle$ and similarly produces the same sigma terms where $l_{i,j}$, $x_{i,j}$, $y_{i,j}$ and $z_{i,j}$ are increased by one (Note that the $\langle\phi_r|\phi_K\rangle$ is itself one of the terms of the sigma). Finally by putting all these terms together, Eq. (3.9) is obtained which completes the proof. □

A.2 Proof of Theorem 15 in Chapter 3

Proof. To show this implication we will show that if a QPT adversary $\mathcal{A}$ can win in qGUU, then $\mathcal{A}$ can also win against μ-qGSU. Although for simplicity we restrict the proof for the case of $\mu = 1$ and the generalisation to any μ is straightforward from the hierarchy of the definition for different μ shown in Theorem 14. Also, we recall that 1-qGSU and 1-qGEU are equivalent. Let $\mathcal{A}$ play the game $\mathcal{G}^{\mathcal{F}}_{\mathsf{qUni}}(\lambda,\mathcal{A})$ by picking a set of learning phase state $\{|\phi_i\rangle\}_{i=1}^{K}$. Let the dimension of the unitary oracle $\mathcal{O}^{\mathcal{E}}$ be $D = 2^n$ and let the subspace of σ_{in} be of dimension $d = poly(n)$. If $\mathcal{A}$ wins the game, then the average probability of $\mathcal{A}$ generating an acceptable output for any $x \in \mathcal{M}$ picked uniformly at random by $\mathcal{C}$ is non-negligible:

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qUni}}(\lambda,\mathcal{A})] = \Pr_{x\in\mathcal{M}}[1 \leftarrow \mathcal{A}(x)] = non\text{-}negl(\lambda). \tag{A.7}$$

where $\Pr_{x\in\mathcal{M}}[1 \leftarrow \mathcal{A}(x)]$ denotes the success probability of the adversary winning the game for input x. Now to be able to translate this game to the 1-qGSU game,

first, we need to make sure that the set of states that $\mathcal{A}$ picks the challenge from them, satisfies the distinguishability condition for $\mu = 1$ i.e. they are orthogonal to all the learning phase states. Let $\mathcal{M}'$ be the set of all the challenges with no overlap with any of the learning phase states ρ_i^{in}. Then we can rewrite the average success probability as follows:

$$\begin{aligned} \Pr_{x\in\mathcal{M}}[1 \leftarrow \mathcal{A}(x)] &= \Pr_{x\in\mathcal{M}'}[1 \leftarrow \mathcal{A}(x)]Pr[x \in \mathcal{M}'] + \Pr_{x\notin\mathcal{M}'}[1 \leftarrow \mathcal{A}(x)]Pr[x \notin \mathcal{M}'] \\ &= \textit{non-negl}(\lambda). \end{aligned} \tag{A.8}$$

since the dimension of the subspace that σ_{in} spans is d and it is polynomial with respect to the size of $\mathcal{M}$ then $\frac{|\mathcal{M}'|}{|\mathcal{M}|} \approx 1$. Hence $Pr[x \in \mathcal{M}'] \approx 1$ but $Pr[x \notin \mathcal{M}'] = 1 - Pr[x \in \mathcal{M}'] = \textit{negl}(\lambda)$. As a result the second term will be negligible and for the whole expression to become non-negligible, the following should hold:

$$\Pr_{x\in\mathcal{M}'}[1 \leftarrow \mathcal{A}(x)] = \textit{non-negl}(\lambda). \tag{A.9}$$

Now let $\mathcal{A}'$ be an adversary who wants to win the game $\mathcal{G}^{\mathcal{F}}_{\mathsf{qSel},\mu}(\lambda, \mathcal{A}')$ by using $\mathcal{A}$. As $\mathcal{A}'$ picks the challenge of their choice, we will show that there is a strategy for $\mathcal{A}'$ to win the game relying on the average success probability of $\mathcal{A}$ being non-negligible over $\mathcal{M}'$. But also as $\mathcal{A}'$ is a QPT, we will show there exist a poly size subspace of $\mathcal{M}'$ in which $\mathcal{A}'$ will win with non-negligible probability. First we assume that $\mathcal{M}'$ is partitioned into K different subsets (or subspaces) S_i with equal size (or dimension in the quantum case) $|S_1| = \cdots = |S_K| = l = \textit{poly}(\lambda)$. Note that this partitioning is only for simplicity and any random partitioning of $\mathcal{M}'$ into the equal size subspace will be enough for our purpose. Now let $\mathcal{A}'$ pick one of the subsets of message space which consists of picking one of the S_i with probability $\frac{1}{K}$. We want to show that if $\mathcal{A}'$ picks the S_i at random and calls $\mathcal{A}$ on that S_i the probability that in the picked subspace the following condition holds is non-negligible:

$$\Pr_{x\in S_i}[1 \leftarrow \mathcal{A}(x)] = \textit{non-negl}(\lambda) \tag{A.10}$$

If this is the case, then by the definition of the average probability there exists at least one x^* for which the $Pr[1 \leftarrow \mathcal{A}(x^*)] = \textit{non-negl}(\lambda)$ and hence the $\mathcal{A}'$ has won the game with a non-negligible probability. Thus we need to find the number of the success probability of $\mathcal{A}'$ picking a desirable subset. This probability is given by:

$$Pr_{succ} = \frac{\#(S_i : \Pr_{x\in S_i}[1 \leftarrow \mathcal{A}(x)] = \textit{non-negl}(\lambda))}{K} = \frac{Q}{K} \tag{A.11}$$

where Q denotes the number of subsets S_i which satisfy the condition and $K = O(|\mathcal{M}'|)$. We then only need to show that $\frac{Q}{K}$ is non-negligible in the security parameter. For simplicity let us replace average probability of $\mathcal{A}$ in wining the game over $\mathcal{M}'$, with the expected value of wining probability of $\mathcal{A}$ over all the different elements of $\mathcal{M}'$ i.e.

$$\Pr_{x\in\mathcal{M}'}[1 \leftarrow \mathcal{A}(x)] = \textit{non-negl}(\lambda) \Rightarrow \mathop{\mathbb{E}}_{\mathcal{M}'}[\mathcal{A}(x)] = \textit{non-negl}(\lambda) \tag{A.12}$$

Then we rewrite the expectation value in terms of all the subsets of $\mathcal{M}'$. As $\mathcal{M}' = S_1 \cup S_2 \cup \cdots \cup S_K$, we have:

$$\mathop{\mathbb{E}}_{\mathcal{M}'}[\mathcal{A}(x)] = \frac{1}{K}\sum_{i=1}^{K}\mathbb{E}_i = \textit{non-negl}(\lambda) \tag{A.13}$$

where $\mathbb{E}_i = \mathop{\mathbb{E}}_{S_i}[\mathcal{A}(x)]$. We then rearrange all the $\mathbb{E}_i$ descending such that the Qth term shows the last smallest $\mathbb{E}_i$ for which the condition is satisfied. Hence we have:

$$\mathop{\mathbb{E}}_{\mathcal{M}'}[\mathcal{A}(x)] = \frac{1}{K}\sum_{i=1}^{Q}\mathbb{E}_i + \frac{1}{K}\sum_{i=Q+1}^{K}\mathbb{E}_i = \textit{non-negl}(\lambda) \tag{A.14}$$

The above equality holds if at least one of the two sums is non-negligible. If the first sum is non-negligible we have:

$$\frac{1}{K}\sum_{i=1}^{Q}\mathbb{E}_i \geqslant \frac{Q\mathbb{E}_Q}{K} \tag{A.15}$$

As $\mathbb{E}_i$s have been ordered and $\mathbb{E}_Q$ is the smallest one which is still non-negligible. Then we can conclude that:

$$\frac{Q}{K} = \textit{non-negl}(\lambda) \tag{A.16}$$

which is what we wanted to show. The second case is when the first sum is negligible and the second sum needs to be non-negligible for the equality to hold. Similar to the previous case due to the descending ordering, we have:

$$\frac{1}{K}\sum_{i=Q+1}^{K}\mathbb{E}_i \leqslant \frac{(K-Q)\mathbb{E}_{Q+1}}{K} \tag{A.17}$$

But followed by our assumption the $\mathbb{E}_{Q+1}$ is itself negligible and $0 < \frac{K-Q}{K} < 1$, thus this sum can never converge to a non-negligible function of λ. Hence we conclude that necessarily the first sum, and as a result, $\frac{Q}{K}$ is non-negligible. Thus we have shown the equation A.10, and there exists a strategy for $\mathcal{A}'$ to win the game by calling $\mathcal{A}$. This concludes that 1-qGSU(μ-qGSU) implies qGUU and the proof is complete. □

A.3 Proof of Theorem 16 in Chapter 3

Proof. We show that 1-qGEU implies BU and vice versa. First, we show that if a scheme is not BU unforgeable against a QPT adversary then it is not 1-qGEU unforgeable either. Let $\mathcal{A}$ be a QPT adversary who forges a scheme $\mathcal{F} = (\mathcal{S}, \mathcal{E}, \mathcal{V})$ with message set $\mathcal{M} = \{0,1\}^n$ in the BU definition. Following the formal definition of BU provided in Definition 23, $\mathcal{A}$ selects an ε for which the blinded region $\mathcal{B}_\varepsilon$ is created by selecting each $m \in \mathcal{M}$ at random with an ε-related probability. Then there exists a non-empty set $\mathcal{B}_\varepsilon$ for which $\mathcal{A}$ interacts with the blinded oracle

associated with it and outputs a pair (m^*, t^*) where $t^* = f(m^*)$ (where f is the classical function of the evaluation $\mathcal{E}$, for instance a $MAC(.)$) such that $\mathcal{V} = Ver_k(m^*, t^*) = acc$, and also the $m^* \in \mathcal{B}_\varepsilon$ with non-negligible probability in $\lambda = poly(n)$. By the definition of the blinding oracle, $\mathcal{A}$ receives a $|\perp\rangle$ for any of the computational basis that is in the blinded region. As a result, we can write $\mathcal{A}$'s input and output queries as follows:

$$\begin{aligned} |\phi_i\rangle &= \sum_{m_i \notin \mathcal{B}_\varepsilon} \alpha_i |m_i, y_i\rangle + \sum_{\overline{m}_j \in \mathcal{B}_\varepsilon} \beta_j |\overline{m}_j, y_j\rangle \\ |\phi_i^{out}\rangle &= \sum_{m_i \notin \mathcal{B}_\varepsilon} \alpha_i |m_i, y_i \oplus f(m_i)\rangle + \sum_{\overline{m}_j \in \mathcal{B}_\varepsilon} \beta_j |\overline{m}_j, y_j \oplus \perp\rangle \end{aligned} \tag{A.18}$$

Now assuming the quantum encoding of the challenge $m^* \in \mathcal{B}_\varepsilon$ to be $|m^*, 0\rangle$ and the tag/output to be $|m^*, t^*\rangle = |m^*, f(m^*)\rangle$, we can see that $\langle m^*, t^*|\phi_i^{out}\rangle = 0$ since m^* will have no overlap with the first part of the superposition, and also to the second part due to the blinding. Now, we show that there exists a unitary non-blinding oracle that generates equivalent queries for this scenario. Let $\mathrm{U}_\mathcal{E}$ be the unitary evaluation oracle such that $|m^*, t^*\rangle = \mathrm{U}_\mathcal{E}|m^*, 0\rangle$, and similarly for all the queries. Due to the unitarity, we have that $\langle m^*, t^*|\phi_i^{out}\rangle = \langle m^*, 0|\mathrm{U}_\mathcal{E}{}^\dagger \mathrm{U}_\mathcal{E}|\phi_i\rangle = 0$. Thus there will also exist an adversary $\mathcal{A}'$ with equivalent queries except that the target forgery will be always orthogonal to the selected challenge. Hence for this adversary, the condition of 1-qGEU is satisfied. Then by calling $\mathcal{A}$, the adversary $\mathcal{A}'$ can generate an output state $|m^*, t^*\rangle$ that passes the test algorithm with also non-negligible probability. Hence we have shown that 1-qGEU implies BU.

To prove the other way of implication we need to show whenever there is an attack on 1-qGEU, then there will also be an attack on BU definition and hence the scheme is also BU insecure. This time we consider $\mathcal{A}$ to be a QPT adversary who wins 1-qGEU by selecting a challenge state $|m^*, y\rangle$ where the m^* is the classical challenge and y is the ancillary register, and querying a set of states $\{|\phi_i\rangle\}_{i=1}^q$ s.t. $\forall|\phi_i\rangle : \langle m^*|\phi_i\rangle = 0$ and $q = poly(n)$. Then by definition, $\mathcal{A}$ can output a $|m^*, t^*\rangle = \mathrm{U}_\mathcal{E}|m^*, y\rangle$ that passes the test algorithm with non-negligible probability. Now an adversary $\mathcal{A}'$ calls $\mathcal{A}$ to win the BU with non-negligible probability.

At this stage we recall the Theorem 10 and we show that an $\mathcal{A}'$ satisfies the conditions of this theorem. Let us write the learning phase queries in the computational basis as follows:

$$|\phi_i^{out}\rangle = \sum_{j=1}^{d} \alpha_{i,j} |b_j\rangle \tag{A.19}$$

where $\{|b_j\rangle\}_{j=1}^d$ is the set of computational bases spanning the effective learning phase subspace. Now we create a non-empty set R by selecting each $x \in \mathcal{M}$ as follows

$$R = \{x \in \mathcal{M} : |x\rangle \neq |b_j\rangle_X\} \tag{A.20}$$

Where $|b_j\rangle_X$ denotes the input register of the full basis. Note that R will always be non-empty as the basis set will only cover a polynomial-size subspace of the whole Hilbert space of messages. Moreover, since $\mathcal{A}'$ includes $\mathcal{A}$ and m^* has no

overlap with any of the input queries, it will also have no overlap with the input register of the output queries. As a result, R has at least one element. Hence the set of all input elements that have non-zero overlap with the queries and the elements included in R have no intersection. This shows that $supp(\mathcal{A}) \cap R = \emptyset$ if the support is defined for the oracle $\mathcal{O}_f$ for a fixed randomly picked classical function f (or key k) during the game. Thus we also have $supp(\mathcal{A}') \cap R = \emptyset$ and $m^* \in R$. Nevertheless, in [AMRS20] it has been mentioned that the support is taken to be the union of the support of all the queries over the choice of the function. In this case, we can also redefine our set and the queries of $\mathcal{A}'$ such that it satisfies the condition of the theorem respectively. We take the set R' to only include one element which is the forgery message m^*. As in the 1-qGEU the function (or the key for the keyed functions) is selected at random in the setup phase, the success probability of $\mathcal{A}$ is inherently taken over the choice of the function. Then $\mathcal{A}'$ queries all the queries of $\mathcal{A}$ for any randomly selected f during the experiment. For any other functions, excludes any queries for which the support will include m^*. Now we can see that $\mathcal{A}'$ can output a valid pair (m^*, t^*) by measuring $|m^*, t^*\rangle$ in the computational basis with probability 1 while $supp(\mathcal{A}') \cap R' = \emptyset$ and $m^* \in R'$. Hence $\mathcal{A}'$ breaks the BU unforgeability and we have shown that BU implies 1-qGEU. This mutual implication shows that these definitions are equivalent and the proof is complete. □

A.4 Alternative model for an adaptive quantum adversary

In this appendix, we introduce an alternative way for capturing full quantum adaptive adversaries. Here we also consider QPT adversaries who have q-query access to the evaluation function of a primitive $\mathcal{F}$, namely $\mathcal{E}$ where q is polynomial in the security parameter. An adaptive adversary can choose and issue any arbitrary query which could also depend on the previous responses received from the black-box oracle. An adaptive quantum adversary is likely to consume the quantum state of the response to be able to pick the next query adaptively. Hence modeling the post-query database of an adaptive quantum adversary is more challenging. In what follows we give a q-query mathematical model for adaptive adversaries.

