Author:
Hance, Jonte R

Title:
The interplay between quantum foundations and quantum technologies
Counterfactual communication, and extensions of quantum mechanics

The interplay between quantum foundations and quantum technologies

Counterfactual communication, and extensions of quantum mechanics

By

JONTE RODERICK HANCE

Department of Electrical and Electronic Engineering
UNIVERSITY OF BRISTOL

A dissertation submitted to the
UNIVERSITY OF BRISTOL
in accordance with the requirements of the degree of
DOCTOR OF PHILOSOPHY
in the
FACULTY OF ENGINEERING

FEBRUARY 2023

Word count: 63722

Abstract

This thesis investigates quantum foundations, evaluating and leveraging this area for the development of quantum technologies. It also demonstrates the usefulness of quantum technologies for experimentally testing foundational hypotheses. It does this by looking at two areas within foundations: counterfactual communication, and extensions of quantum mechanics.

The thesis first focuses on counterfactual communication, and its antecedent, interaction-free measurement. It evaluates the philosophical and foundational issues surrounding these novel developing technologies, such as whether they are counterfactual (by various proposed criteria), and whether they are quantum. It also gives potential practical applications for which we can employ these technologies, such as to transfer quantum information, or to image delicate samples without damaging them.

The thesis then investigates extensions of quantum mechanics: specifically those which allow us to regain Bell-locality by weakening statistical independence. These extensions—which could act as a path to unifying quantum mechanics with general relativity—give predictions differing from standard quantum mechanics on scales only recently made accessible (e.g., using noisy intermediate-scale quantum devices). It then discusses possible experiments which would allow us to test these predictions empirically.

The work in this thesis contributes to the underpinning science of quantum technologies, showing how quantum foundational ideas can be adapted into quantum technological applications. It also shows quantum technologies can benefit quantum foundations—how these technologies can be utilised to test foundational hypotheses, and demonstrate foundational principles. Therefore, this work demonstrates the interplay between quantum foundations and quantum technologies—an often neglected, but vital, part of both of these fields.

Dedication and acknowledgements

I would first like to express my deepest thanks to my two supervisors, Prof John Rarity and Prof James Ladyman. While John specifically led on supervising the technological and applied parts of this thesis, and James the more foundational and philosophical aspects, they both were heavily involved with the interdisciplinary parts of this work. Their support, help, and encouragement has been invaluable, and I am deeply indebted to both of them.

Similarly, I would like to express my appreciation for the support of the Engineering and Physical Sciences Research Council, through the University of York's Doctoral Training Partnership and the UK Quantum Communications Hub. Without their funding, this thesis would not exist.

I also wish to express my deepest thanks to Dr Sabine Hossenfelder and Prof Tim Palmer, with whom I worked closely in the latter stages of this PhD. Part II of this thesis draws heavily on work undertaken in collaboration with Sabine and Tim, and I am deeply honoured to have been able to collaborate with them.

Alongside this, I would like to thank Dr Will McCutcheon and Dr Gary F Sinclair, who were both key influences at the the start of my academic journey, as well as the other coauthors and collaborators I have had the pleasure of working with during my PhD: Mr Hatim Salih, Dr Colm Bracken, and Prof Terry Rudolph.

Further, I wish to express my gratitude to everyone I got to know throughout my PhD, in the Quantum Engineering Technology Laboratories (QETLabs), and across the School of Physics and Departments of Electrical and Electronic Engineering and Philosophy. QETLabs in particular has been an incredibly friendly and welcoming environment, even for an out-of-place theorist/foundations researcher like me, and it has been a privilege to be part of it for the last few years.

Finally, I want to thank my close friends and family, who have supported me through the highs and lows of the last three-and-a-half years. Doing a PhD can often feel like a very solitary endeavour; however, having people around to provide support, a sympathetic ear, or even just terrible puns, helps show that things are never as bad as they may feel, and that you're never truly alone. Were it not for them, I doubt I could have done this. Thank you all.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .... DATE: ....

Manuscripts (Published)

† indicates manuscripts that form part of this thesis.

†**Hance, J. R.**, Ladyman, J., & Rarity, J. (2021). How quantum is quantum counterfactual communication?. *Foundations of Physics*, **51**(1), 1-17.

†Salih, H., **Hance, J. R.**, McCutcheon, W., Rudolph, T., & Rarity, J. (2021). Exchange-free computation on an unknown qubit at a distance. *New Journal of Physics*, **23**(1), 013004.

†**Hance, J. R.**, & Rarity, J. (2021). Counterfactual ghost imaging. *npj Quantum Information*, **7**(1), 88.

Hance, J. R., Sinclair, G. F.,& Rarity, J. (2021). Backscatter and spontaneous four-wave mixing in micro-ring resonators. *Journal of Physics Photonics*, **3**(2), 025003.

†**Hance, J.**, & Rarity, J. (2021). Comment on "Scheme of the arrangement for attack on the protocol BB84". *Optik*, **243**(1), 167451.

†**Hance, J. R.**, Rarity, J., & Ladyman, J. (2022). Could wavefunctions simultaneously represent knowledge and reality. *Quantum Studies: Mathematics and Foundations*, **9**(3), 333-341.

†Salih, H., McCutcheon, W., **Hance, J.**, & Rarity, J. (2022). The laws of physics do not prohibit counterfactual communication. *npj Quantum Information*, **8**(1), 60.

†**Hance, J. R.**, & Hossenfelder, S. (2022). The wave function as a true ensemble. *Proceedings of the Royal Society A,* **478**, 20210705.

†**Hance, J. R.**, Hossenfelder, S., & Palmer, T. N. (2022). Supermeasured: Violating Bell-Statistical Independence without violating physical statistical independence. *Foundations of Physics,* **52**(4), 81.

†**Hance, J. R.**, & Hossenfelder, S. (2022). What does it take to solve the measurement problem?. *Journal of Physics Communications,* **6**(10), 102001.

†**Hance, J.R.**, & Hossenfelder, S. (2022). Bell's theorem allows local theories of quantum mechanics. *Nature Physics,* **18**(12), 1382 (Correspondence).

Manuscripts (Preprint)

† indicates manuscripts that form part of this thesis.

Hance, J. R. (2019). Counterfactuality, Definiteness and Bell's Theorem. *arXiv preprint arXiv:1909.06608.*

†Salih, H., **Hance, J. R.**, McCutcheon, W., Rudolph, T., & Rarity, J. (2020). Deterministic teleportation and universal computation without particle exchange. *arXiv preprint arXiv:2009.05564.*

†**Hance, J. R.**, Palmer, T. N., & Rarity, J. (2021). Experimental Tests of Invariant Set Theory. *arXiv preprint arXiv:2102.07795.*

Hance, J. R., Hossenfelder, S., & Palmer, T. N. (2021). Properties of Invariant Set Theory. *arXiv preprint arXiv:2108.08144.*

†**Hance, J. R.**, Rarity, J., & Ladyman, J. (2021). Do weak values show the past of a quantum particle?. *arXiv preprint arXiv:2109.14060.*

Bracken, C., **Hance, J. R.**, & Hossenfelder, S. (2021). The Quantum Eraser Paradox. *arXiv preprint arXiv:2111.09347.*

Bracken, C., **Hance, J. R.**, & Hossenfelder, S. (2021). Reply to arXiv:2111.13357 ("The Quantum Eraser Non-Paradox"). *arXiv preprint arXiv:2112.00436.*

Hance, J. R., & Hossenfelder, S. (2022). Comment on "Why interference phenomena do not capture the essence of quantum theory" by Catani et al. *arXiv preprint arXiv:2204.01768.*

†**Hance, J. R.**, Rarity, J., & Ladyman, J. (2022). The difference between the dynamical quantum Cheshire cat and counterfactual communication *arXiv preprint arXiv:2204.03374*

Hance, J. R., & Hossenfelder, S. (2022). Comment on "Experimentally adjudicating between different causal accounts of Bell-inequality violations via statistical model selection". *arXiv preprint arXiv:2206.10619.*

Hance, J. R., & Rarity, J. (2022). Interaction-Free Polarimetry. *arXiv preprint arXiv:2207.13539.*

Table of Contents

List of Figures

List of Abbreviations

CHSH Clauser-Horne-Shimony-Holt. xxiv, 7, 8, 104–109, 149, 157, 161, 165

EFEs Einstein Field Equations. 118, 175

HWP Half-Wave Plate. xviii, xx, xxi, xxiii, 21, 29, 30, 52, 54, 56, 59, 68, 69, 76, 78, 90, 96, 97

ICCD Intensified Charge-Coupled Device. xxii, xxiii, 88–91, 94, 96

IST Invariant Set Theory. xxv, xxvi, 8, 9, 114, 129, 155, 157, 160, 161, 163–172, 174, 175, 177, 181–183

MZI Mach-Zehnder Interferometer. xvii, 21, 23, 24, 28, 33

NISQ Noisy Intermediate-Scale Quantum. 3, 168, 169, 177, 183

PBS Polarising Beam-Splitter. xvii–xxiii, 21, 28–31, 46, 52, 53, 56, 68–71, 74–78, 81, 85, 88–91, 96, 97, 109, 187

SNR Signal-to-Noise Ratio. xxiv, 7, 92–94, 96, 97

SPDC Spontaneous Parametric Down-Conversion. xxii–xxv, 87, 88, 90–93, 96, 97, 173

““The difficulty is, of course, that you don’t have to settle these issues,” Weinberg continues. “I’ve had a whole career without knowing what quantum mechanics is. I tell this story in one of the books that my colleague Philip Candelas was referring to a graduate student whose career essentially disintegrated, and I asked what was wrong and he said, ‘He tried to understand quantum mechanics.’ He could have had a perfectly good career without it. But getting into the fundamentals of quantum mechanics is a losing game.””

Sabine Hossenfelder (quoting Steven Weinberg, quoting himself, quoting Philip Candelas),
“Lost in Math” (2018)
([Hos18])

Chapter 1

Introduction

This Chapter goes over the motivation for this thesis, as well as the practical details of its layout.

1.1 Motivation

Quantum foundations is the field of investigating, analysing, and extending the underpinning principles of quantum mechanics and quantum field theory. Work in this field often focuses on those ideas which are paradoxical when considered from a classical (non-quantum) perspective [Adl21b].

Quantum technologies is the field of adapting quantum phenomena towards practical applications. This involves applying the knowledge gained through quantum research to engineer devices or protocols to solve real-world problems, with advantages over classical solutions [DM03]. Such tecxhnologies range from tailor-built quantum components in the lab (e.g. single quantum dots or micro-ring resonators), to full Noisy Intermediate-Scale Quantum (NISQ) devices formed of dozens of interlinked qubits, which can be used to implement quantum circuits but not yet perform full quantum error correction protocols (as is the current state of the art in quantum computing).

This thesis illustrates the interrelation between quantum foundations and quantum technologies: quantum foundations uncovers new phenomena that can be used to develop new quantum technologies, while quantum technologies allow us to experimentally test models proposed in quantum foundations. Most quantum technologies were originally based on quantum foundational work. For instance, quantum key distribution was developed from the uncertainty principle and no-cloning theorem [BB84; WZ82], quantum computation (by Deutsch's own account) was developed from quantum parallelism and interference [Deu85], and quantum metrology was developed from consideration of quantum measurement and back action, and the leveraging of entanglement and squeezing [GLM04]. However, this thesis focuses on two key areas, which act as case studies for the interrelation between quantum foundations and developing quantum technologies: counterfactual communication, and statistical independence-violating extensions of quantum mechanics.

Quantum mechanics is notoriously counterintuitive, and very few quantum effects are as counterintuitive as interaction-free measurement [EV93]. Quantum technology development leverages counterintuitive non-classical effects to achieve advantage over classical solutions to problems. Counterfactual communication [Sal+13b] (a derivative of interaction-free measurement) is a natural candidate on which to base new quantum technology protocols. Existing work applies counterfactual communication for technical benefits, which this thesis extends to both counterfactual communication of quantum information, and counterfactual ghost imaging (a form of metrology). This shows how quantum foundations can provide new and useful forms of quantum technology.

Quantum mechanics has many differing interpretations and extensions. Models which violate statistical independence are often ignored by the scientific community, despite their ability to replicate the predictions of quantum mechanics. An advantage of such models is that they are based on local hidden variables [HP20], even though local hidden variables are often (wrongly) thought prohibited by Bell's theorem [Tea22; HH22c]. These models could also help us resolve two of the biggest problems with quantum mechanics: the measurement problem, and the unification of quantum mechanics with general relativity. These models are often discounted, as they are assumed to make identical predictions to standard quantum mechanics (i.e., only be interpretations), such that there is no way to experimentally test them. However, this is not necessarily true for those that **extend** quantum mechanics [Hos11]. This thesis proposes experiments using current and potential/developing quantum technologies to empirically distinguish the predictions of these extensions from standard quantum mechanics.

By looking at these two areas—counterfactual communication, and statistical independence-violating extensions of quantum mechanics—this thesis demonstrates the interplay between quantum foundations and quantum technologies: quantum foundations seeding the underpinning ideas which can be developed into quantum technologies, and current and hypothetical quantum technologies providing the means to experimentally test foundational ideas.

A recent move in the philosophy of science, referred to as philosophy of science in practice, considers not just the structure and content of scientific theories, but also the methods by which science advances, such as experimentation, model-making, and data analysis [Ank+11]. An approach within this, called philosophy-of-science in practice (distinguished by the hyphens) seeks to bring together philosophers and scientists to consider philosophical problems within a given field, such as background assumptions, or idealisations made when moving from experiments to theory [BL13]. This approach is also known as philosophy **in** science [Pra+21]. Given this thesis was co-supervised by a researcher in experimental quantum optics and a researcher in philosophy of science, it provides a unique example of philosophy-of-science in practice. Within quantum physics, most interaction with philosophers is done by deeply foundational theorists, who often could arguably be called philosophers themselves. This explains why most philosophical analysis of quantum mechanics/quantum field theory examines abstract theory and interpretation, rather than the practical observation of quantum phenomena. By fostering philosophical engagement with a prominent experimentalist, this thesis breaks ground on a new area of collaborative research. We expect philosophical analysis of quantum experimental practice to bring insights to the problems of quantum mechanics, by showing where certain theoretical abstractions and assumptions fall apart.

The use of scientific research within this thesis as case studies supporting a philosophical point also constitutes philosophy of science in practice, as it involves doing science, to then inform research into how science is done. This activity falls within another subdivision of philosophy of science in practice, called philosophy of science-in-practice (again, distinguished by the hyphens); this analyses science-in-the-making and the practice of scientific research, in a way similar to modern history of science. The way this thesis performs substantive scientific research with input from both scientists and philosophers, but also uses that as case studies for a philosophical argument, therefore illustrates both of these trends in philosophy of science in practice.

1.2 Thesis outline

This section gives an overview of the chapters of this thesis. Given academia is inherently collaborative, the articles on which each chapter is based are all multi-author. Therefore, the Declaration of Contribution box at the start of each chapter gives the candidate's specific contribution to the work therein.

This thesis is formed of two parts:

Part I introduces interaction-free measurement and counterfactual communication. It looks at foundational issues affecting these effects, and considers practical applications of these for new quantum technologies (Chapters 2, 3, 4, 5, and 6).

Part II looks at extensions of quantum mechanics—specifically those which violate statistical independence. It first reviews why we might want to extend quantum mechanics, surveying Bell's theorem, then the measurement problem and proposed solutions to it (Chapter 7). It then looks at whether the wavefunction should be interpreted epistemically or ontically, and how this relates to the task of extending quantum mechanics (Chapter 8). Finally, it investigates extensions of quantum mechanics where the wavefunction is interpreted epistemically, as some limited representation of underlying (statistical independence-violating) dynamics (Chapters 9).

Finally, Chapter 10 summarises both of these parts, looks at interesting potential extensions of and future directions for the work presented in this thesis, and relates the work presented back to the interplay between quantum foundations and quantum technologies.

The Appendices consider another area of interplay between developing quantum technologies and quantum foundations—weak values. This builds on work presented in Chapters 2 and 3. However, it moves away from looking at weak values in the context of identifying the path of a quantum particle, to instead look more broadly at the foundational issues around weak values. This work is still very much ongoing, so has been presented as Appendices rather than as part of the main body of this thesis.

1.2.1 Part I: Interaction-Free Quantum Technologies

Chapter 2 introduces the requisite background for the part of this thesis looking at counterfactual communication. It first covers methods proposed to evaluate the path of a quantum particle (Section 2.2); before examining the history of interaction-free measurement and counterfactual communication (Section 2.4) This Chapter serves as a literature review for Part I of this thesis.

Chapter 3 looks more closely at the weak trace approach to the position of a quantum particle first introduced in Subsection 2.2.3, and other similar weak value-based approaches to the path of a quantum particle. Specifically, it shows these approaches do not allow us to see the path of a quantum particle between measurements. It does this by first showing that taking the weak value of the position projection operator will always involve disturbing the system, and so causing some sort of measurement (Section 3.3). It then shows there is no reason to associate a non-zero weak value of the position projection operator along a given path with the quantum particle being present on that path (Section 3.4). This Chapter is based on the preprint paper, *Do weak values show the past of a quantum particle?* [HRL21].

Chapter 4 discusses properties of counterfactual communication. Section 4.2 challenges the claim that counterfactual communication is impossible, even for post-selected quantum particles. It does this by proposing a counterfactual scheme where, by the criteria discussed in Section 2.2, none of the photons that contribute information about the message sent travel between the communicants. It demonstrates counterfactuality experimentally by means of weak measurements (Subsection 4.2.4), and conceptually using consistent histories (Subsection 4.2.3). Section 4.3 then shows that full counterfactual communication is an exclusively quantum effect. It does this by first introducing the logical form of classical counterfactual communication (Subsection 4.3.1), then deriving a sufficient condition for a protocol of this sort to be considered quantum (Subsection 4.3.2. Subsection 4.3.3 then shows that the two protocols which meet our criteria for full counterfactuality both meet our condition for being quantum. This indicates full counterfactual communication is inherently quantum, despite previous assertions to the contrary. This work illustrates how quantum foundations can be useful for the deeper evaluation and analysis of the properties of (hypothetical/developing) quantum technologies. This Chapter is based on the published papers, *The laws of physics do not prohibit counterfactual communication* [Sal+22], and *How quantum is quantum counterfactual communication?* [HLR21].

Chapter 5 is the first chapter introducing new protocols using counterfactual communication for technological benefit. Specifically, Section 5.2 first gives a protocol for Bob to apply a phase between the H and V components of Alice's photon (Subsection 5.2.1). It then uses this as the basis for Bob to apply any single-qubit unitary onto Alice's state, entirely counterfactually, and gives a demonstration of how to use this to apply quantum algorithms (Subsection 5.2.2). Section 5.3 gives a protocol for performing quantum computation counterfactually—this time in the form of a controlled-phase gate, whereby Bob's qubit controls whether or not a phase is applied to Alice's qubit (Subsection 5.3.1). It then uses the entangling properties of the simplest form of this gate (a controlled-Z gate) to show how to perform quantum teleportation and telecloning counterfactually (Subsection 5.3.2). These protocols both serve as examples of quantum technological developments rooted in quantum foundational principles (counterfactual communication, entanglement, and teleportation). This Chapter is based on the published paper, *Exchange-free computation on an unknown qubit at a distance* [Sal+21], and the preprint, *Deterministic teleportation and universal computation without exchanging particles* [Sal+20].

Chapter 6 extends counterfactual communication in a different direction—metrology. Specifically, Section 6.2 gives a protocol for enhancing the efficacy of ghost imaging, by making use of counterfactuality to image the object without any photons hitting it. Section 6.3 then discusses its efficacy, by comparing Signal-to-Noise Ratio (SNR), SNR per photon-object interaction, and visibility to that of ideal ghost imaging, both for ideal and realistically-lossy components. It discusses potential extensions of counterfactual ghost imaging, specifically to computational and phase-resolving ghost imaging. This shows how quantum foundational ideas can be adapted to provide real, useful quantum technological protocols. This Chapter is based on the published paper, *Counterfactual ghost imaging* [HR21b].

1.2.2 Part II: Extensions of Quantum Mechanics

Chapter 7 goes over the background necessary for this part of the thesis, which looks at (statistical independence-violating) extensions of quantum mechanics. For this, Section 7.2 describes Bell's Theorem: Subsection 7.2.1 describes the EPR paradox, Subsection 7.2.2 derives a Bell inequality (the Clauser-Horne-Shimony-Holt (CHSH) inequality) and shows how quantum mechanics violates this inequality, Subsection 7.2.3 describes experimental tests of Bell's theorem and loopholes in those experiments, and Subsection 7.2.4 gives the assumptions necessary to formulate Bell inequalities. Subsection 7.2.5 then looks in more depth at the statistical independence assumption specifically, and extensions of quantum mechanics where this is violated. Secondly, Section 7.3 looks at the measurement problem in quantum mechanics. Subsection 7.3.1 first goes over the axioms of quantum mechanics, then Subsection 7.3.2 shows how these lead to the measurement problem. Subsection 7.3.3 identifies the properties a theory needs to solve the problem. Subsection 7.3.4 shows not only that no current interpretation of quantum mechanics solves the problem, but that, being interpretations rather than extensions of quantum mechanics, they cannot solve it. Subsection 7.3.5 finally discusses why statistical independence-violating extensions of quantum mechanics make good candidates for solving this problem, and why it is worth trying to solve this problem. This chapter serves as a literature review for Part II of this thesis. Parts of the first half of this chapter (Section 7.2) are based on the published paper, *Bell's theorem allows local theories of quantum mechanics* [HH22c]. The second half of this chapter (Section 7.3) is based on the published paper, *What does it take to solve the measurement problem?* [HH22b].

Chapter 8 first argues in Section 8.2 that nothing about the informal ideas of epistemic and ontic interpretations rules out wavefunctions representing both reality and knowledge. This is despite these terms being commonly associated in quantum foundations with the formal definitions given by Harrigan and Spekkens, for the wavefunction in quantum mechanics being ψ-ontic or ψ-epistemic in the context of the ontological models framework. These formal definitions are contradictories, so that the wavefunction can be either ψ-epistemic or ψ-ontic, but not both. Subsection 8.2.1 first goes over the intuitive ideas of the wavefunction being ontic or epistemic, before Subsection 8.2.2 looks at the Harrigan-Spekkens definitions, and the Ontological Models Framework on which they are based, and Subsection 8.2.3 analyses their definitions with respect to the intuitive ideas of ontic and epistemic. Subsection 8.2.4 shows that the implications of no-go theorems on the wavefunction being epistemic, such as the Pusey-Barrett-Rudolph theorem, may need rethinking in light of this analysis. Section 8.3 then argues that the ψ-ontic/epistemic distinction fails to properly identify ensemble interpretations, where the wavefunction predicts probabilities of possible measurement outcomes, but not which individual outcome is realised in each run of an experiment. Subsection 8.3.1 therefore proposes a more useful categorisation, for models where the wavefunction describes an ensemble of states with different values of a hidden variable—ψ-ensemble models. Subsection 8.3.2 illustrates how these are different from Harrigan and Spekkens's psi-ensemble models, then Subsection 8.3.5 shows that all local ψ-ensemble interpretations which reproduce quantum mechanics violate Statistical Independence. Finally, Subsection 8.3.6 explains how this interpretation helps make sense of some otherwise puzzling phenomena in quantum mechanics, such as the delayed choice experiment, the Elitzur-Vaidman Bomb Tester, and the Extended Wigner's Friends Scenario. This Chapter is based on the published papers, *Could wavefunctions simultaneously represent knowledge and reality?* [HRL22], and *The wave function as a true ensemble* [HH22a].

In Chapter 9, Section 9.2 first points out that the common interpretation of Statistical Independence violations, as correlations between the measurement settings and the hidden variables (which determine the measurement outcomes), is at best physically ambiguous and at worst incorrect. The problem with the common interpretation is that Statistical Independence might be violated because of a non-trivial measure in state space, a possibility Subsection 9.2.1 calls "supermeasured" and elaborates on. Subsection 9.2.2 shows how this allows the violation of the CHSH inequality. Subsection 9.2.3 describesInvariant Set Theory (IST), and show it as an example of a supermeasured theory that violates the Statistical Independence assumption in Bell's theorem without requiring correlations between hidden variables and measurement settings (physical statistical independence). Section 9.3 then identifies points of difference between IST and standard quantum theory, and evaluates if these would lead to noticeable differences in predictions between the two theories. From this evaluation, Subsection 9.3.1 gives a number of experiments, based on points of difference between the predictions of IST and standard quantum mechanics around entanglement limits, continuous variables (Subsection 9.3.2), and gravitationally-induced decoherence (Subsection 9.3.3). If undertaken, these experiments would allow us to investigate whether standard quantum mechanics or IST best describes reality. These tests can also be deployed on theories sharing similar properties (e.g. Penrose's gravitational collapse theory). This work shows how developing quantum technologies can be used to test foundational ideas. This chapter is based on the published paper, *Supermeasured: Violating Bell-Statistical Independence without violating physical statistical independence* [HHP22], and the preprint, *Experimental Tests of Invariant Set Theory* [HPR21].

This thesis ends with Chapter 10, which summarises the work presented in each chapter (Section 10.1), discusses future potential directions for research (Section 10.2), and finally ties the thesis back to the general broad theme of the interplay between quantum foundations and quantum technologies (Section 10.3).

1.2.3 Appendix: Weak Values

The appendix looks at another case study area for the interaction between quantum foundations and developing quantum technologies—weak values. While we have refrained from including this work as part of the thesis proper, weak values still serve as a useful demonstration of how a foundational idea can be applied for technological benefit (for weak values, mainly in the fields of metrology and state tomography), and how insights from these practical applications can feed back into our foundational understanding of these deeper underpinning ideas.

Chapter A shows what issues can come about when weak measurement is misinterpreted. In a recent paper (*Scheme of the arrangement for attack on the protocol BB84*, Optik 127(18):7083-7087, Sept 2016), Khokhlov proposed a protocol for using weak measurement to attack BB84. This paper claimed the four basis states typically used could be perfectly discriminated, and so an interceptor could obtain all information carried. We show this attack fails when considered using standard quantum mechanics, as expected - such "single-shot" quantum state discrimination is impossible, even using weak measurement. This Chapter is based on the published paper, *Comment on "Scheme of the arrangement for attack on the protocol BB84"* [HR21a].

Chapter B shows how to experimentally detect the dynamical quantum Cheshire cat proposed by Aharonov et al (*Nat Commun* 12, 4770 (2021)). In practice, to make the effect detectable, the initial state is biased by adding/subtracting a small probability amplitude of the orthogonal state which travels with the disembodied property. This biasing, which can be done either directly or via weakly entangling the state with a pointer, provides a reference with which we can measure the evolution of the state. We thereby show that this effect is different from counterfactual communication, which provably does not require any field to travel between communicating parties when information is sent. This Chapter is based on the preprint, *Is the dynamical quantum Cheshire cat detectable?* [HLR22].

Part I

Interaction-Free Quantum Technologies

"It's real! For real!"

"What? No. That doesn't make sense."

"It's your math, man. Check it."

"It's... whoa. Whoa. Duuuuuuuude."

God learns of the Elitzur-Vaidman bomb tester.

Zach Weinersmith,
SMBC Comics (2019)
("Duuude", [Wei19])

CHAPTER 2

COUNTERFACTUAL COMMUNICATION BACKGROUND

Declaration of contribution: This is an introductory chapter and contains no new research material. This chapter is based on the introduction and appendix of the published article *How Quantum is Quantum Counterfactual Communication?* [HLR21]. This was conceived and written by myself, supervised and edited by Prof James Ladyman and Prof John Rarity. An initial draft of that article was submitted as an extended essay for the Project Unit PHILM0008, a 20-credit unit contributing towards the MSci in Physics and Philosophy at the University of Bristol in 2019. However, it was developed significantly after this.
Subsection 2.2.3 is also based on the preprint article *Do weak values show the past of a quantum particle?* [HRL21]. This was conceived and written by myself, supervised and edited by Prof James Ladyman and Prof John Rarity.
The description of Salih et al's 2013 protocol in Subsection 2.4.3 is also based on the published article *Exchange-Free Computation on an Unknown Qubit at a Distance* [Sal+21]. This was conceived by myself and Mr Hatim Salih in discussion with Prof Terry Rudolph. The manuscript was co-written by myself and Mr Salih, under the supervision of Prof John Rarity. Further discussion and edits were done by all authors (Mr Salih, myself, Dr Will McCutcheon, Prof Rudolph, and Prof Rarity).

2.1 Chapter Introduction

This Chapter lays out necessary background for Part I of the thesis. We focus on counterfactual communication, and other developing interaction-free quantum technologies. We first look at approaches proposed to determine the path(s) via which a quantum particle travelled. These allow us to evaluate whether or not particles cross the communication channel between two communicants, and so whether or not a given protocol for communication is counterfactual. This also involves introducing weak values and Vaidman's weak trace approach to the path of a quantum particle, which we discuss and evaluate in more detail in Chapter 3.

We then describe the history of interaction-free measurement, from Elitzur and Vaidman's initial protocol (the Elitzur-Vaidman Bomb Tester), to Kwiat et al's combination of this with the quantum Zeno effect, to counterfactual computation and imaging protocols based on this. Finally, we look at counterfactual communication proper, introducing and contrasting protocols proposed so far, in preparation for analysis in Chapter 4.

This Chapter traces the development of counterfactual quantum protocols, from Elitzur and Vaidman's initial foundational proposal, to experimentally-demonstrated protocols for communication, computation and imaging. This illustrates how, over thirty years, something initially viewed as a foundational curiosity can be leveraged for technological benefits. Therefore, this chapter serves as an example of the first part of the interplay between quantum foundations and quantum technologies—foundational ideas inspiring new/hypothetical quantum technologies. The rest of Part I evidences this interplay further, as we both foundationally analyse and practically develop these counterfactual and interaction-free quantum technologies.

2.2 Path of a Quantum Particle

This first section describes approaches proposed to determine whether a quantum particle has been in a given spatial region. Such foundational work is intrinsically linked to counterfactual communication: "counterfactual" here referring to the two parties communicating without any quantum particles crossing the channel between them. Analysis of these approaches will therefore allow us to establish whether a protocol is actually counterfactual.

2.2.1 Initial Approaches

When considering the path traversed by a quantum particle, our first instinct is to treat the particle as spatially localised. This models it as travelling along a single path, from its point of origin to its destination. However, quantum phenomena are not solely point-like, but also have wave-like properties, as the Two-Slit Experiment shows [You04]. This means we cannot necessarily treat them as spatially localised. Therefore, we need a criterion for counterfactuality that takes into account quantum phenomena.

Based on this need for a better criterion, a common view posed is to claim communication between two communicants Alice and Bob can only be counterfactual if no information can pass between Alice and Bob. By this criterion, counterfactual communication is inherently impossible, as any successful communication will involve the transfer of information. This view is the one often presented upon first hearing about quantum counterfactual communication. However, as we will show in Chapter 4, there are many classical communication protocols which are considered counterfactual, despite not fulfilling this criterion. Penrose defined a counterfactual as an event which “might have happened, although did not in fact happen" [Pen94]. By this definition, counterfactual communication means inferring from the non-occurrence of something which could have occurred. Such an inference obviously transmits information (see Chapter 4 for examples of this). Therefore, there is no reason to only call communication counterfactual if no information can be transferred. Counterfactual communication is demonstrably possible—we just wish to know whether quantum counterfactual communication is an extension of it.

Our first non-classical approach to a particle’s location involves considering the entire quantum mechanical description of the system. We can do this by looking at the density matrix. This provides all the information that exists (according to standard quantum mechanics) about a particle at a given moment. By associating entries in the density matrix with physical positions, we can see the spread of the particle’s possible locations. However, some parts of the density matrix correspond to paths lost when the wavefunction is collapsed at the protocol’s end, and so information not being sent. We need some way to sort these possible paths into those the particle could have been on when information is sent, and those it was on when information is not sent. This requires post-selection (identification of paths by their final state), rather than solely pre-selection (identification of paths by initial state).

2.2.2 Consistent Histories

To resolve this, we could apply Griffiths’ Consistent Histories approach to the task of identifying the path of a quantum particle [Gri84]. The Consistent Histories approach involves considering histories: tensored chains of projectors, each representing a way the system under consideration could evolve. A family is a set of these histories, that form a projective decomposition of the identity operator over the whole evolution time. By refining certain projectors (replacing them with projectors summing to them), we can model a situation. These histories are consistent if they all mutually commute—if they do not, the Consistent Histories approach claims it is meaningless to ask which history was the ‘correct’ model for the system’s evolution, as we cannot assign positive relative probabilities to the different evolutions [Gri19].

Therefore, we can explicitly say a particle has not gone between Alice and Bob according to the Consistent Histories approach when all histories where it both travels between the two, and information is sent, have probability zero. This requires us to analyse all histories in this family, as Griffiths does for several protocols [Gri16]. In essence, if a particle can go between Alice and Bob when information is transmitted, by Consistent Histories it is not counterfactual; if not all the histories in the family are consistent with one another, then by the Consistent Histories approach it is meaningless to talk about the counterfactuality of the situation. Therefore, a protocol is only explicitly counterfactual by the Consistent Histories criterion, if we have a consistent family of histories, where all histories where the particle travels between Alice and Bob and information is sent have probability zero, and there still exists a history where information is sent between Alice and Bob. We show a protocol which meets this condition in Chapter 4.

2.2.3 Weak Trace

Next, we consider the weak trace approach to the path of a quantum particle, alongside other related approaches (those which use weak values to infer the path of a quantum particle). Weak values are derived from weak measurement [AAV88], which examines the state between measurements, without collapsing it. Weak measurement involves lightly coupling a system to a measuring device. While little information is gathered by one measurement, many measurements on identically-prepared systems will generate a probability distribution. This contrasts with Von Neumann measurements, which cause a system to collapse into an eigenstate of the measured operator. Weak measurement allows us to collect information that would be lost were the system strongly measured [TC13]. We can interpret the probability distribution given by weak measuring the evolution from an initial state as evaluating all forward-evolving paths from that state.

However, rather than working forwards, can also work back from a given result (post-select), to investigate the paths through which the system may have evolved. If we pre- and post-select like this, we obtain a weak value.

Aharonov et al defined the weak value O_w of an operator $\hat{O}$ [AAV88], where

$$O_w = \langle \hat{O} \rangle_w = \frac{\langle \psi_f | \hat{O} | \psi_i \rangle}{\langle \psi_f | \psi_i \rangle} \tag{2.1}$$

As this weak value increases as $\langle \psi_f | \psi_i \rangle$ goes to zero, weak value protocols are used in metrology to amplify signals from delicate results so they can be observed experimentally [Dix+09; JMH14; HBL17]. This is at the expense of postselection reducing success probability. Weak values however also lead to a range of paradoxes [AR05].

To experimentally observe the quantity we define in the expression in Eq. 2.1, we first couple our initial system $|\psi_i\rangle$ to our initial pointer state, $|\phi\rangle$, by weakly measuring them with the probe Hamiltonian $\hat{H} = \lambda \hat{O} \otimes \hat{P}_d / T$ for small coupling constant λ and state-probe interaction time T. This produces the state

$$\begin{aligned} |\psi_w\rangle &= e^{-\frac{i\hat{H}T}{\hbar}} |\psi_i\rangle \otimes |\phi\rangle \\ &= e^{-\frac{i\lambda}{\hbar}\hat{O}\otimes\hat{P}_d} |\psi_i\rangle \otimes |\phi\rangle \end{aligned} \tag{2.2}$$

where $\hat{P}_d$ is the momentum of that pointer. We then strongly measure this weak-measured state $|\psi_w\rangle$ with the operator

$$\hat{F}_1 = |\psi_f\rangle\langle\psi_f| \otimes \hat{I}_d \tag{2.3}$$

Assuming $\hat{P}_d$ has Gaussian distribution around 0 with low variance (so $\hat{X}_d$, the position of the pointer, has Gaussian distribution with high variance),

$$\begin{aligned} e^{-\frac{i\lambda}{\hbar}\hat{O}\otimes\hat{P}_d} &= \sum_{k=0}^{\infty}\big(-\frac{i\lambda}{\hbar}\hat{O}\otimes\hat{P}_d\big)^k/k! \\ &= \mathbb{1}-\frac{i\lambda}{\hbar}\hat{O}\otimes\hat{P}_d+\mathcal{O}(\lambda^2)\approx\mathbb{1}-\frac{i\lambda}{\hbar}\hat{O}\otimes\hat{P}_d \end{aligned} \tag{2.4}$$

so this strong measurement gives the result

$$\begin{aligned} &\left|\psi_f\right\rangle\!\left\langle\psi_f\right| e^{-\frac{i\lambda}{\hbar}\hat{O}\otimes\hat{P}_d}\left|\psi_i\right\rangle\otimes\left|\phi\right\rangle \\ &\approx\left|\psi_f\right\rangle\!\left\langle\psi_f\right|(\mathbb{1}-\frac{i\lambda}{\hbar}\hat{O}\otimes\hat{P}_d)\left|\psi_i\right\rangle\otimes\left|\phi\right\rangle \\ &=\left|\psi_f\right\rangle\otimes\langle\psi_f|\psi_i\rangle(\mathbb{1}-\frac{i\lambda}{\hbar}O_w\hat{P}_d)\left|\phi\right\rangle \\ &\approx\left|\psi_f\right\rangle\otimes\langle\psi_f|\psi_i\rangle e^{-\frac{i\lambda}{\hbar}O_w\hat{P}_d}\left|\phi\right\rangle \end{aligned} \tag{2.5}$$

This means the position of our initial pointer state $|\phi\rangle$, acting as our "readout needle", shifts, with measurement of the pointer position giving a read-out value $(\langle x\rangle - a)$. For an initial Gaussian pointer state, this shift causes the average position to move from $\langle x\rangle$ to $(\langle x\rangle - \overline{a})$ due to the application of the effective operator $\mathbb{1} - i\lambda O_w\hat{P}_d/\hbar$. a is distributed over a wide range of values for many repeats, but the distribution of "a"s will be a Gaussian centred on (the real part of) λO_w. Peculiarly, this weak value O_w can be very far from any of the eigenvalues of $\hat{O}$, or even imaginary [AAV88; TC13; Dre+14]. This is odd, given this weak value appears in exactly the place in the equation that an eigenvalue of $\hat{O}$ would for a Von Neumann measurement (where the variance of $\hat{X}_d$ on the measured state is vanishingly small). This led to the belief that the weak value O_w represented some fundamental value of the operator $\hat{O}$ **between** measurements. (Note that it has recently been shown that weak values can be disconnected from weak measurements, and obtained using strong or projective measurements [CP18; Wag+23] Further, while above we give the conventional way weak values are obtained experimentally, this process is only one of the many ways weak values can be defined.)

Consider a case where the weak value of either the spatial projection operator, or an operator formed by multiplying the spatial projection operator with the operator for some property of a particle, along a certain path was non-zero for a given $|\psi_i\rangle$ and $\langle\psi_f|$. Vaidman argues we should interpret this as meaning a particle was on that path and left a weak trace along it, on two grounds [Vai13a]. First, if it were a strong measurement, an operator containing the spatial projection operator having non-zero eigenvalues would be sufficient for a quantum particle to be present, so correspondingly a non-zero weak value of this operator is a weak trace of particle presence. Secondly, if a particle was present in a region, it should have non-zero interaction with the local environment (as reflected by the subtle decoherence its state would undergo), and so leave a "trace" along its local path. While this interaction is weak, which it must be so that the environmental coupling does not collapse position superpositions in the way a projective

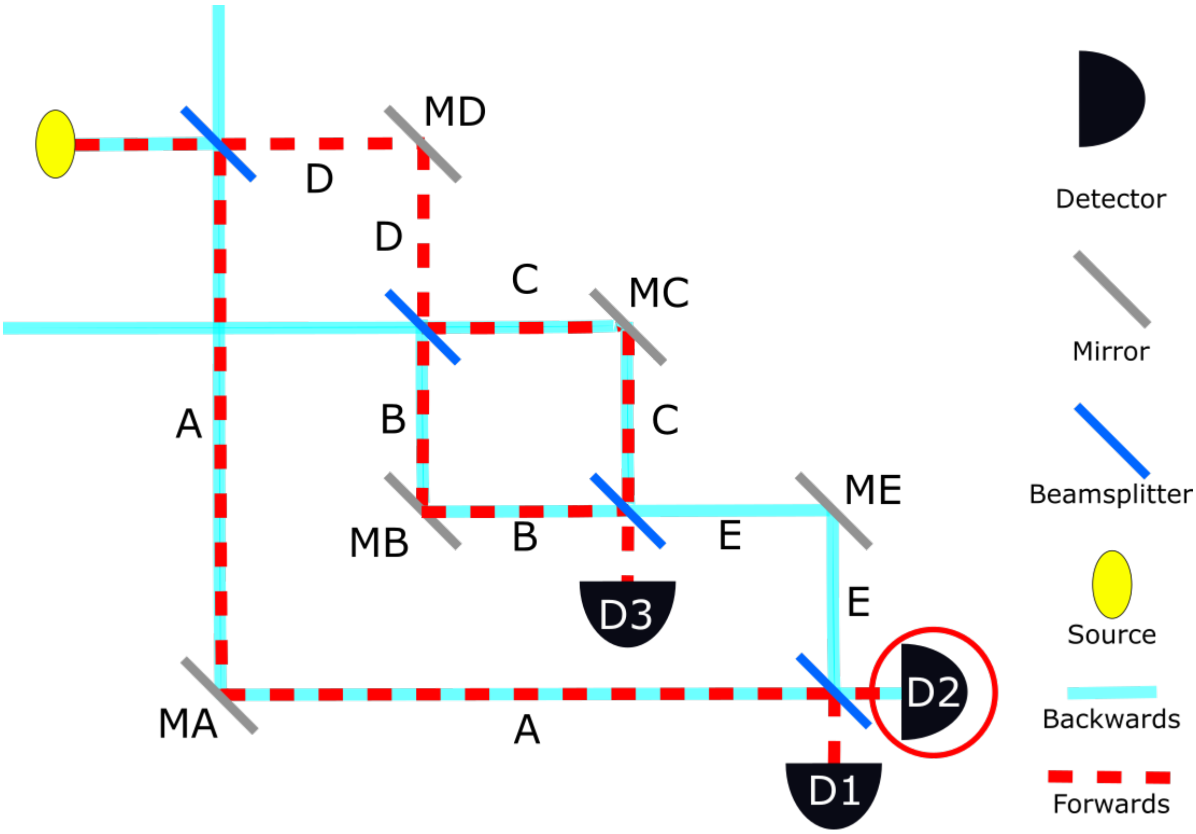

Figure 2.1: Nested-interferometer model, used by supporters of weak value approaches to discuss cases where the Two-State Vector Formalism contradicted "common-sense" continuous-path approaches [Vai13a]. Though no forward or backwards travelling path goes from the source, to Bob (along path C) and into D2, they do overlap over C, meaning there is a weak trace at Bob. This illustrates the peculiar property of the TSVF where particles can jump between regions (e.g. between the inner interferometer and the outer arm) [Vai13a].

measurement would, supporters of the weak trace approach claim it is still in principle detectable (i.e., the interaction would be of the same order of magnitude as λ), and so the interaction would be equivalent in scale to that required to give a non-zero weak value of an operator containing the spatial projection operator along that path.

Due to their interrelatedness, we can graphically show where the spatial projection operator has a non-zero weak value through the Two-State Vector Formalism (TSVF). Aharonov et al developed the TSVF [ABL64], which considers both the backwards and forwards-evolving quantum states, rather than just the forward as in standard quantum mechanics. (A similar intuition led Watanabe to develop the Double Inferential-state Vector Formalism [Wat55].) For some operator $\hat{O}$, the forwards travelling initial state $|\psi_i\rangle$, and the backwards travelling final state $\langle\psi_f|$, the TSVF gives out a conditional probability amplitude $\langle\psi_f|\hat{O}|\psi_i\rangle$, where $\langle\psi_f||\psi_i\rangle$ is referred to as the Two-State Vector. This conditional probability amplitude is of the same form as the numerator of the weak value. Therefore, we can use the TSVF to graphically plot where the weak value of an operator containing the spatial projection operator is nonzero. This involves plotting the forward-evolving state (possible paths the particle could have travelled via from its original position) and the backwards-evolving state (possible locations the particle could have come from to reach its final position).

If an operator returns a non-zero TSVF value, there is to $\mathcal{O}(\lambda)$ a weak trace along the path it describes. If we trace the paths a quantum particle could evolve along from its initial, and those it could have come from to get to its final, state, there is a weak trace where they overlap—so we cannot say the particle was not there [Vai13a].

The nested-interferometer model (Fig. 2.1) was identified as a case where the Two-State Vector Formalism contradicted "common-sense" continuous-path approaches [Vai13a]. The inner interferometer is balanced such that a photon entering from arm D exits to detector $D3$. Consequently, the outer interferometer is unbalanced, so there is an equal probability of a photon introduced from the source ending in $D1$ or $D2$. When the photon ends at $D2$, common sense would tell you it must have travelled via path A. This is as, had it travelled via D into the inner interferometer, it could not have exited onto path E, or reached $D2$. However, supporters of the weak trace approach claim, while the photon never travelled paths D or E, it travelled along paths B and C as well as along path A.

However, this has unintuitive results. If one accepts the weak trace as a valid indicator of a particle's path, this leads to peculiarities. While, with Consistent Histories, a path needs to link the initial and final states, here, it only requires that paths from the initial and final states overlap at some point. This means Bob can have a weak trace on his side of the transmission channel, without any in the channel itself. Danan et al demonstrated this using weak measurements in nested Mach-Zehnder Interferometer (MZI) [Dan+13]. However, Peleg et al claim these peculiar results don't contradict standard quantum theory [PV18].

There are cases where the Weak Trace Approach entails that a particle was present on a path, despite the spatial projection operator for that path having weak value zero. An example of this is Aharonov et al's quantum Cheshire cat (Fig. 2.2).

In the quantum Cheshire cat protocol, an H-polarised single-photon is emitted from the source and passed through a 50:50 beamsplitter, such that it is effectively preselected in the state

$$|\Psi\rangle = (i\,|L\rangle + |R\rangle)\,|H\rangle/\sqrt{2} \tag{2.6}$$

Similarly, a Half-Wave Plate (HWP) and a phase rotator (applying a phase rotation of $\pi/4$—i.e., a factor of i) are put on the right-hand path. After these, the left and right paths recombine at another 50:50 beamsplitter before passing through a Polarising Beam-Splitter (PBS). This transmits H-polarised light and reflects V-polarised light. By postselecting on the photon arriving at detector $D1$, this effectively postselects on the photon being in the state

$$|\Phi\rangle = (|L\rangle\,|H\rangle + |R\rangle\,|V\rangle)/\sqrt{2} \tag{2.7}$$

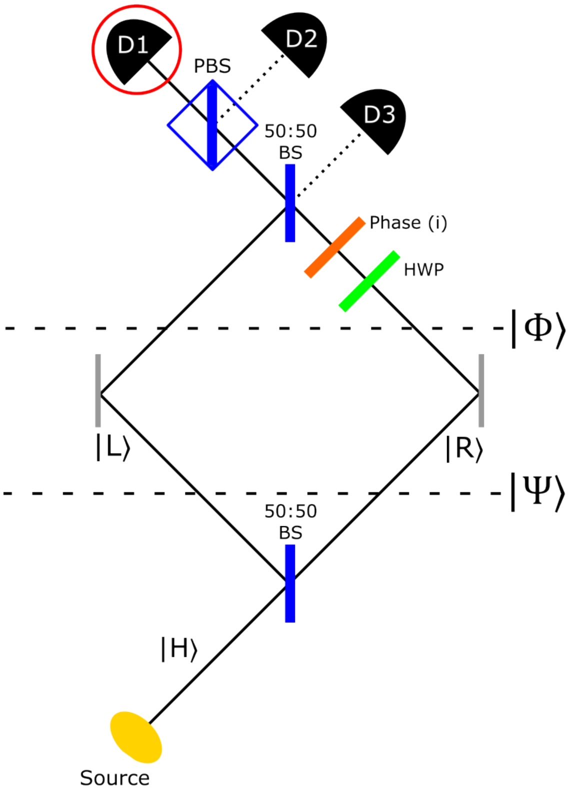

Figure 2.2: Aharonov et al's quantum Cheshire cat. See Subsection 2.2.3 for description of protocol.

Taking the weak value of the spatial projection operators $\Pi^L = |L\rangle\langle L|$ and $\Pi^R = |R\rangle\langle R|$ gives $\Pi^L_w = 1$, $\Pi^R_w = 0$. However, if we define the circular polarisation operator

$$ {}_{\pm}\sigma = |+\rangle\langle +| - |-\rangle\langle -| = i\,|V\rangle\langle H| - i\,|H\rangle\langle V| \tag{2.8} $$

and the compound operators ${}_{\pm}\sigma^R = \Pi^R_{\pm}\sigma$ and ${}_{\pm}\sigma^L = \Pi^L_{\pm}\sigma$, we see that the weak values of these compound operators are ${}_{\pm}\sigma^L_w = 0$, ${}_{\pm}\sigma^R_w = 1$.

Aharonov et al take this to mean, while the photon travels on the left path in the postselected scenario, its polarisation travels on the right path.

However, Vaidman claims a particle is present along a path even when the weak value of the spatial projection operator along that path is zero, so long as the weak value of some other operator multiplied by the spatial projection operator is nonzero. In the quantum Cheshire cat, this is satisfied by the nonzero weak value of the spin along the right-hand path [Vai13a]. This Weak Trace Approach also agrees with an approach recently given, which is informed by Fisher Information [WCV21].

We discuss issues with Vaidman's Weak Trace Approach, and other approaches where weak values are used to infer the path of a quantum particle (as Aharonov et al do in [Aha+13]) in

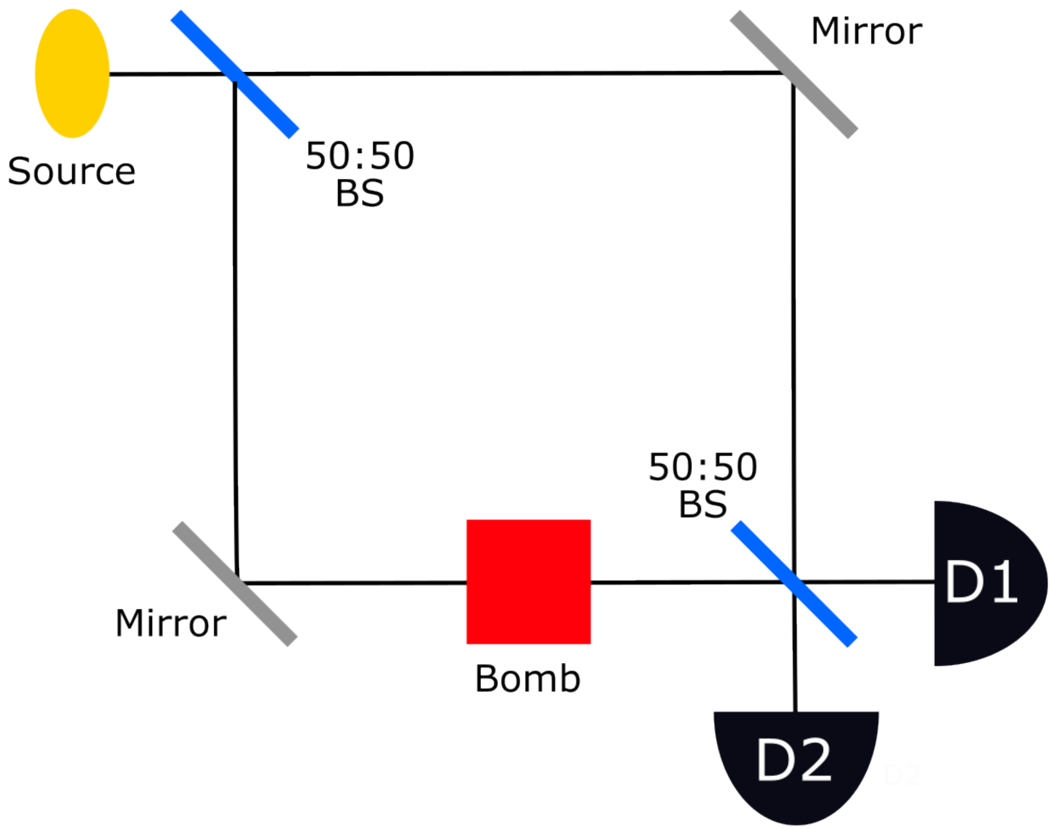

Figure 2.3: The Elitzur-Vaidman Bomb Tester. A photon is emitted from the source (top-left), enters the balanced MZI, and travels in along an equal superposition of both paths. If the bomb is faulty, the photon recombines at the second beam-splitter, and always enters D_1. If the bomb works, and is activated, it destroys the experiment. If the bomb would work, but the photon went down the bomb-free path, the photon has a 50:50 chance of going to either detector.

Chapter 3. We then show counterfactual communication is possible by the weak trace criterion in Chapter 4.

2.3 Interaction-Free Measurement

In this Section, we go through the history of interaction-free measurement, from the Elitzur-Vaidman bomb tester [EV93], via Kwiat et al's combination of this with the quantum Zeno effect [Kwi+99], to interaction-free metrology and counterfactual computation, as proposed by White et al and Hosten et al respectively [Whi+98; Hos+06]. These illustrate the progress made in adapting interaction-free measurement to practical settings, before we introduce counterfactual communication in the next Section.

2.3.1 Elitzur-Vaidman Bomb Tester

All interaction-free measurement protocols stem from the Elitzur-Vaidman Bomb Tester [EV93]. This was based on the idea, proposed by Renninger and Dicke, that a negative measurement could still change the wavefunction of a quantum object, even if there was no mechanism by which the measuring device could interact with the (absent) object [Ren60; Dic81]. Renninger's original thought experiment, based on the Mott Problem [Mot29], had an alpha particle emitted which would go to one of two hemispherical detectors. Assuming the detectors are 100% efficient, the lack of a detection in one tells us the particle would be detected by the other. Therefore, if we remove one of the detectors, the remaining one not clicking causes the particle's wavefunction to go from being spherical to being a hemisphere. The absence of a detection changed the wavefunction of the particle. Dicke later applied the same logic to the case of a particle being imaged optically, where the non-scattering of a photon from the particle appears to transfer momentum and energy to the particle, by changing its wavefunction. However, Elitzur and Vaidman simplified this paradox from continuous-variable position and momentum-space to a simple interferometric set-up, to more effectively illustrate the peculiarity of this behaviour.

In the Elitzur-Vaidman Bomb Tester (Fig. 2.3), an MZI has a potentially faulty bomb along one of its paths. This can only be detonated by a non-demolition single-photon detection on that path.

If the photon goes along the bomb's side of the MZI (and the bomb works), it detonates, and the photon (and everything else) is destroyed. If the bomb is faulty, the photon travels to the merging beamsplitter normally, and is always detected at D_1. However, if the photon goes along the other side, the bomb working changes the interference pattern. This makes the photon able to go to D_2, which it previously couldn't access. This allows us to test if the bomb would have worked, without detonating it, when the bomb is on the path the photon could have, but did not, travel down. Unlike classical counterfactual communication, both options are transmitted counterfactually—the photon's path is from the source to the detectors without going via the object (bomb) under evaluation. However, it is not necessarily always counterfactual. This is as, while the photon *can* carry the information without going via the bomb, it does not necessarily have to. For 50:50 beamsplitters, as in Fig. 2.3, if the bomb works, it has a 50% chance of blowing up. This leaves a 25% chance of the photon going to D_1 (giving an incorrect reading), and only a 25% chance of the photon going to D_0 (giving a correct reading, interaction-free). Therefore the Bomb-Tester is not fully counterfactual.

2.3.2 Chained-Interferometer Interaction-Free Measurement

Soon after Elitzur and Vaidman's initial proposal, Kwiat et al developed an adaption of the Bomb Tester. By using a chain of interferometers to repeatedly probe the bomb, the probability of the bomb detonating could be made arbitrarily small, as the number of interferometers in the chain increased (see Fig. 2.4). Further, this adaption removed any chance of erroneous clicks, so

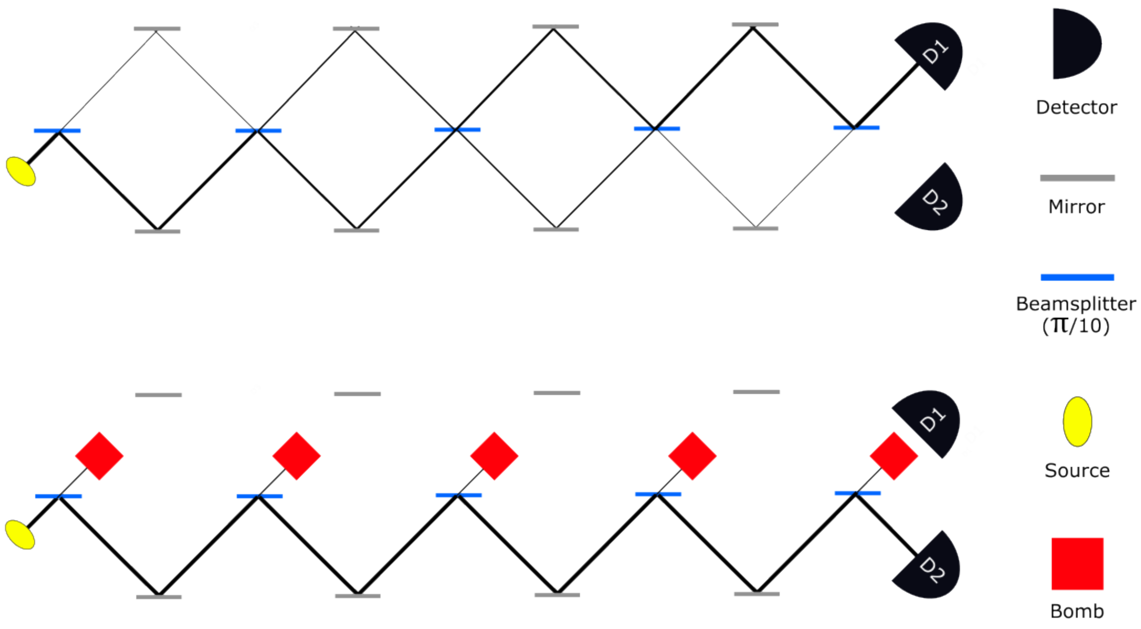

Figure 2.4: Kwiat et al's chained-interferometer version of the Elitzur-Vaidman Bomb Tester [Kwi+95].

long as there were two or more interferometers in the chain. This allows perfect discrimination of the state of the bomb from the detector the photon arrived at.

This adaption was based on combining the Elitzur-Vaidman protocol with the quantum Zeno effect proposed by Misra et al [MS77]. The quantum Zeno effect involves the use of repeated probing of a system (either directly via measurement or indirectly via coupling to a larger environment) to inhibit unitary evolution. They treat the interaction between photon and bomb as a probing measurement, where otherwise the unitary evolution slowly increases the probability amplitude for the photon being on the "bomb's side" of the interferometer from zero by a small amount each interferometer. This interruption ensures (in the limit of a large number of interferometers) that the photon stays on the bomb-free side of the interferometer, if the bomb would work.

2.3.3 Interaction-Free Metrology

A natural use for interaction-free measurement is low-energy metrology—imaging an object, trying to resolve a given level of detail, while reducing light (or optically-transmitted energy) incident on the imaged object to as low as possible. Therefore, it is unsurprising this use for interaction-free measurement was quickly suggested by Kwiat et al [KWZ96], and then rapidly performed by White et al [Whi+98]. While White et al's initial experiment only used a single

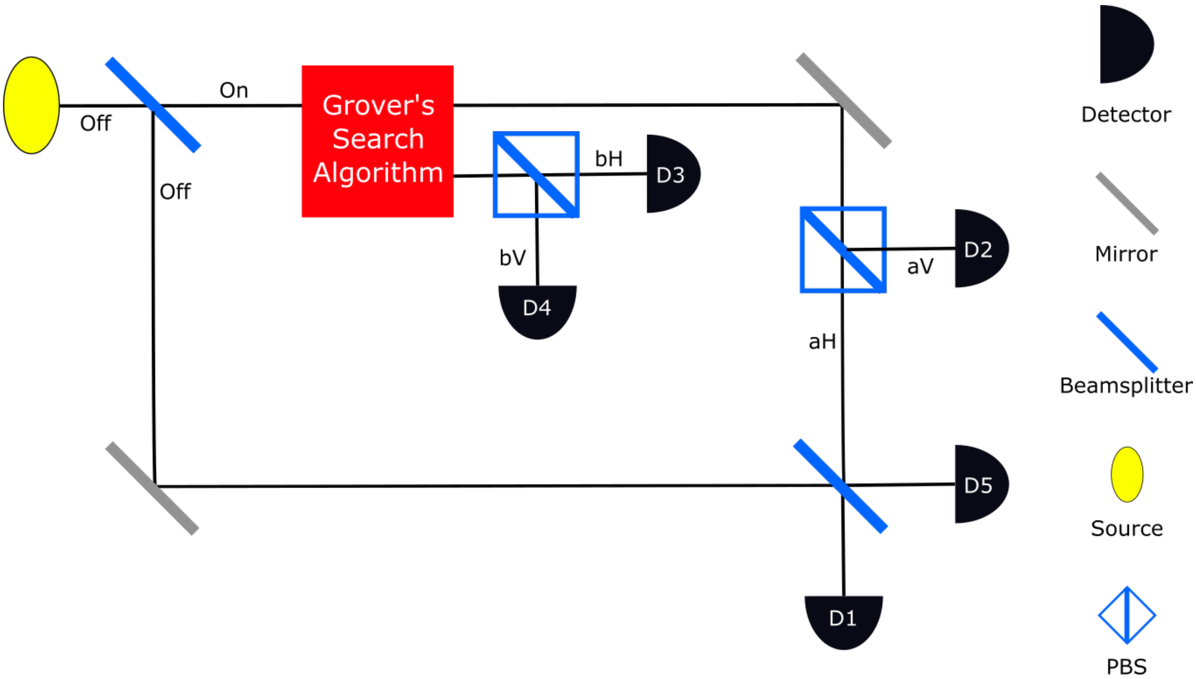

Figure 2.5: Hosten et al's protocol for counterfactual computation [Hos+06].

interferometer, Kwiat et al demonstrated a multi-interferometer Michelson form of the protocol one year later [Kwi+99]. Due to tuning, even the single-interferometer form was effective for low-energy-flux imaging of the one-dimensional profiles (widths) of a large number of objects (e.g., metal wires, cloth threads, and human hairs). The multi-interferometer form had even higher efficiency, despite being limited by the high losses caused by the the imperfect optical components of 25 years ago: 2/3 of the photons imaging the opaque object were measured interaction-free. Considering this was only 6 years after Elitzur and Vaidman's initial paper was published, this shows how quickly focused effort by quantum technology researchers can leverage practical benefit out of a foundational curiosity.

2.3.4 Counterfactual Computation

The next key innovation in interaction-free measurement was obtaining the result of a computation performed by a quantum computer, without said computer ever "running" (e.g., for an optical computer, having photons run through it). This was proposed by Jozsa and Mitchison [Joz99; MJ01] as a theoretical possibility allowed by the counterfactual protocols discussed above. However, it was also conjectured in the latter paper that this form of counterfactual computation could never be leveraged for gain, as the probability of this form of counterfactual inference would be limited to below the information gained by a random guess.

However, counterfactual computation was then experimentally performed by Hosten et al to perform Grover's search algorithm counterfactually [Hos+06]. This used a set-up similar to

Fig. 2.5. At the time, this was the most accurate realisation performed so far of Grover's search algorithm (with only 2.6% error, albeit only 31.9% efficiency). This showed it was possible to violate Jozsa's conjectured limits on information obtainable through counterfactual computation, by proposing extending the experiment to use the quantum Zeno effect. This involved combining multiple interferometers of the form given in Fig. 2.5 to allow the theoretical inference probability of the computation result to go to unity. However, this was not performed experimentally, possibly due to potential losses from imperfect components.

This shows the benefits of even the single-interferometer counterfactual computation protocol, above and beyond other contemporary implementations of Grover's search algorithm. This yet again shows how a foundational curiosity was leveraged for technological gain.

2.4 Quantum Counterfactual Communication Protocols

In this Section, we move from general interaction-free and counterfactual protocols, to look more specifically at Quantum Counterfactual Communication. Since Elitzur and Vaidman first discovered quantum counterfactuality [EV93], and Kwiat et al allowed loss to be made effectively nil [Kwi+95], researchers have tried to exploit it for communication. Despite this, all protocols until recently fell into three broad categories: where communication is counterfactual only for one bit-value; where photons travel between Alice and Bob, but in the opposite direction to the information passed between them; and where no photons pass between Alice and Bob when information flows, but the error/loss rates vary with the bit-value Bob sends.

2.4.1 Counterfactual only for One Bit-Value

Here, the protocol is counterfactual for one bit-value, but the photon goes between Alice and Bob for the other. The first of this type of protocol, and indeed the first quantum counterfactual communication protocol proposed[1], was Noh's (Fig. 2.6) [Noh09]. For matched polarisations, if Alice gets a click, the photon has remained on her side. However, for orthogonal polarisations, it has both been to, and returned from, Bob—so is not counterfactual. Despite this, the work generated a lot of interest [SW10; Yin+10; JSL11; Ren+11; Bri+12; Liu+12; Yin+12; ZWT12a; ZWT12b; SSS13a; SSS13b; LWL13; Zha+13; Li14; Liu+14; SSS14; SS15; WLZ16; Yan+16; SY17; SS13]. While plenty of these focus on reducing loss by reducing the proportion of the photon sent to Bob [SW10], this must be non-zero for the protocol to function. Therefore, the protocol will never be fully counterfactual.

This protocol having a limited maximum efficiency is not a serious drawback for its purpose, since the shared information it provides is random. This means failed attempts can be discarded during post-processing. However, this led others to wonder whether efficient, deterministic communication was possible, without particles crossing the communication channel.

[1]Barring Guo's E-V Bomb Tester adaption, where photons travel between Alice and Bob for both bit-values [GS99].

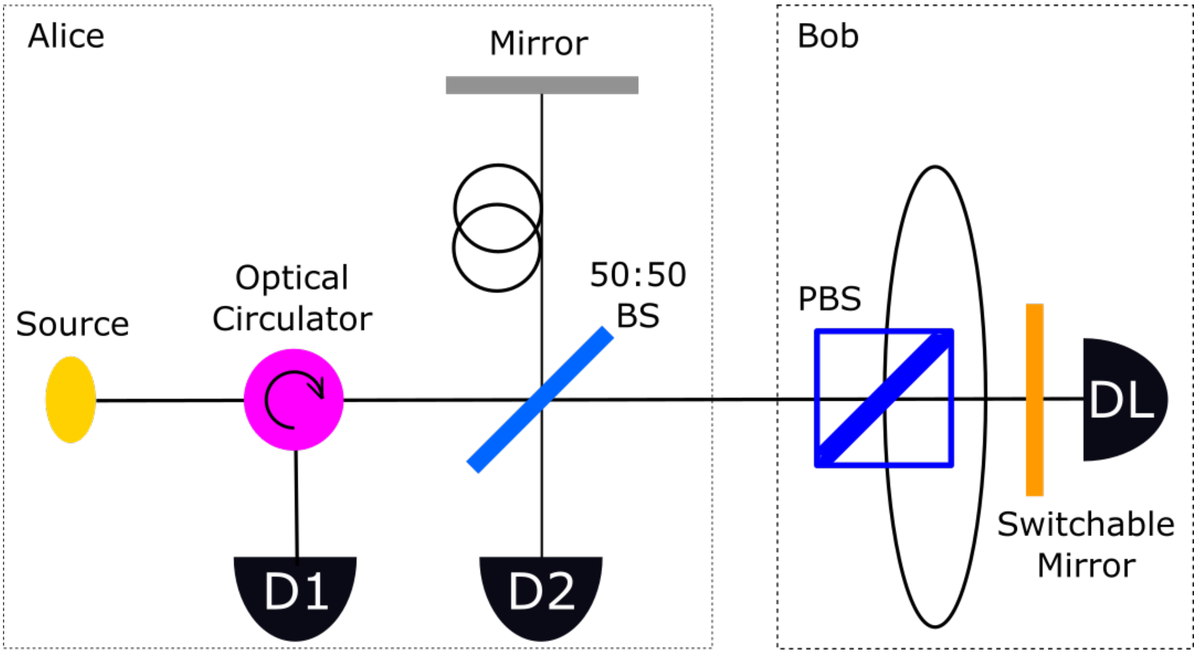

Figure 2.6: Noh's counterfactual cryptography protocol. Alice randomly polarises a photon either H or V-polarised. This photon passes through a beam-splitter, with one of the outputs going to Bob. Bob uses a H/V PBS and delay to time-separate the two possible polarisations arriving at him. He picks one to reflect and one to absorb. If the photon is the same polarisation Bob chose to reflect, it reflects back, and interferes into Alice's D_2. However, if the photon is the polarisation he chose to absorb, it is sent into his detector. If this clicks, the protocol is aborted; if not, the photon goes into D_1 [Noh09].

2.4.2 Information and Photon Travel in Opposite Directions

In the next type of protocol, the photon can cross the channel, but in the opposite direction to the information being sent. For one bit-value, the photon destructively interferes across the quantum channel, keeping it at Alice. For the other, the photon constructively interferes, allowing it through to Bob. Depending on whether she detects a photon, Alice can determine which bit Bob sent.

The only protocol of this type is Arvidsson-Shukur et al's. They propose a device formed of chained MZI. These use the Quantum Zeno effect to keep the photon at Alice if Bob blocks (for many interferometers), and force it to go to Bob if he does not [AB16]—identically to Kwiat et al's Interaction Free Measurement protocol [Kwi+95]. This leads to a high chance of Alice wrongly believing Bob did not block, due to the necessarily finite number of interferometers causing some chance of Bob's blocker absorbing the photon. Further, the photon still travels at the same time as the information. Waves carrying information in the opposite direction to which they are travelling is a well-known classical phenomenon [Vai19]. Therefore, this protocol only seems quantum when you consider light as local—where, for one possibility, the photon travels from Alice to Bob.

This obviously creates a weak trace at Bob, and so is not counterfactual. Arvidsson-Shukur et al attempt to advocate their protocol by saying it tolerates error better than others [AGB17]. Similarly, they call other protocols classical by using a model of classicality where Alice and Bob having extra, non-trivial resources (e.g. a shared clock) [AB19]. However, they give no reason to view their protocol as true quantum counterfactual communication. Therefore, Calafell et al's experimental realisation of this protocol just demonstrates classical counterfactuality [Cal+19].

2.4.3 Photon only Travels Erroneously

The next set of protocols have Alice receiving a photon for both bit values, which has never been to Bob. This means the photon cannot go to Bob when Alice gains information, as then Alice would be unable to see what was sent. When a photon goes to Bob, the protocol is aborted and repeated. This creates a source of loss.

The first subtype of these protocols are those where this loss varies depending on the bit-value Bob is trying to send. This leads us to ask if Alice can, by knowing loss probabilities, guess the bit Bob sends solely based on whether any of her detectors click. This would make these protocols the same as the last category. We can also ask if this post-selection is to blame for any peculiar effects observed, even with post-selection enforced according to the protocol.

The first protocol of this sort proposed was Salih et al's 2013 protocol. This was also the first protocol claiming to be fully counterfactual for both bit values [Sal+13a; Sal+13b]. It is formed of a chain of outer interferometers, each containing a chain of inner interferometers (see Fig. 2.7).

We define the Bloch sphere for polarisation with poles $|H\rangle$ and $|V\rangle$, and rotations

$$\hat{\mathbf{R}}_x(\theta) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -i\sin\left(\frac{\theta}{2}\right) \\ i\sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{pmatrix} = e^{-i\theta\hat{\sigma}_x/2} \tag{2.9}$$

$$\hat{\mathbf{R}}_y(\theta) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -\sin\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{pmatrix} = e^{-i\theta\hat{\sigma}_y/2} \tag{2.10}$$

$$\hat{\mathbf{R}}_z(\theta) = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix} = e^{-i\theta\hat{\sigma}_z/2} \tag{2.11}$$

for Pauli matrices $\hat{\sigma}_{x,y,z}$.

We think of the detectors as being placed far enough, such that they perform no measurement before the photon had had time to exit the protocol. Any photonic component travelling towards either detector can thus be thought of as entering a loss mode. This means, if the photon exits the protocol successfully, it cannot have taken the path towards that detector, and the detector will not register a click.

We first go over the protocol for an arbitrary outer interferometer cycle. A photon of state $a|H\rangle + b|V\rangle$ enters an outer interferometer through a HWP tuned to apply a $\hat{\mathbf{R}}_y(\pi/M)$ rotation. The photon then enters a PBS, which transmits H-polarised light, but reflects V-polarised.

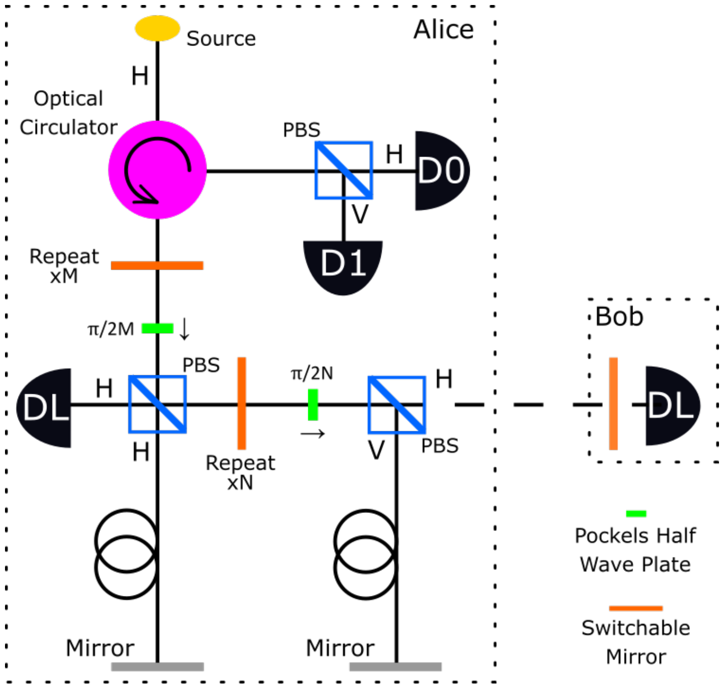

Figure 2.7: Michelson form of Salih et al's protocol for counterfactual communication [Sal+13b]. Each PBS reflects any vertically-polarised (V) light, and transmits any horizontally-polarised (H) light. The HWPs rotate polarisation between horizontal and vertical polarisation unitarily, by an angle of either $\pi/2M$ (the HWPs before each outer interferometer) or $\pi/2N$ (the HWPs before each inner interferometer). To send a '1'-bit, Bob blocks (in this implementation, sets his switchable mirror to transmit, so anything crossing to his side of the channel goes to his loss detector (D_L)). To send a '0'-bit, he doesn't block (in this implementation, sets his switchable mirror to reflect). Switchable mirrors are used to repeat the interferometer for M outer cycles of N inner cycles. In general $M \geq 2$, $N \geq 2$.

The V-polarised component circles through a series of N inner interferometers. In each, it goes through a HWP tuned to apply a $\hat{\mathbf{R}}_y(\pi/N)$ rotation, then through another PBS. The H-polarised component from this PBS passes across the channel, from Alice to Bob, who can choose to block or not block, by switching on or off his switchable mirror. If he blocks, this H-polarised component goes into a loss mode towards Bob's detector D_L; if not, it returns to Alice's side, recombines at another PBS with the V-polarised component, then enters the next inner interferometer. After the chain of N inner interferometers, the resulting components are then passed through one final PBS, sending any H-polarised component that has been to Bob into a loss mode towards Alice's detector D_L. Surviving V-polarised components are recombined at another PBS with the H-polarised component from the delay arm of the outer interferometer. Importantly, neither D_L clicking ensures that the photon has not been to Bob.

Each inner interferometer applies $\hat{\mathbf{R}}_y(\pi/N)$. If Bob doesn't block, the rotations sum to

$$\hat{\mathbf{U}}_{NB}^{N} = (e^{-i\pi\hat{\sigma}_y/2N})^N = e^{-i\pi\hat{\sigma}_y/2} = \hat{\mathbf{R}}_y(\pi) \tag{2.12}$$

Therefore, the state after the inner interferometer chain is

$$|V\rangle_I \rightarrow \hat{\mathbf{U}}_{NB}^N |V\rangle_I = |H\rangle_I \rightarrow \text{Loss} \tag{2.13}$$

This means the V-polarised component becomes H-polarised, and enters the loss mode towards detector D_L after the final PBS. Ergo, if Bob doesn't block, the only component of the wavefunction exiting the outer interferometer is the H-polarised one that went via the outer arm.

Similarly, if Bob blocks for all inner interferometers,

$$\begin{aligned} \hat{\mathbf{A}}_B^N &= \left[e^{-i\pi\hat{\sigma}_y/2N} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right]^N \\ &= \begin{pmatrix} \cos(\frac{\pi}{2N})^N & 0 \\ \cos(\frac{\pi}{2N})^{N-1}\sin(\frac{\pi}{2N}) & 0 \end{pmatrix} \end{aligned} \tag{2.14}$$

Therefore, the state after an outer interferometer is

$$\begin{aligned} |V\rangle_I \rightarrow & \hat{\mathbf{A}}_B^N |V\rangle_I \\ & = \cos\left(\frac{\pi}{2N}\right)^N |V\rangle_I + \cos\left(\frac{\pi}{2N}\right)^{N-1} \sin\left(\frac{\pi}{2N}\right) |H\rangle_I \\ & \rightarrow \cos\left(\frac{\pi}{2N}\right)^N |V\rangle + Loss \end{aligned} \tag{2.15}$$

meaning some V-polarised component exits the outer interferometer.