Definition 54. Let q be a positive integer, and $\mathcal{E} : \mathcal{H}^{d_{\text{in}}} \to \mathcal{H}^{d_{\text{out}}}$ be a quantum evaluation. We model a probabilistic adversary as a CPTP map $\mathcal{A} : \mathcal{R} \times (\mathcal{H}^{d_{\text{in}}})^{\otimes q} \otimes (\mathcal{H}^{d_{\text{out}}})^{\otimes q} \to (\mathcal{H}^{d_{\text{in}}})$. Such an adversary is called an **adaptive** adversary $\mathcal{A}_{ad}$ if for all random coin $r \in \mathcal{R}$ and for any $\bigotimes_{i=1}^{q} \rho_i^{\text{in}} \in (\mathcal{H}^{d_{\text{in}}})^{\otimes q}$ and for $\bigotimes_{i=1}^{q} \rho_i^{\text{out}} \in (\mathcal{H}^{d_{\text{out}}})^{\otimes q}$ (where $\rho_i^{\text{out}} := \mathcal{E}(\rho_i^{\text{in}})$), the mapping $\bigotimes_{i=1}^{q} (\rho_i^{\text{in}} \otimes \rho_i^{\text{out}}) \to \mathcal{A}_{ad}^{r}(\bigotimes_{i=1}^{q} (\rho_i^{\text{in}} \otimes \rho_i^{\text{out}}))$ is dependent on the $\rho_1^{\text{in}} \otimes \rho_1^{\text{out}}, \ldots, \rho_q^{\text{in}} \otimes \rho_q^{\text{out}}$;

Intuitively, an adaptive adversary $\mathcal{A} : \mathcal{R} \times (\mathcal{H}^{d_{\text{in}}})^{\otimes q} \otimes (\mathcal{H}^{d_{\text{out}}})^{\otimes q} \to (\mathcal{H}^{d_{\text{in}}})$ cap-

tures the strategy to choose the query input $\rho_{q+1}^{\text{in}} \in \mathcal{H}^{d_{\text{in}}}$ to $\mathcal{E}$. The adversary can use these query response pairs to predict the output of $\mathcal{E}$. We call the pair $(\bigotimes_{i=1}^{q} \rho_i^{\text{in}}, \bigotimes_{i=1}^{q} \rho_i^{\text{out}})$ that is generated after the q-round of interaction between an adversary $\mathcal{A}$ and $\mathcal{E}$, as a transcript. Note, that the transcripts depend on the choice of the random coins of $\mathcal{A}$.

However, since this model is more complicated to work with, we use our usual notation used in Chapter 3.

A.5 Proof of Theorem 23 in Chapter 3

Proof. Let $\mathcal{A}$ be the QPT adversary playing the game $\mathcal{G}^{\mathcal{F}}_{qGUU-aua,\mu}(\lambda, \mathcal{A})$ and running the algorithm described in Algorithm 4.

Algorithm 4 aua attack on qGUU

- **First learning phase:** null
- **Challenge phase:**
  - prepare qubit $|0\rangle_a$
  - receive $|\psi_m\rangle$ as a challenge
- **Second learning phase:**
  - $|\Psi\rangle_{ca} = CNOT_{c,a}(|\psi_m\rangle|0\rangle)$[1]
  - query register c ($\mathcal{A}$ sends the challenge part of the entangled system, ρ_c as a query.)
  - receive $U_{\mathcal{E}}\rho_c U_{\mathcal{E}}^{\dagger}$ or $(U_{\mathcal{E}} \otimes \mathcal{I})|\Psi\rangle_{ca}$
- **Guess phase:**
  - $|\psi_m^{out}\rangle \otimes |\pm\rangle \leftarrow Measure(|\Psi\rangle_{ca}, \{|\pm\rangle\})$
  - **if** $|\pm\rangle = |+\rangle$
    - **output:** $|t\rangle = |\psi_m^{out}\rangle$
  - **else**
    - **output:** $|t\rangle = CZ^{\otimes n-1}(|\psi_m^{out}\rangle)$
  - $Measure(|\Psi\rangle_{ca}, \{|\pm\rangle\}$ outputs the result of the measurement.

$\mathcal{A}$ does not issue any query during the first learning phase. Then $\mathcal{A}$ receives an unknown challenge state $|\psi_m\rangle = \sum_{i=1}^{D} \alpha_i |b_i\rangle$ where $\{|b_i\rangle\}_{i=1}^{D}$ is a set of complete orthonormal bases for $\mathcal{H}^D$. Now, $\mathcal{A}$ prepares state $|0\rangle$ and performs a CNOT gate on the first qubit of the unknown challenge state and the ancillary qubit ($|0\rangle$) with the control qubit on the challenge state. We can assume the order of the bases is such that in the first half, the first qubit is $|0\rangle$ and in the second half the first

qubit is $|1\rangle$. Then the output entangled state is

$$|\Psi\rangle_{ca} = \sum_{i=1}^{D/2} \alpha_i |b_i\rangle_c \otimes |0\rangle_a + \sum_{i=\frac{D}{2}+1}^{D} \alpha_i |b_i\rangle_c \otimes |1\rangle_a$$

Now we can compute the final state of the two systems after the second learning phase which is:

$$|\Psi^{out}\rangle_{ca} = \sum_{i=1}^{D/2} \alpha_i (\mathsf{U}_{\mathcal{E}} \otimes \mathbb{I})(|b_i\rangle_c \otimes |0\rangle_a) + \sum_{i=\frac{D}{2}+1}^{D} \alpha_i (\mathsf{U}_{\mathcal{E}} \otimes \mathbb{I})(|b_i\rangle_c \otimes |1\rangle_a).$$

By rewriting the first qubit in the $|+\rangle$ basis we have

$$|\psi_m^{out}\rangle = [\mathsf{U}_{\mathcal{E}}(\sum_{i=1}^{D} \alpha_i |b_i\rangle_c)]\frac{|+\rangle}{\sqrt{2}} + [\mathsf{U}_{\mathcal{E}}(\sum_{i=1}^{D/2} \alpha_i |b_i\rangle_c - \sum_{i=\frac{D}{2}+1}^{D} \alpha_i |b_i\rangle_c)]\frac{|-\rangle}{\sqrt{2}}.$$

Then, the adversary measures his local qubit in the $\{|+\rangle, |-\rangle\}$ bases. If he obtains $|+\rangle$, the state collapses to $\mathsf{U}_{\mathcal{E}}(\sum_{i=1}^{D} \alpha_i |b_i\rangle_c) = \mathsf{U}_{\mathcal{E}} |\psi_m\rangle$ that is the desired state with fidelity 1. If the output of the measurement is $|-\rangle$, half of the terms have a minus sign. In this case, $\mathcal{A}$ applies a controlled-Z gate on the second half of the state to obtain again $\mathsf{U}_{\mathcal{E}} |\psi_m\rangle$. As a result, for any κ_1 and κ_2, we have:

$$Pr[1 \leftarrow \mathcal{G}^{\mathcal{F}}_{\mathsf{qUni}-aua,\mu}(\lambda, \mathcal{A})] = Pr[1 \leftarrow \mathcal{T}((\mathsf{U}_{\mathcal{E}} |\psi_m\rangle)^{\otimes \kappa_1}, |t\rangle^{\otimes \kappa_2})] = 1.$$

Now to complete the proof, we show that the μ-distinguishability is satisfied on average. We need to calculate the reduced density matrix of this state and compare it with the density matrix $\rho_\psi = |\psi\rangle\langle\psi|$ in terms of the Uhlmann's fidelity. The reduced density matrix of the challenge state can be calculated as follows:

$$\rho_c = Tr_a[|\psi\rangle\langle\psi|_{ca}] = \sum_{i=1}^{D} |\alpha_i|^2 |b_i\rangle\langle b_i| + \sum_{i=j=1}^{\frac{D}{2}} \sum_{j\neq i, j=\frac{D}{2}+1}^{D} \overline{\alpha_i}\alpha_j |b_i\rangle\langle b_j| +$$

$$\sum_{i=\frac{D}{2}+1}^{D} \sum_{j\neq i, j=1}^{\frac{D}{2}} \overline{\alpha_i}\alpha_j |b_i\rangle\langle b_j|$$

where Tr_a denoted the partial trace taken over the adversary's sub-system. And the first sum shows the diagonal terms of the density matrix. As it can be seen these density matrices are different in half of the non-diagonal terms with the ρ_ψ. According to the Uhlmann's fidelity definition in the preliminary, and the fact that $|\psi\rangle$ is a pure state the fidelity reduce to:

$$F(\rho_\psi, \rho_c) = [Tr(\sqrt{\sqrt{\rho_\psi}\rho_c\sqrt{\rho_\psi}})]^2 = \langle\psi|\rho_c|\psi\rangle = \sum_{i=1}^{D} |\alpha_i|^2 \langle b_i|\rho_c|b_i\rangle.$$

By substituting the ρ_c from above, the result will be as follows:

$$F(\rho_\psi, \rho_c) = \sum_{i=1}^{D} |\alpha_i|^4 + \sum_{i=1}^{\frac{D}{2}} \sum_{j=\frac{D}{2}+1}^{D} 2|\alpha_i\alpha_j|^2 = 1 - \sum_{i=1}^{\frac{D(D-1)}{4}} 2|\gamma_i|^2$$

where $|\gamma_i|^2$ denoted the square of a quarter of the non-diagonal elements of ρ_ψ. This is a positive value and on average over all the state $|\psi\rangle$, non-negligible compared to the dimensionality of the state. Hence:

$$F(\rho_\psi, \rho_c) \leqslant 1 - non\text{-}negl(\lambda)$$

and the distinguishability condition is satisfied and the proof is complete. □

A.6 Calculation of the cost function's gradient for VarQlone

In this appendix we calculate the gradient of the cost functions we have introduced in Section 7.3.2 of Chapter 7, using the techniques introduced in Section 2.6.4.4.

We remind the squared cost:

$$C_{\text{sq}}^{M\to N}(\boldsymbol{\theta}) := \mathop{\mathbb{E}}_{|\psi\rangle\in S}\left[\sum_{i=1}^{N}(1-F_{\text{L}}^{i}(\boldsymbol{\theta}))^2 + \sum_{i<j}^{N}(F_{\text{L}}^{i}(\boldsymbol{\theta}) - F_{\text{L}}^{j}(\boldsymbol{\theta}))^2\right] \tag{A.21}$$

and its partial derivative from Eq. (7.46)

$$\frac{\partial C_{\text{sq}}(\boldsymbol{\theta})}{\partial\theta_l} = 2\mathop{\mathbb{E}}_{|\psi\rangle\in S}\left[\sum_{i=1}^{N}(1-F_{\text{L}}^{i}(\boldsymbol{\theta}))\left[-\frac{\partial F_{\text{L}}^{i}(\boldsymbol{\theta})}{\partial\theta_l}\right] + \sum_{i<j}^{N}(F_{\text{L}}^{i}(\boldsymbol{\theta}) - F_{\text{L}}^{j}(\boldsymbol{\theta}))\left[\frac{\partial F_{\text{L}}^{i}(\boldsymbol{\theta})}{\partial\theta_l} - \frac{\partial F_{\text{L}}^{j}(\boldsymbol{\theta})}{\partial\theta_l}\right]\right] \tag{A.22}$$

Now we assume that each $U(\boldsymbol{\theta}) := U(\theta_d)U(\theta_{d-1})\dots U(\theta_1)$ is composed of unitary gates of the form: $U(\theta_l) = \exp(-i\theta_l\Sigma_l)$, where $\Sigma_l^2 = \mathbb{1}$ (for example, a tensor product of Pauli operators). We can use the *parameter shift rule* from Theorem 12, we get:

$$\frac{\partial U(\boldsymbol{\theta})\rho_{\text{init}}U(\boldsymbol{\theta})^\dagger}{\partial\theta_l} = U^{l+\frac{\pi}{2}}(\boldsymbol{\theta})\rho_{\text{init}}(U(\boldsymbol{\theta})^{l+\frac{\pi}{2}})^\dagger - U^{l-\frac{\pi}{2}}(\boldsymbol{\theta})\rho_{\text{init}}(U(\boldsymbol{\theta})^{l-\frac{\pi}{2}})^\dagger \tag{A.23}$$

Where the notation, $U^{l\pm\frac{\pi}{2}}$, indicates the l^{th} parameter has been shifted by $\pm\frac{\pi}{2}$, i.e. $U^{l\pm\frac{\pi}{2}} := U(\theta_d)U(\theta_{d-1})\dots U(\theta_l \pm \pi/2)\dots U(\theta_1)$. We get:

$$\begin{aligned}\frac{\partial F_{\text{L}}^{j}(\boldsymbol{\theta})}{\partial\theta_l} &= \text{Tr}\left[|\psi\rangle\langle\psi|\,\text{Tr}_{\bar{j}}\left(U^{l+\frac{\pi}{2}}(\boldsymbol{\theta})\rho_{\text{init}}(U(\boldsymbol{\theta})^{l+\frac{\pi}{2}})^\dagger\right)\right] \\ &\quad - \text{Tr}\left[|\psi\rangle\langle\psi|\,\text{Tr}_{\bar{j}}\left(U^{l-\frac{\pi}{2}}(\boldsymbol{\theta})\rho_{\text{init}}(U(\boldsymbol{\theta})^{l-\frac{\pi}{2}})^\dagger\right)\right]\end{aligned} \tag{A.24}$$

$$\begin{aligned}\implies \frac{\partial F_{\text{L}}^{j}(\boldsymbol{\theta})}{\partial\theta_l} &= \text{Tr}\left[|\psi\rangle\langle\psi|\,\rho_j^{l+\frac{\pi}{2}}(\boldsymbol{\theta})\right] - \text{Tr}\left[|\psi\rangle\langle\psi|\,\rho_j^{l-\frac{\pi}{2}}(\boldsymbol{\theta})\right] \\ &= F_{\text{L}}^{(j,l+\frac{\pi}{2})}(\boldsymbol{\theta}) - F_{\text{L}}^{(j,l-\frac{\pi}{2})}(\boldsymbol{\theta})\end{aligned}$$

where we define $F_j^{(l\pm\frac{\pi}{2})}(\boldsymbol{\theta}) := \langle\psi|\,\rho_j^{l\pm\frac{\pi}{2}}(\boldsymbol{\theta})\,|\psi\rangle$ the fidelity of the j^{th} clone, when prepared using a unitary whose l^{th} parameter is shifted by $\pm\frac{\pi}{2}$, with respect to a target input state, $|\psi\rangle$.

Plugging this into Eq. (A.22), we get:

$$\frac{\partial \mathsf{C}_{\mathsf{sq}}(\boldsymbol{\theta})}{\partial \theta_l} = \underset{|\psi\rangle\in S}{2\mathbb{E}} \left[\sum_{i<j}^{N} (F_{\mathsf{L}}^{i} - F_{\mathsf{L}}^{j}) \left[F_{\mathsf{L}}^{(i,l+\frac{\pi}{2})} - F_{\mathsf{L}}^{(i,l-\frac{\pi}{2})} - F_{\mathsf{L}}^{(j,l+\frac{\pi}{2})} + F_{\mathsf{L}}^{(j,l-\frac{\pi}{2})} \right] \right. \tag{A.25}$$

$$\left. - \sum_{i=1}^{N} (1 - F_{\mathsf{L}}^{i}) \left[F_{\mathsf{L}}^{(i,l+\frac{\pi}{2})} - F_{\mathsf{L}}^{(i,l-\frac{\pi}{2})} \right] \right] \tag{A.26}$$

Using the same method, we can also derive the gradient of the local cost, Eq. (7.43) with N output clones as:

$$\frac{\partial \mathsf{C}_{\mathsf{L}}(\boldsymbol{\theta})}{\partial \theta_l} = \mathbb{E} \left(\sum_{i=1}^{N} \left[F_{\mathsf{L}}^{i,l-\pi/2} - F_{\mathsf{L}}^{i,l+\pi/2} \right] \right) \tag{A.27}$$

Finally, similar techniques result in the analytical expression of the gradient of the global cost function:

$$\frac{\partial \mathsf{C}_{\mathsf{G}}(\boldsymbol{\theta})}{\partial \theta_l} = \mathbb{E} \left(F_{\mathsf{G}}^{l-\pi/2} - F_{\mathsf{G}}^{l+\pi/2} \right) \tag{A.28}$$

where $F_{\mathsf{G}} := F(|\psi\rangle\langle\psi|^{\otimes N}, \rho_{\boldsymbol{\theta}})$ is the global fidelity between the parameterised output state and an N-fold tensor product of input states to be cloned.