If Bob doesn't block, the outer cycle applies

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} e^{-i\pi\hat{\sigma}_y/2M} \tag{2.16}$$

If he does block, the outer cycle applies

$$\begin{pmatrix} 1 & 0 \\ 0 & \cos(\frac{\pi}{2N})^N \end{pmatrix} e^{-i\pi\hat{\sigma}_y/2M} \tag{2.17}$$

Salih et al's protocol consists of M chained copies of this outer interferometer. The state entering the first outer interferometer is a H-polarised photon. The output from this first outer interferometer is the input for the second, the output from the second is the input for the third, and so on. After the M^{th} outer interferometer, Salih et al use a final PBS to split the photon into H- and V-polarised components.

As Alice applies a $\hat{\mathbf{R}}_y(\pi/M)$ rotation at the start of each outer interferometer, if Bob doesn't block, the state of the photon after M outer cycles is

$$\cos\left(\frac{\pi}{2M}\right)^M |H\rangle \tag{2.18}$$

Therefore, if the photon isn't lost, it remains H-polarised. However, if Bob blocks, the photon after M outer cycles (as $N \rightarrow \infty$) becomes V-polarised.

If Bob does not block, the chain of N 'π/N' rotations turn the V-polarised component in the inner interferometer chain to H, and so anything in the inner chain is sent to a loss channel D_L. This means Alice can only receive the photon if it went via her outer path, and so arrived at her D_0. If Bob blocks his paths, he absorbs this inner-chain H-polarised light. Therefore, the light is continually reset to V-polarised at the end of each inner interferometer. This stays on Alice's side, and reaches her D_1 as a V-polarised photon, having never travelled to Bob. A small number of photons are absorbed at Bob, reducing the efficiency. The probability of this happening decreases as M and N increase.

The only way Alice's D_1 can click is if Bob blocks; and in the infinite limit of chained outer cycles, the only way her D_0 can click is if he doesn't. Unlike the classical case, both the '0' and '1' bit-values are received without energy transfer across the channel, and so both are sent counterfactually.

However, Vaidman claimed this was not counterfactual, when assessed by the Weak Trace criterion (see Subsection 2.2.3) [Vai13a]. He claimed the TSVF gives a weak trace on Bob's side of the channel when Bob does not block, meaning we cannot say the photon was not there [Vai14a; Vai16b]. However, this is only for the simplest (polarisation-free) form of the protocol. The polarisation-based form (as shown in Fig. 2.7 and [Sal+22]) avoids a weak trace on Bob's side. It does this by ensuring the only waves that go to Bob are H-polarised, which are lost via D_3 on Bob's side, restarting the protocol [Sal+14]. Then, when Bob does not block, the differences in polarisation between the forward- and backward-travelling states keep them separate on Bob's side—giving no Weak Trace there. In Chapter 4, we demonstrate this through weak measurement, by applying Danan et al's method [Dan+13] to a single outer cycle version of Salih et al's protocol, and show there is no weak trace from Bob's side visible at Alice's detectors.

Griffiths also claimed the protocol is not counterfactual by Consistent Histories, as a history with a non-zero probability could be traced to Bob's side and back when Bob does not block [Gri16; Gri84; Gri17]. However, again, Griffiths only considered physical paths, rather than polarisations: these provide an extra degree of freedom [Sal18a]. Griffith later showed, when using more than one outer cycle, the family became inconsistent, and so calling the protocol counterfactual was meaningless [Gri18]. However, Salih notes the final outer cycle is counterfactual, while identical earlier ones are meaningless, which seems paradoxical [Sal18c].

Once Salih et al published their protocol, various demonstrations and implementations began to appear [Cao+14; Guo+14b; Guo+14a; Sal14b; ZZL14; SP14; Zub+14; Che+15a; Che+15b; Guo+15; SS15; ALZ16; Che+16; Sal16; Cao+17; Guo+17; Guo+18; Sal18b; Liu+18; ZJS18; Sal18c; Han+19; HR21b; Sal+21; Sal+20]. While many of these don't make use of polarisation, some do.

Zhang et al proposed a protocol, based on Salih et al's, for probabilistic counterfactual communication. Their protocol is not always counterfactual, but they claim the chance of the photon being at Bob can be reduced to nil, and losses (from noise/blocking) are lower [Zha17]. However, they assume photons only trace one path, which isn't always true—so it is not counterfactual.

Vaidman originally claimed counterfactual communication of both bit-values was impossible. However, at roughly the same time as Salih et al defended their protocol using polarisation, he, alongside Aharonov, released a protocol allowing just this [AV19]. This method is effectively the same as in Salih et al's original protocol. However, to avoid a weak trace in this set-up, without using a polarisation degree of separation, at least two inner interferometers are needed. Alongside this, Aharonov and Vaidman make repeated reference to a double-sided mirror in the protocol. However, all this does is connect the two inner interferometers, and fold the outer path to reduce physical space used. Therefore, it is irrelevant to the protocol's counterfactuality. Also, we show in Subsection 4.2.5 and [Sal+22] that this protocol is not counterfactual by Consistent Histories.

There has only been one protocol published so far where Alice's probability of losing her photon is the same regardless of the bit-value Bob is trying to send: Vaidman's [Vai19].

Shortly after publishing the above protocol with Aharonov, Vaidman created another weak trace-free counterfactual communication protocol. For one outer cycle, this protocol avoids the risk of an erroneous reading that Salih et al's, and their earlier, protocol has. It is again based on a chained MZI set-up, but uses interference from light passing through the inner interferometer chain to alter which detector the photon ends up at when Bob blocks (see Fig. 2.8). This allows Alice, when she receives a bit, to be certain it is the same value Bob sent. Another benefit of the protocol is that, for certain beam-splitter values, losses were the same whether or not Bob blocked. This means Alice cannot infer if Bob blocked, just based on if she receives a photon. Therefore, the protocol cannot be reduced to the information and photon travelling simultaneously, in opposite directions. However, it remains to be seen if it is counterfactual by the Consistent Histories criterion.

2.5 Chapter Conclusion

In this Chapter, we provided the background for the work in the first Part of this thesis. This has involved two threads. First, we looked at the foundational approaches proposed to allow us to evaluate the path travelled by a quantum particle. Some of these approaches allow us to make claims as to the presence or absence of a quantum particle, when standard quantum mechanics remains silent on the matter. Some also make claims which seem to contradict our intuitions as to the path such a particle could have taken (e.g., Vaidman's weak trace approach, which we analyse in Chapter 3).

Secondly, we looked more historically at the evolution of interaction-free and counterfactual developing quantum technologies (or rather devices and protocols which may in future lead to quantum technologies). We spanned from their initial foundational roots in the Elitzur-Vaidman Bomb Tester, to developed, experimentally-tested protocols, such as Hosten et al's 2006 protocol for counterfactual computation, and Salih et al's 2013 protocol for counterfactual computation.

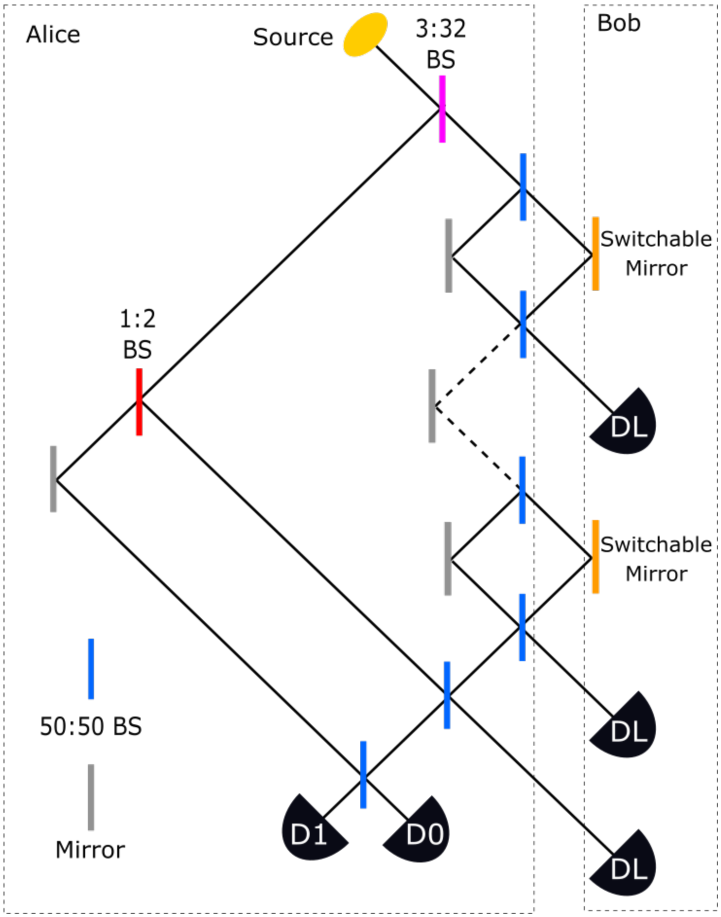

Figure 2.8: Vaidman's protocol for counterfactual communication [Vai19]. Unlike Salih et al's protocol, this protocol does not use polarised light, instead using ordinary (non-polarising) beamsplitters. Through these beamsplitter, when Bob doesn't block his side of the channel, each inner interferometer outputs into loss channels (here, the D_Ls). This means the interferometer on the outer arm (starting with the red beamsplitter) always outputs at D_0, unless the photon is lost. However, when Bob blocks, waves coming out of the inner interferometer chain negatively interfere at the final beamsplitter, causing the photon to go to D_1, if it stayed at Alice. The beam-splitting values given in this example make the losses equal, for two inner interferometers, regardless of whether or not Bob blocks his path (sends bit-value '1' or '0'). We discuss this further in Subsection 2.4.3.

We will move to examine the properties of counterfactual communication in Chapter 4, and Salih et al's protocol will specifically form a key primitive element for the counterfactual quantum information transfer protocols we propose in Chapter 5 and the counterfactual ghost imaging protocol we propose in Chapter 6.

By showing how counterfactual and interaction-free technologies evolved from their foundational beginnings to applied, experimentally-testable protocols ripe to be applied to practical scenarios, this Chapter illustrated how new quantum technologies can be developed from what is initially purely foundational research. This helps show how the fields of quantum foundations and quantum technologies interact, and so the interplay between the two areas. This interaction is continued in the rest of this Part, where we foundationally evaluate and practically develop these counterfactual and interaction-free technologies, reaffirming the interplay between the two fields.

CHAPTER 3

Do weak values show the past of a quantum particle?

Declaration of contribution: This chapter is based on the preprint article *Do weak values show the past of a quantum particle?* [HRL21]. This was conceived and written by myself, supervised and edited by Prof James Ladyman and Prof John Rarity.

3.1 Chapter Introduction

According to weak value approaches to particle presence, non-zero weak values show where a quantum particle has been, even when this would normally be impossible due to the pre- and post-selected states not being eigenstates of spatial projection operators [Dan+13; Vai14b]. Despite weak values only being defined over ensembles, in these approaches, inferred particle presence is attributed to individual particles.

One such approach is the Weak Trace Approach, according to which particles leave traces along the paths they traverse, due to interaction (coupling) with the environment, and such a trace is left even when this interaction is insufficient to count as a projective measurement. A non-zero weak value of an operator on a path (weak trace) is taken to show that an individual particle travelled along that path [Vai13a] even in the limit of this interaction (coupling) vanishing.

According to this approach there is a weak trace in a region when any operator formed by the product of the spatial projection operator for that region with an operator for a property of the particle (e.g. spin) has a non-zero weak value. This can happen even if the weak value of the spatial projection operator itself is zero. Weak traces are supposed to give us more information about the particle than von Neumann measurement of the spatial projection operator would allow. The Weak Trace Approach even allows discontinuous particle trajectories, hinting at a mechanism by which seemingly disconnected events can affect one another, such as in counterfactual communication

[Sal+13b]. Therefore, this approach has been controversial [LAZ13; Vai13b; Eng+17; Gri16; Vai17; Gri17; Sal18a; Gri18; Has+16; Vai16d].

Another approach connecting weak values and particle presence, implicit in the quantum Cheshire cat [Aha+13], requires that the spatial projection operator in a specific location has a non-zero weak value for us to say the particle was present at that location. In the quantum Cheshire cat [Aha+13] (see Subsection 2.2.3 and Fig. 2.2), these approaches come apart since the weak value of the spatial projection operator on a path can be zero, but the weak value of the spin operator times the spatial projection operator non-zero—in these cases, according to the Weak Trace Approach the particle *was* on the path with a zero weak value for the spatial projection operator, contrary to [Aha+13].

This Chapter first looks specifically at the Weak Trace Approach, building on Section 2.2.3. It is argued in [PV18] that the Weak Trace Approach is justified because there is some non-vanishing local interaction between the quantum particle and its environment along whatever path it travels [PV18]. This implies the weak value of the spatial projection operator along a path should always be non-zero whenever a particle travels along that path, and so the particle was not present if the weak value of the spatial projection operator is zero. However, there are cases where according to the Weak Trace Approach a particle is present on a path, despite the spatial projection operator for that path having weak value zero (e.g. in the quantum Cheshire cat set-up).

We then investigate four key issues with weak value approaches.

First, even weak measurements disturb a system, so any approach relying on such a perturbation to determine the location of a quantum particle will only describe this disturbed system, not a hypothetical undisturbed state. We highlight this using the case of a balanced interferometer tuned to have destructive interference (i.e. no light exiting at its dark port) where such a perturbation changes the nature of the system. The unperturbed state is the vacuum, while the perturbed state has light present. While the measurement effect can be made arbitrarily small, this is not the same as removing it entirely. Further, we show attempts to respond to this line of argument by saying there are no completely unperturbed systems, as weak interactions are ubiquitous in nature (as in [PV18]), raises questions about why weak value approaches only associate presence with the part of this interaction which is of the same order of magnitude as the coupling between system and environment.

Secondly, even assuming no disturbance, there is no reason to associate the non-zero weak value of an operator containing the spatial projection operator with the classical idea of 'particle presence'. Indeed, in some situations, just taking the weak value gives features inconsistent with classical ideas associated with a particle being present (e.g., giving discontinuous particle trajectories, or particles not being in coarse-grainings of their location).

Thirdly, weak values are only measurable over ensembles, and so to infer properties of individual particles from values of them is problematic at best.

Finally, weak value approaches to the path of a particle do not give us any new physics beyond standard quantum theory [Vai96], to explain the causes of counterintuitive quantum effects (or even the paradoxes the approach itself creates). They assume a connection between particle presence and weak values without contributing testable new physics.

We also then show that experiments which purportedly support these weak values approaches, do not provide support for these approaches, with their results being replicable using classical electrodynamics. (This is as, due to their design, these experiments can only give results corresponding to the magnitudes of the weak values, rather than showing the phases—negative, imaginary or complex weak values being required for the treatment to be necessarily non-classical).

While this Chapter does not directly introduce any new quantum technological developments, the analysis of weak values approaches to the past of a quantum particle (such as the Weak Trace Approach) relates heavily to how best to quantify if a given counterfactual communication protocol is in fact counterfactual. Further, this Chapter links heavily to weak values, a foundational tool which can provide real technological benefit. We discuss weak values in more depth in the Appendix, as an example of the interplay between quantum foundations and developing quantum technologies.

3.2 The Weak Trace Approach

In this section, we look specifically at issues with the Weak Trace Approach, where there can be a weak trace even if the weak value for the spatial projection operator is zero, so long as the weak value of some compound operator formed from the product of the spatial projection operator with another operator is nonzero.

If the justification for the Weak Trace Approach is the non-vanishing local interaction, as in [PV18], then, if a quantum particle travels along a path, either: the weak value of the spatial projection operator on that path should be non-zero; or we should consider the higher-order terms in λ, as they would indicate additional information about the path on which the particle travelled; or we should say a non-zero weak value for the spatial projection operator is sufficient, but not necessary, to indicate the presence of a quantum particle, as this non-vanishing local interaction can exist along a path without there necessarily being a non-zero weak value for this operator. Hence, the Weak Trace Approach implicitly requires this third option, which raises the question of how it can be used to infer that particles were never present in some specific regions (see the nested interferometer, below). However, it could be that unlike the specific case of the weak value of the spatial projection operator, there existing at least one operator formed of the product of the spatial projection operator for a location and some other operator, with a non-zero weak value, is both a necessary and a sufficient condition for particle presence at that location. Further, it could be that in the nested interferometer case, this condition cannot be met at certain points, so the particle was not there according to the Weak Trace Approach.

While such an argument resolves any formal contradiction in associating a non-zero weak value for the product of some operator with the spatial projection operator with presence, even when the weak value of the spatial projection operator is zero, this still seems peculiar. A key part of the motivation behind weak value approaches is some assumption that the weak value of an operator tells us something about the property associated with that operator, similarly to an eigenvalue obtained by measuring that operator. By this assumption, while measuring the weak value of a spatial projection operator infers some information about whether the particle was at the location corresponding to that operator, measuring the weak value of some product of the spatial projection operator with some other operator corresponding to a property of the particle only infers some information about that property at that location. It requires the separate assumption that a property of a particle cannot be disembodied from that particle to then use the nonzero weak value of such a product operator to then infer about the particle's location.

3.3 Issues with the Weakness Assumption

The process of weak measurement and strong postselection given above leads to a small shift in the position X_d of our pointer system $|\phi(x)\rangle$, with an average value

$$\overline{a} = \lambda \mathrm{Re}(O_w) << \Delta X_d \tag{3.1}$$

Performing the measurement on a pre-and-postselected ensemble of N particles allows measurement of the shift to precision $\Delta X_d/\sqrt{N}$. Therefore, according to the Weak Values Approach, as long as $N > (\Delta X_d)^2 \mathcal{O}(\lambda^{-2})$, the presence of the particle is revealed. This is still however a coupling—so long as $\lambda \neq 0$, there is measurement. This is as it must be, because Busch's Theorem says we cannot gain information about the state of the system without disturbing it [Bus09].

This is in tension with with the Weakness Assumption inherent to weak measurement—that if we measure weakly enough, we can effectively see how the system behaves when it is not measured. As argued above, weak interaction is still interaction and so disturbs the system (albeit negligibly so for some purposes). As long as there is a non-zero coupling, as in all weak value experiments, the 'negligibly' small disturbance is still a disturbance, and there are differences in observable effects between small coupling and no coupling (see [HLR22] for an example of such a difference).

According to popular interpretations of weak values they exist in the absence of measurement and perturbation [Aha+17; Aha+18; Vai96]. However, this is clearly not the case [Sve13; Wie14; Sal15; Sve15; BV17; Kas17; Sve17; Ips22; HLR22]. To defend against this point [PV18] explicitly abandons the Weakness Assumption, saying, "in a hypothetical world with vanishing interaction of the photon with the environment, [the Weak Trace definition] is not applicable, but in the real world there is always some non-vanishing local interaction. Unquestionably, there is an unavoidable interaction of the photon with the mirrors and beam splitters of the interferometer".

This statement, while formally correct and sufficient to avoid the issues associated with the Weakness Assumption, raises questions about why we should only consider the part of the weak trace which is of $\mathcal{O}(\lambda)$, especially when considered in the context of the approach taken practically in obtaining spatial weak values. If the purpose of weak value approaches is to identify the non-vanishing interaction a quantum particle has with its environment, to trace its path, then a definition of this trace which neglects some of these non-vanishing local interactions is only telling us part of the story. Weak value approaches only consider the first-order trace, as per the approximation given in Eq. 2.4, due to the supposedly vanishing nature of the terms of second-order or higher in λ. Despite this, [PV18] specifically says that these environmental terms provide non-vanishing local interaction. A typical photon interaction with a mirror for instance transfers of order 10^{-33} of its energy per reflection (for a 1500nm-wavelength photon and a 1 gram mirror), a suitably weak interaction that is mostly safely ignored. Hence even the higher order terms are many orders of magnitude larger than the strongest perturbation from the environment in optics experiments, and so the invoking in [PV18] of the weak interaction naturally left by the quantum object on its surroundings, while still having the weak trace ignore higher-order terms, seems more an attempt to avoid the issues with the Weakness Assumption than a physically well-motivated argument for believing the weak trace tells us everything about the location of a quantum particle.

It could be argued that this is more an objection to the experimental methods used to detect spatial weak values than the application of the concept itself—however, it is consideration of these experimental methods that is taken to motivate neglecting the higher-order terms (due to their comparative undetectability)—so therefore, if we are instead looking more theoretically at the non-vanishing local interactions, we should associate presence with all terms, rather than just the first order terms in λ.

The need to consider higher-order coefficients of λ has also been discussed in [PC17; ACT18; GC18; Wae+22].

3.4 Do weak values reveal particle presence?

In this Section, we argue that, even assuming no disturbance, there is no reason to associate the non-zero weak value of the spatial projection operator with the classical idea of particle presence. To make this argument, we first give a key assumption for this section—that presence is a property typically ascribed to classical particles, and so any attempt to form a definition of presence for quantum particles should correspond to our intuitions about classical presence, unless we have a good reason for it to deviate from this.

The classical conception of a particle presence—being present at a certain place at a certain time—can be characterised as follows:

i). Every particle is located in space at all times.

ii). Particles cannot be on more than one path simultaneously.

iii). Particle trajectories are continuous (or at least as continuous as space is) so particles cannot get from one place to another without passing through the space in between.

iv). Particles interact with other objects/fields local to their location.

v). If a particle is on a path at a given time, and that path is within some region, then the particle is also located in that region at that time.

vi). If a particle's property is at a location, the particle must be at that location too.

In standard quantum mechanics, when a particle is in an eigenstate of a position operator, then it is attributed a position. However, quantum particles can be in superpositions of position eigenstates. Whether they then have positions at all, or even have two positions at once, is contentious and interpretation-dependent. To require that there is always some location at every time where conditions (i) and (ii) are satisfied is to advocate a hidden variable approach to quantum mechanics, where the hidden variable is the particle's location. However, despite quantum tunnelling and other phenomena, a reformulation of condition (iii) still applies in so far as the probability current of quantum particles evolving according to the Schrödinger equation obeys a continuity equation. (iv-vi) are also compatible with standard quantum mechanics.

Particles that are not in an eigenstate of one of the spatial projection operators demarcating coarse-grained locations of interest (e.g. specific paths in a nested-interferometer experiment) can nonetheless have nonzero weak values for any of those spatial projection operators (or any composite operators formed of any of those spatial projection operators multiplied by some other operator). Perhaps a condition for a particle being present at a location in a quantum context is that it has a non-zero weak value for an operator containing the spatial projection operator at that location. Whenever a particle satisfies condition (ii), it also satisfies this condition, as its forwards and backwards-travelling states always overlap. However, this can only be a necessary condition for particle presence in the sense of (i) to (vi) above, rather than a sufficient condition, as there are cases where:

1). A particle has a non-zero weak value of an operator containing the spatial projection operator at a location, but has no continuous path to/from this location.

2). A particle has a non-zero weak value of an operator containing the spatial projection operator at a location, but a weak value of zero for an operator containing the spatial projection operator at a coarse-graining of that location (e.g. having a non-zero weak value of an operator containing the spatial projection operator on path B, and a non-zero weak value of an operator containing the spatial projection operator on path C, but a weak value of zero for an operator containing the spatial projection operator for the space composed of paths B and C).

3). A particle has a weak value of zero for the spatial projection operator along a certain path, yet has a non-zero weak value for some other operator (e.g. the spin operator in a given direction) multiplied by the spatial projection operator along that path (e.g. in the quantum Cheshire cat set-up).

Using the two-state vector formalism and weak value tools, we can quantitatively analyse the nested interferometer set-up to show Point (2). We first define the forwards-travelling initial vector and backwards-travelling final vector by the paths via which they evolve:

$$\begin{aligned} \left|\psi_i\right\rangle &= \frac{\sqrt{2}\,|A\rangle + |B\rangle + |C\rangle}{2} \\ \left\langle\psi_f\right| &= \frac{\sqrt{2}\,\langle A| + \langle B| - \langle C|}{2} \end{aligned} \tag{3.2}$$

Using these, and defining the spatial projection operator for arm A as $\hat{P}_A = |A\rangle\langle A|$ (similarly for B and C), we get the weak values as [Vai13a]:

$$\begin{aligned} \langle\hat{P}_A\rangle_w &= \frac{\left\langle\psi_f\right| |A\rangle\langle A| \left|\psi_i\right\rangle}{\langle\psi_f|\psi_i\rangle} = 1 \\ \langle\hat{P}_B\rangle_w &= \frac{\left\langle\psi_f\right| |B\rangle\langle B| \left|\psi_i\right\rangle}{\langle\psi_f|\psi_i\rangle} = \frac{1}{2} \\ \langle\hat{P}_C\rangle_w &= \frac{\left\langle\psi_f\right| |C\rangle\langle C| \left|\psi_i\right\rangle}{\langle\psi_f|\psi_i\rangle} = -\frac{1}{2} \end{aligned} \tag{3.3}$$

However, if we want to see if the particle was present in the inner interferometer as a whole (as in either arm B or C), we define $\hat{P}_{BC} = |B\rangle\langle B| + |C\rangle\langle C|$, as we are allowed to do since projectors in standard quantum mechanics are linear. We find

$$\begin{aligned} \langle\hat{P}_{BC}\rangle_w &= \frac{\left\langle\psi_f\right| (|B\rangle\langle B| + |C\rangle\langle C|) \left|\psi_i\right\rangle}{\langle\psi_f|\psi_i\rangle} \\ &= \frac{1}{2} - \frac{1}{2} = 0 \end{aligned} \tag{3.4}$$

If we assumed a non-zero weak trace implies particle presence, this would mean the photon was never in the inner interferometer (made up of arms B or C) overall. ([Vai96; Ben+17] explicitly say that weak values obey the sum rule, and so allows us to say $\langle\hat{P}_{BC}\rangle_w$ must equal $\langle\hat{P}_B\rangle_w + \langle\hat{P}_C\rangle_w$.) Taken separately the summands are taken to mean that the particle was in arm B, and was in arm C, respectively, yet the sum being 0 is taken to mean that the particle is not in their union (i.e., not in the inner interferometer at all); this seems incoherent. (Aharonov et al consider a similar scenario in their three-box experiment, and also discuss this idea of negative weak value in the equivalent of arm C cancelling the positive weak value in the equivalent of arm B [Aha+17]. This usefulness of considering the sign of the weak value for resolving paradoxes is also discussed in [Wae+22].) This illustrates the importance of the sign of the weak values, which the Weak Trace Approach, and the graphical form of the TSVF (as in Fig. 2.1), neglect.

For weak value experiments aiming to distinguish which-path information, we can link this to the Visibility-Distinguishability Inequality [JSV95; Eng96],

$$\mathscr{D}^2 + \mathscr{V}^2 \leq 1 \tag{3.5}$$

where $\mathscr{D}$ is the distinguishability of which path the light travelled, and $\mathscr{V}$ is the fringe-visibility at the output of the interferometer. Therefore, doing anything which would increase the distinguishability between the two paths (e.g. placing different tags on B and C) will affect the interference pattern at the BCE beamsplitter. Given perfect interference is required to ensure all light that enters the inner interferometer from D exits into $D3$, anything causing distinguishability between paths B and C will allow light to leak through onto E, and cause a trace to show in $D2$. Therefore, it will never be showing what would happen in an unperturbed system.

Given Eq. 3.5, $\langle \hat{P}_{BC} \rangle_w$ is a far better measure of whether light reaching $D2$ was ever in the undisturbed inner interferometer. This is as measuring $\langle \hat{P}_{BC} \rangle_w$ does not cause distinguishability between paths B and C in the inner interferometer, and so does not affect the interference pattern required to output all inner interferometer light into $D3$. $\langle \hat{P}_{BC} \rangle_w$ being 0 provides support for a 'common-sense' path (i.e. light only travelling via A to reach $D2$) in an unperturbed system.

An argument against looking at the weak value of the projector over multiple paths (e.g. $\hat{P}_{BC}$) specific to the Weak Trace Approach is that the weak trace is only defined as the effect of the particle on the environment local to it—and combining the two weak values for these two paths is not a local operation, so does not indicate anything about the weak trace. However, this still goes against classical ideas of presence, given it shows that the presence in each of the two arms must have opposite sign, so they can cancel when added, and so implies that being realist about weak values implying presence requires being realist in some way about negative presence. This negative presence is yet another way the kind of presence that is attributed by weak value approaches differs from classical presence.

Point (3) seems to support the Weak Trace Approach to particle presence, whereby a non-zero weak value of any operator corresponding to a property of the particle, multiplied by the spatial projection operator along a path, implies the particle was present along that path. However, as we saw in Section 3.2, taking this to imply presence even when the weak value of the spatial projection operator alone is zero, contradicts the argument in [PV18] that the weak trace comes from the necessarily nonzero interaction between a particle and its environment, given if this were the case, we would expect a particle to necessarily have a nonzero weak value for the spatial projection operator for any path along which it travels. This is a dilemma for weak value approaches—they either contradict the rationale for weak values indicating particle presence, or imply that a particle's properties are not necessarily in the same location as that particle.

From these three points, we see that the "nonzero weak value for an operator containing the spatial projection operator" condition can be satisfied while the conditions for a particle being present at a location are not satisfied. Therefore, it is not a sufficient condition for particle presence in the standard sense of the term.

3.5 Weak values are only defined over ensembles

It is important to note that the weak value approach to particle presence is used to attribute presence to individual particles. This is despite the experimental results of measurements of weak values necessarily being produced by ensembles, because weak values are obtained by post-selection. It is non-obvious that we can go from facts about ensembles to facts about individual elements of that ensemble. As an example, we see that one can have a weak value of the velocity of electrons in a given set-up that is greater than the speed of light in a vacuum [RA02; Sol+04; AP04]. However, this does not necessarily mean than any of the individual electrons in the ensemble used to obtain this weak value travelled at superluminal speeds—as otherwise this would be a violation of special relativity. Instead, similar to how group velocities of wavepackets can be superluminal despite the phase velocity (or velocity of the components) being below c, this superluminal weak value for speed is a fact only about the ensemble, rather than any of its constituents. Similarly, there is no reason to infer facts about the presence of individual quantum particles from weak values of operators containing the spatial projection operator.

Note, recent work has shown anomalous weak values can be observed using single-particle detection [Reb+21]—however, this requires the pre- and postselected states to be protected, which effectively involves re-initialising the preselected state on the particle after coupling. This process is could be described as still considering an ensemble of results, just embodied on the same particle at different times, rather than different particles at the same time.

3.6 Why adopt weak value approaches to particle presence?

Weak value approaches are intended to provide more (interpretational) information about the underlying state of the system than standard quantum mechanics. Vaidman [Vai13a] says that an issue with Wheeler's "common-sense" approach to particle trajectories [Whe78], is that it is entirely operational: not telling us anything about the underlying mechanisms at work, just the final result. Similarly, [Vai13a] says "the von Neumann description of the particle alone is not sufficient to explain the weak trace"—implying weak value approaches provide something more than the standard von Neumann approach does. We however turn this criticism back on weak value approaches – they do not tell us anything about the underlying system either, beyond standard quantum mechanics.

A key motivation for weak value approaches is to explain how the results of interference are affected by changes to disconnected regions. They are intended to show that phenomena such as Wheeler's Delayed Choice, or Salih et al's Counterfactual Communication protocol [Sal+13b; HLR21] and related effects [Sal16; Sal18c; HR21b; Sal+21; Sal+20] are not as "spooky" as they appear [Vai14a; Vai15; Vai16b; Vai19]. Salih et al however have shown that in their communication protocol, the weak value of the spatial projection operator (or any compound operator including the spatial projection operator) at Bob is always zero when Alice receives the

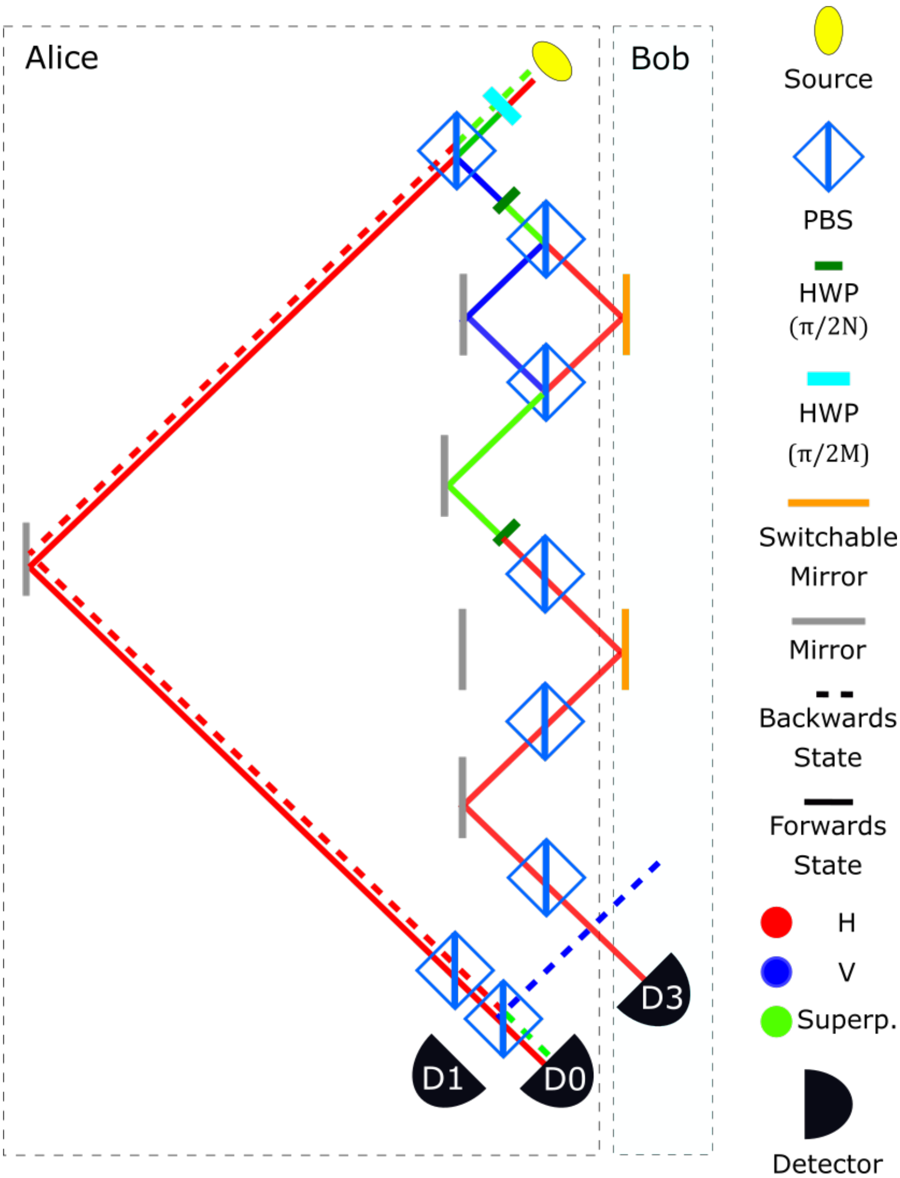

Figure 3.1: Two-State Vector Formalism analysis of Salih et al's polarisation-based single-outer-cycle protocol for counterfactual communication. Bob communicates with Alice by turning off/on his switchable mirrors to determine whether the photon goes to $D1$ or $D0$ respectively. We specify the polarisation, given it determines direction of travel through the PBSs. The forwards- and backwards-travelling states do not overlap anywhere on the inner interferometer chain when there is a detection at $D0$, meaning, by weak value approaches, the particle detected at $D0$ was never at Bob. This means Bob's ability to communicate with Alice is not explained by weak value approaches any more than it is by standard quantum mechanics. We reinforce this analysis experimentally in Section 4.2.

quantum particle (see Fig. 3.1, and Section 4.2 and [Sal+22]). Aharonov and Vaidman have also given an altered protocol for counterfactual communication where the weak value of the spatial projection operator (or any compound operator including the spatial projection operator) at Bob is zero [AV19; WCV21]. Each of these results show that weak value approaches do not explain this phenomenon.

While suggesting a time-symmetry to quantum processes through the TSVF, weak value approaches do not imply any new physics beyond standard quantum theory [Vai96], to explain the causes of counterintuitive quantum effects when there is a nonzero weak value of the spatial projection operator. These approaches simply assume the particle was present wherever the weak value of an operator containing the spatial projection operator is nonzero. Hence they posit particle presence but contribute nothing testable to our physics. This label confuses matters by oversimplifying a complex concept: what it means for a specific particle to be sequentially "present" at a two specific places in quantum field theory. It also causes a number of paradoxes by itself, such as the discontinuous photon trajectories in the nested interferometer set-up above.

3.7 Danan et al's experiment does not support weak value approaches

In this Section, we describe issues with Danan et al's experiment, which they claim provides evidence for weak value approaches to the past of a quantum particle. Danan et al had mirrors MA, MB, MC, MD and ME on Fig. 2.1 each oscillate at a different frequency, and sent weak coherent pulses through the arrangement, to probe where, supposedly, a photon had been. This involved recording the position over time of the beam centroid detected at $D2$, taking the Fourier transform of this, and cross-referencing each frequency peak with its respective mirror [Dan+13]. Zhou et al repeated the experiment and obtained the same results [Zho+17]. Both Danan et al and Zhou et al showed only peaks corresponding to the frequencies of MA, MB and MC, which they claim shows photons never travel by paths D or E, and proves Vaidman's hypothesis. It is debatable however if this is a weak value experiment, or if it shows the validity of weak value approaches.

The nested-interferometer experiment in Fig. 2.1 requires perfect interference at the beam-splitters splitting and recombining paths B and C, to ensure no photons travel along paths D or E. Experimentally, it is vital to preserve the interference pattern (i.e., to ensure paths B and C negatively interfere at the ports leading to path D and E) to ensure photons do not travel along path D and E in the postselected case. We can see this by considering the protocol by standard quantum mechanics (not including postselection) where half of the photons will travel via path D and enter the inner interferometer. In this situation, if B and C don't interfere destructively for port E on the BCE beamsplitter, photons will leak out via path E, and enter detectors $D1$ and $D2$. Obviously photons will then have travelled via paths D and E to get to $D2$. Danan and

Vaidman both claim we know no photons have travelled via paths D and E when we don't observe the mirror oscillation from MD or ME on the output at $D2$, and the intensity detected at $D3$ remains constant.

This ignores however the difference in scale between the effects of oscillations on the combined beam travelling to/from the inner interferometer (that caused by MD or ME), and the interference-altering oscillations when the beam is split (that caused by MB and MC) [Has+16]. By definition, the raw oscillations have to be small, in order for the perturbation to count as a weak measurement, but small compared to the beam diameter at $D2$ ($\mathcal{O}(10^{-3}m)$) is very different to small compared to the wavelength of the photon used ($\mathcal{O}(10^{-6}m)$. (This difference in scale is well known, and is key to homodyne detection). Therefore, the oscillations from mirrors MD and ME are not of the scale of the oscillations from MB and MC—but this does not mean photons didn't travel along paths D or E.

To quantify this, Danan et al say the oscillating mirrors each caused an angular change of 300nrad. They do not give the dimensions of the inner interferometer, but assuming it is of $\mathcal{O}(10^{-1})$m, the change in position on the recombining beamsplitter in the inner interferometer is roughly 30nm. Given the beam had wavelength 785nm, this is a comparatively large change in interference, allowing 1/50 of the light going into the inner interferometer via D to exit via E. One would expect this to show up on detector $D3$. Danan et al unfortunately do not show us the detected spectrum from $D3$.

Further, by showing that classical optics also predicts to first order oscillations from MA, MB and MC, but not from MD or ME, if light went via D and E, Danan et al make our point for us—their experiment gives us no reason to believe light can discontinuously "hop" into the inner interferometer without travelling through arms D and E. Saldanha, and Potǒchek and Ferenczi, both give the same analysis of Danan et al's results being obtainable using classical optics, with light slipping through when the oscillations at MB and MC unbalance the interferometer [Sal14a; PF15; Vai16a].

This analysis is reinforced by that of Salih [Sal15; Dan+15], Wieśniak [Wie14], and Svensson [Sve13; Sve15; BV17; Sve17], who all independently show that, if MB and MC oscillate at the same rate, their respective oscillation patterns on $D2$ disappear, as the interferometer is kept balanced. Similarly, Alonso and Jordan show that adding a Dove prism to the set-up along path C, while still balancing interferometer CD, allows the oscillations at MD to be seen. This yet again proves the visibility of MB and MC, but not MD and ME, to be the result of differences in scale of oscillation effect, rather than light never passing along D or E [AJ15].

Supporters of weak value approaches claim this invalidates the experiment, by conceding that the Danan et al experiment isn't a direct measurement of the trace left by the photon [VT18]. This however makes our point for us: the experiment isn't a real measurement of a weak value, and so it contributes little to confirming or denying weak value approaches.

3.8 Chapter Conclusion

This Chapter investigated four issues with both the Weak Trace Approach, and other approaches where a nonzero weak value of the spatial projection operator is implicitly assumed to indicate the presence of an individual quantum particle. First, we have shown even weak measurements disturb a system, so any approach relying on such a perturbation to determine the location of a quantum particle will only describe this disturbed system, not a hypothetical undisturbed state. Secondly, we have shown even assuming no disturbance, there is no reason to associate the non-zero weak value of an operator containing the spatial projection operator with the classical idea of 'particle presence'. Thirdly, we have discussed that weak values are only measurable over ensembles, and so to infer properties of individual particles from values of them is problematic at best. Finally, we have shown that weak value approaches to the path of a particle do not contribute any new physics—the assumption of a connection between particle presence and weak values does not give us anything testable. This is not to say that weak values are not useful in general—as discussed, they are very useful in metrology [Dix+09; JMH14; HBL17]—but that these approaches specifically are not useful for helping identify the past path of quantum particles.

This Chapter did not directly introduce any new quantum technological developments. However, its analysis of weak values approaches to the past of a quantum particle relates to how to evaluate whether a given counterfactual communication protocol is in fact counterfactual. Alongside this, it discussed weak values, a foundational tool which can provide real technological benefit. We give weak values as an example of the interplay between quantum foundations and developing quantum technologies in the Appendix.

CHAPTER 4

Properties of Counterfactual Communication

Declaration of contribution: This chapter is based on two published articles—*The laws of physics do not prohibit counterfactual communication* [Sal+22] and *How Quantum is Quantum Counterfactual Communication?* [HLR21].
The laws of physics do not prohibit counterfactual communication was conceived by Mr Hatim Salih and Dr Will McCutcheon, under the supervision of Prof John Rarity. My contribution comprised designing and performing the experiment in Section 4.2.4, co-writing and editing the manuscript, and leading in responding to referees' comments while the paper was at review.
How Quantum is Quantum Counterfactual Communication? was conceived and written by myself, supervised and edited by Prof John Rarity and Prof James Ladyman. An initial draft of what would become this article was submitted as an extended essay for the Project Unit PHILM0008, a 20-credit unit contributing towards the MSci in Physics and Philosophy at the University of Bristol in 2019. However, it was developed significantly after this.

4.1 Chapter Introduction

This Chapter looks at foundational issues related to counterfactual communication. First, we investigate whether truly counterfactual communication is actually possible, by considering whether we can design a protocol which meets both the weak trace [Vai15; Vai14a; Sal+14] and consistent histories [Gri84; Gri16] criteria for counterfactuality introduced in Chapter 2. This involves experimentally demonstrating that there is no weak trace crossing the channel between Alice and Bob when information is communicated. Secondly, we look at whether Salih et al's counterfactual communication protocol is inherently quantum, or whether truly counterfactual communication of both bit-values can be performed classically.

This Chapter gives examples of how we can foundationally investigate developing quantum technologies, to evaluate their properties, and whether they meet claims associated with them. Further, it gives an example of how we can use more developed technologies (here, optical technologies such as weak coherent laser light, micro -electro-mechanical system (MEMS) mirrors, and quadrant position-sensing photodetectors) to test foundational properties (here, whether or not there is a weak trace left across the communication channel) of developing quantum technologies (here, counterfactual communication) based on foundational ideas (here, interaction-free measurement). This demonstrates repeated interaction between quantum foundations and quantum technologies.

4.2 The laws of physics do not prohibit counterfactual communication

4.2.1 Setup

Our aim in this Section is not to construct an efficient communication protocol, with each bit carried by a single photon, but rather to construct a communication protocol where:

i). Alice can determine Bob's bit choice with arbitrarily high accuracy, and;

ii). It can be shown unambiguously that Alice's post-selected photons have never been to Bob.

Similarly, our purpose here is not to construct a secure communication protocol—an eavesdropper may be able to exploit our protocol to obtain the information Bob sends without Alice or Bob realising.

Consider our proposed protocol setup in Fig. 4.1, which includes the equivalent of one outer cycle of Salih et al's (Michelson-type) protocol (see 2.4.3) laid-out sequentially in time, as in [Sal18a]. The two underlying principles here are interaction-free measurement [EV93] and the quantum Zeno effect [MS77]. Here's how the setup works. Alice sends a H-polarised photon from photon source **S**, whose polarisation is then rotated by the action of polarisation rotator HWP1, before the PBS passes the H part along arm **A**, while reflecting the V part along arm **D**. (All PBSs transmit H and reflect V.) The V-polarised component in **D** then encounters a series of polarisation rotators HWP2, each affecting a small rotation, and thePBS, whose collective action is to rotate polarisation from V to H, if Bob does not block the transmission channel. In this case, this component is passed straight towards detector D_3. In this case the only way a photon can be detected at Alice is if it has travelled along arm A instead, in which case it passes through two consecutive PBSs and is detected in $D0$ (bit value '0'). What happens if Bob blocks the transmission channel? Provided that the photon is not lost to Bob's blocking devices, the part of the photon superposition that was in arm **D** at time t_1 is now, after last HWP2, in arm **D**, V-polarised. It is then reflected by two consecutive PBSs on its way towards detector D_1. A click

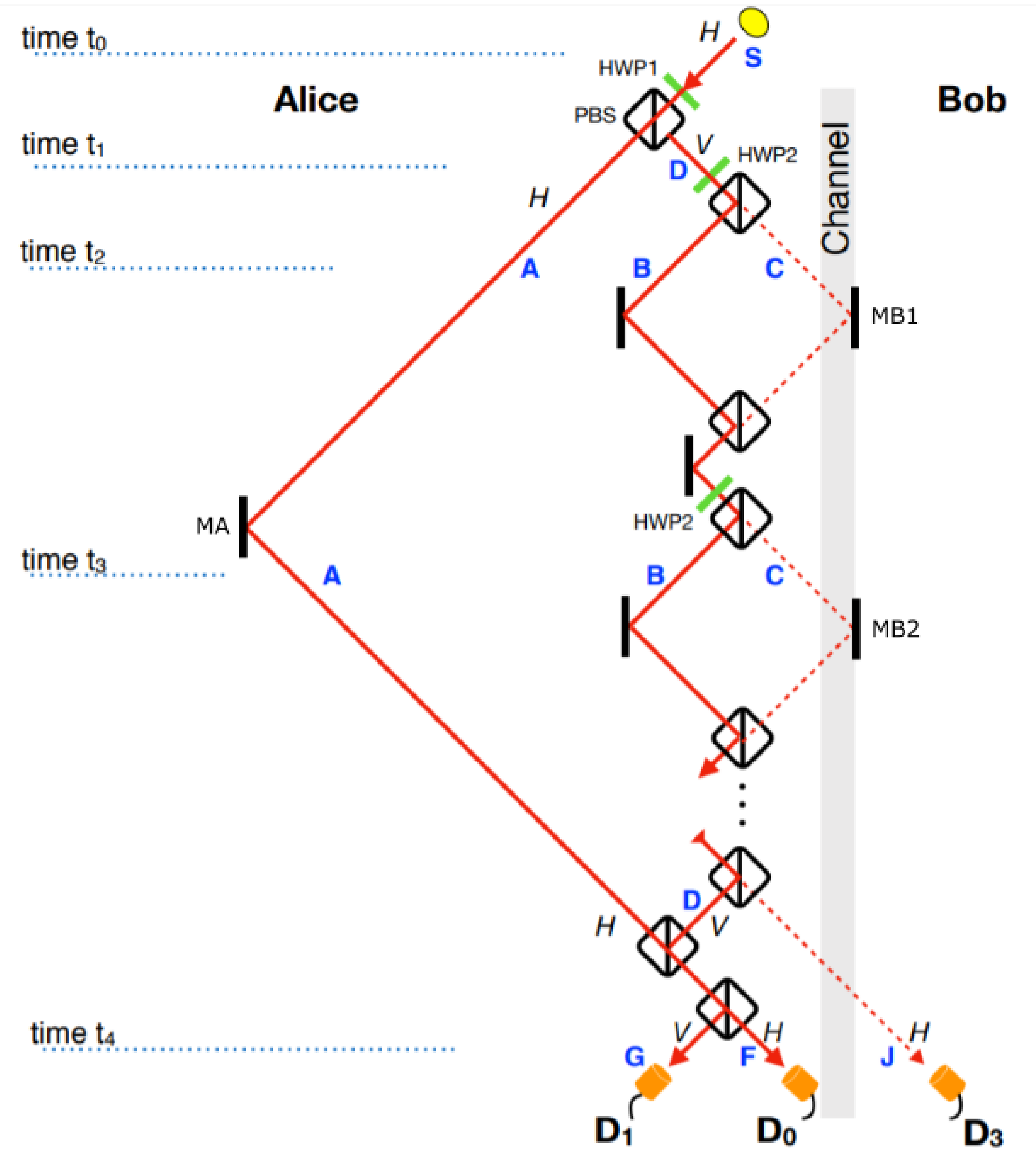

Figure 4.1: Schematic setup for our one cycle implementation of Salih et al's 2013 counterfactual communication protocol (see [Sal+13b] and Subsection 2.4.3). All beamsplitters are PBSs, transmitting H-polarised and reflecting V-polarised light. We want to know if photons detected at Alice's D_0 have been to Bob on the right-hand side. We place detector D_0 at the bottom, rather than immediately after the topmost PBS, so that the setup exactly includes the equivalent of one cycle of Salih et al's Michelson-type protocol laid-out sequentially in time [Sal18a]. This allows the conclusions drawn for this one cycle to be applicable to any of the concatenated cycles from the 2013 counterfactual communication protocol.

at detector D_1 corresponds uniquely to Bob blocking the channel. But there's a chance that the photon component that has travelled along **A** causes detector D_0 to incorrectly click. For example, given that Alice had initially rotated her photon's polarisation after time t_0 such that it is in arm **A** with probability 1/3, and in arm **D** with probability 2/3, and given a large number of HWP2s such that the chance of losing the photon to Bob's blocking device is negligible, then it is straight forward to calculate that the accuracy of detector D_0 is 75%, in contrast to 100% accuracy for D_1, with half the photons being lost on average. Importantly, accuracy can be made arbitrarily close to 100%, by HWP1 initially rotating the photon's polarisation closer to V, at the expense of more photons being lost.

Note, the postselection in the protocol is passive—it happens without any classical communication with Bob. The post-selection process is simply the photons arriving at Alice's detectors D_0 or D_1. The protocol works because when Bob blocks a lot more photons arrive in D_1 than D_0. When Bob doesn't block, all detections are in D_0, having arrived directly from Alice. D_3 is not actually needed for the communication, and is simply a loss channel. In both cases where Alice gets information, the photon stays in her lab, and so a) provably never goes to Bob, and b) has been postselected with no communication coming from Bob.

4.2.2 Demonstrating Counterfactuality

Now we turn to the question of whether Alice's post-selected photon has ever been to Bob. It is accepted that for the case of Bob blocking the transmission channel, Alice's photon could not have been to Bob—otherwise the photon would have been absorbed by Bob's blockers (which act as loss modes from path **C** in this protocol. It is the case of Bob not blocking the channel that is interesting.

4.2.3 Consistent Histories

We first consider the question from a Consistent Histories viewpoint, building on the analysis in [Sal18a]. While we do not unreservedly advocate the Consistent Histories Interpretation, we still engage with it here within the frame of Consistent Histories, as it has been used by Griffiths to question the counterfactuality of counterfactual communication protocols [Gri16]. See [Gri19] for a thorough explanation of Consistent Histories. By constructing a family $\mathscr{Y}$ of consistent histories (which we will shortly explain the meaning of) between an initial state and a final state, that includes histories where the photon takes path **C**, we can ask what the probability of the photon having been to Bob is. Our setup allows us to do just that.

$$\begin{aligned}\mathscr{Y}:&S_0\otimes H_0\odot\{A_1\otimes I_1,D_1\otimes I_1\}\odot\\&\{A_2\otimes I_2,B_2\otimes I_2,C_2\otimes I_2,\}\odot\\&\{A_3\otimes I_3,B_3\otimes I_3,C_3\otimes I_3,\}\odot F_4\otimes H_4\end{aligned}$$

where S_0 and H_0 are the projectors onto arm **S** and polarisation H, respectively, at time t_0. A_1 and I_1 are the projectors onto arm **A** and the identity polarisation I at time t_1, etc. The curly brackets contain different possible projectors at that given time. A history then consists of a sequence of projectors, at successive times. This family of histories therefore consists of a total of 18 histories. For example, the history $(S_0 \otimes H_0) \odot (A_1 \otimes I_0) \odot (A_2 \otimes I_2) \odot (A_3 \otimes I_3) \odot (F_4 \otimes H_4)$ has the photon travelling along arm **A** on its way to detector D_0. Each history has an associated chain ket, whose inner product with itself gives the probability of the sequence of events described by that particular history. Here's the chain ket associated with the history we just stated, $|S_0 \otimes H_0, A_1 \otimes I_1, A_2 \otimes I_2, A_3 \otimes I_3, F_4 \otimes H_4\rangle = (F_4 \otimes H_4)T_{4,3}(A_3 \otimes I_3)T_{3,2}(A_2 \otimes I_2)T_{2,1}(A_1 \otimes I_1)T_{1,0}\,|S_0H_0\rangle$, where $T_{1,0}$ is the unitary transformation between times t_0 and t_1, etc. By applying these unitary transformations and projections, we see that this chain ket is equal to, up to a normalization factor, $|F_4H_4\rangle$.

A family of histories is said to be consistent if all its associated chain kets are mutually orthogonal. It is straight forward to verify that for the family $\mathscr{Y}$ above, each of the other 17 chain kets is zero. For example, the chain ket $|S_0 \otimes H_0, D_1 \otimes I_1, C_2 \otimes I_2, I_3 \otimes I_3, F_4 \otimes H_4\rangle = (F_4 \otimes H_4)T_{4,3}(I_3 \otimes I_3)T_{3,2}(C_2 \otimes I_2)T_{2,1}(D_1 \otimes I_1)T_{1,0}\,|S_0H_0\rangle = (F_4 \otimes H_4)T_{4,3}(I_3 \otimes I_3)T_{3,2}(C_2 \otimes I_2)T_{2,1}\,|D_1V_1\rangle = (F_4 \otimes H_4)T_{4,3}(I_3 \otimes I_3)T_{3,2}\,|C_2H_2\rangle = (F_4 \otimes H_4)T_{4,3}(|C_3H_3\rangle + |B_3V_3\rangle) = (F_4 \otimes H_4)(|G_4V_4\rangle + |J_4H_4\rangle)$, up to a normalization factor. Because projectors F_4, G_4, and J_4 are mutually orthogonal, this chain ket is zero.

Family $\mathscr{Y}$ is therefore consistent.

This means, using $\mathscr{Y}$, we can ask the question of whether the photon has been to Bob. Consistent Histories gives a clear answer: Since every history in this family, except the one where the photon travels along arm **A**, is zero, we can conclude that the photon has never been to Bob.

(Note that, when considering the time evolution of histories ending up at D_1, we can get a non-zero ket for ones where the photon travels to Bob, e.g. a history where the photon is on path **C** at time t_2. However, for the case in question where Bob doesn't block, we know that except for experimental imperfections, only detectors D_0 and D_3 can click—in other words detector D_1 never clicks. This is because the coherent evolution in the inner interferometer-chain ensures that any photon exiting the chain is H-polarised, and as such cannot go to D_1. We therefore apply postselection to "manually" exclude all unphysical histories ending at D_1. We caution that this is only possible as this detection forms a final measurement.)

It can straightforwardly be seen that any history containing any projector C_i, in any family of histories with these start and end points (regardless of coarse/fine-graining), has zero probability due to the final state projection $F_4 \otimes H_4$. Further, the only consistent families starting at $S_0 \otimes H_0$ and ending at $F_4 \otimes H_4$, in which projector $C_i \otimes I_i$ is part of a consistent set at time t_i, are coarse-grainings of family $\mathscr{Y}$ with this same property (i.e. those where we replace $\{A_1 \otimes I_1, D_1 \otimes I_1\}$ with $I_1 \otimes I_1$, or replace $\{A_i \otimes I_i, B_i \otimes I_i, C_i \otimes I_i,\}$ with $\{(I-C)_i \otimes I_i, C_i \otimes I_i,\}$ for times $i = 2$, $i = 3$, or both $i = 2$ and $i = 3$). Therefore, for all relevant consistent families, we can conclude the photon has never been to Bob.

4.2.4 Weak Trace

We now ask the same question in the weak trace language (as introduced in Section 2.2.3 and discussed further in Chapter 3). While we do not unreservedly advocate this point of view (which has been extensively analysed [AJ15; Sal15; Sal18a; HRL21]), we still engage with it here within the frame of weak measurement, as it has been used to challenge the counterfactuality of counterfactual communication protocols [Vai14a; Vai15; Vai16c], .

Let's apply this to our setup. The pre-selected state is that of the photon in arm **S**, H-polarised. And the post-selected state, for the case in question of Bob not blocking, is that of the photon in **F**, also H-polarised. Consider weak measurements where Bob's mirrors, M_{B1} and M_{B2}, are made to vibrate at specific frequencies, before checking if these frequencies show up at a detector D_0 capable of such measurement [Dan+13; Sal15; Dan+15]. The forward evolving state from **S** is clearly present at Bob's, because of the photon component directed by the action of HWP1 and PBS along arm **D**. What about the backward evolving state from **F**? A H-polarised photon travelling from **F** will pass through the two consecutive PBSs along arm **A**, away from Bob. Since, the forward evolving state and the backward evolving state do not overlap at Bob, a weak measurement, at least as a first order approximation, will be zero.

We now show that any weak measurement at Bob will be zero—not just to a first order approximation. Consider a weak measurement where Bob vibrates one or more of his mirrors (as in Danan et al [Dan+13]). This disturbance will cause the part of the photon superposition in arm **D**, after the last HWP2, which can only be V polarised, to be nonzero. This small V component will be reflected by two PBSs towards D_1, and crucially, away from detector D_0. Bob's action has no way of reaching Alice's post-selected state: The photon has never been to Bob.

We performed this experiment using a version of the setup in Fig. 4.1, which we show in Fig. 4.2, with two inner M-Z interferometers within one outer cycle of Salih et al's protocol. The PBSs used are ThorLabs PBS251, and the HWPs used are ThorLabs AHWP05M-600—all three half wave plates are tuned with their fast axis at an angle to the normal such that the polarisation of the light is rotated by $\pi/4$ (i.e. $H \to (H+V)/\sqrt{2}$). In this setup, the single photon source is replaced by a continuous-wave diode laser [635nm ThorLabs LDM635], and the three vibrating mirrors are micro-electro-mechanical systems (MEMS) mirrors [Hamamatsu S12237-03P] weakly oscillating sinusoidally in the horizontal plane—Alice's mirror (M_A) at 29Hz, Bob's first mirror (M_{B1}) at 13Hz, and his second mirror (M_{B2}) at 19Hz (chosen so not to be harmonics of each other). This oscillation is made sufficiently weak so as not to disturb the counterfactual properties of the system. This was done by having maximal 0.01mm movement detected over a 5mm beam diameter at the detectors. The other mirrors in the set-up are all standard ThorLabs MRA25-E02 mirrors. The detectors used for D_0, D_1 and D_3 are segmented quadrant position-sensing photodetectors [ThorLabs PDQ80A], sampled by a LeCroy Wavesurfer 452 at a rate of

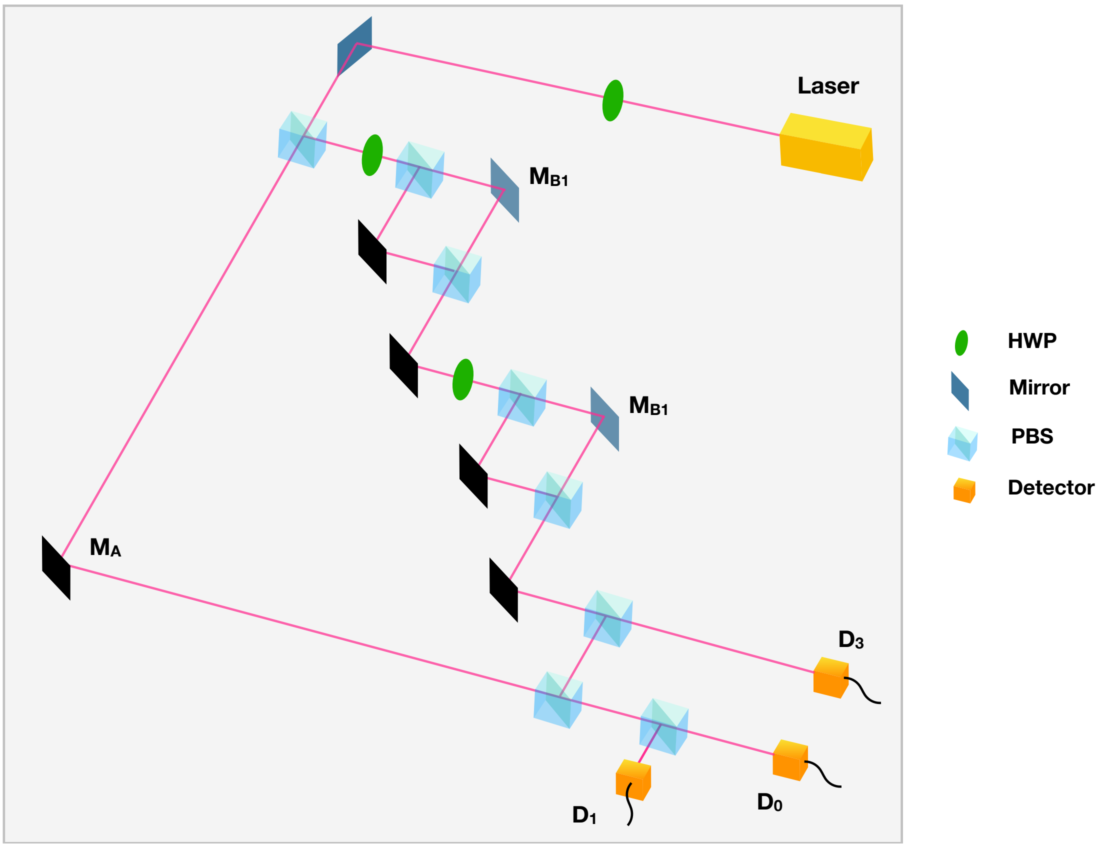

Figure 4.2: 3D depiction of experimental setup. This is derived from the setup in Fig. 4.2. M_A, M_{B1} and M_{B2} are MEMS mirrors oscillating at different frequencies. If a frequency associated with a given mirror is absent from the power spectrum at detector D_0, then according to Vaidman's weak trace approach, we know that photons detected at D_0 have not been near that mirror.

25KHz for 5 seconds. We applied a Fast Fourier transform to the position signal as a function of time, to observe the spectrum of oscillation frequencies at D_0, D_1 and D_3. As can be seen from Fig. 4.3a, there is an oscillation at D_0 from Alice's mirror but not from either of Bob's mirrors. This shows that the weak measurement at Bob is zero, thus demonstrating experimentally the counterfactuality of the protocol.

Note, in Fig. 4.3c, the peak for M_{B1} and its harmonic at D_3 are larger than that for M_{B2}, as M_{B1} is further away from the detectors than M_{B2}. Note also that the peak for M_A in D_1, shown in Fig. 4.3b, is close to the noise level as it is an erroneous signal, caused by light from Alice's path leaking into D_1. However, the fact we can see this error signal at D_1 in Fig. 4.3b, but not the error signals at D_1 from M_{B1} and M_{B2}, shows that in all cases only a *negligible* amount of light (i.e. lower than noise) leaks from Bob back to Alice.

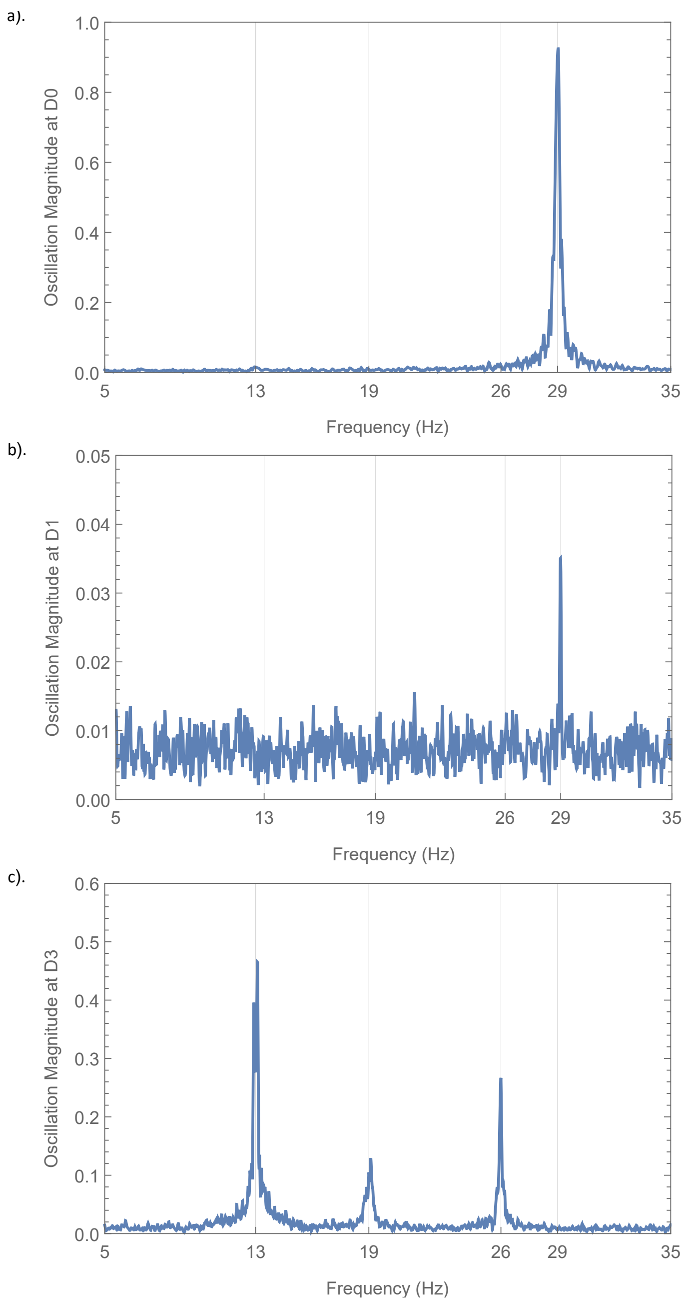

Figure 4.3: Weak measurement tagging showing no weak trace from Bob's mirrors at detector D_0 or D_1. Fourier transform of position with respect to time, of light beam incident on detectors D_0 (a), D_1 (b) and D_3 (c). D_0 and D_1 show the oscillation from Alice's mirror (at 29Hz), but unlike D_3 do not show the oscillations from Bob's two mirrors (at 13Hz and 19Hz, with a second harmonic at 26Hz), proving via weak measurement that no light that goes to Bob's mirrors ends up at either D_0 or D_1.

4.2.5 Discussion

Having shown that any photon detected by Alice at D_0 or D_1 will have never been to Bob, we now look in depth at the probabilities of success, and Alice receiving the correct bit-value. In each round of the proposed experiment Bob chooses a bit, X, he would like to communicate to Alice. He blocks (does not block) his channel when $X = 0$ ($X = 1$). Alice then prepares a single photon, passes it through the system and it is either detected in one of the detectors D_0, D_1 and D_3, or is lost to Bob's blocking device. If Alice detects the photon in either D_0 or D_1, then the round was successful and Alice assigns the estimated values $X_{est} = 0$ and $X_{est} = 1$ to detections in D_0 and D_1 respectively. If the round was not successful another round is performed until she obtains a successful outcome. The post-selected data she obtains displays clear communication from Bob despite the fact that, as we have shown, the postselected photons never passed through the communication channel to Bob. Furthermore, by tuning the initial HWP, the system can be tuned to achieve a postselected success probability arbitrarily approaching unity, at the expense of decreasing the post-selection probability.

We explore the success of the scheme in terms of the free parameter P, the raw probability the photon would be found in the right half of the setup, determined by the setting of HWP1.

The raw conditional probabilities of detection in each of the detectors given Bob Blocking and Not Blocking his side of the channel, for the infinite inner cycle version of the protocol, are: for Blocking, of detection in D_0 of $1-P$, in D_1 of P, and in D_3 (lost) of 0; for Not Blocking, of detection in D_0 of $1-P$, in D_1 of 0, and in D_3 (lost) of P.

We realise that, unlike some other communication protocols, this shows that our protocol has an error chance that varies depending on whether Bob sends a 0-bit or a 1-bit. This, however, is only a feature of the one-outer-cycle model form of counterfactual communication we give here—where two or more outer cycles are used, the error probability in the protocol is 0, and so doesn't vary depending on bits sent (see Chapter 2 for more discussion on unequal bit error rates in counterfactual communication).

Consider for now the limit in which the probability of losing the photon to Bob's blocking apparatus vanishes. Since Bob's strategy for not blocking only succeeds with probability $P(D_0|NB) = (1-P)$, Alice must perform these tests many times to get a successfully postselected event, so the raw probability that Bob was Blocking is greater than a half $P_B \geq 1/2$. This leads to apparent communication in the postselected data since the conditional probability of Blocking given detection events at D_0 decreases.

We assume that on average Bob encodes as many zeros as he does ones. From the raw detection probabilities, the probability of the protocol giving an outcome, that is the post-selection succeeding, is $(2-P)/2$.

We then find the probabilities of the postselected detection events: the probability for detecting in D_0 when Blocking is $(1-P)/2$, and when Not Blocking is $1/2$; the probability for D_1 when Blocking is $P/2$, and when Not Blocking is 0. Therefore after post-selection, the probability of

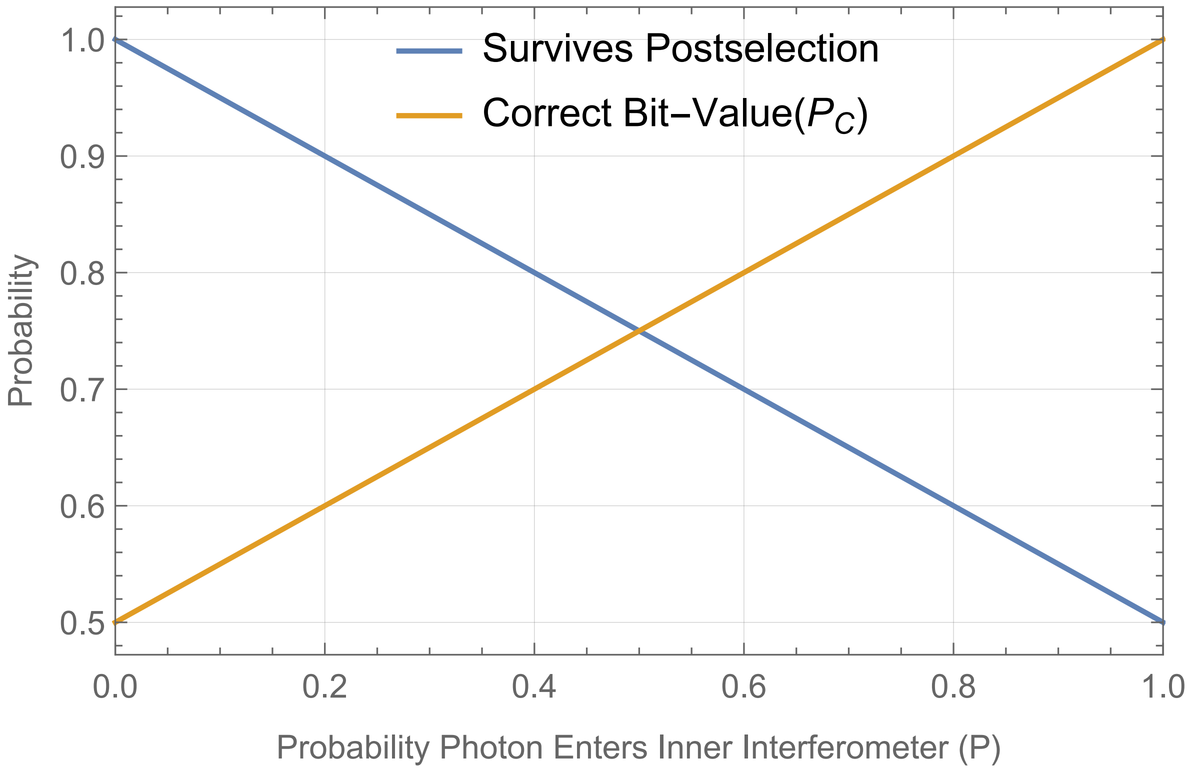

Figure 4.4: Postselection success and accuracy probabilities. The total postselection survival probability (blue), and probability of postselected correct outcome, P_C (orange), for a given bit sent from Bob to Alice in the infinite inner-cycles case of the protocol, plotted against P, the probability of the photon entering the inner interferometer chain.

correct outcome of $P_c = (1+P)/2$.

We see that in the limit $P \to 1$ the protocol becomes deterministic, however the probability of successful postselection goes to zero.

In Fig. 4.4 we plot the overall probability of successful postselected outcome and postselection probability, for different values of the probability P of the photon entering the inner interferometer chain. Notably, for $P = 1/2$ postselection succeeds (photon arrives at Alice) with 3/4 probability, and, if it arrives, is correct (is the bit Bob sent) with 3/4 probability. Increasing P to 2/3, the likelihood of successful postselection drops to 2/3 whilst the probability of being correct increases to 5/6.

Finally, let's illustrate our findings using an amusing scenario. Imagine an outcome-obsessed lab director in charge of this experiment, who is quite happy firing Alice and Bob if a single run of the experiment fails, replacing them with a fresh pair of experimentalists, to start all over, also nicknamed Alice and Bob. The task for any Alice and Bob pair is to communicate a 16-bit message, one bit at a time. Assume the experiment is set up such that the chance of any given run failing is 1/4. Therefore, in order to successfully communicate the 16-bit message, the lab director has to, on average, go through around 100 pairs of experimentalists—which the director secretly enjoys.

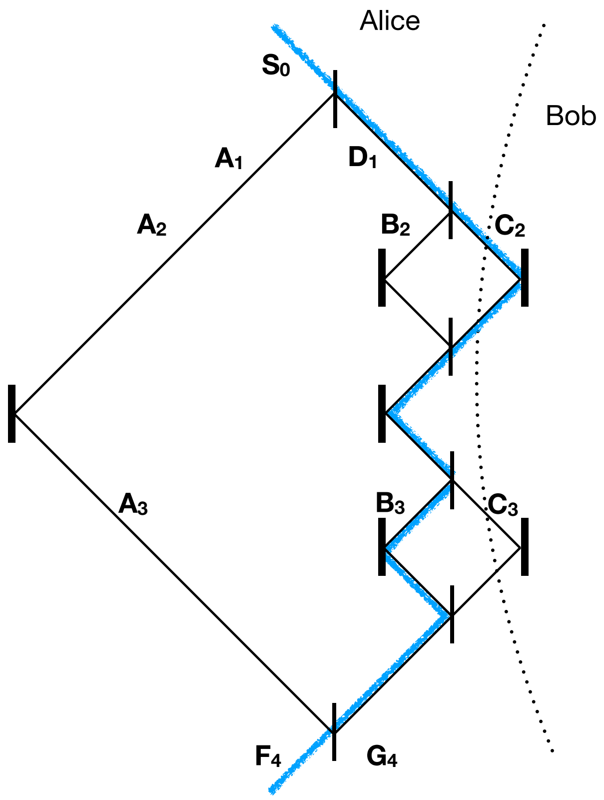

Figure 4.5: Aharonov and Vaidman's protocol overlaid with a counterfactuality-violating history. The modified protocol proposed by Aharonov and Vaidman in [AV19], of which we show one cycle, cannot be said to be counterfactual from a consistent histories viewpoint. This can be seen from the series of projections, or history, highlighted in blue. The thin vertical lines represent non-polarising beamsplitters, thick vertical lines represent mirrors, and the blue path shows an example of a history where, when Bob doesn't block, the photon travels to Bob and back to Alice's relevant detector with nonzero probability. See text for mathematical details.

Each new pair of experimentalists is provided with a new message. Eventually, a lucky Alice and Bob manage to communicate their message (bit accuracy will be 75% on average.) Now the question for the successful pair is: Has any of Alice's photons been to Bob while communicating the message? The answer, as we have shown, is an emphatic no.