Bibliography

[AAB+19] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, October 2019. 2, 172, 241

[AAD11] Graham Allan, Graham R. Allan, and Harold G. Dales. *Introduction to Banach Spaces and Algebras*. Oxford University Press, 2011. Google-Books-ID: e5wVDAAAQBAJ. 19

[Aar09] Scott Aaronson. Quantum Copy-Protection and Quantum Money. In *2009 24th Annual IEEE Conference on Computational Complexity*, pages 229–242, July 2009. ISSN: 1093-0159. 82, 92

[Aar20] Scott Aaronson. Shadow Tomography of Quantum States. *SIAM Journal on Computing*, 49(5):STOC18–368, January 2020. 85, 121

[AC12] Scott Aaronson and Paul Christiano. Quantum money from hidden subspaces. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, STOC '12, pages 41–60, New York, NY, USA, May 2012. Association for Computing Machinery. 81, 92

[AC16] Scott Aaronson and Lijie Chen. Complexity-Theoretic Foundations of Quantum Supremacy Experiments. *arXiv:1612.05903 [quant-ph]*, December 2016. arXiv: 1612.05903. 1

[ACS+19] Andrew Arrasmith, Lukasz Cincio, Andrew T. Sornborger, Wojciech H. Zurek, and Patrick J. Coles. Variational consistent histories as a hybrid algorithm for quantum foundations. *Nature Communications*, 10(1):3438, July 2019. 241

[AD78] Diederik Aerts and Ingrid Daubechies. Physical Justification for Using the Tensor Product to Describe Two Quantum Systems as One Joint System, 1978. 12

[ADDK21] Myrto Arapinis, Mahshid Delavar, Mina Doosti, and Elham Kashefi. Quantum Physical Unclonable Functions: Possibilities and Impossibilities. *Quantum*, 5:475, June 2021. vii, 5, 6

[ADK22] Armando Angrisani, Mina Doosti, and Elham Kashefi. Differential Privacy Amplification in Quantum and Quantum-inspired Algorithms. *arXiv:2203.03604 [quant-ph]*, March 2022. arXiv: 2203.03604. vii

[ADM19] Mohammad Hassan Ameri, Mahshid Delavar, and Javad Mohajeri. Provably secure and efficient PUF-based broadcast authentication schemes for smart grid applications. *International Journal of Communication Systems*, 32(8):e3935, 2019. 125

[ADR02] Jee Hea An, Yevgeniy Dodis, and Tal Rabin. On the Security of Joint Signature and Encryption. In Lars R. Knudsen, editor, *Advances in Cryptology — EUROCRYPT 2002*, pages 83–107, Berlin, Heidelberg, 2002. Springer. 55, 56

[AdW17a] Srinivasan Arunachalam and Ronald de Wolf. Guest Column: A Survey of Quantum Learning Theory. *ACM SIGACT News*, 48(2):41–67, June 2017. 67, 68

[AdW17b] Srinivasan Arunachalam and Ronald de Wolf. Optimal Quantum Sample Complexity of Learning Algorithms. *arXiv:1607.00932 [quant-ph]*, June 2017. arXiv: 1607.00932. 85

[AE07] Andris Ambainis and Joseph Emerson. Quantum t-designs: t-wise Independence in the Quantum World. In *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)*, pages 129–140, June 2007. ISSN: 1093-0159. 134

[AFFS19] Anders Andreassen, Ilya Feige, Christopher Frye, and Matthew D. Schwartz. JUNIPR: a framework for unsupervised machine learning in particle physics. *The European Physical Journal C*, 79(2):102, February 2019. 60

[AHM+14] Joël Alwen, Martin Hirt, Ueli Maurer, Arpita Patra, and Pavel Raykov. Key-Indistinguishable Message Authentication Codes. In Michel Abdalla and Roberto De Prisco, editors, *Security and Cryptography for Networks*, pages 476–493, Cham, 2014. Springer International Publishing. 56

[AK16] Emily Adlam and Adrian Kent. Quantum paradox of choice: More freedom makes summoning a quantum state harder. *Physical Review A*, 93(6):062327, June 2016. 75

[AL13] Hoda A. Alkhzaimi and Martin M. Lauridsen. Cryptanalysis of the SIMON Family of Block Ciphers. *Cryptology ePrint Archive*, 2013. 3

[Ala19] Mohammed M. Alani. Applications of machine learning in cryptography: a survey. In *Proceedings of the 3rd International Conference on Cryptography, Security and Privacy*, ICCSP '19, pages 23–27, New York, NY, USA, January 2019. Association for Computing Machinery. 241

[ALL+21] Scott Aaronson, Jiahui Liu, Qipeng Liu, Mark Zhandry, and Ruizhe Zhang. New Approaches for Quantum Copy-Protection. In Tal Malkin and Chris Peikert, editors, *Advances in Cryptology – CRYPTO 2021*, pages 526–555, Cham, 2021. Springer International Publishing. 82

[ALP21] Prabhanjan Ananth and Rolando L. La Placa. Secure Software Leasing. In Anne Canteaut and François-Xavier Standaert, editors, *Advances in Cryptology – EUROCRYPT 2021*, pages 501–530, Cham, 2021. Springer International Publishing. 82

[AM17] Gorjan Alagic and Christian Majenz. Quantum Non-malleability and Authentication. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology – CRYPTO 2017*, Lecture Notes in Computer Science, pages 310–341, Cham, 2017. Springer International Publishing. 55, 58, 171

[Amb04] Andris Ambainis. A new protocol and lower bounds for quantum coin flipping. *Journal of Computer and System Sciences*, 68(2):398–416, March 2004. 59

[AMR20] Gorjan Alagic, Christian Majenz, and Alexander Russell. Efficient Simulation of Random States and Random Unitaries. In Anne Canteaut and Yuval Ishai, editors, *Advances in Cryptology – EUROCRYPT 2020*, volume 12107, pages 759–787. Springer International Publishing, Cham, 2020. 156

[AMRS20] Gorjan Alagic, Christian Majenz, Alexander Russell, and Fang Song. Quantum-Access-Secure Message Authentication via Blind-Unforgeability. In Anne Canteaut and Yuval Ishai, editors, *Advances in Cryptology – EUROCRYPT 2020*, pages 788–817, Cham, 2020. Springer International Publishing. 45, 47, 56, 57, 58, 92, 105, 113, 294

[AMS+15] Giuseppe Ateniese, Luigi V. Mancini, Angelo Spognardi, Antonio Villani, Domenico Vitali, and Giovanni Felici. Hacking smart machines with smarter ones: How to extract meaningful data from machine learning classifiers. *International Journal of Security and Networks*, 10(3):137–150, January 2015. 241

[AMSY16] Frederik Armknecht, Daisuke Moriyama, Ahmad-Reza Sadeghi, and Moti Yung. Towards a Unified Security Model for Physically Unclonable Functions. In Kazue Sako, editor, *Topics in Cryptology - CT-RSA 2016*, Lecture Notes in Computer Science, pages 271–287, Cham, 2016. Springer International Publishing. 92, 127, 128, 130, 133, 135, 150, 151, 211

[Ang92] Dana Angluin. Computational learning theory: survey and selected bibliography. In *Proceedings of the twenty-fourth annual ACM symposium on Theory of Computing*, pages 351–369. Association for Computing Machinery, New York, NY, USA, July 1992. 67

[APJAD18] Mauricio Araya-Polo, Joseph Jennings, Amir Adler, and Taylor Dahlke. Deep-learning tomography. *The Leading Edge*, 37(1):58–66, January 2018. 61

[AQY22] Prabhanjan Ananth, Luowen Qian, and Henry Yuen. Cryptography from Pseudorandom Quantum States. *arXiv:2112.10020 [quant-ph]*, March 2022. arXiv: 2112.10020. 156

[ARW19] Atul Singh Arora, Jérémie Roland, and Stephan Weis. Quantum weak coin flipping. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, pages 205–216, New York, NY, USA, June 2019. Association for Computing Machinery. 59

[ATSVY00] Dorit Aharonov, Amnon Ta-Shma, Umesh V. Vazirani, and Andrew C. Yao. Quantum bit escrow. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, STOC '00, pages 705–714, New York, NY, USA, May 2000. Association for Computing Machinery. 59, 242, 245, 246, 247, 251, 252, 282

[BB84] Charles Bennett and Gilles Brassard. Quantum cryptography: public key distribution and coin tossing int. volume 175, Bangalore, India, 1984. 244

[BB14] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560:7–11, December 2014. 3, 37, 244, 246, 252

[BBBG09] Guido Berlín, Gilles Brassard, Félix Bussières, and Nicolas Godbout. Fair loss-tolerant quantum coin flipping. *Physical Review A*, 80(6):062321, December 2009. 246, 252

[BBD+97] Adriano Barenco, André Berthiaume, David Deutsch, Artur Ekert, Richard Jozsa, and Chiara Macchiavello. Stabilization of Quantum Computations by Symmetrization. *SIAM Journal on Computing*, 26(5):1541–1557, October 1997. 30

[BBFO+19] M. Bouillard, G. Boucher, J. Ferrer Ortas, B. Pointard, and R. Tualle-Brouri. Quantum Storage of Single-Photon and Two-Photon Fock States with an All-Optical Quantum Memory. *Physical Review Letters*, 122(21):210501, May 2019. 48, 210

[BBGP16] Johannes A. Buchmann, Denis Butin, Florian Göpfert, and Albrecht Petzoldt. Post-Quantum Cryptography: State of the Art. In Peter Y. A. Ryan, David Naccache, and Jean-Jacques Quisquater, editors, *The New Codebreakers: Essays Dedicated to David Kahn on the Occasion of His 85th Birthday*, pages 88–108. Springer, Berlin, Heidelberg, 2016. 43, 47

[BBHB97] V. Bužek, S. L. Braunstein, M. Hillery, and D. Bruß. Quantum copying: A network. *Physical Review A*, 56(5):3446–3452, November 1997. 37, 240, 244, 271, 276

[BCC+15] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma. Simulating Hamiltonian Dynamics with a Truncated Taylor Series. *Physical Review Letters*, 114(9):090502, March 2015. 63

[BCD+09] Alessandro Bisio, Giulio Chiribella, Giacomo Mauro D'Ariano, Stefano Facchini, and Paolo Perinotti. Optimal Quantum Tomography. *IEEE Journal of Selected Topics in Quantum Electronics*, 15(6):1646–1660, November 2009. 60, 79, 85

[BCF+96] Howard Barnum, Carlton M. Caves, Christopher A. Fuchs, Richard Jozsa, and Benjamin Schumacher. Noncommuting Mixed States Cannot Be Broadcast. *Physical Review Letters*, 76(15):2818–2821, April 1996. 34, 37

[BCF+14] Harry Buhrman, Nishanth Chandran, Serge Fehr, Ran Gelles, Vipul Goyal, Rafail Ostrovsky, and Christian Schaffner. Position-Based Quantum Cryptography: Impossibility and Constructions. *SIAM Journal on Computing*, 43(1):150–178, January 2014. 49

[BCHJ+21] Fernando G.S.L. Brandão, Wissam Chemissany, Nicholas Hunter-Jones, Richard Kueng, and John Preskill. Models of Quantum Complexity Growth. *PRX Quantum*, 2(3):030316, July 2021. 156

[BCMDM00] Dagmar Bruß, Mirko Cinchetti, G. Mauro D'Ariano, and Chiara Macchiavello. Phase-covariant quantum cloning. *Physical Review A*, 62(1):012302, June 2000. 36, 242, 244

[BCMdW10] Harry Buhrman, Richard Cleve, Serge Massar, and Ronald de Wolf. Nonlocality and communication complexity. *Reviews of Modern Physics*, 82(1):665–698, March 2010. 30

[BCWdW01] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum Fingerprinting. *Physical Review Letters*, 87(16):167902, September 2001. 30, 179

[BDE+98] Dagmar Bruß, David P. DiVincenzo, Artur Ekert, Christopher A. Fuchs, Chiara Macchiavello, and John A. Smolin. Optimal universal and state-dependent quantum cloning. *Physical Review A*, 57(4):2368–2378, April 1998. 38, 39, 242, 250

[BDF+11] Dan Boneh, Özgür Dagdelen, Marc Fischlin, Anja Lehmann, Christian Schaffner, and Mark Zhandry. Random Oracles in a Quantum World. In Dong Hoon Lee and Xiaoyun Wang, editors, *Advances in Cryptology – ASIACRYPT 2011*, pages 41–69, Berlin, Heidelberg, 2011. Springer. 50, 93

[BDS+18] Anindita Bera, Tamoghna Das, Debasis Sadhukhan, Sudipto Singha Roy, Aditi Sen(De), and Ujjwal Sen. Quantum discord and its allies: a review of recent progress. *Reports on Progress in Physics*, 81(2):024001, February 2018. 16

[Bec15a] Georg T. Becker. The Gap Between Promise and Reality: On the Insecurity of XOR Arbiter PUFs. In Tim Güneysu and Helena Handschuh, editors, *Cryptographic Hardware and Embedded Systems – CHES 2015*, volume 9293, pages 535–555. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. 174, 210

[Bec15b] Georg T. Becker. On the Pitfalls of Using Arbiter-PUFs as Building Blocks. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 34(8):1295–1307, August 2015. 174, 210

[BEM98] Dagmar Bruss, Artur Ekert, and Chiara Macchiavello. Optimal Universal Quantum Cloning and State Estimation. *Physical Review Letters*, 81(12):2598–2601, September 1998. 36, 137

[Ben80] Paul Benioff. The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as

represented by Turing machines. *Journal of Statistical Physics*, 22(5):563–591, May 1980. 2, 25

[Ben82] Paul Benioff. Quantum mechanical hamiltonian models of turing machines. *Journal of Statistical Physics*, 29(3):515–546, November 1982. 25

[Ber07] János A Bergou. Quantum state discrimination and selected applications. *Journal of Physics: Conference Series*, 84:012001, October 2007. 29

[Ber10] János A. Bergou. Discrimination of quantum states. *Journal of Modern Optics*, 57(3):160–180, February 2010. 29

[BFH06] János A. Bergou, Edgar Feldman, and Mark Hillery. Optimal unambiguous discrimination of two subspaces as a case in mixed-state discrimination. *Physical Review A*, 73(3):032107, March 2006. 29

[BFK09] Anne Broadbent, Joseph Fitzsimons, and Elham Kashefi. Universal Blind Quantum Computation. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 517–526, October 2009. ISSN: 0272-5428. 37, 172, 173

[BFNV19] Adam Bouland, Bill Fefferman, Chinmay Nirkhe, and Umesh Vazirani. Quantum Supremacy and the Complexity of Random Circuit Sampling. *Nature Physics*, 15(2):159–163, February 2019. arXiv: 1803.04402. 55

[BFSK11] Christina Brzuska, Marc Fischlin, Heike Schröder, and Stefan Katzenbeisser. Physically Uncloneable Functions in the Universal Composition Framework. In Phillip Rogaway, editor, *Advances in Cryptology – CRYPTO 2011*, Lecture Notes in Computer Science, pages 51–70, Berlin, Heidelberg, 2011. Springer. 124, 127, 135

[BFV19] Adam Bouland, Bill Fefferman, and Umesh Vazirani. Computational pseudorandomness, the wormhole growth paradox, and constraints on the AdS/CFT duality. *arXiv:1910.14646 [gr-qc, physics:hep-th, physics:quant-ph]*, October 2019. arXiv: 1910.14646. 156, 170

[BGM04] Mihir Bellare, Oded Goldreich, and Anton Mityagin. The Power of Verification Queries in Message Authentication and Authenticated Encryption. *Cryptology ePrint Archive*, 2004. 56

[BGR95] Mihir Bellare, Roch Guérin, and Phillip Rogaway. XOR MACs: New Methods for Message Authentication Using Finite Pseudorandom Functions. In Don Coppersmith, editor, *Advances in Cryptology — CRYPT0' 95*, pages 15–28, Berlin, Heidelberg, 1995. Springer. 55, 56

[BH96] V. Bužek and M. Hillery. Quantum copying: Beyond the no-cloning theorem. *Physical Review A*, 54(3):1844–1852, September 1996. 35, 36

[BHH16] Fernando G.S.L. Brandão, Aram W. Harrow, and Michał Horodecki. Efficient Quantum Pseudorandomness. *Physical Review Letters*, 116(17):170502, April 2016. 54

[BHT98] Gilles Brassard, Peter HØyer, and Alain Tapp. Quantum cryptanalysis of hash and claw-free functions. In Cláudio L. Lucchesi and Arnaldo V. Moura, editors, *LATIN'98: Theoretical Informatics*, pages 163–169, Berlin, Heidelberg, 1998. Springer. 3

[Bia21] Jacob Biamonte. Universal variational quantum computation. *Physical Review A*, 103(3):L030401, March 2021. 241

[BIS+20] Ville Bergholm, Josh Izaac, Maria Schuld, Christian Gogolin, M. Sohaib Alam, Shahnawaz Ahmed, Juan Miguel Arrazola, Carsten Blank, Alain Delgado, Soran Jahangiri, Keri McKiernan, Johannes Jakob Meyer, Zeyue Niu, Antal Száva, and Nathan Killoran. PennyLane: Automatic differentiation of hybrid quantum-classical computations. *arXiv:1811.04968 [physics, physics:quant-ph]*, February 2020. arXiv: 1811.04968. 172

[BJ95] Nader H. Bshouty and Jeffrey C. Jackson. Learning DNF over the uniform distribution using a quantum example oracle. In *Proceedings of the eighth annual conference on Computational learning theory*, COLT '95, pages 118–127, New York, NY, USA, July 1995. Association for Computing Machinery. 68

[BJL+21] Anne Broadbent, Stacey Jeffery, Sébastien Lord, Supartha Podder, and Aarthi Sundaram. Secure Software Leasing Without Assumptions. In Kobbi Nissim and Brent Waters, editors, *Theory of Cryptography*, pages 90–120, Cham, 2021. Springer International Publishing. 82

[BK10] Robin Blume-Kohout. Optimal, reliable estimation of quantum states. *New Journal of Physics*, 12(4):043034, April 2010. 61

[BK15] Joonwoo Bae and Leong-Chuan Kwek. Quantum state discrimination and its applications. *Journal of Physics A: Mathematical and Theoretical*, 48(8):083001, February 2015. 28, 253

[BKOV17] Saikrishna Badrinarayanan, Dakshita Khurana, Rafail Ostrovsky, and Ivan Visconti. Unconditional UC-Secure Computation with (Stronger-Malicious) PUFs. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *Advances in Cryptology – EUROCRYPT 2017*, Lecture Notes in Computer Science, pages 382–411, Cham, 2017. Springer International Publishing. 124

[BKR00] Mihir Bellare, Joe Kilian, and Phillip Rogaway. The Security of the Cipher Block Chaining Message Authentication Code. *Journal of Computer and System Sciences*, 61(3):362–399, December 2000. 56

[BL06] Dagmar Bruß and G. Leuchs, editors. *Lectures on Quantum Information*. Wiley, 1 edition, November 2006. 11, 12, 14, 16, 34, 35, 39

[BL17] Daniel J. Bernstein and Tanja Lange. Post-quantum cryptography. *Nature*, 549(7671):188–194, September 2017. 47

[BLL+18] Darius Bunandar, Anthony Lentine, Catherine Lee, Hong Cai, Christopher M. Long, Nicholas Boynton, Nicholas Martinez, Christopher DeRose, Changchen Chen, Matthew Grein, Douglas Trotter, Andrew Starbuck, Andrew Pomerene, Scott Hamilton, Franco N.C. Wong, Ryan Camacho, Paul Davids, Junji Urayama, and Dirk Englund. Metropolitan Quantum Key Distribution with Silicon Photonics. *Physical Review X*, 8(2):021009, April 2018. 238

[BLSF19] Marcello Benedetti, Erika Lloyd, Stefan Sack, and Mattia Fiorentini. Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, 4(4):043001, November 2019. 241

[Blu83] Manuel Blum. Coin flipping by telephone a protocol for solving impossible problems. *ACM SIGACT News*, 15(1):23–27, January 1983. 58, 246

[BL13] Karol Bartkiewicz, Karel Lemr, Antonín Černoch, Jan Soubusta, and Adam Miranowicz. Experimental Eavesdropping Based on Optimal Quantum Cloning. *Physical Review Letters*, 110(17):173601, April 2013. 240, 284

[BM84] Manuel Blum and Silvio Micali. How to Generate Cryptographically Strong Sequences of Pseudorandom Bits. *SIAM Journal on Computing*, 13(4):850–864, November 1984. 155

[BM99] Dagmar Bruß and Chiara Macchiavello. Optimal state estimation for d-dimensional quantum systems. *Physics Letters A*, 253(5-6):249–251, March 1999. 242

[BM06] Dagmar Bru and Chiara Macchiavello. Approximate Quantum Cloning. In Dagmar Bru and G. Leuchs, editors, *Lectures on Quantum Information*, pages 53–71. Wiley-VCH Verlag GmbH, Weinheim, Germany, November 2006. 247

[BOHL+05] Michael Ben-Or, Michał Horodecki, Debbie W. Leung, Dominic Mayers, and Jonathan Oppenheim. The Universal Composable Security of Quantum Key Distribution. In Joe Kilian, editor, *Theory of Cryptography*, pages 386–406, Berlin, Heidelberg, 2005. Springer. 238

[BPLC+20] Carlos Bravo-Prieto, Ryan LaRose, M. Cerezo, Yigit Subasi, Lukasz Cincio, and Patrick J. Coles. Variational Quantum Linear Solver. *arXiv:1909.05820 [quant-ph]*, June 2020. arXiv: 1909.05820. 241, 261, 270

[BRA+19] C.E. Bradley, J. Randall, M.H. Abobeih, R.C. Berrevoets, M.J. Degen, M.A. Bakker, M. Markham, D.J. Twitchen, and T.H. Taminiau. A Ten-Qubit Solid-State Spin Register with Quantum Memory up to One Minute. *Physical Review X*, 9(3):031045, September 2019. 48, 210

[Bro19] Harvey R. Brown. The Reality of the Wavefunction: Old Arguments and New. In Alberto Cordero, editor, *Philosophers Look at Quantum Mechanics*, pages 63–86. Springer International Publishing, Cham, 2019. 75

[BS16] Anne Broadbent and Christian Schaffner. Quantum cryptography beyond quantum key distribution. *Designs, Codes and Cryptography*, 78(1):351–382, January 2016. 46, 49, 171, 237

[BS19] Zvika Brakerski and Omri Shmueli. (Pseudo) Random Quantum States with Binary Phase. In Dennis Hofheinz and Alon Rosen, editors, *Theory of Cryptography*, volume 11891, pages 229–250. Springer International Publishing, Cham, 2019. 162

[BS20] Zvika Brakerski and Omri Shmueli. Scalable Pseudorandom Quantum States. In Daniele Micciancio and Thomas Ristenpart, editors, *Advances in Cryptology – CRYPTO 2020*, volume 12171, pages 417–440. Springer International Publishing, Cham, 2020. 156, 162

[BSW06] Dan Boneh, Emily Shen, and Brent Waters. Strongly Unforgeable Signatures Based on Computational Diffie-Hellman. In Moti Yung, Yevgeniy Dodis, Aggelos Kiayias, and Tal Malkin, editors, *Public Key Cryptography - PKC 2006*, pages 229–240, Berlin, Heidelberg, 2006. Springer. 55, 56

[BWM21] S. Blinov, B. Wu, and C. Monroe. Comparison of Cloud-Based Ion Trap and Superconducting Quantum Computer Architectures. *arXiv:2102.00371 [quant-ph]*, January 2021. arXiv: 2102.00371. 172

[BWP+17] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, September 2017. 69, 85, 241

[BZ13a] Dan Boneh and Mark Zhandry. Quantum-Secure Message Authentication Codes. In Thomas Johansson and Phong Q. Nguyen, editors, *Advances in Cryptology – EUROCRYPT 2013*, Lecture Notes in Computer Science, pages 592–608, Berlin, Heidelberg, 2013. Springer. 45, 47, 49, 56, 92, 93, 107, 113, 171

[BZ13b] Dan Boneh and Mark Zhandry. Secure Signatures and Chosen Ciphertext Security in a Quantum Computing World. In Ran Canetti and Juan A. Garay, editors, *Advances in Cryptology – CRYPTO 2013*, Lecture Notes in Computer Science, pages 361–379, Berlin, Heidelberg, 2013. Springer. 45, 47, 48, 49, 56, 135

[B+17] Karol Bartkiewicz, Antonín Černoch, Grzegorz Chimczak, Karel Lemr, Adam Miranowicz, and Franco Nori. Experimental quantum forgery of quantum optical money. *npj Quantum Information*, 3(1):1–8, March 2017. 240, 284

[CAB+21] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, September 2021. 70, 71, 241

[Can01] R. Canetti. Universally composable security: a new paradigm for cryptographic protocols. In *Proceedings 42nd IEEE Symposium on Foundations of Computer Science*, pages 136–145, October 2001. ISSN: 1552-5244. 45, 238

[CB98] Anthony Chefles and Stephen M. Barnett. Optimum unambiguous discrimination between linearly independent symmetric states. *Physics Letters A*, 250(4):223–229, December 1998. 29

[CBTW17] Patrick J. Coles, Mario Berta, Marco Tomamichel, and Stephanie Wehner. Entropic uncertainty relations and their applications. *Reviews of Modern Physics*, 89(1):015002, February 2017. 21, 232, 233

[CC07] Lin Chen and Yi-Xin Chen. Mixed qubits cannot be universally broadcast. *Physical Review A*, 75(6):062322, June 2007. 37

[CCB18] Marcello Caleffi, Angela Sara Cacciapuoti, and Giuseppe Bianchi. Quantum internet: from communication to distributed computing! In *Proceedings of the 5th ACM International Conference on Nanoscale Computing and Communication*, NANOCOM '18, pages

1–4, New York, NY, USA, September 2018. Association for Computing Machinery. 171

[CCL19] Iris Cong, Soonwon Choi, and Mikhail D. Lukin. Quantum convolutional neural networks. *Nature Physics*, 15(12):1273–1278, December 2019. 241

[CCT+20] Angela Sara Cacciapuoti, Marcello Caleffi, Francesco Tafuri, Francesco Saverio Cataliotti, Stefano Gherardini, and Giuseppe Bianchi. Quantum Internet: Networking Challenges in Distributed Quantum Computing. *IEEE Network*, 34(1):137–143, January 2020. 171

[CCW18] Song Cheng, Jing Chen, and Lei Wang. Information Perspective to Probabilistic Modeling: Boltzmann Machines versus Born Machines. *Entropy*, 20(8):583, August 2018. 60

[CDKK22] Brian Coyle, Mina Doosti, Elham Kashefi, and Niraj Kumar. Progress toward practical quantum cryptanalysis by variational quantum cloning. *Physical Review A*, 105(4):042604, April 2022. vii, 8, 258, 266, 274

[CDM+18] Ulysse Chabaud, Eleni Diamanti, Damian Markham, Elham Kashefi, and Antoine Joux. Optimal quantum-programmable projective measurement with linear optics. *Physical Review A*, 98(6):062318, December 2018. 31, 179

[CDM+21] Kaushik Chakraborty, Mina Doosti, Yao Ma, Myrto Arapinis, and Elham Kashefi. Quantum Lock: A Provable Quantum Communication Advantage. *arXiv:2110.09469 [quant-ph]*, November 2021. arXiv: 2110.09469. vii, 7, 214, 216, 225, 228

[CDP08] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti. Optimal Cloning of Unitary Transformation. *Physical Review Letters*, 101(18):180504, October 2008. 76, 134

[Cer00a] Nicolas J. Cerf. Asymmetric quantum cloning in any dimension. *Journal of Modern Optics*, 47(2-3):187–209, February 2000. 35

[Cer00b] Nicolas J. Cerf. Pauli Cloning of a Quantum Bit. *Physical Review Letters*, 84(19):4497–4500, May 2000. 35

[CETU21] Tore Vincent Carstens, Ehsan Ebrahimi, Gelo Noel Tabia, and Dominique Unruh. Relationships Between Quantum IND-CPA Notions. In Kobbi Nissim and Brent Waters, editors, *Theory of Cryptography*, pages 240–272, Cham, 2021. Springer International Publishing. 48

[CEV20] Céline Chevalier, Ehsan Ebrahimi, and Quoc-Huy Vu. On Security Notions for Encryption in a Quantum World. Technical Report 237, 2020. 48, 49

[CF01] Ran Canetti and Marc Fischlin. Universally Composable Commitments. In Gerhard Goos, Juris Hartmanis, Jan van Leeuwen, and Joe Kilian, editors, *Advances in Cryptology — CRYPTO 2001*, volume 2139, pages 19–40. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001. Series Title: Lecture Notes in Computer Science. 124

[CGC+12] Jerry M. Chow, Jay M. Gambetta, A. D. Córcoles, Seth T. Merkel, John A. Smolin, Chad Rigetti, S. Poletto, George A. Keefe, Mary B. Rothwell, J. R. Rozen, Mark B. Ketchen, and M. Steffen. Universal Quantum Gate Set Approaching Fault-Tolerant Thresholds with Superconducting Qubits. *Physical Review Letters*, 109(6):060501, August 2012. 25

[CGH+19] Joshua Combes, Kyle V. Gulshen, Matthew P. Harrigan, Peter J. Karalekas, Marcus P. da Silva, M. Sohaib Alam, Amy F. Brown, Shane Caldwell, Lauren C. Capelluto, Gavin E. Crooks, Daniel Girshovich, Blake R. Johnson, Eric C. Peterson, Anthony M. Polloreno, Nicholas C. Rubin, Colm A. Ryan, Alexa N. Staley, Nikolas A. Tezak, and Joseph A. Valery. Forest Benchmarking: QCVV using PyQuil, September 2019. 276

[CGK+16] Véronique Cortier, David Galindo, Ralf Küsters, Johannes Müller, and Tomasz Truderung. SoK: Verifiability Notions for E-Voting Protocols. In *2016 IEEE Symposium on Security and Privacy (SP)*, pages 779–798, May 2016. ISSN: 2375-1207. 95

[CHS+15] Jacques Carolan, Christopher Harrold, Chris Sparrow, Enrique Martín-López, Nicholas J. Russell, Joshua W. Silverstone, Peter J. Shadbolt, Nobuyuki Matsuda, Manabu Oguma, Mikitaka Itoh, Graham D. Marshall, Mark G. Thompson, Jonathan C. F. Matthews, Toshikazu Hashimoto, Jeremy L. O'Brien, and Anthony Laing. Universal linear optics. *Science*, 349(6249):711–716, August 2015. 156

[CHZ+09] Wei Chen, Zheng-Fu Han, Tao Zhang, Hao Wen, Zhen-Qiang Yin, Fang-Xing Xu, Qing-Lin Wu, Yun Liu, Yang Zhang, Xiao-Fan Mo, You-Zhen Gui, Guo Wei, and Guang-Can Guo. Field Experiment on a "Star Type" Metropolitan Quantum Key Distribution Network. *IEEE Photonics Technology Letters*, 21(9):575–577, May 2009. 236

[CI93] Richard Cleve and Russell Impagliazzo. Martingales, collective coin flipping and discrete control processes (Extended Abstract), 1993. 59

[CIVA02] N.J. Cerf, S. Iblisdir, and G. Van Assche. Cloning and cryptography with quantum continuous variables. *The European Physical Journal D - Atomic, Molecular, Optical and Plasma Physics*, 18(2):211–218, February 2002. 241, 272

[CK18] Bob Coecke and Aleks Kissinger. Picturing Quantum Processes. In Peter Chapman, Gem Stapleton, Amirouche Moktefi, Sarah Perez-Kriz, and Francesco Bellucci, editors, *Diagrammatic Representation and Inference*, pages 28–31, Cham, 2018. Springer International Publishing. 3

[CKDK21] Federico Centrone, Niraj Kumar, Eleni Diamanti, and Iordanis Kerenidis. Experimental demonstration of quantum advantage for NP verification with limited information. *Nature Communications*, 12(1):850, February 2021. 241

[Cle86] R Cleve. Limits on the security of coin flips when half the processors are faulty. In *Proceedings of the eighteenth annual ACM symposium on Theory of computing - STOC '86*, pages 364–369, Berkeley, California, United States, 1986. ACM Press. 59

[CM97] Christian Cachin and Ueli Maurer. Unconditional security against memory-bounded adversaries. In Gerhard Goos, Juris Hartmanis, Jan van Leeuwen, and Burton S. Kaliski, editors, *Advances in Cryptology — CRYPTO '97*, volume 1294, pages 292–306. Springer Berlin Heidelberg, Berlin, Heidelberg, 1997. 48

[CMDK20] Brian Coyle, Daniel Mills, Vincent Danos, and Elham Kashefi. The Born supremacy: quantum advantage and training of an Ising Born machine. *npj Quantum Information*, 6(1):1–11, July 2020. 60, 241

[Cou16] Rachel Courtland. China's 2,000-km quantum link is almost complete [News]. *IEEE Spectrum*, 53(11):11–12, November 2016. 238

[Coy22] Brian Coyle. Machine learning applications for noisy intermediate-scale quantum computers. *arXiv:2205.09414 [quant-ph]*, May 2022. arXiv: 2205.09414. 69, 70

[CRAG18] Y. Cao, J. Romero, and A. Aspuru-Guzik. Potential of quantum computing for drug discovery. *IBM Journal of Research and Development*, 62(6):6:1–6:20, November 2018. 3

[Cro18] Andrew Cross. The IBM Q experience and QISKit open-source quantum computing software. 2018:L58.003, January 2018. ADS Bibcode: 2018APS..MARL58003C. 172

[CRO+19] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim

Menke, Borja Peropadre, Nicolas P. D. Sawaya, Sukin Sim, Libor Veis, and Alán Aspuru-Guzik. Quantum Chemistry in the Age of Quantum Computing. *Chemical Reviews*, 119(19):10856–10915, October 2019. 3

[CSSC18] Lukasz Cincio, Yiğit Subaşı, Andrew T Sornborger, and Patrick J Coles. Learning the quantum algorithm for state overlap. *New Journal of Physics*, 20(11):113022, November 2018. 273

[CSU+20] D. Chivilikhin, A. Samarin, V. Ulyantsev, I. Iorsh, A. R. Oganov, and O. Kyriienko. MoG-VQE: Multiobjective genetic variational quantum eigensolver. *arXiv:2007.04424 [cond-mat, physics:quant-ph]*, July 2020. arXiv: 2007.04424. 273

[CSV+21] M. Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J. Coles. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nature Communications*, 12(1):1791, March 2021. 70, 259, 260, 269, 274

[CZSD07] Hongwei Chen, Xianyi Zhou, Dieter Suter, and Jiangfeng Du. Experimental realization of $\mathbf{1}\mathbf{\ensuremath{\rightarrow}}\mathbf{2}$ asymmetric phase-covariant quantum cloning. *Physical Review A*, 75(1):012317, January 2007. 240