Aharonov and Vaidman posted an interesting paper, around the same time this work was first posted on the arXiv, conceding that counterfactual communication was possible, and suggesting a way to modify a version of Salih et al's 2013 protocol that does not use polarisation which satisfies Vaidman's weak trace criterion for communication to be counterfactual [AV19]. Their modification has the effect of satisfying Vaidman's weak trace criterion for counterfactuality for both bits. The scheme has the advantage over ours of reducing the chance of a communication error caused by imperfect interference in the inner interferometers when Bob encodes bit 0, thus satisfying the weak trace criterion even for erroneous bit-clicks, which we are not concerned about. However, unlike the scheme we give above, their modification fails the Consistent Histories criterion for counterfactuality [Gri84]. In Fig. 4.5 we give an example of a history where the photon goes to Bob, where the associated chain-ket is nonzero. More precisely, take the histories family,

$$\begin{aligned}\mathscr{Y}' :& S_0 \odot \{A_1, D_1\} \odot \\ & \{A_2, B_2, C_2,\} \odot \\ & \{A_3, B_3, C_3,\} \odot F_4\end{aligned} \tag{4.1}$$

The chain ket associated with the path highlighted in blue in Fig. 4.5 is

$$\begin{aligned}& |S_0, D_1, C_2, B_3, F_4\rangle = \\ & (F_4)T_{4,3}(B_3)T_{3,2}(C_2)T_{2,1}(D_1)T_{1,0}\,|S_0\rangle = \\ & (F_4)T_{4,3}(B_3)T_{3,2}(C_2)T_{2,1}\,|D_1\rangle = \\ & (F_4)T_{4,3}(B_3)T_{3,2}\,|C_2\rangle = (F_4)T_{4,3}\,|B_3\rangle = |F_4\rangle\end{aligned} \tag{4.2}$$

up to a nonzero normalization factor. Similarly, the chain ket associated with path A

$$\begin{aligned}& |S_0, A_1, A_2, A_3, F_4\rangle = \\ & (F_4)T_{4,3}(A_3)T_{3,2}(A_2)T_{2,1}(A_1)T_{1,0}\,|S_0\rangle = \\ & (F_4)T_{4,3}(A_3)T_{3,2}(A_2)T_{2,1}\,|A_1\rangle = \\ & (F_4)T_{4,3}(A_3)T_{3,2}\,|A_2\rangle = (F_4)T_{4,3}\,|A_3\rangle = |F_4\rangle\end{aligned} \tag{4.3}$$

is also nonzero up to a nonzero normalization factor. The two chain kets are not orthogonal, which means that the family of histories is not consistent. The original Aharonov-Vaidman setup therefore cannot be said to be counterfactual from a consistent histories point of view. Nonetheless, we agree Aharonov and Vaidman's modification presents important progress, as it robustly passes the weak trace criterion for counterfactuality (as shown in [WCV21]). It has thus been included as an additional element in various recent protocols proposed for counterfactual communication (e.g. [Sal18c; Sal+21; Sal+20; HR21b]).

Likewise, Aharonov-Vaidman's modification can be incorporated straightforwardly in our present setup. The way to do this is by simply repeating the inner-cycles sequence, between the arms marked D, twice. This would eliminate the weak trace even for erroneous detector-clicks while still passing the consistent histories test.

In summary, we have shown both theoretically and experimentally that, given post-selection, sending a message without exchanging any physical particles is allowed by the laws of physics.

4.3 How quantum is quantum counterfactual communication?

In Section 2.4, we went over the history of protocols for counterfactual communication, from Noh's initial attempt, to later proposals by Salih et al, Aharonov and Vaidman, and Vaidman.

However, we are still left with one key question—what, if anything, makes these protocols essentially quantum [Gis13]?

To answer this, we need to determine the underlying structure of classical counterfactual communication for comparison (Subsection 4.3.1), and give a sufficient condition for a protocol to be quantum (Subsection 4.3.2).

Subsection 4.3.3 then examines the quantum counterfactual communication protocols proposed so far, to assess their non-classicality. We show what separates them from classical counterfactual communication protocols and how they meet the condition for being quantum.

We identify two essential differences between classical and quantum counterfactual communication. The first is that only one bit-value (e.g. ‘0’) can be sent in a classical protocol without matter or energy transfer associated with the bit being sent. The second is that the two-value quantum protocols require wave-particle duality to be able to send either bit value of each bit sent.

4.3.1 Classical Counterfactuality

Counterfactual communication long predates quantum mechanics. For instance, in the Sherlock Holmes story, *Silver Blaze*, Holmes infers a racehorse was abducted by its own trainer, as the stable dog didn’t bark. As Holmes puts it, “the curious incident of the dog in the night-time" was that the dog did nothing [Doy94]. A more recent fictional example is the Bat-Signal. If there were a major crime being committed the Bat-Signal would appear in the sky, and so the Bat-Signal’s absence counterfactually communicates to Bruce Wayne that all is well. Whenever we receive information from a sign’s absence we are being signalled to counterfactually (e.g. the signal that an engine’s components are functioning as they should is that the warning light is off).

Obviously in each of these cases a single bit is transmitted, and the bit value is signalled without the transfer of matter or energy. However, only one bit-value can be communicated by an absence in this way. Had a stranger kidnapped the racehorse, the dog would have barked, and energy would have have been transferred through the communication channel; correspondingly of course the Bat-Signal and other warning lights involve the transmission of energy when they are on. A sign’s absence can transmit one value of a bit, only if the sign always occurs for the other bit value [Mau02]. This is counterfactual communication based on counterfactual inference.[1] Counterfactual inference is not rare but ubiquitous in everyday life and in science. For example, if there was an ether then the Michelson-Morley experiment would not have a null result.

The structure of classical counterfactual communication as above is as follows. Were A to happen, B would happen. B did not happen. Therefore A did not happen. B not happening is a signal that A did not happen, only because B happening signals that A happened.

Formally,

[1] It is not required that the signalled event’s non-occurrence directly causes the the sign’s absence, as there could be other common causal factors. Also a channel may be more reliable for signalling one bit value than for another.

$$A \supset B; \neg B; \therefore \neg A \tag{4.4}$$

Any instance of this structure in which B doesn't happen can be thought of as counterfactual communication of A's not happening. However, typically, we want a one-to-one correspondence between the signalling event A and the inferred event B, so we always recognise the inferred event's absence. For this, we need the further condition that, were A not to happen, B would not happen ($\neg A \supset \neg B$).

4.3.2 Quantum as Non-Classical

Next, to evaluate the proposed protocols we need a sufficient condition for something being quantum. There are many differences between classical and quantum physics. For optics (which all protocols so far have used), classical physics is everything up to and including Maxwell's equations. These formulate light as the evolution of waves whose intensity can be split continuously [Gri05]. In contrast to this, in the quantum optics needed for many situations, we must consider light as photons [Ein05], which are quanta of the electromagnetic field that are detected as discrete packets of absorbed energy. Despite this discrete particle-like behaviour, in propagation photons retain wave-like properties such as interference. Therefore, in the context of optics it is appropriate to take a protocol to be quantum if it requires using both wave- and particle-like features by combining interference with single photon detection. The latter nullifies the splitting of light intensity across different detectors, and forces it to end in a single location. Therefore, the sufficient definition of "quantum" we shall use in this paper if a protocol requires both interference effects and discrete detection events to work (i.e. each run of the protocol only ever ends in a single detection event, as it uses for instance a single photon).

4.3.3 Protocol Evaluation

Of the protocols proposed so far, only one has been shown counterfactual by both Weak Trace and Consistent Histories—Salih et al's [Sal+13b; Sal+22]. We show this protocol in Fig. 2.7 and give a detailed description in the associated caption. Above we argued that a sufficient condition for a protocol to be quantum is that it requires both discrete detection events and path-interference—which, for light, only single photons can do. We now consider whether this protocol has to meet this condition in order to be counterfactual.

While the limit of many single photons may generate the same results as coherent states, the way in which they produce them differs. This is due to the discreteness discussed, which is not considered when using coherent states, but is when using Fock states (i.e. single photons).

In the quantum case, beamsplitters split a photon's probability amplitude between the two eigenstates that correspond to the photon going in each direction; in the classical case, they split the beam intensity (and field). As interference still occurs, when Bob does not block, waves on both sides still destructively interfere, so the light never returns to Alice. However, Bob's D_3 and Alice's D_0 both detect light simultaneously. Similarly, when he blocks, light goes to his blockers and Alice's D_1 simultaneously. Therefore, in both cases, as light goes between Alice and Bob, it is not counterfactual. While the amount of light going to Bob's D_3 may be infinitesimal for an infinite number of outer cycles, and that going to his blockers infinitesimal for infinite inner cycles, this is not the same as no light going there in either case—so, regardless the number of cycles, with classical light, the protocol isn't counterfactual. This may seem obvious, but many have not realised this and claimed this protocol could be performed classically (e.g. [Gis13]).

The only way to avoid this is to force the light to end at only one point—to postselect, with information only travelling when nothing goes between Alice and Bob. Only single photons can do this. Therefore, the only way to make the protocol counterfactual is to use these, and so make the protocol quantum.

Alongside the proven protocol of Salih et al, Vaidman recently proposed one [Vai19] which, while not yet proven valid by Consistent Histories, has been shown to be valid by the Weak Trace criterion. We show this protocol in Fig. 2.7 and give a detailed description in the associated caption. Similarly to Salih et al's, it relies on both interference and single-photon detection—without the use of single photons, when Bob blocks the paths on his side of the inner interferometers, light could reach both Bob's blocker and Alice's detector D_1. Further, when Bob doesn't block, light could reach both the loss channels (marked 'D_L' in Fig. 2.8) and Alice's D_0—in both these cases, the protocol would definitely not be counterfactual.

Extrapolating from these protocols, any counterfactual protocol (valid or not) whereby the interference from Bob blocking or not blocking his side of an interferometer affects the destination of light on Alice's side, without both Alice and Bob simultaneously detecting that light, relies on both wave-like and particle-like properties. This means, all these protocols, from the Elitzur-Vaidman Bomb Tester, to Noh's counterfactual cryptography scheme, to even Arvidsson-Shukur et al's proposal (despite only sending one bit-value counterfactually), are essentially quantum by our definition above.

We have shown Quantum Counterfactual Communication is essentially quantum. This confirms that both particle-like behaviour and path interference are necessary for schemes where *both* bit-values are sent counterfactually. In all schemes experimentally demonstrated so far, this is the only way it is quantum (though protocols which send quantum information have been proposed theoretically [Sal16; Sal18c; Sal+21; Sal+20]). Quantum Counterfactual Communication allows us to look at principles at the heart of the foundations of quantum physics—such as single-particle self-interference and the quantum Zeno effect—in a new and exciting way, and will hopefully motivate new thought experiments and experimental work based on this seemingly impossible phenomenon.

4.4 Chapter Conclusion

This Chapter considered foundational issues related to counterfactual communication. First, we investigated whether truly counterfactual communication is actually possible, by considering whether we can design a protocol which meets both weak trace [Vai15; Vai14a; Sal+14] and consistent histories [Gri84; Gri16] criteria for counterfactuality (as introduced in Chapter 2). This involved showing experimentally that the protocol meets the weak trace criterion for there being no particles crossing the channel between Alice and Bob when information is communicated, by performing an experiment similar to Danan et al's [Dan+13]. Secondly, we showed that Salih et al's counterfactual communication protocol is inherently quantum, and that counterfactual communication of this sort (where both bit values are transmitted with no particle crossing the channel between communicants) cannot be performed classically.

This Chapter gave examples of how we can foundationally investigate developing quantum technologies, especially those themselves based on foundational ideas, to evaluate their properties, and whether they meet the claims associated with them. Further, it acted as an example of how we can use technology to test foundational properties of developing quantum technologies themselves based on foundational ideas. This elegantly illustrated the interplay between quantum foundations and quantum technologies.

CHAPTER 5

Counterfactual Communication of Quantum Information

Declaration of contribution: This chapter is based the published article *Exchange-Free Computation on an Unknown Qubit at a Distance* [Sal+21], and the preprint article *Deterministic Teleportation and Universal Computation Without Particle Exchange* [Sal+20]. Both were conceived by myself and Mr Hatim Salih, in discussion with Prof Terry Rudolph. I then worked with Mr Salih to create the full protocols given here. I performed all simulation of the protocols in Wolfram Mathematica. The manuscripts were co-written by myself and Mr Salih, under the supervision of Prof John Rarity. Further discussion and edits were done by all authors (including Dr Will McCutcheon, who also wrote an appendix to [Sal+21] not included in this thesis).

5.1 Chapter Introduction

This chapter is the first in this thesis where a foundational idea (here, counterfactual communication) is used for the development of a new form of quantum technology. Specifically, this chapter first gives a protocol where counterfactual communication is used as a primitive to allow the direct application of an arbitrary single-qubit unitary by one party onto another party's qubit. This protocol occurs without any particles being exchanged by the two parties. Alongside this, this Chapter then gives another protocol by which a controlled-Z gate can be applied counterfactually by one party onto the other party's qubit. This is then used as the basis by which quantum information transfer protocols making use of controlled-Z gates, such as teleportation and telecloning, can be performed counterfactually.

5.2 Exchange-Free Computation on an Unknown Qubit

This Section presents a protocol allowing a remote Bob to prepare any qubit he wishes at Alice without any particles passing between them. More generally, Bob can directly apply any arbitrary Bloch-sphere rotation to an unknown qubit at Alice. This allows him to perform any single-qubit quantum computation. While we describe an optical realisation using photon polarisation, the scheme is in principle applicable to other physical implementations.

5.2.1 Protocol

Our protocol consists of a number of nested outer interferometers, each containing a number of inner interferometers, as in Salih et al's 2013 protocol (see [Sal+13b] and Subsection 2.4.3). We combine these interferometers into a device that we call a Phase Unit, allowing Bob to apply a relative phase to Alice's photonic qubit (Fig. 5.1). We pair two Phase Units such that one applies some phase to Alice's H-polarised component, while the other applies an equal but opposite-sign phase to her V-polarised component, resulting in a $\hat{\mathbf{R}}_z(\theta)$ rotator. By chaining three such $\hat{\mathbf{R}}_z(\theta)$ rotators, interspersed with appropriate wave-plates, Bob can apply any arbitrary unitary to Alice's qubit, exchange-free (Fig. 5.2).

We now go into the detail of how the $\hat{\mathbf{R}}_z(\theta)$ rotation is achieved, using the device in Fig. 5.1, where the H-polarised and V-polarised components, are separated out.

We put the H-polarised component through one run of Salih et al's 2013 protocol (see Chapter 2), with Bob either always blocking or not blocking his channel. If he blocks, and the component exits V-polarised, the PBS sends it through a HWP that flips it to H-polarised, and it is kicked out of the device. However, if it is H-polarised, it goes through a phase plate (gaining a phase increase of $\pi/2L$), hits a mirror, goes back through the phase plate (gaining another phase increase of $\pi/2L$, for a total increase of π/L), and re-enters the device for another run.

This is repeated L times, with Bob blocking or not blocking for all outer cycles in a given run. After each run, the component goes into a PBS: if it is H-polarised, it gains a phase of π/L; if V-polarised, it is flipped to H-polarised and sent out from the unit. Bob first doesn't block for k runs, applying a phase of $k\pi/L$, then blocks, applying the transformation

$$|H\rangle \rightarrow e^{ik\pi/L}|H\rangle \tag{5.1}$$

When N is finite, the rotations applied by each outer cycle when Bob blocks are not complete, meaning one run (M outer cycles) doesn't fully rotate the state from H to V. However, given Bob only blocks after the component has had a phase applied to it, to kick the component out of the device, any erroneous H-polarised component can be kept in the device by Bob not blocking for the remaining $L-k$ full runs afterwards. This lets us treat the erroneous H-component as loss.

While coarse-grained for finite L, as L goes to infinity (with $0 \le k/L \le 1$), Bob can generate any relative phase for Alice's qubit, from 0 to π. Further, by moving the $\pi/2$ HWP from its location

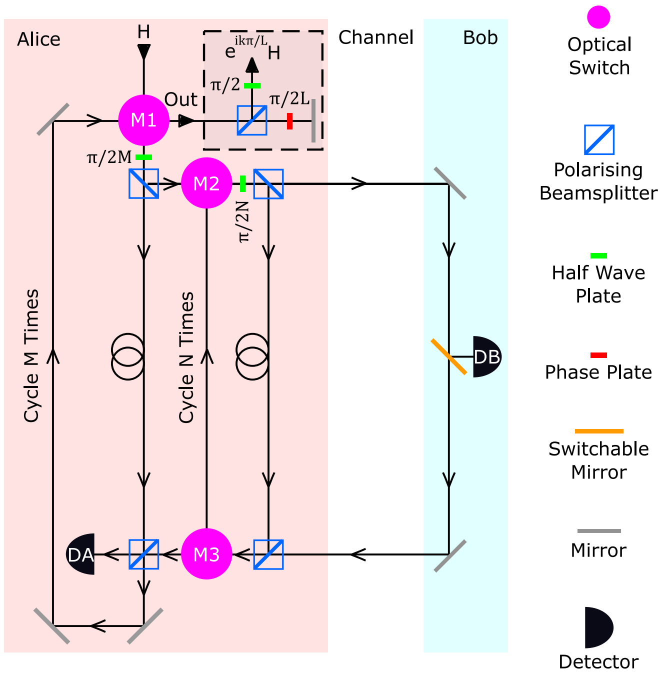

Figure 5.1: Our exchange-free Phase Unit, which applies a phase determined by Bob to Alice's H-polarised input photon. The Phase Unit comprises an equivalent setup to that of Salih et al's 2013 protocol [Sal+13b], but with an added phase-module in the dashed box. The optical switches each alter the paths at different times in the protocol to allow the photon to do the correct number of cycles. Optical switch $M1$ inserts the photon into the device, and keeps it in for M outer cycles; optical switches $M2$ and $M3$ cycle the photon around for N inner cycles per outer cycle. The PBSs transmit H-polarised light, and reflect V-polarised light. The HWPs are tuned to implement $\hat{\mathbf{R}}_y(\theta)$ rotations on polarisation with θ of π, π/M and π/N, as shown in the figure. Detectors D_A and D_B not clicking ensure that the photon has not been to Bob. After M outer cycles, the photon is sent by $M1$ to the right. The photon only exits the Phase Unit if its polarisation had been flipped to V as a result of Bob blocking the channel (which he does by switching his Switchable Mirror on) because of the action of the PBS in the dashed box. The phase plate (tuned to enact a $\hat{\mathbf{R}}_z$, or phase, rotation, and realisable using a tiltable glass plate) adds a phase of $\pi/2L$ to the photon every time it passes through it, summing to π/L every time it is sent H-polarised to the right by M1. Bob doesn't block for k runs (out of a maximum L), then blocks, allowing him to set the final phase of the photon, $k\pi/L$, anywhere from 0 to π, in increments of π/L. An initially V-polarised photon can be put through an altered version of this device to add a phase to it (identical, except for the $\pi/2$ HWP being moved to above $M1$). The unit rotates Alice's qubit by $\hat{\mathbf{R}}_z(k\pi/L)$.

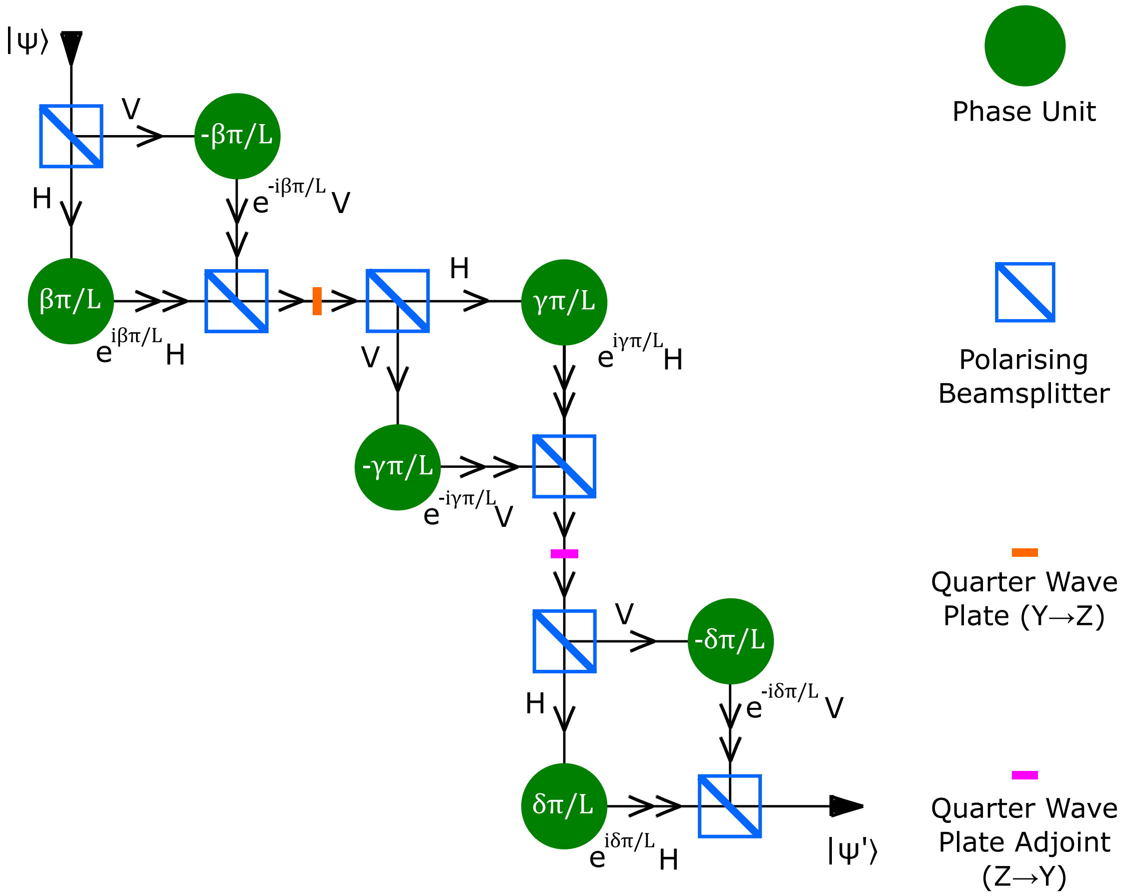

Figure 5.2: The overall protocol, incorporating multiple Phase Units from Fig. 5.1, as well as PBSs (which transmit horizontally-polarised, and reflect vertically-polarised, light), as well as a quarter wave plate and its adjoint (conjugate-transpose). The setup allows Bob to implement any arbitrary unitary on any initial pure state $|\psi\rangle$ Alice inserts, entirely exchange-free.

in Fig. 5.1 to the input, a similar phase can be added to a V-polarised component, relative to a H-polarised component.

This Phase Unit can be constructed to include Aharonov and Vaidman's modification of Salih et al's 2013 protocol [AV19], satisfying their weak-measurement criterion for exchange-free communication. We do this by running the inner cycles for $2N$ cycles rather than N, except that for the case of Bob not blocking, he instead blocks for one of the $2N$ inner cycles, namely the Nth inner cycle. This removes any lingering V component exiting the inner interferometer of Fig. 5.1 due to imperfections in practical implementation.

We now use our Phase Unit as the building block for a protocol where Bob can implement any arbitrary unitary onto Alice's qubit, exchange-free.

Any arbitrary 2×2 unitary matrix can be written as

$$\begin{aligned} \hat{\mathbf{U}} &= e^{i(2\alpha' - \beta'\hat{\sigma}_z - \gamma'\hat{\sigma}_y - \delta'\hat{\sigma}_z)/2} \\ &= e^{i\alpha'}\hat{\mathbf{R}}_z(\beta')\hat{\mathbf{R}}_y(\gamma')\hat{\mathbf{R}}_z(\delta') \end{aligned} \tag{5.2}$$

Note, the factor of $e^{i\alpha'}$ can be ignored, as it provides global rather than relative phase, which is unphysical for a quantum state [NC10].

We can apply the $\hat{\mathbf{R}}_z(\theta)$ rotations using the Phase Unit, and make a $\hat{\mathbf{R}}_y(\theta)$ rotation by sandwiching a $\hat{\mathbf{R}}_z(\theta)$ rotation between a $-\pi/4$-aligned Quarter Wave Plate, $\hat{\mathbf{U}}_{QWP}$, and its adjoint, $\hat{\mathbf{U}}^{\dagger}_{QWP}$, where

$$\begin{aligned} \hat{\mathbf{U}}_{QWP} &= \hat{\mathbf{R}}_x(-\pi/2) = e^{i\pi\hat{\sigma}_x/4} \\ \hat{\mathbf{U}}^{\dagger}_{QWP} &= \hat{\mathbf{R}}_x(\pi/2) = e^{-i\pi\hat{\sigma}_x/4} \end{aligned} \tag{5.3}$$

We set

$$\beta' = 2\pi\beta/L,\ \gamma' = 2\pi\gamma/L,\ \delta' = 2\pi\delta/L \tag{5.4}$$

where, for the three Phase Unit runs, k is β, γ and δ.

The Phase Units form components of the overall protocol, as shown in Fig. 5.2. Here, Alice first splits her input state $|\psi\rangle$ into H- and V-polarised components with a PBS, before putting each component through a Phase Unit, to generate equal and opposite phases on each. She recombines these at another PBS. Afterwards, she puts the components through a quarter wave plate, then through another run of PBS, Phase Unit, and PBS, then through the conjugate-transpose of the quarter-wave plate, tuned to convert the partial $\hat{\mathbf{R}}_z$ rotation (phase rotation) into a partial $\hat{\mathbf{R}}_y$ rotation. Finally, she applies another run of PBS, Phase Unit, and PBS to implement a second $\hat{\mathbf{R}}_z$ rotation.

Using these $\hat{\mathbf{R}}_z$ and $\hat{\mathbf{R}}_y$ rotations, Bob can implement any arbitrary rotation on the surface of the Bloch sphere on Alice's state. This can be used either to allow Bob to prepare an arbitrary pure state at Alice (if she inserts a known state, such as $|H\rangle$), or to perform any arbitrary unitary transformation on Alice's qubit, without Bob necessarily knowing that input state.

5.2.2 Discussion

Because the Phase Units output their respective photon components after Bob blocks for a run, the timing of which depends on the phase Bob wants to apply, there is a time-binning (a grouping of exit times into discrete bins) of the components from each Phase Unit correlated with the phase Bob applies in that unit. Bob can, on his side, compensate for the time-binning (given he knows the phase he is applying). Further, in order to locate the photon in time, Alice can detect the time of exit using a non-demolition single photon detector.

Alternatively, we could add a final pair of Phase Units with the value of k set to $3L' - \beta - \gamma - \delta$ (where L' is the value of L for each of the first three Phase Unit pairs, and β, γ and δ are their

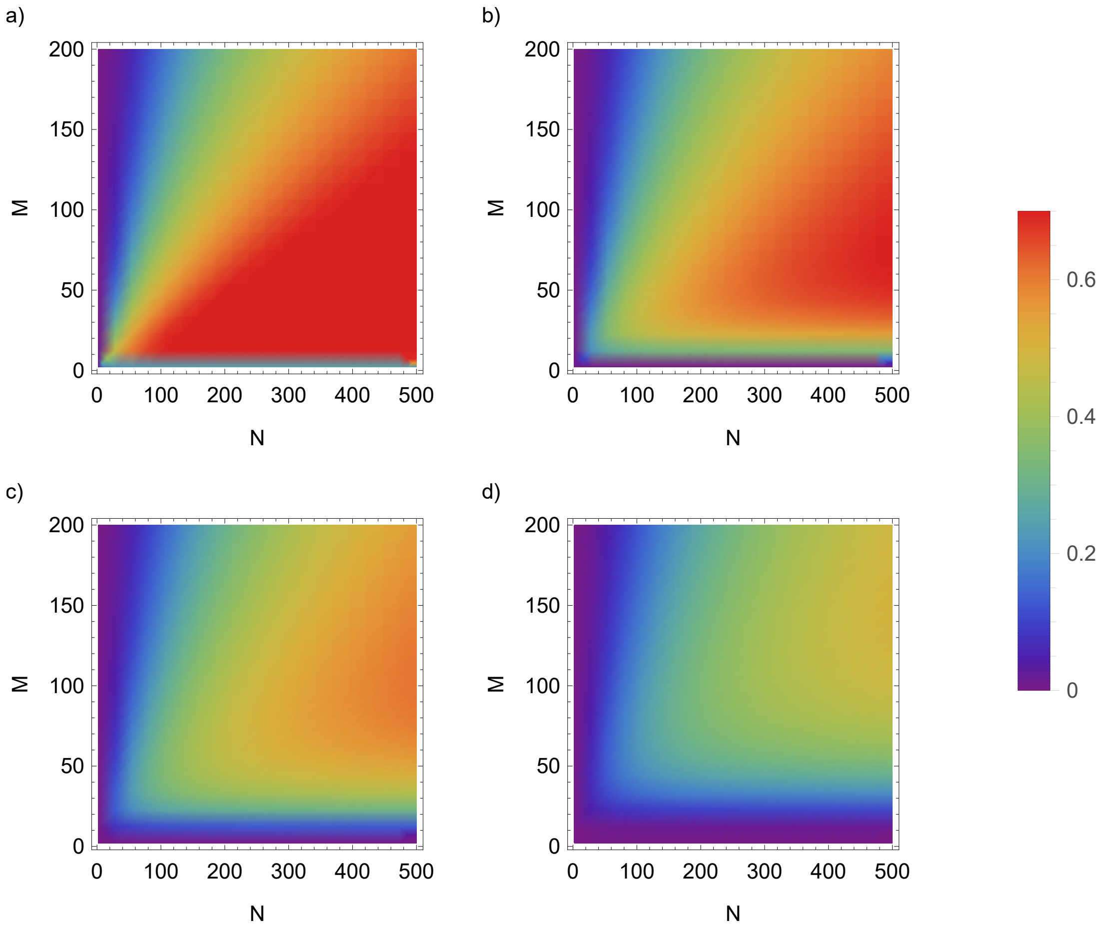

Figure 5.3: The survival probability of a photon going through a Phase Unit (Fig. 5.1) of given M (number of outer cycles) and N (number of inner cycles). This is shown for the unit imparting phase $ik\pi/L$, where k, the number of runs of the protocol before the photon is emitted from the unit, is 1 for (a), 5 for (b), 10 for (c) and 20 for (d). Note there is no dependence on L, the maximum number of runs.

respective k-values), but without phase plates (see Fig. 5.1). This means that while no phase is applied, a time delay is still added to the components, meaning the photon always exits the overall device at a time proportional to $3L$, rather than β, γ and/or δ as before. This makes the time of exit uncorrelated to Bob's unitary, which means Alice can know in advance the expected exit time of her photon from the protocol (without needing to perform a non-demolition measurement to find it).

When considering a finite number of outer and inner cycles, there is a nonzero probability of the photon not returning to Alice, which reduces the protocol's efficiency. The survival probability of a photon going through a Phase Unit is plotted in Fig. 5.3. The survival probability for the

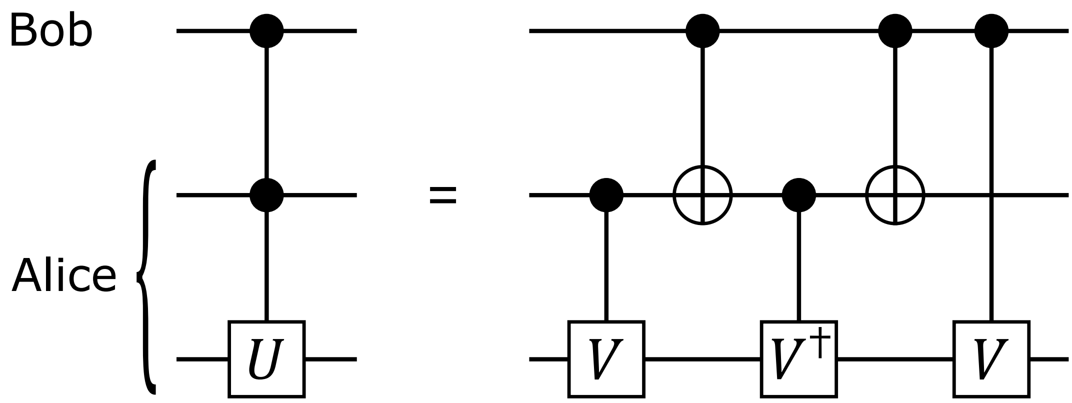

Figure 5.4: Quantum circuit diagram, showing how a 3-qubit gate applying a controlled-controlled unitary U can be constructed from two-qubit gates along with single-qubit gates, where U is some unitary transformation, and $V^2 = U$ [Bar+95]. Using our exchange-free single-qubit gate, a classical Bob can directly simulate the control action on Alice's photonic qubits. Since any quantum circuit can be constructed using 2-qubit gates along with single-qubit ones, our exchange-free single-qubit gate allows Bob in principle to directly program any quantum algorithm at Alice, without exchanging any photons.

overall protocol is the product of the survival probability for the three Phase Units:

$$P(Tot)_{Sv} = P(\beta)_{Sv} \cdot P(\gamma)_{Sv} \cdot P(\delta)_{Sv} \tag{5.5}$$

As expected, as $\{M, N\} \to \infty$, the survival probability goes to one.

Regardless, postselection renormalises Alice's output state such that if Alice receives an output photon, it will be in a pure state. Thus, for our set-up, given ideal optical components, the rotation enacted on Alice's qubit is always the rotation Bob has applied, not just for any L, but also for any N, M, and k.

Interestingly, a phase Unit, which outputs a photon into one of L different time bins depending on the number of runs Bob blocks, could be adapted to sending, exchange-free, a classical logical state of dimension d greater than two—a "dit", rather than a bit. We do this by removing the phase plate in the Phase Unit (see Fig. 5.1). Bob first doesn't block for k runs, then blocks for the remaining $L - k$, meaning the photon's output occurs in the k^{th} time-bin of L possible time-bins. This encodes a dit of dimension L into the photon, which Alice can read via non-demolition single photon detector.

We now show how our exchange-free protocol enabling arbitrary single-qubit operations, can in principle allow a classical Bob to directly enact any quantum algorithm he wishes on Alice's qubits, without exchanging any particles with her. This is based on the fact any quantum algorithm can be efficiently constructed from 2-qubit operations (such as CNOT) and single-qubit ones. Our protocol already enables exchange-free single-qubit operations, i.e. gates. Thus, if Bob can directly activate or not, a 2-qubit gate at Alice, exchange-free, then directly programming an entire quantum algorithm at Alice using these two building blocks becomes possible. The quantum network of Fig. 5.4 shows how a 3-qubit controlled-controlled gate, applying some unitary U to the target qubit at Alice, can be constructed from 2-qubit controlled gates [Bar+95]. (Some controlled-controlled gates can have more implementable circuits than the general one given here [NC10].) A classical Bob, at the top end of Fig. 5.4, uses our exchange-free single-qubit gate to simulate the control action on Alice's photonic qubits. For simulating the CNOT gate, he can choose to either apply the identity transformation, representing control-bit $|0\rangle$, or apply an X transformation, representing control-bit $|1\rangle$. For the controlled-V gate, he can choose to either apply the identity transformation, again representing control-bit $|0\rangle$, or apply a V transformation, representing control-bit $|1\rangle$. In this scenario, we envisage an optical programmable circuit, with exchange-free single-qubit gates acting on Alice's qubits that Bob can directly program, and 2-qubit gates acting on Alice's qubits that he can directly choose to activate, exchange-free.

While considering how Bob could prepare an arbitrary qubit at Alice exchange-free, we found a simpler protocol for Bob to prepare exchange-free a qubit with real, positive probability amplitudes.

Consider Fig. 5.1, without the phase module. Starting with Alice's H-polarised photon, instead of Bob blocking or not blocking every cycle, he instead doesn't block for the first $M-k$ outer cycles, then blocks for the rest. In order to eliminate the error resulting from a finite number of blocked inner cycles, Alice introduces loss, attenuating the outer arm of the interferometer on her side by a factor of $\cos(\pi/2N)^N$ for each outer cycle. This means, before the final PBS, the state is

$$|\Psi\rangle = \cos\left(\frac{\pi}{2N}\right)^{MN} \cos\left(\frac{\pi}{2M}\right)^{M-k} \left(\cos\left(\frac{k\pi}{2M}\right)|H\rangle + \sin\left(\frac{k\pi}{2M}\right)|V\rangle\right) \tag{5.6}$$

By postselecting on Alice's photon successfully exiting the protocol, she receives the state

$$|\Psi\rangle_{PS} = \cos\left(\frac{k\pi}{2M}\right)|H\rangle + \sin\left(\frac{k\pi}{2M}\right)|V\rangle \tag{5.7}$$

By choosing k, Bob directly applies a $\hat{\mathbf{R}}_y$ rotation to Alice's $|H\rangle$ input state. Now, in order to allow Bob to apply such a rotation to an arbitrary input polarisation state, Alice's photon is initially split into H and V-components using a PBS. The desired rotation is applied separately. In the case of the V-component, its polarisation is first flipped to H before the rotation is applied, followed by a phase flip and a polarisation flip upon exit. The separate components can then be combined using a 50:50 beamsplitter, with the correct state obtained 50% of the time. The advantage, however, is that, assuming perfect optical components and a large number of cycles, only two runs of the protocol are needed on average.

In summary, this Section presented a protocol allowing Bob to directly perform any computation on a remote Alice's qubit, without exchanging any photons between them. We use this to show how, in principle, Bob can directly enact any quantum algorithm at Alice, exchange-free.

5.3 Counterfactual Quantum Computation

Counterfactual communication was generalised to sending quantum information exchange-free for the first time in the Salih14 protocol [Sal16], also known as counterportation, proposing an exchange-free quantum CNOT gate as a new computing primitive. The exchange-free CNOT was later employed by Zaman et al to propose exchange-free Bell analysis, albeit with a 50% theoretical efficiency limit [ZJS18].

This Section instead proposes a controlled $\hat{R}_z$-rotation, based on this above-mentioned CNOT gate, which is universal and has no theoretical limit on efficiency. We then combine quantum teleportation with exchange-free computation at a distance to propose deterministic exchange-free teleportation. The core of this gate is set up by the entangling operation enabled by a one-dimensional cavity atom system. The ground state of an atom in a cavity can be put into a superposition of being on-resonant ('0') and reflecting, or off-resonant ('1') and transmitting, a photon [Hu+09; RR15]. We then construct a counterfactual way of probing whether Bob is blocking/not blocking (transmitting/reflecting) using the standard counterfactual communication-style protocol.