[CZZ17] Chip-Hong Chang, Yue Zheng, and Le Zhang. A Retrospective and a Look Forward: Fifteen Years of Physical Unclonable Function Advancement. *IEEE Circuits and Systems Magazine*, 17(3):32–62, 2017. 125, 137

[DCEL09] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Physical Review A*, 80(1):012304, July 2009. 54, 134, 163

[DDKA21] Mina Doosti, Mahshid Delavar, Elham Kashefi, and Myrto Arapinis. A Unified Framework For Quantum Unforgeability. *arXiv:2103.13994 [quant-ph]*, October 2021. arXiv: 2103.13994. vii, 5

[Del19] Jeroen Delvaux. Machine-Learning Attacks on PolyPUFs, OB-PUFs, RPUFs, LHS-PUFs, and PUF–FSMs. *IEEE Transactions on Information Forensics and Security*, 14(8):2043–2058, August 2019. 174, 210

[Deu83] David Deutsch. Uncertainty in Quantum Measurements. *Physical Review Letters*, 50(9):631–633, February 1983. 21

[Deu85] David Deutsch. Quantum theory, the Church–Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 400(1818):97–117, July 1985. 25

[DF07] Gui-Fang Dang and Heng Fan. Optimal broadcasting of mixed states. *Physical Review A*, 76(2):022323, August 2007. 37

[DF17] Persi Diaconis and Peter J. Forrester. Hurwitz and the origins of random matrix theory in mathematics. *Random Matrices: Theory and Applications*, 06(01):1730001, January 2017. 40

[DFC05] Thomas Durt, Jaromír Fiurášek, and Nicolas J. Cerf. Economical quantum cloning in any dimension. *Physical Review A*, 72(5):052322, November 2005. 35

[DFSS05] I.B. Damgard, S. Fehr, L. Salvail, and C. Schaffner. Cryptography In the Bounded Quantum-Storage Model. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 449–458, Pittsburgh, PA, USA, 2005. IEEE. 48, 49

[DG97] Lu-Ming Duan and Guang-Can Guo. Two non-orthogonal states can be cloned by a unitary-reduction process. *arXiv:quant-ph/9704020*, April 1997. arXiv: quant-ph/9704020. 35

[DG98] Lu-Ming Duan and Guang-Can Guo. Probabilistic Cloning and Identification of Linearly Independent Quantum States. *Physical Review Letters*, 80(22):4999–5002, June 1998. 35

[DGJ+20] Yfke Dulek, Alex B. Grilo, Stacey Jeffery, Christian Majenz, and Christian Schaffner. Secure Multi-party Quantum Computation with a Dishonest Majority. In Anne Canteaut and Yuval Ishai, editors, *Advances in Cryptology – EUROCRYPT 2020*, pages 729–758, Cham, 2020. Springer International Publishing. 171

[DHLT20] Yuxuan Du, Min-Hsiu Hsieh, Tongliang Liu, and Dacheng Tao. Expressive power of parametrized quantum circuits. *Physical Review Research*, 2(3):033125, July 2020. 241

[DHY+20] Yuxuan Du, Tao Huang, Shan You, Min-Hsiu Hsieh, and Dacheng Tao. Quantum circuit architecture search: error mitigation and trainability enhancement for variational quantum solvers. *arXiv:2010.10217 [quant-ph]*, November 2020. arXiv: 2010.10217. 273

[Dia19] Eleni Diamanti. Demonstrating Quantum Advantage in Security and Efficiency with Practical Photonic Systems. In *2019 21st International Conference on Transparent Optical Networks (ICTON)*, pages 1–2, July 2019. ISSN: 2161-2064. 171

[Die82] D. Dieks. Communication by EPR devices. *Physics Letters A*, 92(6):271–272, November 1982. 33

[Die88] D. Dieks. Overlap and distinguishability of quantum states. *Physics Letters A*, 126(5):303–306, January 1988. 29

[DiV00] David P. DiVincenzo. The Physical Implementation of Quantum Computation. *Fortschritte der Physik*, 48(9-11):771–783, September 2000. 24

[DJ99] A. C. Doherty and K. Jacobs. Feedback control of quantum systems using continuous state estimation. *Physical Review A*, 60(4):2700–2711, October 1999. 28

[DKDK21] Mina Doosti, Niraj Kumar, Mahshid Delavar, and Elham Kashefi. Client-server Identification Protocols with Quantum PUF. *ACM Transactions on Quantum Computing*, 2(3):12:1–12:40, September 2021. vii, 7

[DKK17] Mina Doosti, Farzad Kianvash, and Vahid Karimipour. Universal superposition of orthogonal states. *Physical Review A*, 96(5):052318, November 2017. 34, 92, 140

[DKKC22] Mina Doosti, Niraj Kumar, Elham Kashefi, and Kaushik Chakraborty. On the connection between quantum pseudorandomness and quantum hardware assumptions. *Quantum Science and Technology*, 7(3):035004, July 2022. vii, 6, 7

[DKLP02] Eric Dennis, Alexei Kitaev, Andrew Landahl, and John Preskill. Topological quantum memory. *Journal of Mathematical Physics*, 43(9):4452–4505, September 2002. 48, 210

[DKPW12] Yevgeniy Dodis, Eike Kiltz, Krzysztof Pietrzak, and Daniel Wichs. Message Authentication, Revisited. In David Pointcheval and Thomas Johansson, editors, *Advances in Cryptology – EUROCRYPT 2012*, pages 355–374, Berlin, Heidelberg, 2012. Springer. 56

[DLP01] G. M. D'Ariano and P. Lo Presti. Quantum Tomography for Measuring Experimentally the Matrix Elements of an Arbitrary Quantum Operation. *Physical Review Letters*, 86(19):4195–4198, May 2001. 62, 85

[DM03] Giacomo Mauro D'Ariano and Chiara Macchiavello. Optimal phase-covariant cloning for qubits and qutrits. *Physical Review A*, 67(4):042306, April 2003. 38

[DMAM17] Mahshid Delavar, Sattar Mirzakuchaki, Mohammad Hassan Ameri, and Javad Mohajeri. PUF-based solutions for secure communications in Advanced Metering Infrastructure (AMI). *International Journal of Communication Systems*, 30(9):e3195, 2017. 125, 135

[DMM16] Mahshid Delavar, Sattar Mirzakuchaki, and Javad Mohajeri. A Ring Oscillator-Based PUF With Enhanced Challenge-Response Pairs. *Canadian Journal of Electrical and Computer Engineering*, 39(2):174–180, 2016. 125

[DS94] Persi Diaconis and Mehrdad Shahshahani. On the eigenvalues of random matrices. *Journal of Applied Probability*, 31(A):49–62, 1994. 42

[DWT+19] J. F. Dynes, A. Wonfor, W. W.-S. Tam, A. W. Sharpe, R. Takahashi, M. Lucamarini, A. Plews, Z. L. Yuan, A. R. Dixon, J. Cho, Y. Tanizawa, J.-P. Elbers, H. Greißer, I. H. White, R. V. Penty, and A. J. Shields. Cambridge quantum network. *npj Quantum Information*, 5(1):1–8, November 2019. 171

[EA5] Joseph Emerson, Robert Alicki, and Karol Życzkowski. Scalable noise estimation with random unitary operators. *Journal of Optics B: Quantum and Semiclassical Optics*, 7(10):S347–S352, October 2005. 54, 62

[ECBY21] Suguru Endo, Zhenyu Cai, Simon C. Benjamin, and Xiao Yuan. Hybrid Quantum-Classical Algorithms and Quantum Error Mitigation. *Journal of the Physical Society of Japan*, 90(3):032001, March 2021. 241

[EHW+20] Jens Eisert, Dominik Hangleiter, Nathan Walk, Ingo Roth, Damian Markham, Rhea Parekh, Ulysse Chabaud, and Elham Kashefi. Quantum certification and benchmarking. *Nature Reviews Physics*, 2(7):382–390, July 2020. 55

[EWA+21] Prashant S. Emani, Jonathan Warrell, Alan Anticevic, Stefan Bekiranov, Michael Gandal, Michael J. McConnell, Guillermo Sapiro, Alán Aspuru-Guzik, Justin T. Baker, Matteo Bastiani, John D. Murray, Stamatios N. Sotiropoulos, Jacob Taylor, Geetha Senthil, Thomas Lehner, Mark B. Gerstein, and Aram W. Harrow. Quantum computing at the frontiers of biological sciences. *Nature Methods*, 18(7):701–709, July 2021. 3

[Fan57] U. Fano. Description of States in Quantum Mechanics by Density Matrix and Operator Techniques. *Reviews of Modern Physics*, 29(1):74–93, January 1957. 11

[FGG14] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A Quantum Approximate Optimization Algorithm. *arXiv:1411.4028 [quant-ph]*, November 2014. arXiv: 1411.4028. 241

[Fit17] Joseph F. Fitzsimons. Private quantum computation: an introduction to blind quantum computing and related protocols. *npj Quantum Information*, 3(1):1–11, June 2017. 171

[Fiu03] Jaromír Fiurášek. Optical implementations of the optimal phase-covariant quantum cloning machine. *Physical Review A*, 67(5):052314, May 2003. 240

[FKF00] Dietmar G. Fischer, Stefan H. Kienle, and Matthias Freyberger. Quantum-state estimation by self-learning measurements. *Physical Review A*, 61(3):032306, February 2000. 28

[FL12] Agnes Ferenczi and Norbert Lütkenhaus. Symmetries in quantum key distribution and the connection between optimal attacks and optimal cloning. *Physical Review A*, 85(5):052310, May 2012. 277

[FLD+17] Bernd Fröhlich, Marco Lucamarini, James F. Dynes, Lucian C. Comandar, Winci W.-S. Tam, Alan Plews, Andrew W. Sharpe, Zhiliang Yuan, and Andrew J. Shields. Long-distance quantum key distribution secure against coherent attacks. *Optica*, 4(1):163–167, January 2017. 238

[FMWW01] Heng Fan, Keiji Matsumoto, Xiang-Bin Wang, and Miki Wadati. Quantum cloning machines for equatorial qubits. *Physical Review A*, 65(1):012304, December 2001. 37, 38, 244

[FNAF19] Lukas Fladung, Georgios M. Nikolopoulos, Gernot Alber, and Marc Fischlin. Intercept-Resend Emulation Attacks against a Continuous-Variable Quantum Authentication Protocol with Physical Unclonable Keys. *Cryptography*, 3(4):25, December 2019. 127, 151, 152, 174

[FWJ+14] Heng Fan, Yi-Nan Wang, Li Jing, Jie-Dong Yue, Han-Duo Shi, Yong-Liang Zhang, and Liang-Zhu Mu. Quantum Cloning Machines and the Applications. *arXiv:1301.2956 [quant-ph]*, August 2014. arXiv: 1301.2956. 35, 37, 240, 244, 276

[Gav12] Dmitry Gavinsky. Quantum Money with Classical Verification. In *2012 IEEE 27th Conference on Computational Complexity*, pages 42–52, June 2012. ISSN: 1093-0159. 81

[GBC+18] Edward Grant, Marcello Benedetti, Shuxiang Cao, Andrew Hallam, Joshua Lockhart, Vid Stojevic, Andrew G. Green, and Simone Severini. Hierarchical quantum classifiers. *npj Quantum Information*, 4(1):1–8, December 2018. 241

[GC99] Daniel Gottesman and Isaac L. Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature*, 402(6760):390–393, November 1999. 25

[GCvDD02] Blaise Gassend, Dwaine Clarke, Marten van Dijk, and Srinivas Devadas. Silicon physical random functions. In *Proceedings of the 9th ACM conference on Computer and communications security*, CCS '02, pages 148–160, New York, NY, USA, November 2002. Association for Computing Machinery. 124, 174, 210

[GD18] Daniel Greenbaum and Zachary Dutton. Modeling coherent errors in quantum error correction. *Quantum Science and Technology*, 3(1):015007, January 2018. 130

[GEBM19] Harper R. Grimsley, Sophia E. Economou, Edwin Barnes, and Nicholas J. Mayhall. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nature Communications*, 10(1):3007, July 2019. 273

[GGM86] Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to construct random functions. *Journal of the ACM*, 33(4):792–807, August 1986. 45, 155

[GHM+14] Sebastianus A. Goorden, Marcel Horstmann, Allard P. Mosk, Boris Skorić, and Pepijn W. H. Pinkse. Quantum-secure authentication of a physical unclonable key. *Optica*, 1(6):421–424, December 2014. 127, 151, 152, 174

[GHS16] Tommaso Gagliardoni, Andreas Hülsing, and Christian Schaffner. Semantic Security and Indistinguishability in the Quantum World. In Matthew Robshaw and Jonathan Katz, editors, *Advances in Cryptology – CRYPTO 2016*, Lecture Notes in Computer Science, pages 60–89, Berlin, Heidelberg, 2016. Springer. 47, 48, 49, 50, 84, 92

[GI20] Laszlo Gyongyosi and Sandor Imre. Optimizing High-Efficiency Quantum Memory with Quantum Machine Learning for Near-Term Quantum Devices. *Scientific Reports*, 10(1):135, January 2020. 48, 210

[Gis98] N Gisin. Quantum cloning without signaling. *Physics Letters A*, 242(1-2):1–3, May 1998. 34

[GKB20] Giulio Gianfelici, Hermann Kampermann, and Dagmar Bruß. Theoretical framework for physical unclonable functions, including quantum readout. *Physical Review A*, 101(4):042337, April 2020. 127, 174, 176

[GKS21] Tommaso Gagliardoni, Juliane Krämer, and Patrick Struck. Quantum Indistinguishability for Public Key Encryption. In Jung Hee Cheon and Jean-Pierre Tillich, editors, *Post-Quantum Cryptography*, pages 463–482, Cham, 2021. Springer International Publishing. 48, 49, 50, 84

[GKST07] Jorge Guajardo, Sandeep S. Kumar, Geert-Jan Schrijen, and Pim Tuyls. FPGA Intrinsic PUFs and Their Use for IP Protection. In Pascal Paillier and Ingrid Verbauwhede, editors, *Cryptographic Hardware and Embedded Systems - CHES 2007*, pages 63–80, Berlin, Heidelberg, 2007. Springer. 125, 174, 210

[GLN05] Alexei Gilchrist, Nathan K. Langford, and Michael A. Nielsen. Distance measures to compare real and ideal quantum processes. *Physical Review A*, 71(6):062310, June 2005. 16

[GLRS16] Markus Grassl, Brandon Langenberg, Martin Roetteler, and Rainer Steinwandt. Applying Grover's Algorithm to AES: Quantum Resource Estimates. In Tsuyoshi Takagi, editor, *Post-Quantum Cryptography*, pages 29–43, Cham, 2016. Springer International Publishing. 3

[GM97] N. Gisin and S. Massar. Optimal Quantum Cloning Machines. *Physical Review Letters*, 79(11):2153–2156, September 1997. 36, 79

[GMR88] Shafi Goldwasser, Silvio Micali, and Ronald L. Rivest. A Digital Signature Scheme Secure Against Adaptive Chosen-Message Attacks. *SIAM Journal on Computing*, 17(2):281–308, April 1988. 55, 56

[Gol96] D. Gollmann. What do we mean by entity authentication? In *Proceedings 1996 IEEE Symposium on Security and Privacy*, pages 46–54, May 1996. ISSN: 1081-6011. 171

[GR18] Sudip Ghosh and Suvrat Raju. Quantum information measures for restricted sets of observables. *Physical Review D*, 98(4):046005, August 2018. 16

[Gro96] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing - STOC '96*, pages 212–219, Philadelphia, Pennsylvania, United States, 1996. ACM Press. 46

[GTFS16] Fatemeh Ganji, Shahin Tajik, Fabian Fäßler, and Jean-Pierre Seifert. Strong Machine Learning Attack Against PUFs with No Mathematical Model. In Benedikt Gierlichs and Axel Y. Poschmann, editors, *Cryptographic Hardware and Embedded Systems – CHES 2016*, Lecture Notes in Computer Science, pages 391–411, Berlin, Heidelberg, 2016. Springer. 125, 151

[GYZ17] Sumegha Garg, Henry Yuen, and Mark Zhandry. New Security Notions and Feasibility Results for Authentication of Quantum Data. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology – CRYPTO 2017*, pages 342–371, Cham, 2017. Springer International Publishing. 58

[Haa33] Alfred Haar. Der Massbegriff in der Theorie der Kontinuierlichen Gruppen. *The Annals of Mathematics*, 34(1):147, January 1933. 40

[Hal18] Halak. *Physically unclonable functions*. Springer Berlin Heidelberg, New York, NY, 2018. 125

[Har13] Lucien Hardy. Are quantum states real? *International Journal of Modern Physics B*, 27(01n03):1345012, January 2013. 75

[Hay05] Masahito Hayashi. *Asymptotic Theory Of Quantum Statistical Inference: Selected Papers*. World Scientific, February 2005. Google-Books-ID: jfjICgAAQBAJ. 85

[HBC+21] Hsin-Yuan Huang, Michael Broughton, Jordan Cotler, Sitan Chen, Jerry Li, Masoud Mohseni, Hartmut Neven, Ryan Babbush, Richard Kueng, John Preskill, and Jarrod R. McClean. Quantum advantage in learning from experiments. *arXiv:2112.00778 [quant-ph]*, December 2021. arXiv: 2112.00778. 156

[HBR21] Hsin-Yuan Huang, Kishor Bharti, and Patrick Rebentrost. Near-term quantum algorithms for linear systems of equations with regression loss functions. *New Journal of Physics*, 23(11):113021, November 2021. 241

[Hel69] Carl W. Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1(2):231–252, June 1969. 250

[HFGW18] A. K. Hashagen, S. T. Flammia, D. Gross, and J. J. Wallman. Real Randomized Benchmarking. *Quantum*, 2:85, August 2018. 55

[HHL09] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum Algorithm for Linear Systems of Equations. *Physical Review Letters*, 103(15):150502, October 2009. 60, 69

[HILL99] Johan HÅstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. A Pseudorandom Generator from any One-way Function. *SIAM Journal on Computing*, 28(4):1364–1396, January 1999. 155

[HKP20] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, October 2020. 61, 79, 85, 121

[HKP21] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Information-Theoretic Bounds on Quantum Advantage in Machine Learning. *Physical Review Letters*, 126(19):190505, May 2021. 152, 238

[Hol73] A. S Holevo. Statistical decision theory for quantum systems. *Journal of Multivariate Analysis*, 3(4):337–394, December 1973. 219, 250

[HP07] Patrick Hayden and John Preskill. Black holes as mirrors: quantum information in random subsystems. *Journal of High Energy Physics*, 2007(09):120–120, September 2007. 55

[HPS21] Mohsen Heidari, Arun Padakandla, and Wojciech Szpankowski. A Theoretical Framework for Learning from Quantum Data. In *2021 IEEE International Symposium on Information Theory (ISIT)*, pages 1469–1474, July 2021. 121

[HR16] Fumio Hiai and Mary Beth Ruskai. Contraction coefficients for noisy quantum channels. *Journal of Mathematical Physics*, 57(1):015211, January 2016. 19

[Hra97] Z. Hradil. Quantum-state estimation. *Physical Review A*, 55(3):R1561–R1564, March 1997. 28

[HS21] Jack K. Horner and John F. Symons. What Have Google's Random Quantum Circuit Simulation Experiments Demonstrated About Quantum Supremacy? In Hamid R. Arabnia, Leonidas Deligiannidis, Fernando G. Tinetti, and Quoc-Nam Tran, editors, *Advances in Software Engineering, Education, and e-Learning*, pages 411–419, Cham, 2021. Springer International Publishing. 2

[HSG13] Chris Heunen, Mehrnoosh Sadrzadeh, and Edward Grefenstette. *Quantum Physics and Linguistics: A Compositional, Diagrammatic Discourse*. OUP Oxford, February 2013. Google-Books-ID: mvEXAwAAQBAJ. 3

[HSNF18] Kentaro Heya, Yasunari Suzuki, Yasunobu Nakamura, and Keisuke Fujii. Variational Quantum Gate Optimization. *arXiv:1810.12745 [quant-ph]*, October 2018. arXiv: 1810.12745. 241

[HYKD14] Charles Herder, Meng-Day Yu, Farinaz Koushanfar, and Srinivas Devadas. Physical Unclonable Functions and Applications: A Tutorial. *Proceedings of the IEEE*, 102(8):1126–1141, August 2014. 125, 230

[IAC+05] S. Iblisdir, A. Acín, N. J. Cerf, R. Filip, J. Fiurášek, and N. Gisin. Multipartite asymmetric quantum cloning. *Physical Review A*, 72(4):042328, October 2005. 35

[Iva87] I. D. Ivanovic. How to differentiate between non-orthogonal states. *Physics Letters A*, 123(6):257–259, August 1987. 29

[JB22] Tyson Jones and Simon C. Benjamin. Robust quantum compilation and circuit optimisation via energy minimisation. *arXiv:1811.03147 [quant-ph]*, January 2022. arXiv: 1811.03147. 241

[JJB+19] Jan Jašek, Kateřina Jiráková, Karol Bartkiewicz, Karol Bartkiewicz, Antonín Černoch, Tomáš Fürst, and Karel Lemr. Experimental hybrid quantum-classical reinforcement learning by boson sampling: how to train a quantum cloner. *Optics Express*, 27(22):32454–32464, October 2019. 259

[JL18] Stephen P. Jordan and Yi-Kai Liu. Quantum Cryptanalysis: Shor, Grover, and Beyond. *IEEE Security Privacy*, 16(5):14–21, September 2018. 3

[JLS18] Zhengfeng Ji, Yi-Kai Liu, and Fang Song. Pseudorandom Quantum States. In Hovav Shacham and Alexandra Boldyreva, editors, *Advances in Cryptology – CRYPTO 2018*, volume 10993, pages 126–152. Springer International Publishing, Cham, 2018. 6, 53, 54, 55, 78, 116, 155, 156, 162, 165, 170

[Kal21] Gil Kalai. The Argument against Quantum Computers, the Quantum Laws of Nature, and Google's Supremacy Claims. *arXiv:2008.05188 [quant-ph]*, March 2021. arXiv: 2008.05188. 2

[Kar21] N. Karimi. Optimal unambiguous discrimination of pure quantum states using SDP method. *Chinese Journal of Physics*, 72:681–687, August 2021. 29

[Kat07] Jonathan Katz. Universally Composable Multi-party Computation Using Tamper-Proof Hardware. In Moni Naor, editor, *Advances in Cryptology - EUROCRYPT 2007*, Lecture Notes in Computer Science, pages 115–128, Berlin, Heidelberg, 2007. Springer. 124

[KB17] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. *arXiv:1412.6980 [cs]*, January 2017. arXiv: 1412.6980. 258

[KBA+19] Nathan Killoran, Thomas R. Bromley, Juan Miguel Arrazola, Maria Schuld, Nicolás Quesada, and Seth Lloyd. Continuous-variable quantum neural networks. *Physical Review Research*, 1(3):033063, October 2019. 241

[KDK17] Niraj Kumar, Eleni Diamanti, and Iordanis Kerenidis. Efficient quantum communications with coherent state fingerprints over multiple channels. *Physical Review A*, 95(3):032337, March 2017. 30

[KDW20] Wojciech Kozlowski, Axel Dahlberg, and Stephanie Wehner. Designing a quantum network protocol. In *Proceedings of the 16th International Conference on emerging Networking EXperiments and Technologies*, pages 1–16. Association for Computing Machinery, New York, NY, USA, November 2020. 171

[KG19] Mahmoud Khalafalla and Catherine Gebotys. PUFs Deep Attacks: Enhanced modeling attacks using deep learning techniques to break the security of double arbiter PUFs. In *2019 Design, Automation Test in Europe Conference Exhibition (DATE)*, pages 204–209, March 2019. ISSN: 1558-1101. 125, 151

[KHH+18] Min-Sung Kang, Jino Heo, Chang-Ho Hong, Hyung-Jin Yang, Sang-Wook Han, and Sung Moon. Controlled mutual quantum entity authentication with an untrusted third party. *Quantum Information Processing*, 17(7):159, May 2018. 171

[Kit97] A. Yu Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191, December 1997. 27

[KJ09] Max S. Kaznady and Daniel F. V. James. Numerical strategies for quantum tomography: Alternatives to full optimization. *Physical Review A*, 79(2):022109, February 2009. 85

[KKR06] Julia Kempe, Alexei Kitaev, and Oded Regev. The Complexity of the Local Hamiltonian Problem. *SIAM Journal on Computing*, 35(5):1070–1097, January 2006. 269

[KKVB02] Elham Kashefi, Adrian Kent, Vlatko Vedral, and Konrad Banaszek. Comparison of quantum oracles. *Physical Review A*, 65(5):050304, May 2002. 49, 50

[KL18] Younghyun Kim and Yongwoo Lee. CamPUF: physically unclonable function based on CMOS image sensor fixed pattern noise. In *Proceedings of the 55th Annual Design Automation Conference*, DAC '18, pages 1–6, New York, NY, USA, June 2018. Association for Computing Machinery. 174, 210

[KL20] Jonathan Katz and Yehuda Lindell. *Introduction to Modern Cryptography*. CRC Press, December 2020. Google-Books-ID: RsoOEAAAQBAJ. 44, 45, 51, 52

[KLLNP16] Marc Kaplan, Gaëtan Leurent, Anthony Leverrier, and María Naya-Plasencia. Quantum Differential and Linear Cryptanalysis. *IACR Transactions on Symmetric Cryptology*, pages 71–94, December 2016. arXiv: 1510.05836. 3, 47

[KLP+19] Sumeet Khatri, Ryan LaRose, Alexander Poremba, Lukasz Cincio, Andrew T. Sornborger, and Patrick J. Coles. Quantum-assisted quantum compiling. *Quantum*, 3:140, May 2019. 241, 242, 259, 260, 261

[KMF+16] Mario Krenn, Mehul Malik, Robert Fickler, Radek Lapkiewicz, and Anton Zeilinger. Automated Search for new Quantum Experiments. *Physical Review Letters*, 116(9):090405, March 2016. 241

[KMK21] Niraj Kumar, Rawad Mezher, and Elham Kashefi. Efficient Construction of Quantum Physical Unclonable Functions with Unitary t-designs. *arXiv:2101.05692 [quant-ph]*, January 2021. arXiv: 2101.05692. 152, 166, 168, 170

[KMY03] Hirotada Kobayashi, Keiji Matsumoto, and Tomoyuki Yamakami. Quantum Merlin-Arthur Proof Systems: Are Multiple Merlins More Helpful to Arthur? In Toshihide Ibaraki, Naoki Katoh, and Hirotaka Ono, editors, *Algorithms and Computation*, pages 189–198, Berlin, Heidelberg, 2003. Springer. 30

[Kni95] E. Knill. Approximation by Quantum Circuits. *arXiv:quant-ph/9508006*, August 1995. arXiv: quant-ph/9508006. 40, 156

[Kop18] Dawid Kopczyk. Quantum machine learning for data scientists. *arXiv:1804.10068 [quant-ph]*, April 2018. arXiv: 1804.10068. 241

[KPB00] Arun Kumar Pati and Samuel L. Braunstein. Impossibility of deleting an unknown quantum state. *Nature*, 404(6774):164–165, March 2000. 34

[Kre21] William Kretschmer. Quantum Pseudorandomness and Classical Complexity. In Min-Hsiu Hsieh, editor, *16th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2021)*, volume 197 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 2:1–2:20, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 156, 165

[KRS09] Robert Konig, Renato Renner, and Christian Schaffner. The Operational Meaning of Min- and Max-Entropy. *IEEE Transactions on Information Theory*, 55(9):4337–4347, September 2009. 23, 235

[KTP20] Isaac Kim, Eugene Tang, and John Preskill. The ghost in the radiation: robust encodings of the black hole interior. *Journal of High Energy Physics*, 2020(6):31, June 2020. 156

[KV94] Michael J. Kearns and Umesh Vazirani. *An Introduction to Computational Learning Theory*. MIT Press, August 1994. Google-Books-ID: vCA01wY6iywC. 67

[KW99] M. Keyl and R. F. Werner. Optimal cloning of pure states, testing single clones. *Journal of Mathematical Physics*, 40(7):3283–3299, July 1999. 271

[KZZ15a] Aggelos Kiayias, Thomas Zacharias, and Bingsheng Zhang. DEMOS-2: Scalable E2E Verifiable Elections without Random Oracles. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, CCS '15, pages 352–363, New York, NY, USA, October 2015. Association for Computing Machinery. 95

[KZZ15b] Aggelos Kiayias, Thomas Zacharias, and Bingsheng Zhang. End-to-End Verifiable Elections in the Standard Model. In Elisabeth Oswald and Marc Fischlin, editors, *Advances in Cryptology - EUROCRYPT 2015*, pages 468–498, Berlin, Heidelberg, 2015. Springer. 95

[LaR19] Ryan LaRose. Overview and Comparison of Gate Level Quantum Software Platforms. *Quantum*, 3:130, March 2019. 242

[LC98] Hoi-Kwong Lo and H.F. Chau. Why quantum bit commitment and ideal quantum coin tossing are impossible. *Physica D: Nonlinear Phenomena*, 120(1-2):177–187, September 1998. 246

[Lev15] Anthony Leverrier. Composable Security Proof for Continuous-Variable Quantum Key Distribution with Coherent States. *Physical Review Letters*, 114(7):070501, February 2015. 238

[LFC+20] Li Li, Minjie Fan, Marc Coram, Patrick Riley, and Stefan Leichenauer. Quantum optimization with a novel Gibbs objective function and ansatz architecture search. *Physical Review Research*, 2(2):023074, April 2020. 273

[LJBR12] Peter G. Lewis, David Jennings, Jonathan Barrett, and Terry Rudolph. Distinct Quantum States Can Be Compatible with a Single State of Reality. *Physical Review Letters*, 109(15):150404, October 2012. 75

[LLSHB02] Antıa Lamas-Linares, Christoph Simon, John C. Howell, and Dik Bouwmeester. Experimental Quantum Cloning of Single Photons. *Science*, 296(5568):712–714, April 2002. 240

[LM17] Gregor Leander and Alexander May. Grover Meets Simon – Quantumly Attacking the FX-construction. In Tsuyoshi Takagi and Thomas Peyrin, editors, *Advances in Cryptology – ASIACRYPT 2017*, pages 161–178, Cham, 2017. Springer International Publishing. 3

[LMR14] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum principal component analysis. *Nature Physics*, 10(9):631–633, September 2014. 65, 86

[LMR+17] Norbert M. Linke, Dmitri Maslov, Martin Roetteler, Shantanu Debnath, Caroline Figgatt, Kevin A. Landsman, Kenneth Wright, and Christopher Monroe. Experimental comparison of two quantum computing architectures. *Proceedings of the National Academy of Sciences*, 114(13):3305–3310, March 2017. 236

[LR88] Michael Luby and Charles Rackoff. How to Construct Pseudorandom Permutations from Pseudorandom Functions. *SIAM Journal on Computing*, 17(2):373–386, April 1988. 155

[LRGR+21] Dario Lago-Rivera, Samuele Grandi, Jelena V. Rakonjac, Alessandro Seri, and Hugues de Riedmatten. Telecom-heralded entanglement between multimode solid-state quantum memories. *Nature*, 594(7861):37–40, June 2021. 48, 210

[LS20] Matteo Lostaglio and Gabriel Senno. Contextual advantage for state-dependent cloning. *Quantum*, 4:258, April 2020. 39

[LST09] Alexander I. Lvovsky, Barry C. Sanders, and Wolfgang Tittel. Optical quantum memory. *Nature Photonics*, 3(12):706–714, December 2009. 48, 175, 210

[LSY18] Hanxiao Liu, Karen Simonyan, and Yiming Yang. DARTS: Differentiable Architecture Search. September 2018. 273

[LTOJ+19] Ryan LaRose, Arkin Tikku, Étude O'Neel-Judy, Lukasz Cincio, and Patrick J. Coles. Variational quantum state diagonalization. *npj Quantum Information*, 5(1):1–10, June 2019. 259

[LWG+10] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White. Towards quantum chemistry on a quantum computer. *Nature Chemistry*, 2(2):106–111, February 2010. 3

[LZZ+19] Weiqiang Liu, Lei Zhang, Zhengran Zhang, Chongyan Gu, Chenghua Wang, Maire O'neill, and Fabrizio Lombardi. XOR-Based Low-Cost Reconfigurable PUFs for IoT Security. *ACM Transactions on Embedded Computing Systems*, 18(3):25:1–25:21, April 2019. 125

[Mae13] Roel Maes. Physically Unclonable Functions: Properties. In Roel Maes, editor, *Physically Unclonable Functions: Constructions, Properties and Applications*, pages 49–80. Springer, Berlin, Heidelberg, 2013. 125, 174, 210

[MAK+18] Charis Mesaritakis, Marialena Akriotou, Alexandros Kapsalis, Evangelos Grivas, Charidimos Chaintoutis, Thomas Nikas, and Dimitris Syvridis. Physical Unclonable Function based on a Multi-Mode Optical Waveguide. *Scientific Reports*, 8(1):9653, June 2018. 126

[man80] *Computable and uncomputable*. 128. Sovetskoye Radio, Moscow, 1980. 2

[Mau93] Ueli M. Maurer. Protocols for Secret Key Agreement by Public Discussion Based on Common Information. In Ernest F. Brickell, editor, *Advances in Cryptology — CRYPTO' 92*, volume 740, pages 461–470. Springer Berlin Heidelberg, Berlin, Heidelberg, 1993. 48

[Mau05] Ueli Maurer. Abstract Models of Computation in Cryptography. In Nigel P. Smart, editor, *Cryptography and Coding*, pages 1–12, Berlin, Heidelberg, 2005. Springer. 45, 238