5.3.1 Protocol

We first go through the chained quantum Zeno effect (CQZE) unit, as given in Fig. 5.5. This is based on Salih's exchange-free CNOT gate, which has Bob enacting a superposition of blocking and not blocking his side of the communication channel [Sal16; Sal18c]. The switchable mirror, SM1, is first switched off to allow the photon into the outer interferometer, before being switched on again. The switchable polarisation rotator, SPR1, rotates the photon's polarisation from H to V, by a small angle $\frac{\pi}{2M}$:

$$\begin{aligned} |H\rangle &\to \cos\frac{\pi}{2M}|H\rangle + \sin\frac{\pi}{2M}|V\rangle \\ |V\rangle &\to \cos\frac{\pi}{2M}|V\rangle - \sin\frac{\pi}{2M}|H\rangle \end{aligned} \tag{5.8}$$

PBS2 passes the H-polarised component towards the mirror below it, while reflecting the small V-polarised component towards the inner interferometer. The switchable mirror, SM2, is then switched off to allow the V-polarised component into the inner interferometer, before being switched on again. The switchable polarisation rotator, SPR2, rotates the V-polarised component by a small angle $\frac{\pi}{2N}$:

$$|V\rangle \to \cos\frac{\pi}{2N}|V\rangle - \sin\frac{\pi}{2N}|H\rangle \tag{5.9}$$

PBS3 then reflects the V-polarised component up, towards a mirror while passing the H-polarised component towards Bob's trapped atom, which is in a superposition, $\alpha|0\rangle + \beta|1\rangle$, of reflecting back any photon, and transmitting it through to the loss detector. If the atom reflects (is in state $|0\rangle$), the H-polarised component survives—if the atom transmits ($|1\rangle$), the component

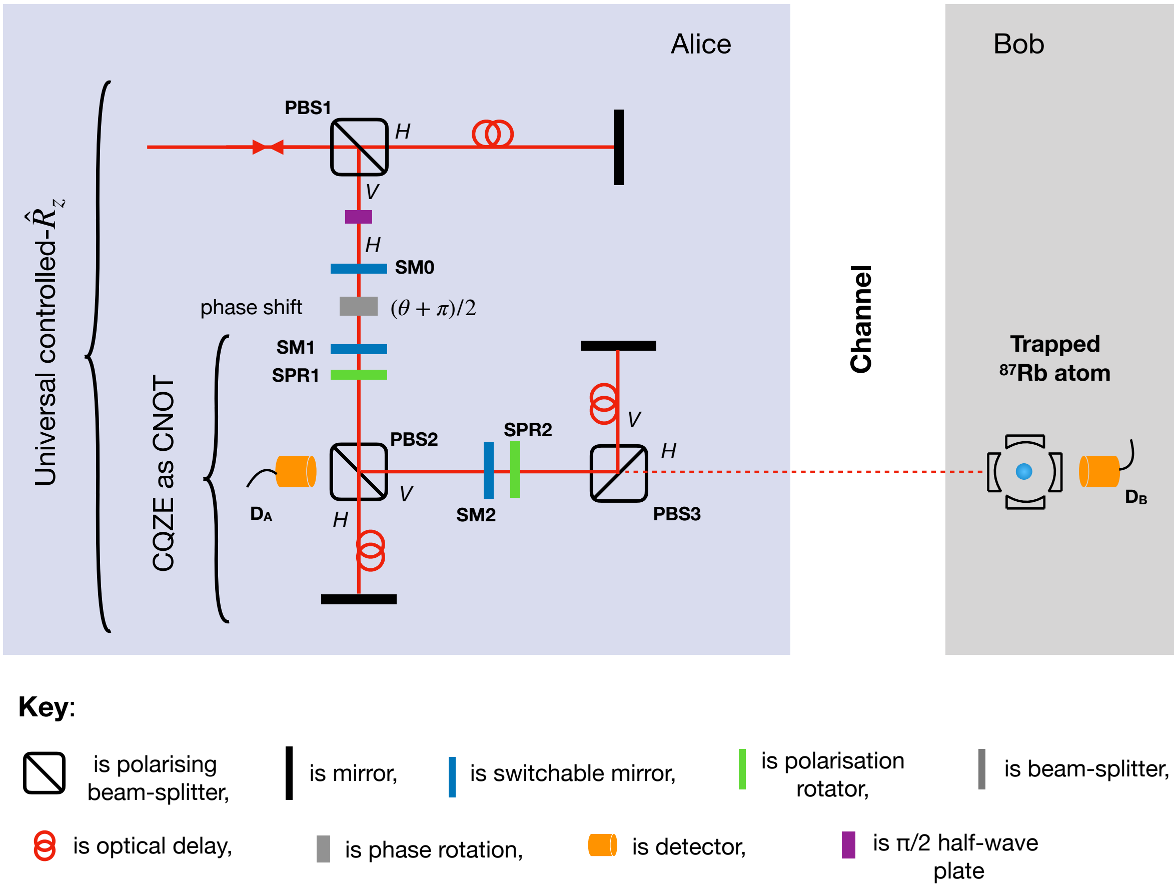

Figure 5.5: Our setup for an experimentally feasible, exchange-free controlled-$\hat{R}_z$, universal gate. This is based on Salih's exchange-free CNOT gate which has Bob enacting a superposition of blocking and not blocking the communication channel by means of a trapped atom [Sal16; Sal18c]. With the addition of a phase-shift plate applying a $(\theta+\pi)/2$ rotation, switchable mirror SM0, a $\pi/2$ HWP that flips polarisation, PBS1, and an optical delay loop, the chained quantum Zeno effect unit becomes the basis of our controlled phase-rotation universal gate, entangling the states of Alice's photonic qubit and Bob's trapped atom qubit.

is lost. This means the action of the first inner interferometer, assuming the photon is not lost to Bob's detector D_B, is

$$|V\rangle(\alpha|0\rangle+\beta|1\rangle)\rightarrow\alpha(\cos\frac{\pi}{2N}|V\rangle-\sin\frac{\pi}{2N}|H\rangle)|0\rangle+\beta\cos\frac{\pi}{2N}|V\rangle|1\rangle \tag{5.10}$$

After N such cycles the photonic superposition has now been brought back together by PBS3 towards SM2. Because the $|0\rangle$ term coherently rotates from $|V\rangle$ to $|H\rangle$, after N such cycles the state is,

$$|V\rangle(\alpha|0\rangle+\beta|1\rangle)\rightarrow\alpha|H\rangle|0\rangle+\beta\cos^N\frac{\pi}{2N}|V\rangle|1\rangle \tag{5.11}$$

The switchable mirror, SM2, is then switched off to let the photonic component inside the inner interferometer out. Since for large N, $\cos^N\frac{\pi}{2N}\rightarrow 1$, the implemented transformation becomes,

$$|V\rangle(\alpha|0\rangle+\beta|1\rangle)\rightarrow\alpha|H\rangle|0\rangle+\beta|V\rangle|1\rangle \tag{5.12}$$

Note that this state is entangled, but the $|H\rangle|0\rangle$ term has been to/from Bob, so this wouldn't be counterfactual. Consequently, we send this term to loss channel D_A via PBS2.

The first outer cycle, starting with the photon at SM1 (assuming the photon is neither lost to to Alice's detector D_A nor to Bob's D_B) implements first the SPR1, then N inner cycles, giving

$$\begin{aligned}
&|H\rangle(\alpha|0\rangle+\beta|1\rangle)\rightarrow\left(\cos\frac{\pi}{2M}|H\rangle+\sin\frac{\pi}{2M}|V\rangle\right)(\alpha|0\rangle+\beta|1\rangle)\\
&\rightarrow\cos\frac{\pi}{2M}|H\rangle\left(\alpha|0\rangle+\beta|1\rangle\right)+\sin\frac{\pi}{2M}|V\rangle\cos^N\frac{\pi}{2N}|1\rangle\\
&=\alpha\cos\frac{\pi}{2M}|H\rangle|0\rangle+\beta(\cos\frac{\pi}{2M}|H\rangle+\cos^N\frac{\pi}{2N}\sin\frac{\pi}{2M}|V\rangle)|1\rangle
\end{aligned} \tag{5.13}$$

as after the initial SPR1 rotation, the H-polarised component goes to the delay, while the V-polarised component goes into N inner cycles, and so evolves as per Eq. 5.11, with the $|H\rangle|0\rangle$ component generated by this sent to loss mode D_A.

This represents one outer cycle, containing N inner cycles. The photonic superposition has now been brought back together by PBS2 towards SM1. After M such cycles (setting $\cos^N\frac{\pi}{2N}\rightarrow 1$), the coherent rotation of the $|1\rangle$ part of the state from $|H\rangle$ to $|V\rangle$ means the protocol applies

$$|H\rangle(\alpha|0\rangle+\beta|1\rangle)\rightarrow\alpha\cos^M\frac{\pi}{2M}|H\rangle|0\rangle+\beta|V\rangle|1\rangle \tag{5.14}$$

Since for large M, $\cos^M\frac{\pi}{2M}$ approaches 1, Eq. 5.14 goes to,

$$|H\rangle(\alpha|0\rangle+\beta|1\rangle)\to\alpha|H\rangle|0\rangle+\beta|V\rangle|1\rangle \tag{5.15}$$

The switchable mirror SM1 is now switched off to let the photon out. Crucially, this last equation describes the action of a quantum CNOT gate with Bob's as the control qubit, acting on Alice's H-polarised photon.

Ignoring the large-M limit, let is illustrate this for $M=2$. Taking the state at the end of Eq. 5.13, setting $\cos^N\frac{\pi}{2N}\to 1$, and putting it through SPR1 again, we get

$$\begin{aligned}
&\cos\frac{\pi}{2M}|H\rangle\left(\alpha|0\rangle+\beta|1\rangle\right)+\sin\frac{\pi}{2M}|V\rangle|1\rangle\to\\
&\cos\frac{\pi}{2M}\left(\cos\frac{\pi}{2M}|H\rangle+\sin\frac{\pi}{2M}|V\rangle\right)\left(\alpha|0\rangle+\beta|1\rangle\right)+\sin\frac{\pi}{2M}\left(\cos\frac{\pi}{2M}|V\rangle-\sin\frac{\pi}{2M}|H\rangle\right)|1\rangle\\
&=_{\text{For }M=2}\frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}|H\rangle+\frac{1}{\sqrt{2}}|V\rangle\right)\left(\alpha|0\rangle+\beta|1\rangle\right)+\frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}|V\rangle-\frac{1}{\sqrt{2}}|H\rangle\right)|1\rangle\\
&=\frac{|H\rangle}{2}\alpha|0\rangle+|V\rangle\left(\frac{\alpha}{2}|0\rangle+\beta|1\rangle\right)
\end{aligned} \tag{5.16}$$

If we now send this through N more inner interferometers, due to the $|H\rangle$ component being kept in the delay below PBS2, and the $|V\rangle|0\rangle$ component coherently rotating to $|H\rangle|0\rangle$ and so being lost to loss channel D_A, we get the transformation for the entire two outer-cycle ($M=2$) protocol as

$$|H\rangle(\alpha|0\rangle+\beta|1\rangle)\to\frac{|H\rangle}{2}\alpha|0\rangle+|V\rangle\beta|1\rangle \tag{5.17}$$

We now explain the rest of the setup, which uses the CQZE unit to implement a universal, general-input controlled-$\hat{R}_z$ rotation gate (Fig. 5.5). We begin with a superposition state at Alice, $a|V\rangle+b|H\rangle$. This is split at PBS1, with the H-polarised component going into an optical loop, and the V-polarised component going through a $\pi/2$ HWP flipping its polarisation to H, before being admitted into the CQZE unit by turning off the switchable mirror SM0, before turning it on again. Upon exiting the CQZE unit, it is reflected back by SM0, having a phase of $\theta+\pi$ if V-polarised (0 if H) applied to it by the phase shifter, before going through another run of the CQZE unit. This always produces an H-polarised state, as noted in [Li+19]. Note that the π term in the phase shifter is a correction term. The photonic component now exits through SM0, which is switched off, before being flipping back to V-polarisation at the $\pi/2$ HWP, having acquired a θ phase shift. It then recombines with the H-polarised component at PBS1.

Given the initial state of the overall system is

$$(a|V\rangle+b|H\rangle)\otimes(\alpha|0\rangle+\beta|1\rangle) \tag{5.18}$$

and that only the initially V-polarised component "interacts" with the trapped atom, we get the final state

$$a|V\rangle(\alpha|0\rangle+\beta e^{i\theta}|1\rangle)+b|H\rangle(\alpha|0\rangle+\beta|1\rangle) \tag{5.19}$$

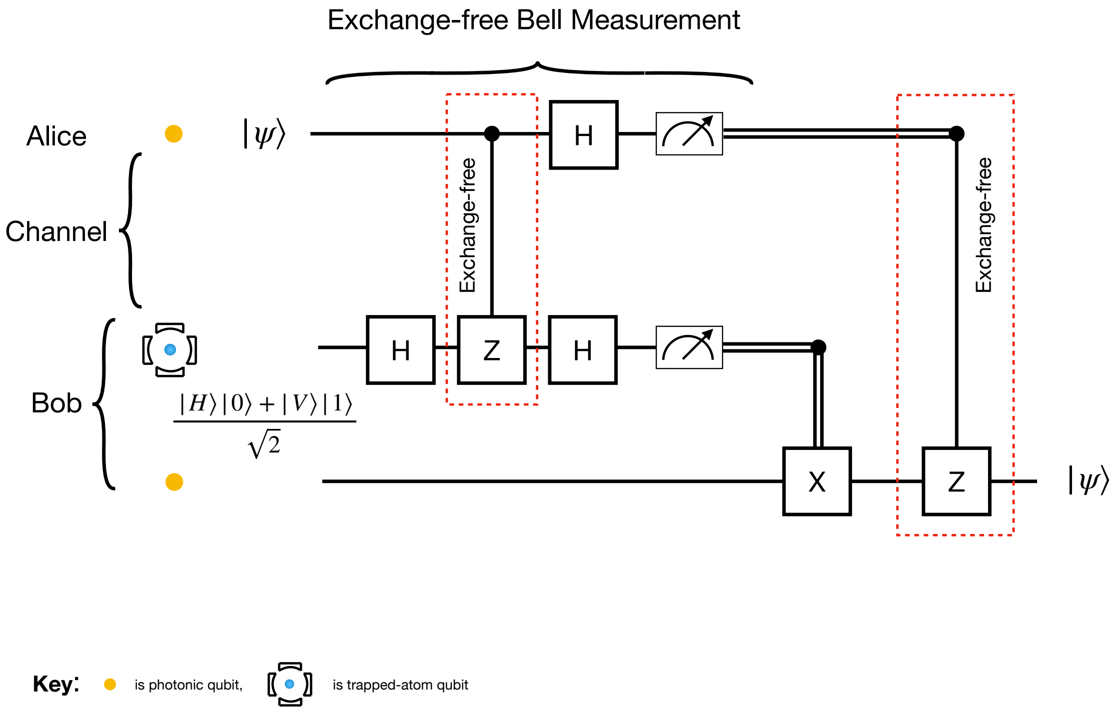

Figure 5.6: Our protocol for exchange-free teleportation. In this protocol, Alice has a photon-polarisation qubit, and Bob has a maximally-entangled pair of qubits, one implemented as a trapped-atom enacting a superposition of blocking and not blocking the communication channel, and the other as photon polarisation. Alice's qubit begins in the state to be teleported, $|\psi\rangle$. Importantly, complete Bell detection takes place without Alice and Bob exchanging any particles, and instead of classical communication, Alice directly applies a controlled-Z (phase flip) operation on Bob's photonic qubit. The two exchange-free Controlled-Z gates, marked by dashed red-boxes, are instances of the set-up of Fig. 5.5.

This is an entangled state between Alice's polarisation qubit and Bob's trapped ion qubit: a controlled-phase rotation, with Alice's as the control qubit and Bob's as the target. Due to the symmetry of control and target qubits for this type of rotation, it can also be factorised as

$$\alpha |0\rangle (a|V\rangle + b|H\rangle) + \beta |1\rangle (ae^{i\theta}|V\rangle + b|H\rangle) \tag{5.20}$$

the same controlled-phase rotation expressed differently, now with Bob's as the control qubit and Alice's as the target. Taking the special case when $\theta = \pi$, we get a controlled-Z gate.

On universality, our exchange-free controlled-$\hat{R}_z$, as a two-qubit gate, allows efficient implementation of any quantum circuit when combined with local operations. But there's another sense in which it is universal. As explained later, this gate can be operated differently, allowing one party with classical action to enact any desired operation on a second party's remote photonic qubit, exchange-free. This classical action can even control a two-qubit gate at the second party, as shown in [Sal+21]. Our controlled-$\hat{R}_z$ gate can therefore be thought of as a universal set in its own right.

Bob needs a way to implement a superposition of reflecting, bit "0", and blocking, bit "1", Alice's photon. There are many ways to go about this; however, recent breakthroughs in trapped atoms inside optical cavities [RR15], such as the demonstration of light-matter quantum logic gates [Tie+14; Rei+14], make trapped atoms an obvious choice.

Bob's qubit is a single ^{87}Rb atom trapped inside a high-finesse optical resonator by a three-dimensional optical lattice [Rei+14; Rei14]. Depending on which of its two internal ground states the ^{87}Rb atom is in, a photon impinging on the cavity in Fig. 5.5 will either be reflected as a result of strong coupling, or otherwise enter the cavity on its way towards detector D_B. For this, it needs to have mirror reflectivities such that a photon entering the cavity exists towards detector D_B, similar to [Müc+10]. By placing the ^{87}Rb in a superposition of its two ground states, by means of Raman transitions applied through a pair of Raman lasers, Bob implements the desired superposition of reflecting Alice's photon back and blocking it. Note that coherence time for such a system is of $\mathcal{O}(10^{-4}s)$ [Rei14], with longer times possible. Therefore, if the protocol is completed within $\mathcal{O}(10^{-5}s)$, lower-bounded by the $\mathcal{O}(10^{-9}s)$ switching speed of switchable components, then decoherence effects can be ignored.

(Experimental tricks can ensure the correct number of cycles without having to use switchable optical elements, as in Cao et al's experimental implementation of Salih et al's 2013 protocol [Cao+17].)

We now move to an exchange-free implementation of teleportation. This is based on the quantum teleportation first devised by Bennett et al [Ben+93]. In Bennett et al's protocol, any arbitrary quantum state can be transferred from one party to another, so long as the two parties share an entangled state, and can send classical information between them. However, in our protocol, teleportation is recast such that there is no need for previously-shared entanglement between Alice and Bob, nor classical communication between them. Our teleportation scheme is shown in Fig. 5.6.

In this protocol, we have a photon-polarisation qubit at Alice, and an entangled pair of qubits, one trapped-atom and the other photon-polarisation, at Bob. Alice's qubit is instantiated in the state to be teleported, e.g. by a third party, while Bob's two modes are in the maximally entangled state $|H\rangle|0\rangle + |V\rangle|1\rangle/\sqrt{2}$ (which can be created by Bob sending an H-polarised photon into the cavity, when it is in the superposition $(|0\rangle + |1\rangle)/\sqrt{2}$, and then recombining the photon's components from the transmitted and reflected outputs at a PBS, after applying a $H \to V$ rotation on the path from the transmitted side).

To enact teleportation, Bob first applies a Hadamard gate to his trapped-atom qubit, before Alice applies an exchange-free controlled-Z gate, with her photonic qubit as the control and and Bob's trapped-atom qubit as the target. Bob and Alice then apply Hadamard gates onto their respective qubits, before measuring the states in the computational basis for Bob, and in the H/V basis for Alice, together performing a complete Bell measurement. Bob then either flips or doesn't flip the polarisation of his photonic qubit based on the classical measurement outcome of his trapped-atom qubit. Alice then, based on the classical measurement outcome of her qubit, either performs an exchange-free controlled-Z on Bob's photonic qubit with her control set to $|1\rangle$ by blocking both runs, or else sets her control to $|0\rangle$ by not blocking both runs. These last two steps by Bob and Alice respectively act as the feed-forward step of teleportation (which next-generation trapped atoms are expected to allow) leaving Bob's photonic qubit in the state of Alice's original qubit.

5.3.2 Discussion

Our exchange-free protocol bears all the hallmarks of teleportation as given by Pirandola et al [Pir+15]. Alice's input state is unknown to her, and can be provided by a third party who also verifies the teleported state at Bob. The protocol allows complete Bell detection, which in our case is jointly carried out by Alice and Bob exchange-free. The protocol allows the possibility of real-time correction on Bob's photonic qubit. Lastly, even for the smallest number of cycles, achievable fidelity for our protocol exceeds the 2/3 limit of "classical teleportation", which comes from the no-cloning theorem [WZ82]. Fig. 5.7 gives the average fidelity $F(\theta,\phi)$, where

$$F(\theta,\phi) = \langle\Psi_{in}|\Psi_{out}\rangle\langle\Psi_{out}|\Psi_{in}\rangle \tag{5.21}$$

and Ψ_{in} and Ψ_{out} are the input and output states of the protocol, θ and ϕ parameterise the input state's Bloch sphere (azimuthal and radial angle respectively), and the average fidelity is $F(\theta,\phi)$ averaged over θ and ϕ [OLL02]. The average efficiency of the protocol is 30% for a number of cycles M equal to 10 and N equal to 100, but improves for larger numbers of cycles.

Since quantum telecloning combines approximate cloning with teleportation to transport multiple approximate copies of a states, one would think that our exchange-free teleportation protocol might allow telecloning to be carried out exchange-free. In fact the telecloning scheme of Murao et al [Mur+99], which employs a Bell measurement, along with local operations at the

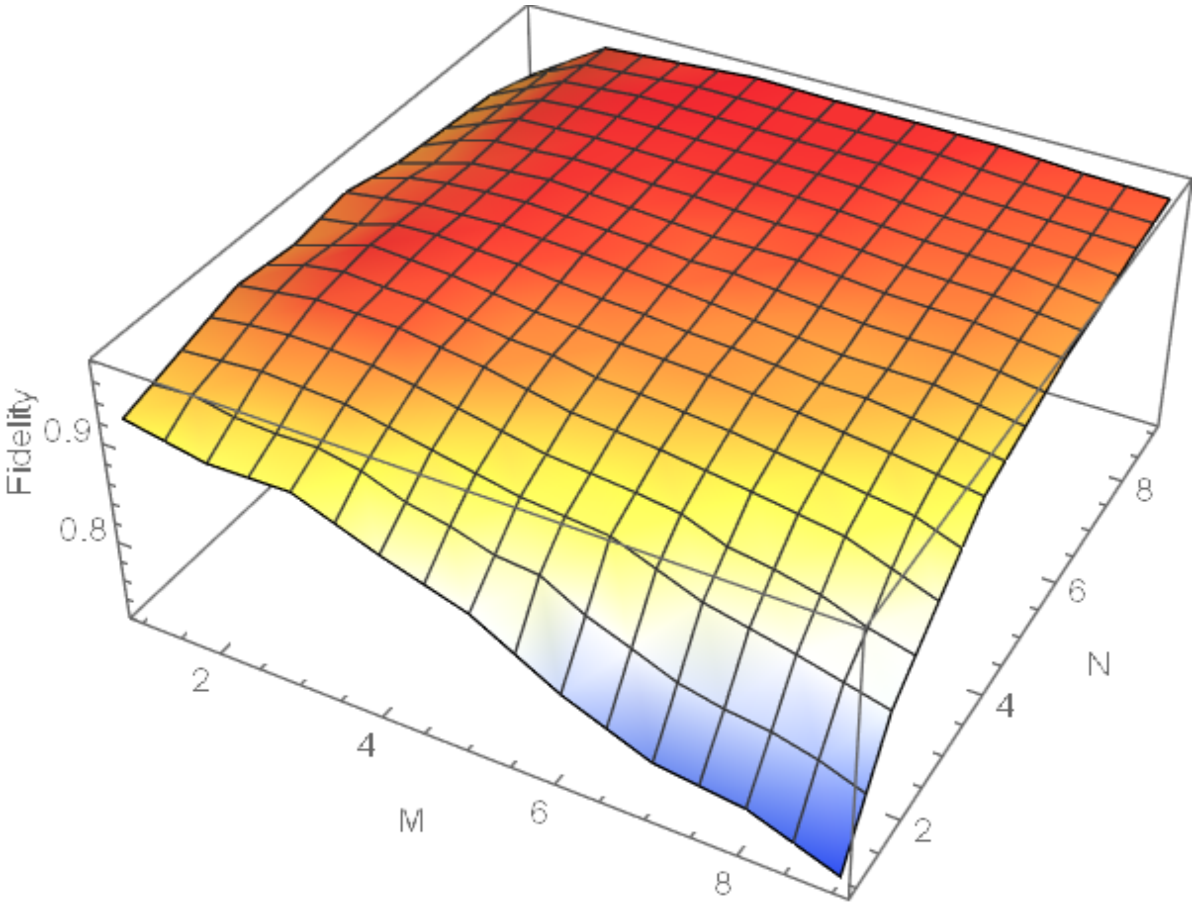

Figure 5.7: Average fidelity of our exchange-free teleportation protocol as a function of the number of outer and inner cycles, M and N. This is for an imperfect trapped-atom at Bob that fails to reflect an incident photon 34% of the time when it should reflect, and fails to block the photon 8% of the time when it should block. Fidelity is averaged over 100 points evenly distributed over possible states Alice could send.

receiver based on the Bell detections, can be made exchange-free in a similar manner to how we made teleportation exchange-free. Their scheme starts with an already prepared multipartite entangled state [Mur+99], which for our purposes we take to be located at Bob, with one of the entangled qubits in the form of say a trapped-atom, and the output qubits where the approximate copies appear, all photonic. Alice and Bob jointly perform an exchange-free Bell measurement between Alice's photonic input qubit, and Bob's trapped-atom 'port' qubit, as we show in Fig. 5.8. Based on the classical outcomes of the Bell measurement, Alice applies suitable exchange-free controlled-rotations (Pauli operations) to recover the approximate copies at Bob. The fidelity of these copies is limited by the no-cloning theorem to

$$\gamma = \frac{2q+1}{3q} \tag{5.22}$$

where q is the number of approximate copies of our state that we want to send. For the protocol,

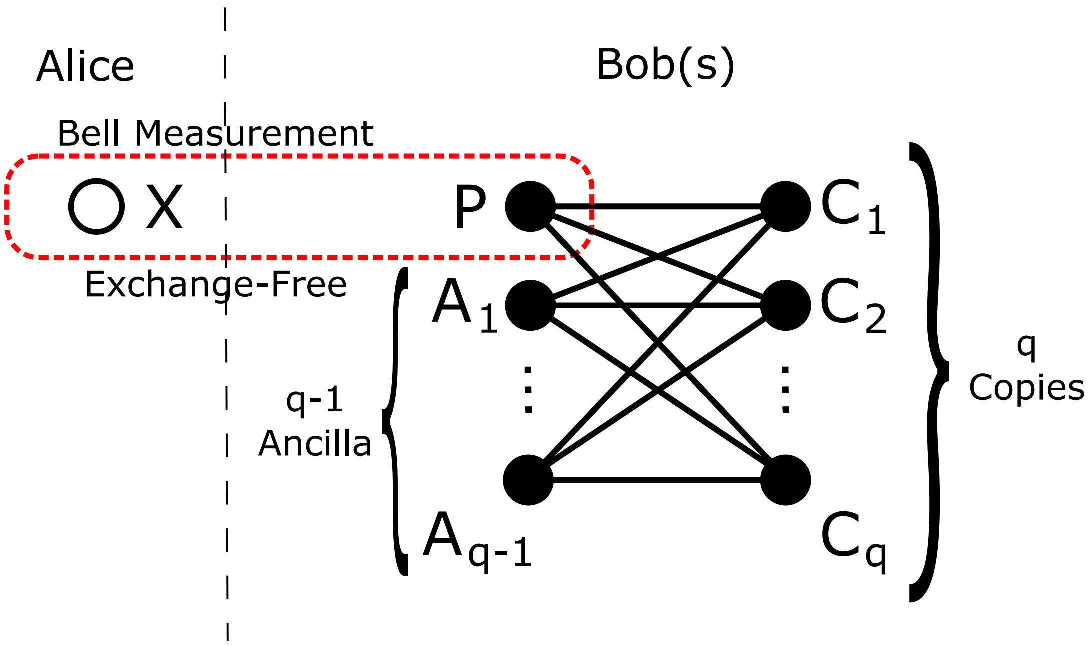

Figure 5.8: An entanglement diagram for exchange-free telecloning. The diagram shows the initial entangled state between port-qubit P, copy-qubits C_q, and ancilla-qubits A_{q-1}, all at Bob(s). Also shown is the Bell Measurement done on Bob's port qubit P and Alice's initial state qubit X. The thick black lines mark entanglement, while the red dashed box indicates an exchange-free Bell measurement. This forces the system into one of four states. Alice applies suitable exchange-free controlled-rotations (Pauli operations) to recover the approximate copies at Bob.

when the trapped-atom interaction is ideal, we always reach this limit of fidelity (5/6 for two copies) even for the smallest number of cycles. In Fig. 5.9, we give the fidelity for an imperfect trapped-atom at Bob that fails to reflect an incident photon 34% of the time when it should reflect, and fails to block the photon 8% of the time when it should block, for different values of M and N. Average efficiency is 14% for M equal to 10 and N equal to 100, but improves for larger numbers of cycles.

Above (and in [Sal+21]), we gave a protocol that allows Bob to implement any phase on Alice's qubit, exchange-free [Sal+21]. This then forms the basis of a device that we called a phase unit, allowing Bob to apply any arbitrary single-qubit unitary to the qubit, exchange-free. However, an issue that phase unit displayed was that the time the photon exited the device was correlated with the phase applied by Bob. While we provided a way for Bob to undo this time-binning after the fact, it is generally desirable to remove it altogether.

By adapting the controlled phase-rotation above, a phase unit can be constructed that doesn't exhibit this time-binning. We use the set-up in Fig. 5.5, but instead with a classical Bob either blocking or not-blocking, and with SM0 keeping Alice's photon in the device for 2L runs (rather than 2). Bob sets θ to π/L, blocking for $2k$ of the runs and not blocking for $2(L-k)$, in units of 2 runs where he either blocks for both or does not block for both. This allows Bob to set a

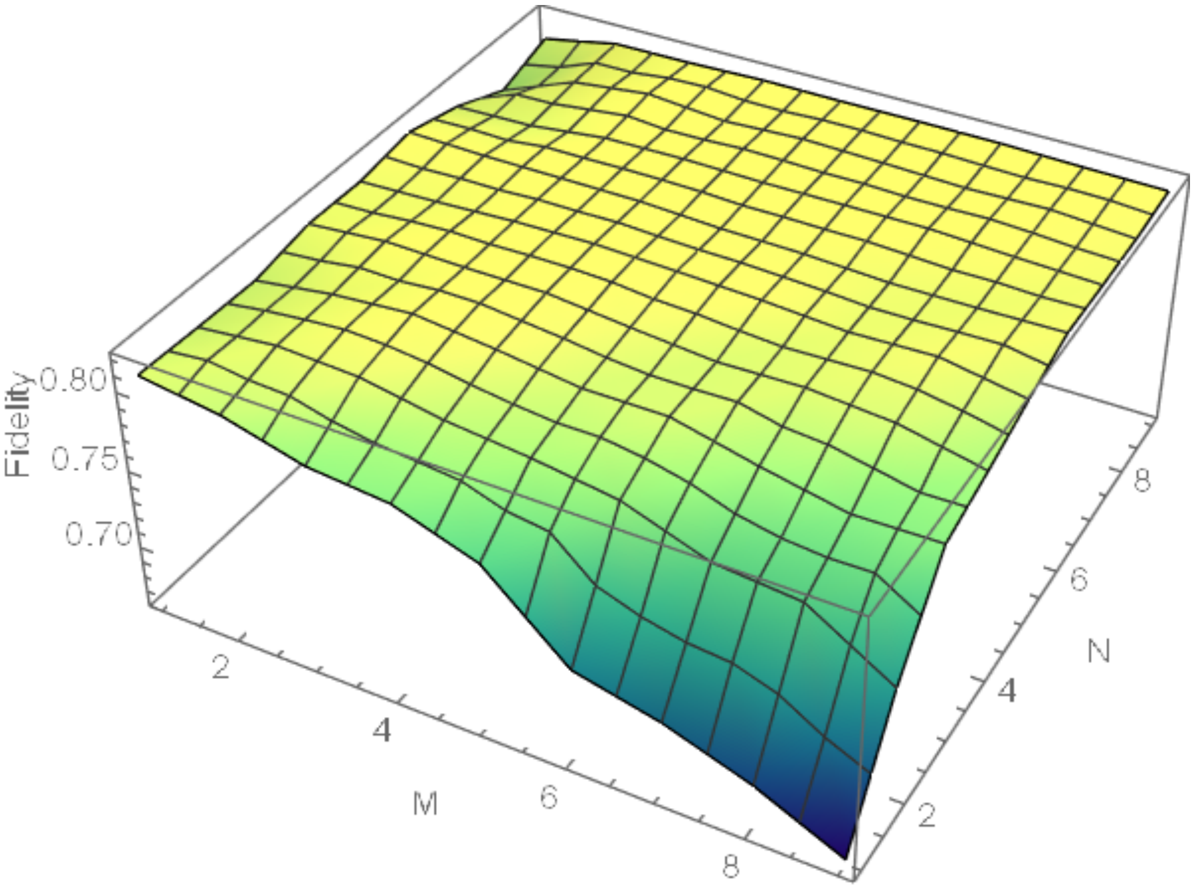

Figure 5.9: Average fidelity for exchange-free telecloning using the exchange-free controlled-Z gate described above, plotted for different numbers of outer (M) and inner (N) cycles. Fidelity is calculated for an imperfect trapped-atom at Bob that fails to reflect an incident photon 34% of the time when it should reflect, and fails to block the photon 8% of the time when it should block. Fidelity is averaged over 100 points evenly distributed on the Bloch sphere of possible states Alice could send.

phase on Alice's photon of $2\pi k/L$. We place three of these devices in series, interspersed with a $-\pi/4$-aligned Quarter Wave Plate, $\hat{\mathbf{U}}_{QWP}\hat{\mathbf{R}}_x(-\pi/2)$, and its adjoint, $\hat{\mathbf{U}}^{\dagger}_{QWP}\hat{\mathbf{R}}_x(\pi/2)$, to create a chained-$\hat{R}_z\hat{R}_x\hat{R}_z$a set of rotations, into which any single qubit unitary can be decomposed. Bob can thus apply any arbitrary single-qubit unitary to Alice's qubit—both exchange-free, and without time-binning. This also, as we show in [Sal+21], allows us to classically control of a universal two-qubit gate, which enables Bob to directly enact in principle any desired algorithm on a remote Alice's programmable quantum circuit.

An interesting modification of Salih et al's 2013 protocol was recently proposed by Aharonov and Vaidman, satisfying their criterion, based on weak trace, for exchange-free communication [AV19]. Following Salih's 2018 paper on counterportation [Sal18c], we now show how to implement it in our protocol. In the CQZE module, after applying SPR2 inside the inner interferometer for the Nth cycle, Alice now makes a measurement by blocking the entrance to channel leading to Bob. (She may alternatively flip the polarisation and use a PBS to direct the photonic component away from Bob.) Instead of switching SM2 off, it is kept turned on for a duration corresponding to N more inner cycles, after which SM2 is switched off as before. One has to compensate for the added time by means of optical delays. The idea here is that, for the case of Bob not blocking, any lingering V component inside the inner interferometer after N inner cycles (because of weak measurement or otherwise) will be rotated towards H over the extra N inner cycles. This has the effect that, at least as a first order approximation, any weak trace in the channel leading to Bob is made negligibly small.

While alternative proposals have been given for counterfactual communication by Arvidsson-Shukur et al [AB16; Cal+19], the protocols' definition of counterfactuality have been the subject of debate [Vai19; HLR21; WCV21]. Their adoption of Fisher Information as a tool for analysing counterfactuality is interesting nonetheless.

As Vaidman points out in [Vai16b], Salih's 2014 protocol—also known as counterportation [Sal18c]—achieves the same end goal of Bennett et al.'s (1993) teleportation [Ben+93]: the disembodied transport of an unknown quantum state, albeit over a large number of protocol cycles. It is an entirely different protocol though, as can be seen from their respective quantum-circuit diagrams. Our current protocol, by contrast to counterportation, is directly based on the 1993 teleportation protocol, but implemented using our universal, newly proposed counterfactual Z-gate, with the aim of exploring the foundations of this most central of quantum information protocols.

In this Section, we have shown how the chained quantum Zeno effect can be employed to construct an experimentally feasible, exchange-free controlled-$\hat{R}_z$ operation, which is not only a universal gate, but can be considered a universal set. This allowed us to propose a protocol for deterministic teleportation of an unknown quantum state between Alice and Bob, without exchanging particles.

5.4 Chapter Conclusion

This Chapter presented a protocol allowing Bob to directly perform any computation on a remote Alice's polarisation qubit, without exchanging any photons between them. We use this to show how, in principle, Bob can directly enact any quantum algorithm at Alice, exchange-free. We then showed how the chained quantum Zeno effect can be employed to construct an experimentally feasible, exchange-free controlled-$\hat{R}_z$ operation, which is not only a universal gate, but can be considered a universal set. This allowed us to propose a protocol for deterministic teleportation of an unknown quantum state between Alice and Bob, without exchanging particles.

This is the first example in this thesis of new research where a foundational idea (here, counterfactual communication) is used for the development of a novel form of quantum technology.

This illustrates the interplay between quantum foundations and quantum technologies, insofar as foundational ideas can be leveraged for technological gain.

Chapter 6

Counterfactual Ghost Imaging

Declaration of contribution: This chapter is based on the published article *Counterfactual Ghost Imaging* [HR21b], which was conceived, simulated, and written by myself, supervised and edited by Prof John Rarity.

6.1 Chapter Introduction

Ghost Imaging exploits the the position-momentum entanglement between correlated photon pairs to derive image information. When one photon of the pair travels via the object and is focused into a bucket detector, an image can still be formed using coincident detection of the partner photon in a high-resolution pixel detector [Mal+10; SB12; Asp+13; PB17] (as shown in Fig.6.1).

Pittman et al originally conceived ghost imaging to illustrate the power of quantum correlations [Pit+95]. However, it has since been shown thermal light can be used to replicate this effect classically [BBB02; CZ05; Fer+05; DDX13]. Doing so however, rather than using an entangled photon pair, removes some benefits of the original protocol. Entangled photon pairs created by Spontaneous Parametric Down-Conversion (SPDC) are both position-correlated (so can image in the near-field), and momentum anti-correlated, meaning they are position anti-correlated in the far-field, and can image there too. Classically, they can only be (anti-)correlated for one of the two conjugate variables, so can only image in one of these regimes [PB17]. Therefore, despite the comparative ease of using thermal light, entangled photon pairs still provide an advantage.