[Mau12] Ueli Maurer. Constructive Cryptography – A New Paradigm for Security Definitions and Proofs. In Sebastian Mödersheim and Catuscia Palamidessi, editors, *Theory of Security and Applications*, pages 33–56, Berlin, Heidelberg, 2012. Springer. 238

[MBK21] Andrea Mari, Thomas R. Bromley, and Nathan Killoran. Estimating the gradient and higher-order derivatives on quantum hardware. *Physical Review A*, 103(1):012405, January 2021. 71

[MBM+18] Cédric Marchand, Lilian Bossuet, Ugo Mureddu, Nathalie Bochard, Abdelkarim Cherkaoui, and Viktor Fischer. Implementation and Characterization of a Physical Unclonable Function for IoT: A Case Study With the TERO-PUF. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 37(1):97–109, January 2018. 125

[MBS+18] Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9(1):4812, November 2018. 55, 259

[MEAG+20] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan. Quantum computational chemistry. *Reviews of Modern Physics*, 92(1):015003, March 2020. 3

[Mec19] Elizabeth S. Meckes. *The Random Matrix Theory of the Classical Compact Groups*. Cambridge University Press, August 2019. Google-Books-ID: BqCkDwAAQBAJ. 41, 169

[MGdF+21] Konstantinos Meichanetzidis, Stefano Gogioso, Giovanni de Felice, Nicolò Chiappori, Alexis Toumi, and Bob Coecke. Quantum Natural Language Processing on Near-Term Quantum Computers. *Electronic Proceedings in Theoretical Computer Science*, 340:213–229, September 2021. arXiv: 2005.04147. 3

[MGL22] José D. Martín-Guerrero and Lucas Lamata. Quantum Machine Learning: A tutorial. *Neurocomputing*, 470:457–461, January 2022. 60, 69

[ML16] Iman Marvian and Seth Lloyd. Universal Quantum Emulator. *arXiv:1606.02734 [quant-ph]*, June 2016. arXiv: 1606.02734. 62, 63, 65, 66, 85, 86, 87

[MNKF18] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii. Quantum circuit learning. *Physical Review A*, 98(3):032309, September 2018. 241

[Moc07] Carlos Mochon. Quantum weak coin flipping with arbitrarily small bias. *arXiv:0711.4114 [quant-ph]*, November 2007. arXiv: 0711.4114. 59

[Mos18] Michele Mosca. Cybersecurity in an Era with Quantum Computers: Will We Be Ready? *IEEE Security Privacy*, 16(5):38–41, September 2018. 46

[MPNK+18] Alexey A. Melnikov, Hendrik Poulsen Nautrup, Mario Krenn, Vedran Dunjko, Markus Tiersch, Anton Zeilinger, and Hans J. Briegel. Active learning machine learns to create new quantum experiments. *Proceedings of the National Academy of Sciences*, 115(6):1221–1226, February 2018. 241

[MPO22] Filip B. Maciejewski, Zbigniew Puchała, and Michał Oszmaniec. Exploring Quantum Average-Case Distances: proofs, properties, and examples. *arXiv:2112.14284 [quant-ph]*, January 2022. arXiv: 2112.14284. 16

[MPP16] Houssem Maghrebi, Thibault Portigliatti, and Emmanuel Prouff. Breaking Cryptographic Implementations Using Deep Learning Techniques. Technical Report 921, 2016. 241

[MPS+10] Kavan Modi, Tomasz Paterek, Wonmin Son, Vlatko Vedral, and Mark Williamson. Unified View of Quantum and Classical Correlations. *Physical Review Letters*, 104(8):080501, February 2010. 16

[MQR09] Jörn Müller-Quade and Renato Renner. Composability in quantum cryptography. *New Journal of Physics*, 11(8):085006, August 2009. 45

[MR09] Daniele Micciancio and Oded Regev. Lattice-based Cryptography. In Daniel J. Bernstein, Johannes Buchmann, and Erik Dahmen, editors, *Post-Quantum Cryptography*, pages 147–191. Springer, Berlin, Heidelberg, 2009. 43

[MR16] Ueli Maurer and Renato Renner. From Indifferentiability to Constructive Cryptography (and Back). In Martin Hirt and Adam Smith, editors, *Theory of Cryptography*, pages 3–24, Berlin, Heidelberg, 2016. Springer. 45

[MRBAG16] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2):023023, February 2016. 69, 241

[MRL08] M. Mohseni, A. T. Rezakhani, and D. A. Lidar. Quantum-process tomography: Resource analysis of different strategies. *Physical Review A*, 77(3):032322, March 2008. 62, 80

[MSB04] Masoud Mohseni, Aephraim M. Steinberg, and János A. Bergou. Optical Realization of Optimal Unambiguous Discrimination for Pure and Mixed Quantum States. *Physical Review Letters*, 93(20):200403, November 2004. 29

[MSCK99] Dominic Mayers, Louis Salvail, and Yoshie Chiba-Kohno. Unconditionally Secure Quantum Coin Tossing. *arXiv:quant-ph/9904078*, April 1999. arXiv: quant-ph/9904078 version: 1. xvi, 59, 245, 246, 247, 248, 278

[MSCK18] Dominic Mayers, Louis Salvail, and Yoshie Chiba-Kohno. Unconditionally Secure Quantum Coin Tossing. *arXiv:quant-ph/9904078*, February 2018. arXiv: quant-ph/9904078. 242

[MTB18] Mauro E. S. Morales, Timur Tlyachev, and Jacob Biamonte. Variational learning of Grover's quantum search algorithm. *Physical Review A*, 98(6):062333, December 2018. 241

[MTdFC20] Konstantinos Meichanetzidis, Alexis Toumi, Giovanni de Felice, and Bob Coecke. Grammar-Aware Question-Answering on Quantum Computers. *arXiv:2012.03756 [quant-ph]*, December 2020. arXiv: 2012.03756. 3

[MU88] Hans Maassen and J. B. M. Uffink. Generalized entropic uncertainty relations. *Physical Review Letters*, 60(12):1103–1106, March 1988. 21

[MU17] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data*

Analysis. Cambridge University Press, July 2017. Google-Books-ID: E9UIDwAAQBAJ. 189

[Muk16] Debdeep Mukhopadhyay. PUFs as Promising Tools for Security in Internet of Things. *IEEE Design Test*, 33(3):103–115, June 2016. 125

[MY21] Tomoyuki Morimae and Takashi Yamakawa. Quantum commitments and signatures without one-way functions. Technical Report 1691, 2021. 156

[NC03] Patrick Navez and Nicolas J. Cerf. Cloning a real d-dimensional quantum state on the edge of the no-signaling condition. *Physical Review A*, 68(3):032313, September 2003. 34

[NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, Cambridge, 2010. 1, 10, 13, 14, 15, 16, 17, 18, 20, 25, 26, 27, 79, 80, 131, 137, 138, 163

[ND17] Georgios M. Nikolopoulos and Eleni Diamanti. Continuous-variable quantum authentication of physical unclonable keys. *Scientific Reports*, 7(1):46047, April 2017. 127, 134, 149, 151, 152, 174, 176

[NG99] Chi-Sheng Niu and Robert B. Griffiths. Two-qubit copying machine for economical quantum eavesdropping. *Physical Review A*, 60(4):2764–2776, October 1999. 37

[Nik18] Georgios M. Nikolopoulos. Continuous-variable quantum authentication of physical unclonable keys: Security against an emulation attack. *Physical Review A*, 97(1):012324, January 2018. 127, 151, 152, 174

[Nik21] Georgios M. Nikolopoulos. Remote Quantum-Safe Authentication of Entities with Physical Unclonable Functions. *Photonics*, 8(7):289, July 2021. 153, 174

[NMR+19] Rosanna Nichols, Lana Mineh, Jesús Rubio, Jonathan C F Matthews, and Paul A Knott. Designing quantum experiments with a genetic algorithm. *Quantum Science and Technology*, 4(4):045012, October 2019. 241

[NZO+21] Yoshifumi Nakata, Da Zhao, Takayuki Okuda, Eiichi Bannai, Yasunari Suzuki, Shiro Tamiya, Kentaro Heya, Zhiguang Yan, Kun Zuo, Shuhei Tamate, Yutaka Tabuchi, and Yasunobu Nakamura. Quantum Circuits for Exact Unitary t-Designs and Applications to Higher-Order Randomized Benchmarking. *PRX Quantum*, 2(3):030339, September 2021. 55, 156

[OGB21] Mateusz Ostaszewski, Edward Grant, and Marcello Benedetti. Structure optimization for parameterized quantum circuits. *Quantum*, 5:391, January 2021. arXiv: 1905.09692. 273

[OGHW16] Michał Oszmaniec, Andrzej Grudka, Michał Horodecki, and Antoni Wójcik. Creating a Superposition of Unknown Quantum States. *Physical Review Letters*, 116(11):110403, March 2016. 34, 140

[OIO+12] Ryo Okamoto, Minako Iefuji, Satoshi Oyama, Koichi Yamagata, Hiroshi Imai, Akio Fujiwara, and Shigeki Takeuchi. Experimental Demonstration of Adaptive Quantum State Estimation. *Physical Review Letters*, 109(13):130404, September 2012. 28

[Omn02] Roland Omnès. *Quantum Philosophy: Understanding and Interpreting Contemporary Science*. Princeton University Press, February 2002. 75

[ONK19] L. O'Driscoll, R. Nichols, and P. A. Knott. A hybrid machine learning algorithm for designing quantum experiments. *Quantum Machine Intelligence*, 1(1):5–15, May 2019. 241

[OSS+21] Carlos Outeiral, Martin Strahm, Jiye Shi, Garrett M. Morris, Simon C. Benjamin, and Charlotte M. Deane. The prospects of quantum computing in computational molecular biology. *WIREs Computational Molecular Science*, 11(1), January 2021. 3

[OSVW13] Rafail Ostrovsky, Alessandra Scafuro, Ivan Visconti, and Akshay Wadia. Universally Composable Secure Computation with (Malicious) Physically Uncloneable Functions. In Thomas Johansson and Phong Q. Nguyen, editors, *Advances in Cryptology – EUROCRYPT 2013*, Lecture Notes in Computer Science, pages 702–718, Berlin, Heidelberg, 2013. Springer. 135

[OW16] Ryan O'Donnell and John Wright. Efficient quantum tomography. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, STOC '16, pages 899–912, New York, NY, USA, June 2016. Association for Computing Machinery. 85

[Par70] James L. Park. The concept of transition in quantum mechanics. *Foundations of Physics*, 1(1):23–33, March 1970. 33

[PB16] Stefano Pirandola and Samuel L. Braunstein. Physics: Unite to build a quantum Internet. *Nature*, 532(7598):169–171, April 2016. 236

[PBR12] Matthew F. Pusey, Jonathan Barrett, and Terry Rudolph. On the reality of the quantum state. *Nature Physics*, 8(6):475–478, June 2012. 75

[Per88] Asher Peres. How to differentiate between non-orthogonal states. *Physics Letters A*, 128(1):19, March 1988. 29

[PHH08] Marco Piani, Paweł Horodecki, and Ryszard Horodecki. No-Local-Broadcasting Theorem for Multipartite Quantum Correlations. *Physical Review Letters*, 100(9):090502, March 2008. 34

[PM22a] Arun Padakandla and Abram Magner. PAC Learning of Quantum Measurement Classes : Sample Complexity Bounds and Universal Consistency. In *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, pages 11305–11319. PMLR, May 2022. 68

[PM22b] Arun Padakandla and Abram Magner. PAC Learning of Quantum Measurement Classes : Sample Complexity Bounds and Universal Consistency. In *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*, pages 11305–11319. PMLR, May 2022. 85, 121

[PMS+14] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O'Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(1):4213, July 2014. 241

[PMSW16] Nicolas Papernot, Patrick McDaniel, Arunesh Sinha, and Michael Wellman. Towards the Science of Security and Privacy in Machine Learning. *arXiv:1611.03814 [cs]*, November 2016. arXiv: 1611.03814. 241

[PPA+20] S. Pirandola, S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. Shamsul Shaari, J. Shamsul Shaari, M. Tomamichel, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden. Advances in quantum cryptography. *Advances in Optics and Photonics*, 12(4):1012–1236, December 2020. 171

[PPM08] A. Poppe, M. Peev, and O. Maurhart. Outline of the secoqc quantum-key-distribution network in vienna. *International Journal of Quantum Information*, 06(02):209–218, April 2008. 238

[PPS21] Anna Pappa, Niklas Pirnay, and Jean-Pierre Seifert. Learning Classical Readout Quantum PUFs based on single-qubit gates. *arXiv:2112.06661 [quant-ph]*, December 2021. arXiv: 2112.06661. 153

[PR04] Matteo Paris and Jaroslav Rehacek. *Quantum State Estimation*. Springer Science & Business Media, August 2004. Google-Books-ID: Grr25VFtGgUC. 28, 61, 85

[PR21] Christopher Portmann and Renato Renner. Security in Quantum Cryptography. *arXiv:2102.00021 [quant-ph]*, August 2021. arXiv: 2102.00021. 46, 49

[Pre18] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018. 2

[Pre21] John Preskill. Quantum computing 40 years later. *arXiv:2106.10522 [quant-ph]*, June 2021. arXiv: 2106.10522. 2

[PRTG02] Ravikanth Pappu, Ben Recht, Jason Taylor, and Neil Gershenfeld. Physical One-Way Functions. *Science*, September 2002. 125, 126, 176

[PSA+21] Koustubh Phalak, Abdullah Ash Saki, Mahabubul Alam, Rasit Onur Topaloglu, and Swaroop Ghosh. Quantum PUF for Security and Trust in Quantum Computing. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 11(2):333–342, June 2021. 153

[PT21] Mohammad Pirhooshyaran and Tamás Terlaky. Quantum circuit design search. *Quantum Machine Intelligence*, 3(2):25, October 2021. 273

[QHL+13] Bo Qi, Zhibo Hou, Li Li, Daoyi Dong, Guoyong Xiang, and Guangcan Guo. Quantum State Tomography via Linear Regression Estimation. *Scientific Reports*, 3(1):3496, December 2013. 61

[RB01] Robert Raussendorf and Hans J. Briegel. A One-Way Quantum Computer. *Physical Review Letters*, 86(22):5188–5191, May 2001. 25

[Ren08] Renato Renner. Security of quantum key distribution. *International Journal of Quantum Information*, 06(01):1–127, February 2008. 20, 23, 24, 238

[RH14] Ulrich Rührmair and Daniel E. Holcomb. PUFs at a glance. In *2014 Design, Automation Test in Europe Conference Exhibition (DATE)*, pages 1–6, March 2014. ISSN: 1558-1101. 124, 125, 151

[Ric88] Richard Phillips Feynman. *QED*. Princeton University Press, October 1988. 1

[Rig] Rigetti. Welcome to the Docs for pyQuil! — pyQuil 3.1.0 documentation. 172

[RKKN19] Tomasz Rymarczyk, Edward Kozłowski, Grzegorz Kłosowski, and Konrad Niderla. Logistic Regression for Machine Learning in Process Tomography. *Sensors*, 19(15):3400, January 2019. 61

[Rom90] J. Rompel. One-way functions are necessary and sufficient for secure signatures. In *Proceedings of the twenty-second annual ACM symposium on Theory of computing - STOC '90*, pages 387–394, Baltimore, Maryland, United States, 1990. ACM Press. 155

[RS14] Ulrich Rührmair and Jan Sölter. PUF modeling attacks: An introduction and overview. In *2014 Design, Automation Test in Europe Conference Exhibition (DATE)*, pages 1–6, March 2014. ISSN: 1558-1101. 127, 151

[RSK21] Yosef Rinott, Tomer Shoham, and Gil Kalai. Statistical Aspects of the Quantum Supremacy Demonstration. *arXiv:2008.05177 [quant-ph, stat]*, July 2021. arXiv: 2008.05177. 2

[RSS09] Ulrich Rührmair, Jan Sölter, and Frank Sehnke. On the Foundations of Physical Unclonable Functions. Technical Report 277, 2009. 156

[RSS+10] Ulrich Rührmair, Frank Sehnke, Jan Sölter, Gideon Dror, Srinivas Devadas, and Jürgen Schmidhuber. Modeling attacks on physical unclonable functions. In *Proceedings of the 17th ACM conference on Computer and communications security*, CCS '10, pages 237–249, New York, NY, USA, October 2010. Association for Computing Machinery. 125, 151, 174, 210

[RST03] Terry Rudolph, Robert W. Spekkens, and Peter S. Turner. Unambiguous discrimination of mixed states. *Physical Review A*, 68(1):010301, July 2003. 29

[RWR+18] Alexander Radovic, Mike Williams, David Rousseau, Michael Kagan, Daniele Bonacorsi, Alexander Himmel, Adam Aurisano, Kazuhiro Terao, and Taritree Wongjirad. Machine learning at the energy and intensity frontiers of particle physics. *Nature*, 560(7716):41–48, August 2018. 60