Despite claims in [Asp+15], having short-wavelength photons going to the high-resolution detector and longer wavelengths to the object does not allow an increase in imaging resolution above and beyond standard diffraction limits [Mor+18]. Regardless, the ability to ghost image

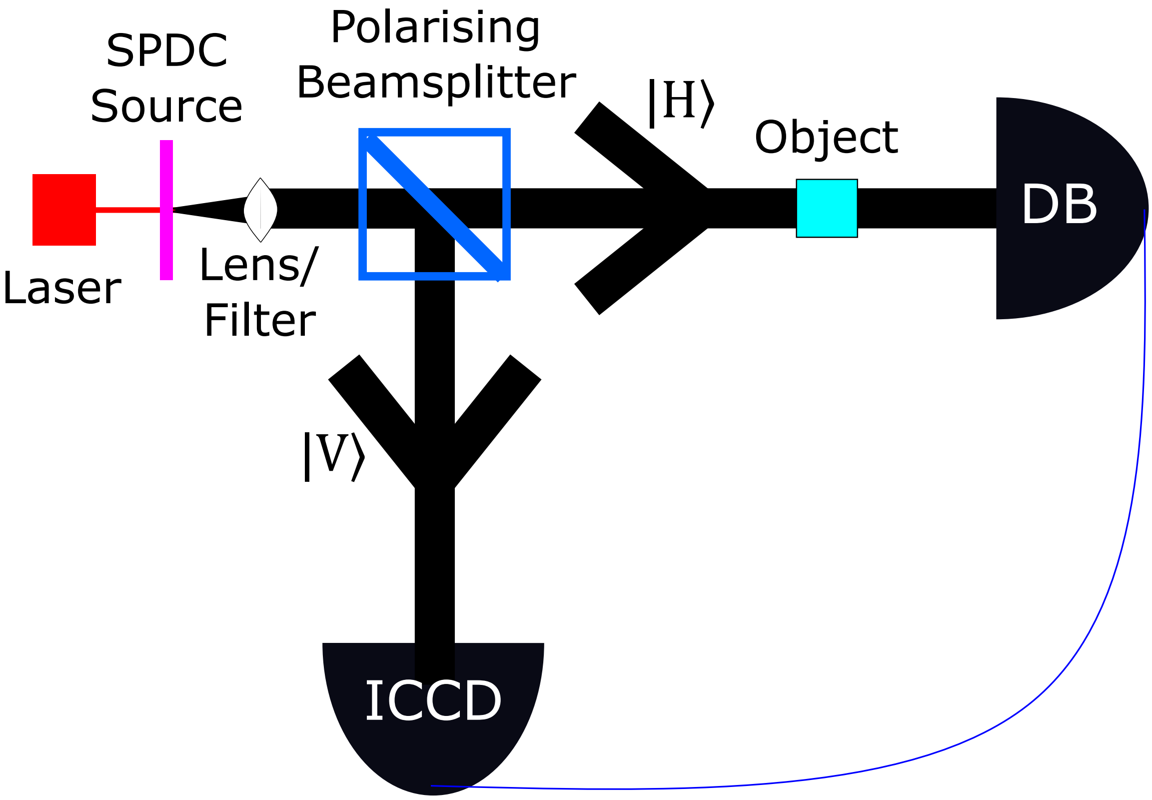

Figure 6.1: A standard set-up for ghost imaging. Position-and-momentum-entangled photons are generated in pairs at the SPDC source from a laser beam, with conjugate polarisations (i.e. always in pairs of horizontal and vertical polarisation). A PBS sends the vertically polarised photon to an Intensified Charge-Coupled Device (ICCD) camera (which records in high resolution its arrival position), and sends the horizontally-polarised photon via a sample to be imaged, to a bucket detector (DB). By recording coincidences between the detections at the ICCD and the bucket detector, the sample can be "ghost imaged".

while reducing the energy going to the object under investigation, to reduce potential damage, could be massively beneficial [Gen16; Gil+19].

Zhang et al combined Ghost Imaging with the Elitzur-Vaidman Bomb Tester [EV93] to create a form of Ghost Imaging where there is a chance that information is still received about the imaged object without any photons being absorbed by it [Zha+19]. However, the Elitzur-Vaidman Bomb Tester is not always necessarily counterfactual—there is a (reasonably high) chance the photon can go via the object being investigated [HLR21].

By replacing the Elitzur-Vaidman object-detection system in Zhang's protocol with Salih et al's method for counterfactual communication [Sal+13b], we create a protocol for ghost imaging that is always counterfactual: whenever information is received about the imaged object, no photons have interacted with that object. Further, even in runs where no information travels, far fewer photons go to the object than in either the standard ghost imaging or in the interaction-free ghost imaging case. This reduces the energy absorbed by the object, and so potentially damage done to that object by the imaging process. This shows the benefit of leveraging quantum foundational ideas—here, counterfactual communication and entanglement—for quantum technological gain, illustrating the interplay between these two fields.

6.2 Protocol

For Counterfactual Ghost Imaging, we make use of a version of Salih et al's counterfactual communication protocol as a primitive. See Chapter 2.4.3 and Fig. 3.1 and 4.2 for the details of this protocol. However, we use a common-path interferometer form of this protocol—PBSs are replaced with polarisation separators, which subtly divert horizontally-polarised light, and directly transmit vertically-polarised light. We also use switchable mirrors to allow the repeated use of a single inner/outer interferometer, rather than a spatially-extended chain of interferometers as in the Michelson set-up described previously.

We now describe how to adapt Salih et al's protocol to use it for Counterfactual Ghost Imaging (as shown in Fig.6.2). A pair of conjugately-polarised photons, entangled in position and momentum, are split at a PBS. The vertically-polarised photon is sent through an image-preserving optical delay line to an Intensified Charge-Coupled Device (ICCD) camera (both of which have previously been used to allow multi-mode ghost imaging, rather than single-mode raster-scanning [Asp+13; Asp+15; Mor+15]). The horizontally-polarised photon is sent through one run of Salih et al's protocol, where the object to be imaged is put in place of Bob's blocker. If the object doesn't block the path, the photon, in that spatial mode, remains horizontally-polarised, and so goes to $D0$, leading to a coincidence measurement between $D0$ and that pixel of the ICCD camera; however, if the object does block the path, the photon becomes V-polarised and so goes to $D1$, causing a coincidence measurement between $D1$ and that pixel of the ICCD camera.

Note, because arrival in $D0$ ($D1$) is correlated far more closely in Salih et al's protocol with the object not blocking (blocking) the path (see Fig.6.3b) than in the Elitzur-Vaidman Bomb Tester (and so Zhang et al's protocol [Zha+19]), we only **need** to resolve coincidences between $D0$ and the ICCD camera (unlike Zhang et al, who need to form two images—one based on $D0$-ICCD coincidences, and the other on $D1$-ICCD coincidences—and subtract one from the other, due to the interference patterns produced by the Elitzur-Vaidman Bomb Tester [EV93]). However, by resolving both $D0$-ICCD and $D1$-ICCD coincidences, and subtracting one from the other, we can image with high accuracy even for low N—so we do this.

A similar protocol could be constructed by replacing Salih et al's protocol in the imaging set-up with either Aharonov and Vaidman's modified protocol [AV19], or Vaidman's later adaptation [Vai19]—however, such a protocol would not be counterfactual by the consistent histories criterion (as shown by Salih [Sal18c; Sal+22]).

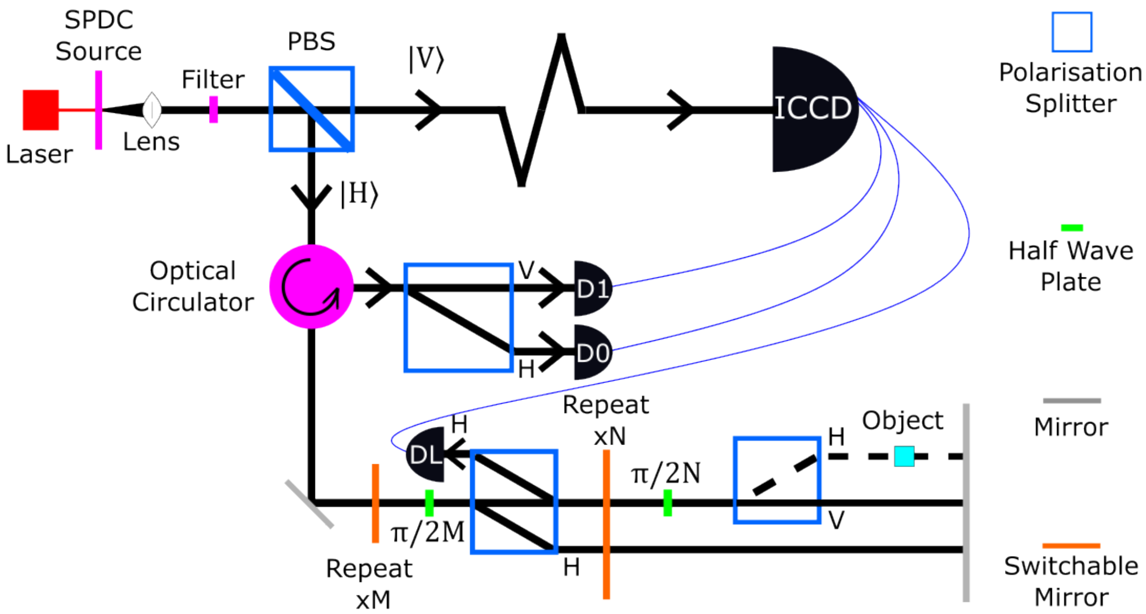

Figure 6.2: Our Counterfactual Ghost Imaging Protocol. This is based on the combination of standard Ghost Imaging (Fig.6.1) and a common-path interferometer version of Salih et al's counterfactual communication protocol [Sal+13b]. We create a pair of position-and-momentum-entangled photons, one horizontally polarised and one vertically polarised, by passing a pulsed pump laser through a SPDC crystal, before collimating the beam, and filtering out the pump. The photon pair is split at a PBS, with the V-polarised photon going through a long optical delay to an ICCD camera, and the H-polarised photon going through a run of Salih et al's protocol, adapted so the object to be investigated is put in place of Bob's blocker. The switchable mirrors allow the photon to cycle the correct number of times: the first for M outer cycles; and the second for N inner cycles per outer cycle. The polarisation separators subtly divert horizontally-polarised light, and directly transmit vertically-polarised light. The HWPs are tuned to implement a $\hat{\mathbf{R}}_y(\theta)$ polarisation-mode rotation with θ of $\pi/2M$ and $\pi/2N$ respectively. The detector D_L acts as our loss channels (which we postselect against). After M outer cycles, the switchable mirror sends the photon to the optical circulator, which sends it to the PBS. The path not being blocked by the object leads the photon to remain H-polarised, and so go to $D0$, leading to a coincidence measurement between that and the ICCD camera; however, the path being blocked leads to the photon becoming V-polarised and so going to $D1$, so coincidence measurement between that and the ICCD camera. The use of multi-mode interferometers and (position-momentum) correlations between the entangled photons enables multi-mode ghost imaging in this counterfactual set-up. Note, the polarisation separators ensure a common path length for both $H-$ and V-polarised components, while generating beam separations of the several millimetres. An optimisation we mention in the discussion has photons going to DL can trigger a coincidence measurement with the ICCD camera, treated as if it was a detection at $D0$, which does not affect the chance of photons interacting with the object (photons only go to DL if the object doesn't block the channel), and allows us to lower the number of outer cycles to the minimum required (2) with no increase in loss.

For the original Ghost Imaging protocol, photon pairs are generated by SPDC [Pit+95]. This makes use of second-order nonlinearities in an optical medium to generate conjugately-polarised photon pairs entangled in position and with frequencies that sum to the frequency of an input pump laser. By using a low-pass filters, the photons can be split off from the pump laser, and then split from one another using a PBS, sending one to the high-resolution detector, and the other to the object and bucket detector. We propose using the same source for our protocol.

Also note, for the V-polarised photon going to the ICCD camera, rather than using an optical delay, the ICCD camera can detect the photon earlier, but record the time of arrival as well as the position, which can be used with post-processing to determine effective coincidences with the bucket detectors, avoiding issues with long optical delays.

6.3 Discussion

Counterfactual communication has been controversial, with many debating whether given protocols are counterfactual—most notably from the weak trace [Vai13a] and consistent histories [Gri16] approaches to the path of a quantum particle. However, unlike other protocols [Noh09; AB16; LAZ15; AV19; Vai19] Salih et al have recently shown their protocol for counterfactual communication is counterfactual by both these criteria [Sal+22]—whenever Alice detects the photon at either $D0$ or $D1$, she can be sure it has never been at Bob. However, when the number of cycles isn't infinite, there is a chance the photon could end at Bob, rather than Alice, in which case the protocol is aborted and restarted. Given a use of Ghost Imaging is to image photosensitive samples [Gen16; Gil+19] (which could be easily damaged by high-energy photons), we want to reduce the chance of any photons going to/via the object as much as possible. We plot this probability in Fig.6.3a. Note, as N goes to infinity, this probability goes to zero.

Interestingly, Aharonov and Rohrlich have recently shown counterfactual communication conveys modular angular momentum $L_z \mathrm{mod} 2\hbar$ of $\hbar$ from Alice's photon into Bob's blocker whenever Bob blocks [AR20]—however, there is no energy associated with this, meaning there is no chance of this damaging photosensitive samples; therefore, we can ignore this in our analysis.

In Salih et al's 2013 protocol, there is a probability of erroneous $D0$ clicks, as they take $\cos(\pi/2N)^N \to 1$ for large N. This probability, for $M = 2$, is

$$P(D0|Block)_{M=2} = \left(\cos\left(\frac{\pi}{2N}\right)^N - 1\right)^2 /4 \tag{6.1}$$

We plot the probability for given values of M and N in Fig. 6.3b. By increasing the rotation slightly at the start of each outer cycle, this error could be avoided—future work will consider the exact value needed, and the specific benefits of this optimisation.

We now discuss measures of efficacy for the protocol.

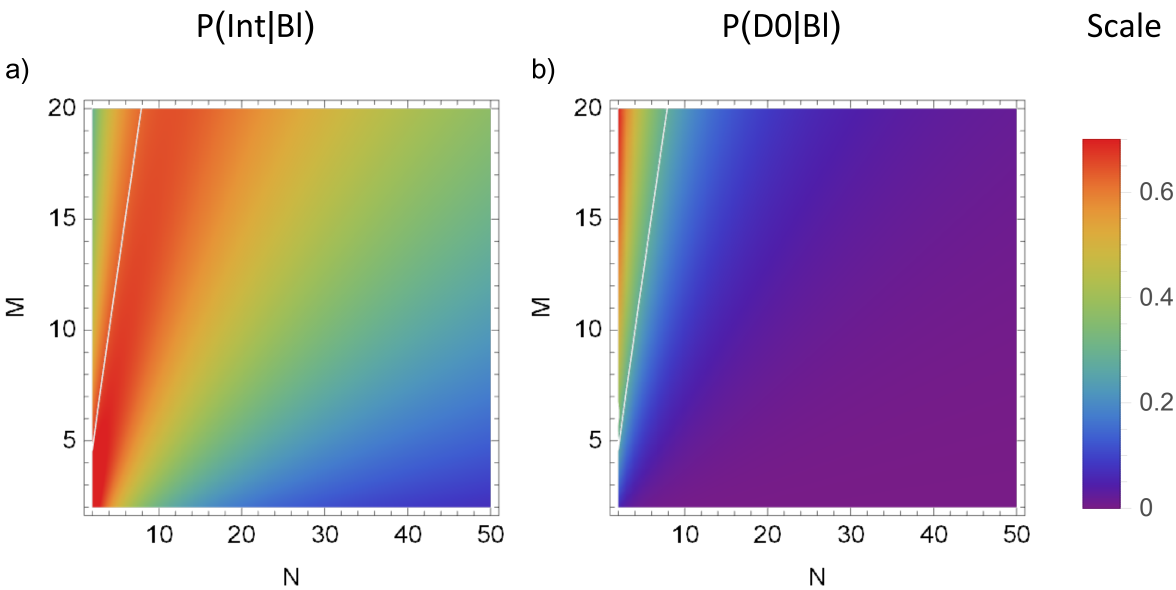

Figure 6.3: Loss probabilities when the object blocks the path. Probability, when the object blocks the channel, of the photon: interacting with the object being imaged (P_{Int}) (a); or erroneously ending up in $D0$ (b). We plot these for given numbers of outer (M) and inner (N) interferometer cycles. Note, the photon only goes via the object erroneously—in any case when $D1$ clicks, the detection of the object will have been fully counterfactual. Further, both the interaction and erroneous-$D0$ probabilities go to 0 as N goes to ∞.

The Signal-to-Noise Ratio (SNR) [OCB10; Bri+11; Gen16], a useful measure of the efficacy of an imaging system, is given by

$$\text{SNR} = \frac{|\Delta \bar{I}|}{\sigma(|\Delta \bar{I}|)} \tag{6.2}$$

where $\Delta \bar{I}$ is the difference in average intensity values observed by a detector between inside and outside the object, and $\sigma(|\Delta \bar{I}|)$ is the standard deviation in this difference.

For standard ghost imaging, when an average of $\bar{N}$ photons (those generated in a given time interval by a SPDC source) interrogate an object, none of the $\bar{N}$ photons will reach the detector, giving a change in photon detection number at the detector of $\Delta N_{GI} = -\bar{N}$. Given spontaneous parametric down-conversion has thermal statistics for rate of emission (which look Poissonian averaged over many temporal modes), the SNR is

$$\text{SNR}_{GI} = \bar{N}/\sqrt{\bar{N}} = \sqrt{\bar{N}} \tag{6.3}$$

For counterfactual ghost imaging, we define ΔN_{D0} (ΔN_{D1}) as the difference in photon numbers received at $D0$ ($D1$) between the object blocking and not blocking the channel (which in each case is $\bar{N}$ times the difference in probability of a photon reaching that detector in each of those two

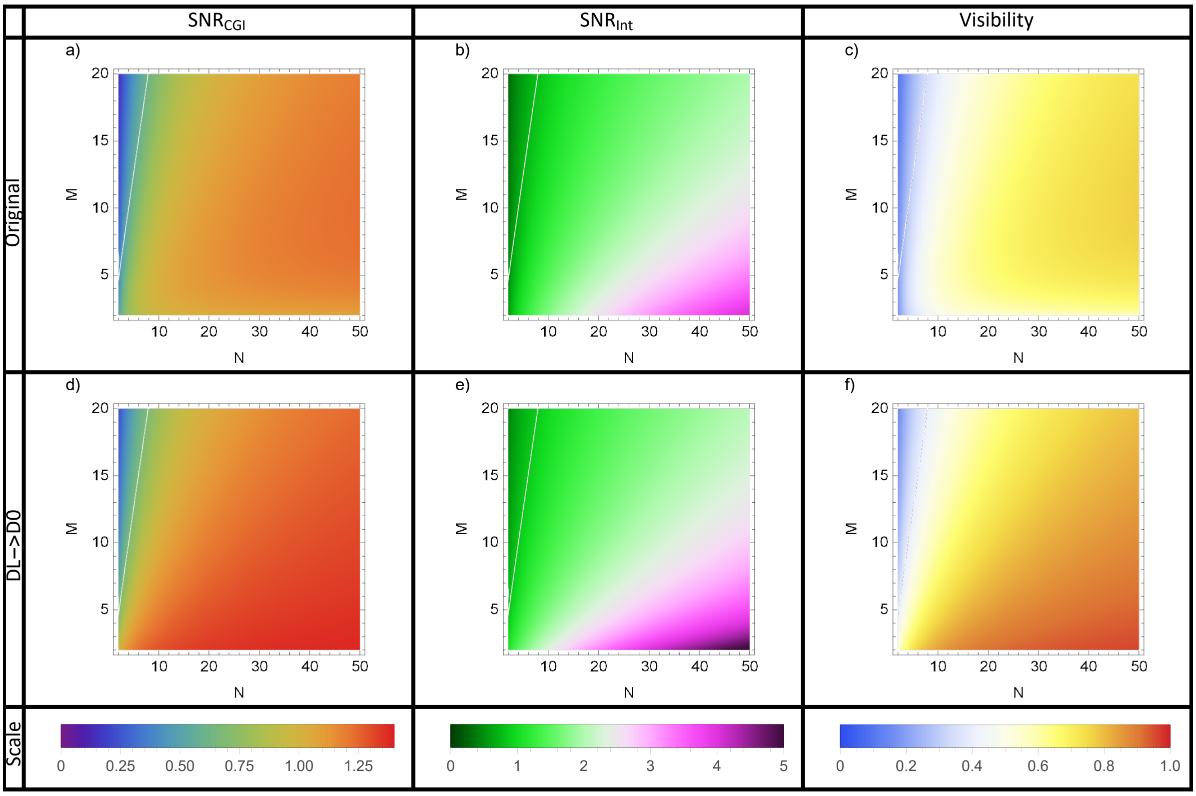

Figure 6.4: Plots of the SNR. This is for: equal numbers of photon pairs generated in a given time by the SPDC source (a); and equal numbers of photons absorbed by the object (b)—and the visibility V of the protocol (c)—for our original protocol. d,e and f give these values for when photons that would go to DL also count for coincidence measurements as if they went to $D0$. These are as functions of the number of outer (M) and inner (N) interferometer cycles.

cases). Note, ΔN_{D0} and ΔN_{D1} will have opposite signs. Therefore,

$$
\begin{aligned}
\mathrm{SNR}_{CGI} &= \frac{|\Delta N_{D0} - \Delta N_{D1}|}{\sigma(|\Delta N_{D0} - \Delta N_{D1}|)} \\
&= f(M,N)\sqrt{\bar{N}} = f(M,N)\mathrm{SNR}_{GI}
\end{aligned}
\tag{6.4}
$$

which we plot in Fig. 6.4a (as a multiple of SNR_{GI}, the SNR of standard ghost imaging). For values of M and N where $\mathrm{SNR}_{CGI} = 1$, the protocol is just as good at imaging as standard ghost imaging. In these cases, Fig. 6.3b shows that in our protocol the probability of a photon interacting with the object is much less than the 73.5% limit from previous protocols [Zha+19].

However, rather than looking at the SNR for the photons generated by a SPDC source in a given period of time, a more apt comparison would be the SNR for which the same number of photons are absorbed by the object as in standard ghost imaging. Given the average number of photons interacting with the object is P_{Int} times $\bar{N}$, we get

$$
\mathrm{SNR}_{Int} = f(M,N)\sqrt{\bar{N}/P_{Int}} = \frac{f(M,N)}{\sqrt{P_{Int}}}\mathrm{SNR}_{GI}
\tag{6.5}
$$

which we plot in Fig. 6.4b (again in terms of SNR_{GI}).

Even for low numbers of outer and inner cycles (M and N), our protocol gives a vast improvement over the signal-to-noise ratio of standard ghost imaging—for instance, two outer cycles of 13 inner cycles gives double the equal-photon-absorption SNR of standard ghost imaging. This is much higher than the 118% improvement in equal-photon-absorption SNR that Zhang et al's protocol gives [Zha+19]. Note, as $N \to \infty$, the probability of a photon interacting with the object goes to 0, meaning SNR_{Int} becomes infinitely larger than the SNR available with standard ghost imaging (if we're willing to wait that long).

Another key measure of an imaging protocol's efficacy is its visibility, V, defined as

$$V = \frac{|\bar{N}_{In} - \bar{N}_{Out}|}{\bar{N}_{In} + \bar{N}_{Out}} \tag{6.6}$$

Visibility gives how responsive to a difference in presence/absence of an object is, defined on a scale from 0 to 1. For our protocol, visibility is given as

$$V_{CGI} = \frac{|\Delta N_{D0} - \Delta N_{D1}|}{2\bar{N}} \tag{6.7}$$

(the changes in intensity for $D0$ and $D1$ over the maximal possible changes in their intensities, remembering their opposite signs). This gives reasonable values (i.e. between 0 and 1) for our protocol, as given in Fig.6.4c. For instance, 5 outer cycles of 12 inner cycles gives a visibility of 0.569. This shows higher visibility than the maximal we calculate for Zhang's protocol (0.5625) [Zha+19].

We now discuss a method for increasing the efficacy of the protocol, at the expense of it no longer being truly counterfactual.

When the object does not block the photon's path in the adapted Salih et al device, for finite numbers of outer interferometers (M), there is a chance the photon will cross through the unblocked gap, in which case it goes to the loss detector DL rather than the coincidence-linked detector $D0$. This occurs with probability

$$P(DL|NB) = 1 - \cos\left(\frac{\pi}{2M}\right)^{2M} \tag{6.8}$$

which we plot in Fig.6.5. This goes to 0 as M goes to infinity.

Note, if we weaken our definition of counterfactuality to be that the photon never goes via the object's path when the object is there (rather than the photon never goes via that path at all), we could link detector DL as well as $D0$ to the ICCD camera, treating coincidences between DL and the ICCD camera as if they were between $D0$ and the ICCD camera. This would let us use the minimum number of outer cycles the protocol works for, two, rather than requiring higher values of M to avoid us erroneously ignoring ICCD detections. We do a version of this in Fig.6.2, having photons which would go to DL treated as if going to $D0$.

This led us to plot altered values for SNRs and visibility, taking into account DL now going to $D0$. These are plotted in Fig.6.4 (d, e and f), and give even better values for all three measures.

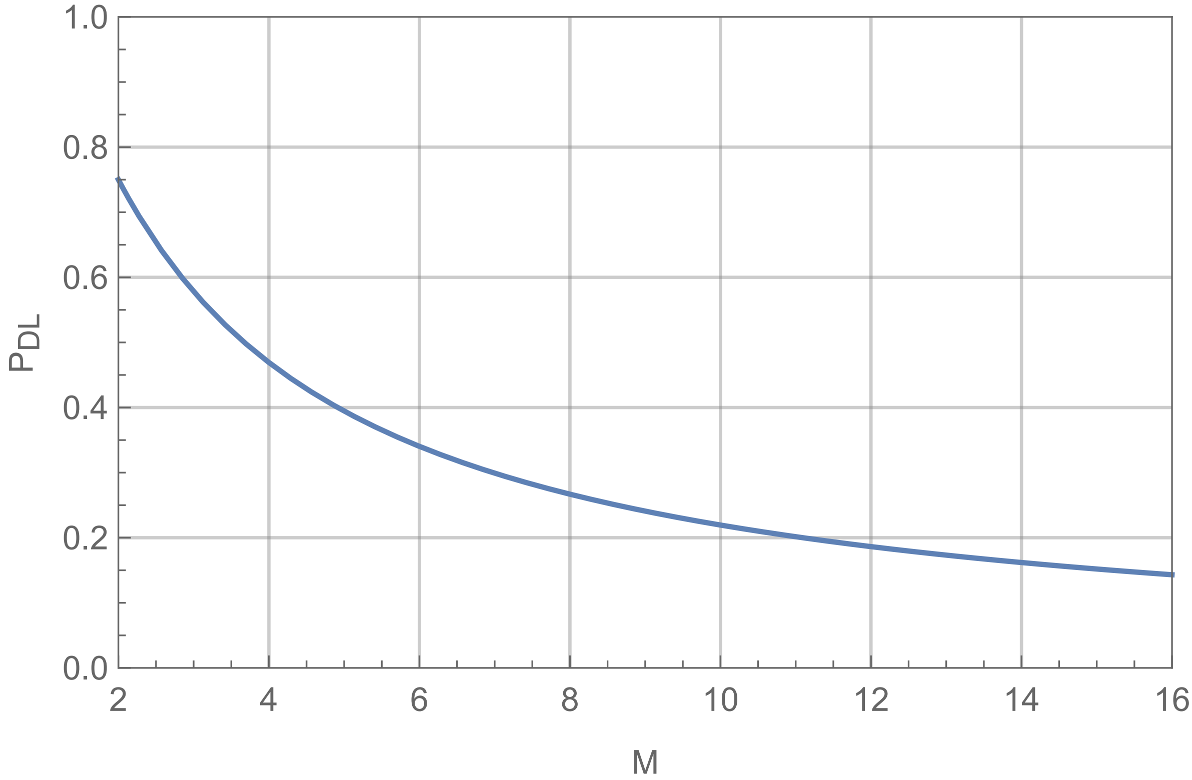

Figure 6.5: Probability the photon goes to DL rather than $D0$. This is the probability that, when the object does not block the photon's path, the photon travels via that path (and so goes to DL rather than $D0$. This is as a function of number of outer cycles (M).

Traditional ghost imaging can only distinguish between whether or not a photon could have passed through a given region (i.e. whether, at that point, a mask would transmit that photon, or whether it would absorb/reflect it). This makes this style of ghost imaging bad for imaging transparent/translucent objects—which is unfortunate, given the many applications for the detection of low-contrast objects. However, Abouraddy et al, and later Gong et al, proposed [Abo+04; GH10] (and Zhang et al experimentally demonstrated [Zha+14]) schemes which make use of phase-sensitivity to allow materials to be detected which, while translucent, create a change to the phase/polarisation of transmitted light. Zhang et al's interaction-free ghost imaging also demonstrated this sensitivity [Zha+19]. Further, given it relies on interference to direct the photon to one of two bucket detectors, our scheme is also sensitive to changes in phase induced by transparent objects—presenting yet another benefit of our protocol over standard ghost imaging.

Alongside standard ghost imaging, which makes use of entangled photons, alternative forms have been created which instead just use classical correlations. Given this is classically simulable, a version has been demonstrated which is referred to as computational ghost imaging—where rather than sending a correlated photon to a high-resolution/scanning detector, a spatial light modulator (SLM) applies an effective reference pattern to the photon probing the object. Given

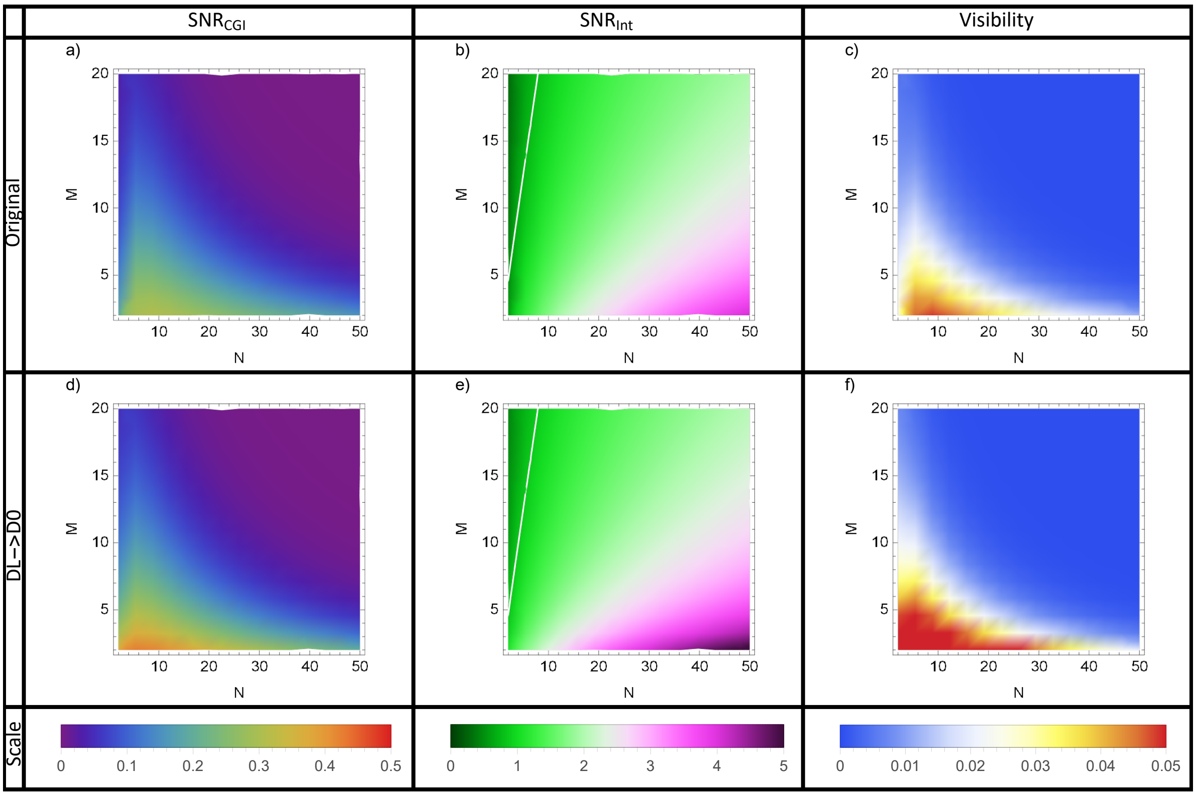

Figure 6.6: Plots (assuming realistic component loss) of the SNR. These are for: equal numbers of photon pairs generated in a given time by the SPDC source (a), and equal numbers of photons absorbed by the object (b)—and the visibility V of the protocol (c)—for our original protocol. d,e and f give these values for when photons that would go to DL also count for coincidence measurements as if they went to $D0$. These are as functions of the number of outer (M) and inner (N) interferometer cycles.

counterfactual ghost imaging allows us to image an object counterfactually while preserving quantum correlations between the signal and idler photons, it can clearly preserve the classical correlations necessary for computational ghost imaging. In such a set-up, we replace the SPDC, beamsplitter and ICCD camera with a single photon source, a spatial light modulator (SLM), a pseudo-random illumination pattern, and computational analysis. While it remains to be seen the effect the loss of quantum correlations would have on fidelity, loss, SNR and visibility, this shows the flexibility of counterfactual alterations to ghost imaging.

The analysis presented above assumes ideal components. Sadly, no component is ideal. In our protocol, the four key components which could through loss affect the protocol are the HWPs, the PBS, the switchable mirrors, and the detectors. In this Subsection, we model the protocol with these at experimentally realistic values, to show that, even with these limitations considered, the protocol still provides a significant advantage over both standard ghost imaging, and classical metrology.

At designed-for wavelengths, HWPs can achieve loss (through reflection) of $\mathcal{O}(0.1\%)$; PBSs can achieve loss (through absorption) of less than 1%, and a typical heralding efficiency for a SPDC/SPAD set-up like ours is 18%. Practically, switchable mirrors pose an issue (given typical switching times for these are of $\mathcal{O}(10^{-6}s)$, while the switching we need has to be of $\mathcal{O}(10^{-9}s)$). However, Cao et al showed how to adapt the protocol to use fixed components that alter a degree of freedom on the photon to effectively 'count' the number of cycles it has travelled, and transmit it after the right number [Cao+17]. This however adds loss in their demonstration of 15/16 per outer cycle. Despite this, as we show in Fig.6.6 (which shows the same quantities as in Fig.6.4 albeit adjusted to take into account these losses), the SNR per photon absorbed is still far higher than for standard ghost imaging.

6.4 Chapter Conclusion

We have given a protocol for ghost imaging in a way that is always counterfactual—while imaging the object, no light interacts with that object. This extends the idea of counterfactuality beyond communication, showing how this interesting phenomenon can be used for metrology. Given, in the infinite limit, no photons ever go to the imaged object, it presents a method of imaging even the most light-sensitive of objects without damaging them. A future direction would be separately leveraging the chained quantum Zeno effect to see if this performance can be improved further. Another future direction, which we discuss in Chapter 10, is leveraging counterfactual/interaction-free metrology for the task of polarimetry—identifying the effect a given sample would have on the polarisation of any arbitrary light ray transmitted through it, while sending vanishingly small quantities of light through the sample itself.

This work gives another promising example of how quantum foundational work can be leveraged for the development of quantum technological tools. Given the large number of potential use cases for optical metrology with reduced light flux through the imaged sample (e.g., the imaging of optically-sensitive biological or archaeological samples), and the impact the use of such a tool could have in each of these cases, the work presented in this Chapter serves to reinforce just how powerful the adaptation of quantum foundational ideas for quantum technologies can be, in allowing us to create tools to benefit society beyond just the quantum community.

Part II

Extensions to Quantum Mechanics

“The concept of ‘measurement’ becomes so fuzzy on reflection that it is quite surprising to have it appearing in physical theory at the most fundamental level... does not any analysis of measurement require concepts more fundamental than measurement? And should not the fundamental theory be about these more fundamental concepts?”

J. S. Bell,
“Quantum Mechanics for Cosmologists” (1981)
([Bel81])

Chapter 7

Extensions of Quantum Mechanics Background

Declaration of contribution: This chapter acts as a literature review for Part II of this thesis. Section 7.2 is partly based on the published paper, *Bell's theorem allows local theories of quantum mechanics* [HH22c], and Section 7.3 is adapted from the published paper *What does it take to solve the measurement problem?* [HH22a]. Both of these papers were conceived and written by myself and Dr Sabine Hossenfelder. These papers were written at a later stage of my PhD, when I had been developing independent collaborations, and my supervisors were happy for this independent work to be included in my thesis.

7.1 Chapter Introduction

In this chapter, we lay out the necessary background for this second Part of the thesis, which looks at (specifically statistical independence-violating) extensions of quantum mechanics. This involves first introducing Bell's Theorem, and the assumptions necessary to formulate a Bell inequality. We then focus on the statistical independence assumption specifically, and proposed extensions of quantum mechanics where this assumption is violated. After this, we move onto the measurement problem in quantum mechanics, illustrating how it arises from the axioms of quantum mechanics, requirements a solution to the problem must possess, and why currently-proposed interpretations and extensions of quantum mechanics fail to solve the problem.

This Chapter, and Part overall, allows us to survey cases where quantum devices and technologies can be used to test foundational ideas—the second part of the interplay between quantum foundations and quantum technologies.

7.2 Bell's Theorem

In this section, we go over Bell's Theorem. This states that, given certain (reasonable-sounding) classical assumptions, one can generate an inequality from the expected results of an experimental scenario, which can be violated by expected results when represented using quantum mechanics [Bel71]. This programme, developing experiments where quantum mechanics gives differing predictions to those expected from reasonable assumptions about the nature of the world, originated with Einstein, along with Podolsky and Rosen (EPR) [EPR35]. They first came up with a case in which interpreting quantum mechanics as fully describing a given situation seemingly leads to a paradox. In this section, we first go through this EPR paradox, and Bohm's simplification of it, before showing how to use this EPR-Bohm set-up to derive a Bell inequality: the Clauser-Horne-Shimony-Holt (CHSH) inequality. We then describe experimental attempts to violate this inequality (e.g. Freedman and Clauser's [FC72], and Aspect et al's [AGR81; AGR82]), loopholes in those experiments, and how those loopholes were resolved. Finally, we consider the assumptions needed to formulate a Bell inequality, and the consequences in each case for that assumption being violated.

7.2.1 EPR Paradox

Einstein, Podolsky and Rosen came to a paradox by combining conjugacy (one variable's uncertainty increasing as the other's decreases) and entanglement (two particles' states being so correlated they cannot be written independently) [EPR35]. By measuring one of a set of conjugate variables for one of a pair of entangled particles, quantum mechanics says the conjugate variable for the other particle becomes uncertain. This is despite the two particles potentially being spatially separated, and standard quantum mechanics providing no mechanism for a signal to propagate between them.

Bohm simplified this paradox, doing away with non-commuting observables, and focusing on joint states—those which quantum mechanics says cannot be written as the (tensor) product of single-particle states [Boh51]. He gave the example of the joint spin state of two entangled electrons, where the total spin of the two particles is 0 (see Fig 7.1).