[Ré61] Alfréd Rényi. On Measures of Entropy and Information. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, 4.1:547–562, January 1961. 23

[SBG+19] Maria Schuld, Ville Bergholm, Christian Gogolin, Josh Izaac, and Nathan Killoran. Evaluating analytic gradients on quantum hardware. *Physical Review A*, 99(3):032331, March 2019. 71, 259

[SBPC+09] Valerio Scarani, Helle Bechmann-Pasquinucci, Nicolas J. Cerf, Miloslav Dušek, Norbert Lütkenhaus, and Momtchil Peev. The security of practical quantum key distribution. *Reviews of Modern Physics*, 81(3):1301–1350, September 2009. 238

[SD07] G. Edward Suh and Srinivas Devadas. Physical Unclonable Functions for Device Authentication and Secret Key Generation. In *2007 44th ACM/IEEE Design Automation Conference*, pages 9–14, June 2007. ISSN: 0738-100X. 125, 230

[SEG+17] P. Sibson, C. Erven, M. Godfrey, S. Miki, T. Yamashita, M. Fujiwara, M. Sasaki, H. Terai, M. G. Tanner, C. M. Natarajan, R. H. Hadfield, J. L. O'Brien, and M. G. Thompson. Chip-based quantum key distribution. *Nature Communications*, 8(1):13984, February 2017. 238

[SFI+11] M. Sasaki, M. Fujiwara, H. Ishizuka, W. Klaus, K. Wakui, M. Takeoka, S. Miki, T. Yamashita, Z. Wang, A. Tanaka, K. Yoshino, Y. Nambu, S. Takahashi, A. Tajima, A. Tomita, T. Domeki, T. Hasegawa, Y. Sakai, H. Kobayashi, T. Asai, K. Shimizu, T. Tokura, T. Tsurumaru, M. Matsui, T. Honjo, K. Tamaki, H. Takesue, Y. Tokura, J. F. Dynes, A. R. Dixon, A. W. Sharpe, Z. L. Yuan, A. J. Shields, S. Uchikoga, M. Legré, S. Robyr, P. Trinkler, L. Monat, J.-B. Page, G. Ribordy, A. Poppe, A. Allacher, O. Maurhart, T. Länger, M. Peev, and A. Zeilinger. Field test of quantum key distribution in the Tokyo QKD Network. *Optics Express*, 19(11):10387–10409, May 2011. 238

[SG04] Rocco A. Servedio and Steven J. Gortler. Equivalences and Separations Between Quantum and Classical Learnability. *SIAM Journal on Computing*, 33(5):1067–1092, January 2004. 67, 68, 69, 238

[Sha83] Adi Shamir. On the generation of cryptographically strong pseudorandom sequences. *ACM Transactions on Computer Systems*, 1(1):38–44, February 1983. 155

[Sho94] P.W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, November 1994. 3, 46, 241

[SIGA05] Valerio Scarani, Sofyan Iblisdir, Nicolas Gisin, and Antonio Acín. Quantum cloning. *Reviews of Modern Physics*, 77(4):1225–1256, November 2005. 35, 36, 37, 79, 240, 243, 244, 260, 271, 272, 277

[SJS16] Vladimir Soukharev, David Jao, and Srinath Seshadri. Post-Quantum Security Models for Authenticated Encryption. In Tsuyoshi Takagi, editor, *Post-Quantum Cryptography*, pages 64–78, Cham, 2016. Springer International Publishing. 92

[SK75] David Sherrington and Scott Kirkpatrick. Solvable Model of a Spin-Glass. *Physical Review Letters*, 35(26):1792–1796, December 1975. 60

[SKCC20] Kunal Sharma, Sumeet Khatri, M Cerezo, and Patrick J Coles. Noise resilience of variational quantum compiling. *New Journal of Physics*, 22(4):043006, April 2020. 259

[Sko10] Boris Skorić. Quantum Readout of Physical Unclonable Functions. In Daniel J. Bernstein and Tanja Lange, editors, *Progress in Cryptology – AFRICACRYPT 2010*, Lecture Notes in Computer Science, pages 369–386, Berlin, Heidelberg, 2010. Springer. 127, 149, 151, 152, 174

[Sko12] Boris Skorić. Quantum readout of Physical Unclonable Functions. *International Journal of Quantum Information*, 10(01):1250001, February 2012. 151, 152, 174

[SLB+11] D Stucki, M Legré, F Buntschu, B Clausen, N Felber, N Gisin, L Henzen, P Junod, G Litzistorf, P Monbaron, L Monat, J-B Page, D Perroud, G Ribordy, A Rochas, S Robyr, J Tavares, R Thew, P Trinkler, S Ventura, R Voirol, N Walenta, and H Zbinden. Long-term performance of the SwissQuantum quantum key distribution network in a field environment. *New Journal of Physics*, 13(12):123001, December 2011. 238

[SMP13] Boris Skorić, Allard P. Mosk, and Pepijn W. H. Pinkse. SECURITY OF QUANTUM-READOUT PUFs AGAINST QUADRATURE-BASED CHALLENGE-ESTIMATION ATTACKS. *International Journal of Quantum Information*, 11(04):1350041, June 2013. 127, 151, 152

[Son14] Fang Song. A Note on Quantum Security for Post-Quantum Cryptography. In Michele Mosca, editor, *Post-Quantum Cryptography*, pages 246–265, Cham, 2014. Springer International Publishing. 46, 47

[SP18a] Maria Schuld and Francesco Petruccione. Machine Learning. In Maria Schuld and Francesco Petruccione, editors, *Supervised Learning with Quantum Computers*, pages 21–73. Springer International Publishing, Cham, 2018. 241

[SP18b] Maria Schuld and Francesco Petruccione. Prospects for Near-Term Quantum Machine Learning. In Maria Schuld and Francesco Petruccione, editors, *Supervised Learning with Quantum Computers*, pages 273–279. Springer International Publishing, Cham, 2018. 69, 241

[SPM17] Boris Skorić, Pepijn W. H. Pinkse, and Allard P. Mosk. Authenticated communication from quantum readout of PUFs. *Quantum Information Processing*, 16(8):200, July 2017. 127

[SS17] Thomas Santoli and Christian Schaffner. Using Simon's Algorithm to Attack Symmetric-Key Cryptographic Primitives. *arXiv:1603.07856 [quant-ph]*, January 2017. arXiv: 1603.07856. 47

[SSH+20] Henry Semenenko, Philip Sibson, Andy Hart, Mark G. Thompson, John G. Rarity, and Chris Erven. Chip-based measurement-device-independent quantum key distribution. *Optica*, 7(3):238–242, March 2020. 238

[SSM21] Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Physical Review A*, 103(3):032430, March 2021. 238

[SSP15] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. An introduction to quantum machine learning. *Contemporary Physics*, 56(2):172–185, April 2015. 60, 69

[Stu05] E. Study. Kürzeste Wege im komplexen Gebiet. *Mathematische Annalen*, 60(3):321–378, September 1905. 17

[SW22] Or Sattath and Shai Wyborski. Uncloneable Decryptors from Quantum Copy-Protection. *arXiv:2203.05866 [quant-ph]*, March 2022. arXiv: 2203.05866. 82

[TL17] Marco Tomamichel and Anthony Leverrier. A largely self-contained and complete security proof for quantum key distribution. *Quantum*, 1:14, July 2017. 20, 238

[TMC+18] Giacomo Torlai, Guglielmo Mazzola, Juan Carrasquilla, Matthias Troyer, Roger Melko, and Giuseppe Carleo. Neural-network quantum state tomography. *Nature Physics*, 14(5):447–450, May 2018. 61

[TPI19] Lars Tebelmann, Michael Pehl, and Vincent Immler. Side-Channel Analysis of the TERO PUF. In Ilia Polian and Marc Stöttinger, editors, *Constructive Side-Channel Analysis and Secure Design*, Lecture Notes in Computer Science, pages 43–60, Cham, 2019. Springer International Publishing. 125

[TR11] Marco Tomamichel and Renato Renner. Uncertainty Relation for Smooth Entropies. *Physical Review Letters*, 106(11):110506, March 2011. 233

[Unr11] Dominique Unruh. Concurrent Composition in the Bounded Quantum Storage Model. In Kenneth G. Paterson, editor, *Advances in Cryptology – EUROCRYPT 2011*, pages 467–486, Berlin, Heidelberg, 2011. Springer. 49

[Unr13] Dominique Unruh. Everlasting Multi-party Computation. In Ran Canetti and Juan A. Garay, editors, *Advances in Cryptology – CRYPTO 2013*, pages 380–397, Berlin, Heidelberg, 2013. Springer. 171

[UWG+19] Ravitej Uppu, Tom A W Wolterink, Sebastianus A Goorden, Bin Chen, Boris Skorić, Allard P Mosk, and Pepijn W H Pinkse. Asymmetric cryptography with physical unclonable keys. *Quantum Science and Technology*, 4(4):045011, October 2019. 127

[Val84] Leslie Valiant. A Theory of the Learnable. *Communications of the ACM*, 27(11), 1984. 67

[Ver19] VeriQloud. Quantum Protocol Zoo, 2019. 4, 81, 171, 286

[Wat03] John Watrous. On the complexity of simulating space-bounded quantum computations. *computational complexity*, 12(1):48–84, June 2003. 175

[WCL+21] Pidong Wang, Feiliang Chen, Dong Li, Song Sun, Feng Huang, Taiping Zhang, Qian Li, Kun Chen, Yongbiao Wan, Xiao Leng, and Yao Yao. Authentication of Optical Physical Unclonable Functions Based on Single-Pixel Detection. *Physical Review Applied*, 16(5):054025, November 2021. 153

[WCY+14] Shuang Wang, Wei Chen, Zhen-Qiang Yin, Hong-Wei Li, De-Yong He, Yu-Hu Li, Zheng Zhou, Xiao-Tian Song, Fang-Yi Li, Dong Wang, Hua Chen, Yun-Guang Han, Jing-Zheng Huang, Jun-Fu Guo, Peng-Lei Hao, Mo Li, Chun-Mei Zhang, Dong Liu, Wen-Ye Liang, Chun-Hua Miao, Ping Wu, Guang-Can Guo, and Zheng-Fu Han. Field and long-term demonstration of a wide area quantum key distribution network. *Optics Express*, 22(18):21739–21756, September 2014. 238

[WEH18] Stephanie Wehner, David Elkouss, and Ronald Hanson. Quantum internet: A vision for the road ahead. *Science*, 362(6412):eaam9288, October 2018. 4, 171, 215

[Wer98] R. F. Werner. Optimal cloning of pure states. *Physical Review A*, 58(3):1827–1832, September 1998. 271

[WHT15] Dave Wecker, Matthew B. Hastings, and Matthias Troyer. Progress towards practical quantum variational algorithms. *Physical Review A*, 92(4):042303, October 2015. 241

[Wie83] Stephen Wiesner. Conjugate coding. *ACM SIGACT News*, 15(1):78–88, January 1983. 81, 212, 245

[Wie02] K. Wieand. Eigenvalue distributions of random unitary matrices. *Probability Theory and Related Fields*, 123(2):202–224, June 2002. 42

[Wik22] Wikipedia. Bloch sphere, April 2022. Page Version ID: 1084268363. xv, 11

[Wis21] Nils Wisiol. nils-wisiol/pypuf: None, August 2021. 228

[WK19] Petros Wallden and Elham Kashefi. Cyber security in the quantum era. *Communications of the ACM*, 62(4):120, March 2019. 3, 43, 46

[WLZ+19] Yunfei Wang, Jianfeng Li, Shanchao Zhang, Keyu Su, Yiru Zhou, Kaiyu Liao, Shengwang Du, Hui Yan, and Shi-Liang Zhu. Efficient quantum memory for single-photon polarization qubits. *Nature Photonics*, 13(5):346–351, May 2019. 48, 210

[WM20] Logan G. Wright and Peter L. McMahon. The Capacity of Quantum Neural Networks. In *Conference on Lasers and Electro-Optics*, page JM4G.5, Washington, DC, 2020. OSA. 241

[WMDB20] Julius Wallnöfer, Alexey A. Melnikov, Wolfgang Dür, and Hans J. Briegel. Machine Learning for Long-Distance Quantum Communication. *PRX Quantum*, 1(1):010301, September 2020. 241

[WMH+20] Andreas Wallucks, Igor Marinković, Bas Hensen, Robert Stockill, and Simon Gröblacher. A quantum memory at telecom wavelengths. *Nature Physics*, 16(7):772–777, July 2020. 48, 210

[WSK+21] Daochen Wang, Aarthi Sundaram, Robin Kothari, Ashish Kapoor, and Martin Roetteler. Quantum algorithms for reinforcement learning with a generative model. In *Proceedings of the 38th International Conference on Machine Learning*, pages 10916–10926. PMLR, July 2021. 85

[WST08] Stephanie Wehner, Christian Schaffner, and Barbara M. Terhal. Cryptography from Noisy Storage. *Physical Review Letters*, 100(22):220502, June 2008. 49

[WZ82] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, October 1982. 33

[XAW+15] Feihu Xu, Juan Miguel Arrazola, Kejin Wei, Wenyuan Wang, Pablo Palacios-Avila, Chen Feng, Shihan Sajeed, Norbert Lütkenhaus, and Hoi-Kwong Lo. Experimental quantum fingerprinting with weak coherent pulses. *Nature Communications*, 6(1):8735, October 2015. 30

[XSE+21] Xiaosi Xu, Jinzhao Sun, Suguru Endo, Ying Li, Simon C. Benjamin, and Xiao Yuan. Variational algorithms for linear algebra. *Science Bulletin*, 66(21):2181–2188, November 2021. 241

[XSW+12] Zhao-Xi Xiong, Han-Duo Shi, Yi-Nan Wang, Li Jing, Jin Lei, Liang-Zhu Mu, and Heng Fan. General quantum key distribution in higher dimension. *Physical Review A*, 85(1):012334, January 2012. 240

[XX18] Qian Xu and Shuqi Xu. Neural network state estimation for full quantum state tomography. *arXiv:1811.06654 [quant-ph]*, November 2018. arXiv: 1811.06654. 61

[Yao82] Andrew C. Yao. Theory and application of trapdoor functions. In *23rd Annual Symposium on Foundations of Computer Science (sfcs 1982)*, pages 80–91, November 1982. ISSN: 0272-5428. 155

[YGLZ16] Yao Yao, Ming Gao, Mo Li, and Jian Zhang. Quantum cloning attacks against PUF-based quantum authentication systems. *Quantum Information Processing*, 15(8):3311–3325, August 2016. 127, 174

[YHD+16] Meng-Day Yu, Matthias Hiller, Jeroen Delvaux, Richard Sowell, Srinivas Devadas, and Ingrid Verbauwhede. A Lockdown Technique to Prevent Machine Learning on PUFs for Lightweight Authentication. *IEEE Transactions on Multi-Scale Computing Systems*, 2(3):146–159, July 2016. 174, 213

[YWC+19] Quanming Yao, Mengshuo Wang, Yuqiang Chen, Wenyuan Dai, Yu-Feng Li, Wei-Wei Tu, Qiang Yang, and Yang Yu. Taking Human out of Learning Applications: A Survey on Automated Machine Learning. *arXiv:1810.13306 [cs, stat]*, December 2019. arXiv: 1810.13306. 273

[Zha12] Mark Zhandry. How to Construct Quantum Random Functions. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 679–687, October 2012. ISSN: 0272-5428. 52

[Zha15] Mark Zhandry. Secure identity-based encryption in the quantum random oracle model. *International Journal of Quantum Information*, 13(04):1550014, June 2015. 113

[Zha19] Mark Zhandry. How to Record Quantum Queries, and Applications to Quantum Indifferentiability. In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology – CRYPTO 2019*, pages 239–268, Cham, 2019. Springer International Publishing. 50

[Zha21] Mark Zhandry. Quantum Lightning Never Strikes the Same State Twice. Or: Quantum Money from Cryptographic Assumptions. *Journal of Cryptology*, 34(1):6, January 2021. 81

[ZHZY21] Shi-Xin Zhang, Chang-Yu Hsieh, Shengyu Zhang, and Hong Yao. Differentiable Quantum Architecture Search. *arXiv:2010.08561 [quant-ph]*, October 2021. arXiv: 2010.08561. 273

[ZS05] Karol Zyczkowski and Hans-Jürgen Sommers. Average fidelity between random quantum states. *Physical Review A*, 71(3):032313, March 2005. 145

[ZWD+20] Han-Sen Zhong, Hui Wang, Yu-Hao Deng, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Jian Qin, Dian Wu, Xing Ding, Yi Hu, Peng Hu, Xiao-Yan Yang, Wei-Jun Zhang, Hao Li, Yuxuan Li, Xiao Jiang, Lin Gan, Guangwen Yang, Lixing You, Zhen Wang, Li Li, Nai-Le Liu, Chao-Yang Lu, and Jian-Wei Pan. Quantum computational advantage using photons. *Science*, 370(6523):1460–1463, December 2020. 241