Using Bra-Ket notation, we write the spin-states of the two electrons as the Bell state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle_1 \otimes |\downarrow\rangle_2 + |\downarrow\rangle_1 \otimes |\uparrow\rangle_2) \tag{7.1}$$

where indices 1 and 2 label the spin states of particles 1 and 2 respectively. Note, while normally we need to define a basis for the result of a spin measurement (e.g., for a spin-1/2 particle, spin-up or spin-down in the X, Y or Z direction), the Bell state in Eq. 7.1 takes the same form regardless of the basis in which we represent it, so long as we represent both particles in the state in the

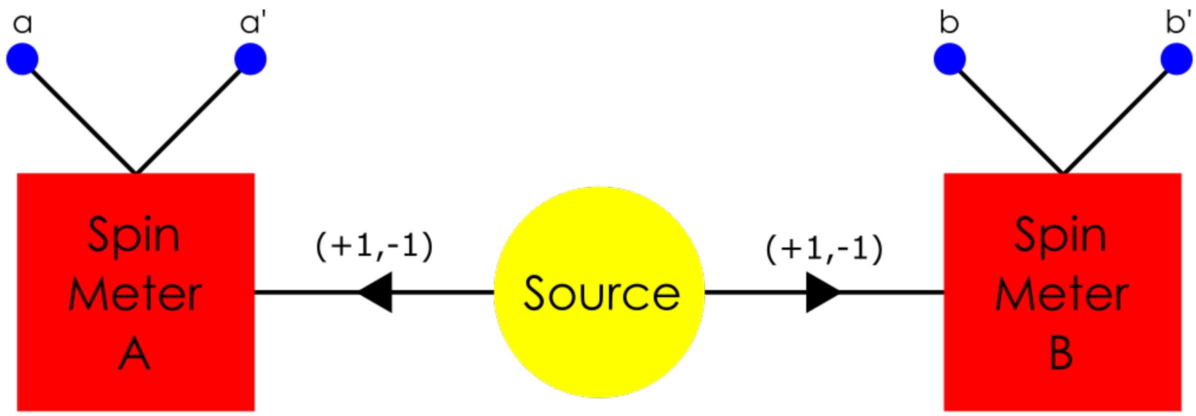

Figure 7.1: The EPR-Bohm experiment [Boh51], where a source emits pairs of spin-entangled particles in opposite directions. The joint spin state of the two particles is $|\psi\rangle = (|\uparrow\rangle_1 \otimes |\downarrow\rangle_2 + |\downarrow\rangle_1 \otimes |\uparrow\rangle_2)/\sqrt{2}$. The left-hand particle is subjected to either test a or a', and the right-hand one, to either test b or b'. All tests have result either $+1$ or -1. This set-up (alongside various assumptions, discussed below) was used to derive the CHSH inequality [Cla+69], which was experimentally violated by Aspect et al [AGR82].

same basis—i.e.,

$$
\begin{aligned}
|\Psi\rangle &= \frac{1}{\sqrt{2}}(|\uparrow_X\rangle_1 \otimes |\downarrow_X\rangle_2 + |\downarrow_X\rangle_1 \otimes |\uparrow_X\rangle_2) \\
&= \frac{1}{\sqrt{2}}(|\uparrow_Y\rangle_1 \otimes |\downarrow_Y\rangle_2 + |\downarrow_Y\rangle_1 \otimes |\uparrow_Y\rangle_2) \\
&= \frac{1}{\sqrt{2}}(|\uparrow_Z\rangle_1 \otimes |\downarrow_Z\rangle_2 + |\downarrow_Z\rangle_1 \otimes |\uparrow_Z\rangle_2)
\end{aligned}
\tag{7.2}
$$

(and similarly for any other choice of representation direction).

As this state cannot be written as the tensor product of the two individual electrons' states, quantum mechanics says we cannot describe these two electrons' spins independently of one another. This means, if this description is the most complete we can have, when one of the electrons' spins is measured, it collapses the overall state, pushing the other electron into its corresponding spin (and making its spin in any of the other two mutually unbiased bases uncertain). If we are being fully realist about both the wavefunction (i.e. saying there is nothing in this system not described by the wavefunction as given/no hidden variables) and this collapse, this just happens instantaneously, and so information must travel superluminally. To reinforce this, the state into which the first electron collapses is completely random according to standard quantum mechanics, so there is no way (by this account) that the second electron can be said to be 'pre-prepared' into the corresponding state in advance, to avoid the need for superluminal information transfer.

Einstein dismissed this effect as "spooky action at a distance" (spukhafte Fernwirkung) [Ein71]—information seemingly passing from one place to another, instantly and without mechanism. To him, this meant there must be a deeper description of what was happening in the system than standard quantum mechanics provided: a possibility referred to as “hidden-variables” theories.

7.2.2 Bell Inequalities

For nearly thirty years after the EPR paper, quantum mechanics was considered either incomplete, failing to account for some ’hidden variables’ which further explain its peculiarities, or to violate some fundamental assumption about the nature of the universe. However, in 1964 Bell proposed an experiment to settle the debate [Bel64]. For the EPR-Bohm experiment, Bell derived the upper limit for measurable correlations between the two particles, assuming they obeyed a local hidden-variable model (among other assumptions).

These local hidden variable models posit a variable λ, which, from the moment an entangled state is generated, holds information as to the final state its constituents will end up in, above and beyond that given by the wavefunction. For instance, for Eq. 7.1, this information would be whether the spins of the two electrons will collapse to $|\uparrow\rangle_1 \otimes |\downarrow\rangle_2$ or $|\downarrow\rangle_1 \otimes |\uparrow\rangle_2$). This variable is hidden (i.e., not necessarily accessible to the experimenter), and local (i.e., cannot be used as a channel to superluminally transmit information between the two particles). The space Λ spans all possible values for this local hidden variable (i.e. $\lambda \in \Lambda$, $\forall \lambda$).

By using properties of these local hidden variable theories alongside certain assumptions, Bell (and others) derived inequalities for sums of correlation functions. The most famous of these Bell inequalities is the CHSH inequality [Cla+69] (of which we use the Wigner form [Wig70; Red87]). As in the EPR-Bohm set-up, a source emits pairs of particles in opposite directions. The left-hand particle is subjected to test a (either a_0 or a_1), and the right-hand one to test b (either b_0 or b_1). All tests have result either +1 or -1 (see Fig. 7.1). For example, for tests of single-photon polarisation, where a_0 and a_1 (b_0 and b_1) are differently oriented polarisers, +1 corresponds to the photon being transmitted, and -1 to it being absorbed.

For a given trial i, we call the result of the experiment of the left A_i either $A_{0,i}$ or $A_{1,i}$ (depending on whether we choose test a to be a_0 or a_1), and similarly B_i as either $B_{0,i}$ or $B_{1,i}$. Therefore, we can write the four pairs of products of the results as $A_{0,i}B_{0,i}$, $A_{1,i}B_{0,i}$, $A_{0,i}B_{1,i}$, and $A_{1,i}B_{1,i}$.

We use these 4 result combinations to give

$$\gamma_i(a_0,a_1,b_0,b_1) = A_{0,i}B_{0,i} + A_{1,i}B_{0,i} + A_{0,i}B_{1,i} - A_{1,i}B_{1,i} \tag{7.3}$$

As $A_{0,i}, A_{1,i}, B_{0,i}, B_{1,i} \in \{-1,1\}$,

$$\begin{aligned} \gamma_i(a_0,a_1,b_0,b_1) &= A_{0,i}(B_{0,i}+B_{1,i}) + A_{1,i}(B_{0,i}-B_{1,i}) \\ &= \pm 2 \end{aligned} \tag{7.4}$$

For N events, given $\gamma_i = \pm 2$,

$$\left|\frac{1}{N}\sum_{i=1}^{N}\gamma_i\right| = \left|\frac{1}{N}\sum_{i=1}^{N}A_{0,i}B_{0,i} + \frac{1}{N}\sum_{i=1}^{N}A_{1,i}B_{0,i} + \frac{1}{N}\sum_{i=1}^{N}A_{0,i}B_{1,i} - \frac{1}{N}\sum_{i=1}^{N}A_{1,i}B_{1,i}\right| \le 2 \tag{7.5}$$

As $A_i \in \{A_{0,i}, A_{1,i}\}$ and $B_i \in \{B_{0,i}, B_{1,i}\}$, the expectation value for the product of the results of a given pair of measurements a, b is

$$E(a,b) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} A_i B_i \tag{7.6}$$

so Eq. 7.5 becomes

$$S = |E(a_0,b_0) + E(a_1,b_0) + E(a_0,b_1) - E(a_1,b_1)| \leq 2 \tag{7.7}$$

This is the CHSH inequality. Note, the CHSH inequality is typically used as, unlike Bell's 1964 inequality, it doesn't require perfect anticorrelation between the measurement outcomes of the two particles when measured using the same setting. This allows it to better account for experimental errors; with perfect anticorrelation, it simplifies to Bell's initial inequality [Bel71].

Quantum mechanics can violate this limit. Returning to the state defined in Eq. 7.1, the quantum mechanical expectation value for the product of the results of two tests a, b on the two parts of the entangled state (i.e., test a on the part of the state labelled subscript 1, and test b on the part labelled subscript 2) is

$$E(a,b) = \langle\Psi|(\sigma_{1,a} \otimes \sigma_{2,b}|\Psi\rangle \tag{7.8}$$

where $\sigma_{1,a}$ is the spin operator on particle 1 in direction a (and $\sigma_{2,b}$ the spin operator on particle 2 in direction b), such that for instance

$$\sigma_{1,Z}|\uparrow_Z\rangle_1 = |\uparrow_Z\rangle_1\,;\ \sigma_{1,Z}|\downarrow_Z\rangle_1 = -|\downarrow_Z\rangle_1 \tag{7.9}$$

As spin lives on the Bloch sphere, we can simplify Eq. 7.8 by decomposing spin measurements in the b direction, into the component of the measurement parallel to a, and the component perpendicular to a (denoted $a\perp$):

$$E(a,b) = \langle\Psi|\left(\sigma_{1,a} \otimes \left(\cos\theta_{ab}\sigma_{2,a} + \sin\theta_{ab}\sigma_{2,a\perp}\right)\right)|\Psi\rangle \tag{7.10}$$

where θ_{ab} is the angular distance along the Bloch sphere between a and b.

As $|\Psi\rangle$ is the state given in Eq. 7.1, which keeps the same form regardless of the basis we represent it in, Eq. 7.8 simplifies further to

$$E(a,b) = -\cos\theta_{ab} \tag{7.11}$$

This means, given

$$S = |E(a_0,b_0) + E(a_1,b_0) + E(a_0,b_1) - E(a_1,b_1)| \tag{7.12}$$

in the quantum case,

$$S = \left|-\cos\theta_{a_0b_0} - \cos\theta_{a_1b_0} - \cos\theta_{a_0b_1} + \cos\theta_{a_1b_1}\right| \tag{7.13}$$

This is maximal when $\theta_{a_0b_0} = \theta_{a_1b_0} = \theta_{a_0b_1} = \pi/4$ and $\theta_{a_1b_1} = 3\pi/4$, such as when

$$a_0 = \sigma_X;\ a_1 = \sigma_Z;\ b_0 = \frac{\sigma_Z + \sigma_X}{\sqrt{2}} \text{ and } b_1 = \frac{\sigma_Z - \sigma_X}{\sqrt{2}} \tag{7.14}$$

and gives

$$S_{max} = \left| -\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} - + \frac{1}{\sqrt{2}} \right| = 2\sqrt{2} \tag{7.15}$$

This is a proof of the Tsirelson bound [Cir80]: the maximal value of S achievable using quantum mechanics. (Note, for the EPR-Bohm scenario, Popescu and Rohrlich showed S can go above this bound, up to $S \leq 4$, for general nonlocal theories [PR94], though Carmi and Cohen showed models where $S > 2\sqrt{2}$ must at the least violate a subtle form of relativistic causality). While there are more formal proofs, generalised for all quantum systems (rather than just dimension-2, like the spin-1/2 qubits above), this proof serves our purpose. It shows the CHSH inequality can be violated using quantum logic: local-hidden-variable models of the sort used to generate the CHSH inequality above fail to account for predicted quantum correlations between the two particles. Therefore, an experiment giving $S > 2$ would show the Universe violates one of the assumptions used to form this inequality.

7.2.3 Experimental Tests and Loopholes

The first experimental attempt to violate a Bell inequality was performed by Freedman and Clauser in 1972 [FC72]. This used polarisation-entangled photon pairs, generated through the atomic cascade of Calcium, to try and violate the CHSH inequality that Clauser et al had proposed two years earlier. However, this experiment left a number of loopholes—ways a local hidden-variable theory could still explain the correlations observed, given practical details of the experiment. In this Subsection, we discuss these loopholes, and how later Bell tests managed to avoid them, such that loophole-free violations of Bell inequalities have now been performed. We follow Larsson et al's categorisation of these loopholes into the Locality loophole, the Memory loophole, the Detection loophole, and the Coincidence loophole [Lar14].

Locality: If local (i.e., subluminal) communication between the two measurement sites is possible in the time between a measurement choice being decided at one measurement site, and the measurement being performed at the other measurement site, then it is perfectly consistent that we could observe correlations which violate our Bell inequality even if the Universe obeys a local hidden-variable theory. This was the Locality loophole found in Freedman and Clauser's experiment—we could imagine some mechanism by which the measurement settings at the first site (the polariser direction being set) locally affect the state of the particle at the second site, in such a way as to generate these correlations. This loophole was first closed in Aspect et al's tests, proposed in 1976 [Asp76], and performed in the early 1980s [AGR81; AGR82]. These experiments also used polarisation-entangled single photons, but used time-varying analysers to perform measurement. These analysers are effectively variable polarisers, which jump between two polarisation orientations in a time far shorter than the time it takes each photon to travel from the source to its respective polariser.

Memory: While Aspect et al's experiment closed the locality loophole, it raised a new loophole—the memory loophole. This loophole comes from the analysers in Aspect et al's experiment shifting between the two polarisations deterministically, at a fixed frequency—the state of the polariser at some given time in the future was predictable. One could imagine, while unlikely, a way the system could extrapolate what the state would be at a given time to allow the generation of Bell inequality-violating coincidences. Aspect himself noted this loophole in his proposal [Asp76], saying the set-up required the supplemental assumption that the polarisers have no "memory" as to what state they had been in previously, so they couldn't extrapolate from any regularity in the settings what their state would be in the future. This assumption however is not needed (and the memory loophole closed) if the measurement settings for each experimental run are chosen in some unpredictable way, so there is no way to infer what the measurement setting will be for a given run. The first Bell tests to close this loophole were those of Weihs et al [Wei+98]. Here, random number generators at each measurement site determined which of the two possible measurement settings to use for a given run, after the photons were emitted from the source. (This also avoids the locality loophole.)

Detection: next, we look at the detection, or efficiency loophole, whereby the possibility of losing or not detecting particles during the experiment can affect the correlations observed. This can occur in two ways. Firstly, when using polarisers (like Freedman and Clauser), one of the measurement results (-1) is recorded when there is no detection, as one assumes the photon was absorbed by the polariser, having had an orthogonal polarisation to the polariser's measurement setting. However, there are many other reasons the photon might be lost during the experiment. For instance, the detector might not have perfect efficiency, the lenses and transmissive optics between the source and the polariser might not be perfectly transmissive, or polarisers might not be perfectly transmissive even at their transmission polarisation. Therefore, -1 would be recorded far more often in practice than we would expect with a theoretically perfect experiment. This could overshadow any correlation from entanglement, and lead to the data not violating the inequality. To avoid this (and not receiving a click when observing a -1 detection), Freedman and Clauser adapted the CHSH inequality to give an inequality in terms of rates of coincidence detection as a function of angle between the two polarisers, divided by rate of coincidences when both polarisers were removed. This however relies on an assumption of fair sampling—that the detection efficiency is the same for different polarisations. This could be violated if, for instance, the detectors had a bias towards detecting certain polarisations more efficiently than others, or the lenses used subtly act as polarisers.

The detection loophole also appears in experiments such as Aspect et al's. These use PBSs of variable basis rather than variable polarisers, so there is a detection associated with both a +1 and a -1 click. However, they still neglects runs (photon pairs) where one or both photons are absorbed. Therefore, forming a Bell inequality for these experiments still requires some assumption of fair sampling—that observed runs are a fair sample of all runs. This loophole is only closed by tests where this loss can be bounded and accounted for. Initially, this was only possible in tests using solid-state platforms rather than optics, such as Rowe et al's [Row+01], which then had the locality loophole. However recent experiments have been performed which close both detection and locality loopholes simultaneously—e.g. Hensen et al's [Hen+15], which use a combination of optics and nitrogen-vacancy centre qubits, and Giustina et al's [Giu+15] and Shalm et al's, which both use entangled photon pairs.

Coincidence: A related loophole is the coincidence loophole: given we rely on simultaneous detection times to tell us which particles where initially generated in the same pair, we could imagine a local hidden variable model whereby state/setting-dependent delays cause correlations which violate the Bell inequality [LG04]. This loophole is most problematic in photon-pair experiments, such as Freedman and Clauser's, and Aspect et al's, but can again be mitigated in solid-state experiments. In these, it is easier to assign measurement results to a given run, and so they don't require coincidence detection. Larsson showed that later photonic Bell tests (e.g., Giustina et al's and Shalm et al's) also avoid this loophole [Lar+14].

Bell tests have been performed which close all these loopholes. These experiments show one of the assumptions which go into generating a theoretical Bell inequality must be false, rather than there potentially being some convoluted experimental reason why we observe such a violation. Over the last fifty years, the quantum community has worked hard to evaluate and mitigate against any way an observed Bell inequality violation could be due to one of these loopholes, rather than due to one of these assumptions being false. This shows the benefits of interaction between foundational and technological minds, and illustrates the necessity of interaction between the quantum foundations and quantum experimentalist/technological communities when performing work such as this.

7.2.4 Assumptions

Above, we showed that observed Bell inequality violations are not due to experimental loopholes, but must instead necessarily be due to one of the assumptions used to generate such a Bell inequality being false. We now look in more detail at these assumptions. In this thesis, we divide the assumptions into Hidden Variables, the Matching Condition, and Factorisability (which we subdivide into No Superluminal Interaction and Statistical Independence). Note there are many other categorisations (e.g. Wiseman and Cavalcanti's [WC17]), which subdivide many of these assumptions further. However this simple categorisation is sufficient for our purposes, given we are mainly describing these assumptions to provide background before looking more specifically at the statistical independence assumption (which is universal to these categorisations).

Hidden variables The first assumption needed for the Bell inequality derivation above is the existence of hidden variables, and the result of a given measurement being a function both of the measurement setting, and some hidden variables. These hidden variables are typically viewed as localised at each particle in some way (i.e. as hidden system properties of each particle). This becomes important when combined with the No Superluminal Signalling assumption mentioned later. Note, these variables need not completely determine the measurement result (i.e., there can still be some classical randomness to the process of measurement result generation). However, given these variables are hidden, we can just add an additional hidden variable to the model which differentiates between the different random options. This allows us to return to a deterministic relationship between the combination of hidden variables and measurement setting, and measurement results, without affecting the conclusions of the model.

We can represent this dependence on hidden variables by writing our product result AB as

$$AB \equiv A(a,b,\lambda)B(a,b,\lambda) \tag{7.16}$$

where a and b are our two measurement setting choices ($a \in \{a_0, a_1\}$, $b \in \{b_0, b_1\}$), and λ represents our hidden variables. Here, for greatest generality, we have each measurement result A, B written as a function of both the measurement choice applied on its side, and the measurement choice applied to the other side. This is restricted further by the Factorisability assumption discussed below.

The hidden-variables assumption is often claimed to be rebutted by the violation of Bell inequalities. Those who hold to this interpretation claim that what Bell's Theorem shows is that quantum mechanics provides a complete description of the world, and hidden variable models are unnecessary classical baggage we bring with us.

Matching Condition The next assumption required to generate a Bell inequality is the matching condition. This is a strong version of the assumption of counterfactual definiteness—that there is a matter of fact about the results of unperformed measurements. Looking back at Subsection 7.2.2, we define γ_i for a given trial i by summing the four different pairs of measurement results. However, this requires there to be a matter of fact about what result we would get, had we performed each of the four possible pairs of measurements on the particle pair. For instance, imagine for a given trial that we had measured the two particles with measurement settings a_0 and b_0, and so got measurement result pair $A_{0,i}B_{0,i}$. To generate a value of γ for that trial, we are now required to also imagine what the results of a measurement would be, for that trial, had we instead measured a_1 instead of a_0 (to get a value of $A_{1,i}B_{0,i}$). Further, we are required to both imagine what the measurement result would be had we measured b_1 rather than b_0 (to get $A_{0,i}B_{1,i}$), and what would happen if we measured both a_1 instead of a_0 **and** b_1 instead of b_0 (to get $A_{1,i}B_{1,i}$). We can only ever measure one pair of measurement settings on each particle, so we have to assume it is meaningful to consider these counterfactual options. Note, this is not the same as assuming that changing measurement settings on one side does not affect the measurement results on the other, which is instead part of the locality assumption below. All the matching condition assumption requires is that the measurement results of unperformed measurements are counterfactually definite for all four pairs of measurement choices.

There are ways to generate Bell inequalities without making explicit use of the Matching Condition, as in stochastic hidden variables models. However, this involves considering probability distributions over possible measurement result pairs, which requires assuming Kolmogorov's axioms [KB18]. However, these axioms are far stronger than the matching condition, and (depending on how one interprets these probabilities) such an approach can itself be said to rely on counterfactual definiteness as a sub-assumption.

Factorisability The next assumption we look at, which we use to form Bell inequalities, is factorisability. This is the assumption we can write the product of the measurement results in the form

$$AB \equiv A(a,\lambda_A)B(b,\lambda_B) \tag{7.17}$$

where there are two implicit assumptions. The first is that the hidden variables λ_A are only able to be influenced by events in the past or future light cone of the measurement event giving A, and the same for λ_B and B. The second is that the measurement settings a and b can be treated as completely free variables—they are not themselves influenced by each other, or the hidden variables, nor can they influence the hidden variables.

This notion of factorisability therefore can be decomposed more fully into these two assumptions—no superluminal interaction, and statistical independence.

No superluminal interaction The No Superluminal Interaction assumption is that there can be no way for spacelike-separated events (those outside each others' past or future light-cones) to affect each other. By having the two measurement events spacelike-separated (as is required to avoid the locality loophole), this assumption prevents the measurement setting on one side of the set-up directly affecting the measurement result on the other side of the set-up (as in, affecting the result beyond such correlation as would be allowed by the shared part of the hidden variables). Therefore, it allows us to partially-factorise the product of the measurement results into

$$AB \equiv A(a(\lambda),\lambda)B(b(\lambda),\lambda) \tag{7.18}$$

However, this does not necessarily mean this assumption requires the measurement settings on one side cannot affect the measurement results from the other, as this correlation between the two could be carried through a correlation between the measurement settings and the hidden variables. (This possibility is only blocked by combining the No Superluminal Interaction assumption with the Statistical Independence assumption below.)

A concern with models which break this assumption is that we haven't observed any other systems where interaction occurs superluminally. Shimony therefore proposed the idea of quantum interactions exhibiting superluminal 'passion', rather than action, at a distance [Shi93]. This is as, even if factorisability was violated by some superluminal effect for entangled particles, no observer would be able to extract transmit information superluminally through this mechanism, due to the No-Signalling Theorem [GRW80; PT04]. Therefore, the violation of this assumption doesn't necessarily allow observations which would contradict Special Relativity.

That this assumption is the one violated is one of the most common interpretations of Bell's Theorem, to the extent that Bell's Theorem is often claimed to show that nature is nonlocal [Tea22], and EPR-correlation (or at least the violation of factorisability) is often referred to in quantum information theory as nonlocality [PR94; Pop14; Bru+14]. However, while it respects No-Signalling, it still seems to violate our intuitions about Special Relativity, given this prima facie prohibits any signal, not just observable information, travelling superluminally.

7.2.5 Statistical Independence

Statistical independence is the assumption that there is no correlation between our measurement choices and our hidden variables. This can be represented in multiple equivalent ways. Following the notation we use above, we can write this assumption as

$$a \neq a(\lambda);\ b \neq b(\lambda);\ \lambda \neq \lambda(a);\ \lambda \neq \lambda(b) \tag{7.19}$$

Alternatively, if we imagine some probability distribution over our hidden variables describing their likelihood, ρ, then we can write this assumption as

$$\rho(\lambda, a, b) = \rho(\lambda),\ \forall a \in \{a_0, a_1\},\ b \in \{b_0, b_1\} \tag{7.20}$$

It is a common misconception that this assumption is not needed, and so Bell's Theorem rules out all local hidden-variable models [SKZ13; SN16; Tea22]. However, the assumption is mathematically necessary to formulate a Bell inequality [Che21; HH22c]—otherwise, the measurement result on one side of the experiment could trivially depend on the measurement setting on the other side of the experiment through the hidden variables, which would violate factorisability without requiring superluminal interaction.

We can imagine three cases (ignoring just repeated coincidence) where this assumption is not true (i.e., where there is some correlation between the hidden variables and measurement settings). First, this correlation could be due to the hidden variables in some way causing the choice of measurement settings; secondly, the measurement settings could in some way cause the value of the hidden variables; or thirdly, the choice of measurement settings and value of the hidden variables could have some common cause. Models following the first or third option are commonly referred to as superdeterministic [HP20; Hos20] (although in Chapter 9 and [HHP22] we discuss another set of models which fall under the third option). Models following the second option are termed retrocausal: we do not discuss these further in this thesis, but recommend [WA20] for a well-written review of these models.

Examples of superdeterministic models include Brans's model [Bra88], 't Hooft's cellular automaton model [Hoo16], Ciepielewski, Okon and Sudarsky's model [COS20], Donadi and Hossenfelder's toy model [DH22], and Palmer's Invariant Set Theory (IST) [Pal20a]. We discuss IST more in Chapter 9, as an example of a model which fits more neatly in a new subcategory of statistical independence-violating model we propose—supermeasured models.

While authors often talk about classical superdeterministic models [DRS22], all the models mentioned above use entanglement, superpositions, density matrices and wavefunctions just like standard quantum mechanics. This illustrates why it is worth investigating local hidden-variable models which avoid Bell's Theorem: not just to return to something which looks like classical mechanics, but to resolve issues which the standard formalism of quantum mechanics cannot. An example of such an issue is the Measurement Problem.

7.3 The Measurement Problem

This section looks at the measurement problem in quantum mechanics. We go over how this problem arises, the issues it causes, what we can say about possible solutions to it, and issues with all previous attempts to resolve it.

Quantum mechanics, in its standard formulation (often referred to as the Copenhagen Interpretation), has two different axioms for its time evolution. The one is the deterministic, linear Schrödinger equation, the other the non-deterministic, non-linear, and generically non-local collapse of the wave function. The latter is sometimes also referred to as the "reduction" or "update" of the wave function.

The collapse of the wave function must be mathematically applied in the event of a measurement, yet the theory leaves unspecified just what constitutes a measurement. While this problem is a century old, it is still hotly debated [Mer22]. We show here that this is more than just unsatisfactory, but is a severe shortcoming that requires a solution. Previous accounts of some of the aspects discussed here can be found in [Mau95; Leg05; Wei16].

7.3.1 The Axioms

Below, we take an instrumental perspective on quantum mechanics. For this Section, quantum mechanics is a mathematical machine. Into this machine we insert some known properties of a system that we have prepared in the laboratory. Then we do the maths, get out a prediction for measurement outcomes, and compare the prediction to the observation.

The axiomatic framework of this mathematical machine can roughly be summarised as:

Axiom 1. *The state of a system is described by a vector $|\Psi\rangle$ in a Hilbert space $\mathscr{H}$.*

Axiom 2. *Observables are described by Hermitian operators $\hat{O}$. Possible measurement outcomes correspond to one of the mutually orthogonal eigenvectors of the measurement observable $|O_I\rangle$.*

Axiom 3. *In the absence of a measurement, the time-evolution of the state is determined by the Schrödinger equation* $i\partial_t|\Psi\rangle = \hat{H}|\Psi\rangle$.

Axiom 4 (Collapse Postulate)**.** *In the event of a measurement, the state of the system is updated to the eigenvector that corresponds to the measurement outcome* $|\Psi\rangle \to |O_I\rangle$.

Axiom 5 (Born's Rule)**.** *The probability of obtaining outcome* $|O_I\rangle$ *is given by* $|\langle\Psi|O_I\rangle|^2$.

Axiom 6. *The state of a composite system is described by a vector* $|\Psi\rangle$ *in the tensor product of the Hilbert-spaces of the individual systems.*

This brief summary doesn't do justice to all the subtleties of quantum mechanics. Among other things, it doesn't specify what the Hamiltonian operator is or how one gets the operators corresponding to the measurement observables. However, the form of those operators will not concern us in the rest of this Section.

Of course, there are many different ways to approach quantum mechanics axiomatically, e.g., that proposed by Hardy [Har01]. We will use the above set of axioms because it is how quantum mechanics is typically taught to students, and we believe that this familiarity will make our argument more accessible.

It is sometimes questioned whether the Collapse Postulate is actually necessary (e.g. in [Zur18]). Without it, quantum mechanics would still correctly predict average values for large numbers of repetitions of the same experiment. This is the statistical interpretation suggested by Ballentine [Bal70].

However, we do not merely observe averages of many experiments: we also observe the outcomes of individual experiments. And we know from observations that the outcome of an experiment is never a superposition of detector eigenstates, nor is it ever a mixed state (whatever that would look like)—a detector either detects a particle or it doesn't, but not both. As Maudlin put it [Mau95], "it is a plain *physical fact* that some individual cats are alive and some dead" (emphasis original). Without the Collapse Postulate, the mathematical machinery of quantum mechanics just does not describe this aspect of physical reality correctly.

This means quantum mechanics without the Collapse Postulate is not wrong, but it describes less of what we observe. The Collapse Postulate is hence useful, and part of the axioms because it increases the explanatory power of the theory. It cannot simply be discarded.

The necessity of the Collapse Postulate to describe observations is not a problem in itself, but it gives rise to the problems discussed below.

7.3.2 The Problems

The Heisenberg Cut

The most obvious problem with the axioms of quantum mechanics is the term "measurement," which remains undefined.

The need to refer to a measurement strikes one as suspect right away, because quantum mechanics is commonly believed to be a fundamental theory for the microscopic constituents of matter. But if quantum mechanics was fundamental, then the behaviour of macroscopic objects (like measurement devices) should be derivable from it. If we are to be realist about quantum mechanics, rather than just instrumentalist, the theory should explain what a measurement is, rather than require it in one of its axioms.

This problem has been known since the earliest days of quantum mechanics and is often referred to as the "Heisenberg Cut", alluding to where to make the "cut" between the unitary Schrödinger evolution and the non-unitary measurement update [Hei52].

One may object that it is a rather inconsequential problem, because in practice we know that, roughly speaking, measurements are caused by large things. This is how we have used the axioms of quantum mechanics so far, and it has worked reasonably well. Us not knowing just how large a device needs to be to induce a measurement hasn't really been an issue. However, the smaller the measurement devices we can manufacture become, the more pressing the question becomes.

That we cannot answer this question has practical consequences already. A few years ago, Frauchinger and Renner argued that quantum mechanics cannot consistently describe the use of itself [FR18]. But as was pointed out in [Rel18; ŻM21], the origin of the inconsistency is that Frauchinger and Renner did not specify what a measurement device is. They treated a measurement merely as a sufficiently strong correlation, which leads to a basis ambiguity that allows mutually contradictory results. The problem was directly created by not making a Heisenberg Cut.

This alleged inconsistency was later experimentally tested with a setup that, stunningly enough, used single photons as stand-ins for observers that supposedly make measurements [Pro+19]. Now, it may be a matter of debate just exactly where to apply the Heisenberg Cut, but Heisenberg would probably be surprised to learn that by 2018 physicists would have confused themselves so much over quantum mechanics that they came to believe single photons are observers. What the Frauchinger-Renner paradox therefore establishes is that quantum mechanics can result in inconsistent predictions so long as we do not add a definition for what a measurement device is to the axioms of quantum mechanics.

This problem could easily remedied—after all, we would just need to write down a definition. However, the definition for a measurement of course should not just remove the risk of inconsistent predictions but also agree with observations, and this just returns us to considering where to apply the cut.

It was argued in [Rel18; ŻM21] that the Frauchinger-Renner paradox can be resolved by taking into account decoherence. But decoherence is still a unitary and linear process that is described by the Schrödinger equation. It can therefore not give rise to the measurement update, so this still has to be added to the axioms.

Decoherence can to some extent be used to identify the circumstances under which the measurement update should be applied, but this idea has its problems too. We will comment on this in more detail in Section 7.3.4. For now, let us just note that decoherence alone simply will not evolve a system into a single detector eigenstate, and hence does not agree with what we observe. Tracing out the environment gives us a mixed state, but that is still not what we observe, not to mention that taking this trace is not a physical process, and therefore doesn't change anything about the state of the system.

The Classical Limit

Before quantum mechanics, there was classical mechanics, and classical mechanics still describes most of our observations correctly. Unfortunately, quantum mechanics doesn't correctly reproduce it.

It has long been known that recovering the classical time-evolution for suitably defined expectation values in quantum mechanics works properly only for integrable systems [BZ78; Zas81]. For chaotic systems, on the other hand, the quantum-classical correspondence breaks down after a finite amount of time [CR97; BGP99]. As pointed out by Zurek [Zur03] (see also [Ber01]), this time may be long, but not so long that we can't observe it. Zurek estimates that the chaotic motion of Hyperion (a moon of Saturn) would last less than 20 years if we used the Schrödinger evolution for its constituents. Alas, it has lasted hundreds of millions of years.

Again, decoherence allegedly solves the problem. If one includes the interaction of Hyperion with dust and photons in its environment, then one sees that the moon becomes entangled with its environment much faster than its motion could significantly deviate from the classical limit.

However, we have to note again that tracing out the environment is not a physical process. Therefore, all entanglement gives us is a very big entangled state. What we would have to do to get a classical non-linear motion of a localised object is to actually include the Collapse Postulate into the dynamical law. This shouldn't be so surprising: the non-linearity has to come from somewhere.

To do this, we would have to know when and how the collapse happens, but we don't. Do the photons detect the moon? Or does the moon detect the photons? If there's neither photons nor dust, does the moon detect itself? And if the state collapses, then just exactly when do we update what part of the moon's wave function? These are not philosophical questions; these are questions about how to apply the axioms of our quantum machinery, and they are questions that we simply do not have an answer to.

Another way to look at this problem was summarised by Klein [Kle11]: the $\hbar \to 0$ limit of quantum mechanics just does not reproduce classical mechanics, unless one restricts oneself to special states (generalised coherent states) and specific types of potentials.

Locality and Causality

The trouble with Hyperion brings us to the next problem. The collapse of the wave function in quantum mechanics is instantaneous; it happens at the same time everywhere in space.

This "spooky action at a distance" [EPR35] understandably worried Einstein because it seems incompatible with the speed-of-light limit. We know now [GRW80] that no information can be exchanged with the collapse of the wave function, but this doesn't explain how to apply the collapse postulate.

Consequently, people have debated for decades how to make the collapse compatible with relativistic invariance, and whether it requires backwards causation [Per98; MR02; Myr02]. No resolution has been reached.

We acknowledge, however, that the non-locality of the collapse is not a problem for the instrumentalist because, in the Copenhagen Interpretation, collapse is not necessarily a physical process, and is not related to any observable. So, in just which reference frame it happens does not matter; there are no predictions tied to this frame anyway.

The reason we mention locality and causality is that these matter when we cross from special to general relativity, as we discuss next.

Conservation Laws

In the Einstein Field Equations (EFEs)

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = T_{\mu\nu} \ , \tag{7.21}$$

where the entries of the stress-energy tensor, $T_{\mu\nu}$ are $\mathbb{C}$-valued functions [Ein15; MTW73]. In quantum field theory, however, the stress-energy tensor is operator-valued, $\hat{T}_{\mu\nu}$, as it derives from operator-valued observables. To insert the stress-energy tensor from quantum field theory into Einstein's field equations, one thus has to give quantum properties to the left side of the field equations, which would require us to develop a complete theory of quantum gravity. Attempts to develop such a theory have been made since the 1950s, but have yet to provide such a complete theory.

Another option is to instead convert the operator-valued stress-energy tensor into a $\mathbb{C}$-function. The most obvious way of doing this is to take its expectation value, $\langle \Psi | \hat{T}_{\mu\nu} | \Psi \rangle$, with respect to some quantum state $|\Psi\rangle$ that suitably describes the particle content of space-time. This is often referred to as semi-classical gravity.

Most important for our purposes is that the semi-classical approximation is generally believed to be at least approximately correct in the weak-field limit and if fluctuations of the stress-energy tensor are small [For82; KF93].

But the expectation value in the stress-energy tensor generically has to be updated upon measurement with the collapse of the wave function. And since this update is non-local, it violates the (contracted) Bianchi-identities which the left side of Einstein's field equations do fulfil and that are usually associated with local stress-energy conservation [Wei72; MTW73].

Take for example a photon (γ) which passes through a beam splitter (S). Due to momentum conservation, this creates an entangled state

$$\frac{1}{\sqrt{2}} \left(|\gamma(p)\rangle |S(-p)\rangle + |\gamma(-p)\rangle |S(p)\rangle \right) \ , \tag{7.22}$$

where p is the momentum and we will assume that it is really the mean value of a suitably localised wave-packet. The expectation value of this entangled state is a sum of localised momentum distributions going out in opposite directions from the beam splitter, while the mean momentum transferred to the beam splitter is zero.

According to the axioms of the quantum machine, when we measure the photon we have to update the wave function. In this moment, the momentum of one detector suddenly increases by p $(-p)$ and that of the beam splitter switches to $-p$ (p). The momentum is conserved, but how did it get to the detector? Without general relativity there is no observable tied to this question, but in general relativity there is: even though it's too small to measure, something must have happened with the space-time curvature. But what?

Again this is a problem for which we simply do not have an answer. We do not have the mathematics to describe what happens with the gravitational field of a particle if its wave function collapses, not even approximately.

However, this shows an issue with approaches to the measurement problem which claim that wavefunction collapse is just a nonlocal process (e.g. Shimony's 'passion-at-a-distance' [Shi93]): while they may work to some extent for non-relativistic many-particle quantum mechanics, they are difficult to reconcile with both relativistic quantum mechanics and gravity. And just letting go of relativistic covariance is not an option either since it is experimentally extremely well-confirmed [Mat05; KR11]. That is to say, while it cannot strictly speaking be ruled out that a nonlocal approach can be made "local enough" to agree with all available evidence, it seems like a stretch. The more pragmatic approach is to just look for an approach that is local and relativistically covariant to begin with.

7.3.3 What can we say about the solution?

Solution Requirements

From the previous section, we see that a satisfactory solution of the measurement problem must achieve the following:

Requirement 1. *Agree with all existing data.*

Requirement 2. *Reproduce quantum mechanics, including the Collapse Postulate (Axiom 4) and Born's Rule (Axiom 5), in a well-defined limit.*

Requirement 3. *Determine unambiguously what a measurement device is, at least in principle.*

Requirement 4. *Reproduce classical physics in a well-defined limit.*

Requirement 5. *Resolve the inconsistency between the non-local measurement collapse and local stress-energy conservation.*

Requirement 1 must be fulfilled by any scientifically adequate theory and we just add it for completeness. Requirement 2 recognises that within its domain, standard quantum mechanics is incredibly well-supported by data and a new theory would not become accepted without reproducing the achievements of its predecessor. Requirement 3 is necessary to resolve the problem of the Heisenberg Cut, Requirement 4 the problem of recovering the classical limit, and Requirement 5 the problems of locality, causality and local stress-energy conservation.

Requirements 2 and 4 have the same form—they both consist of ensuring the solution reproduces an established, well-proven theory in some well-defined limit. Their resolution might well be related; however, this does not have to be so, which is why we list them separately. Likewise, we would expect that Requirements 2 and 4 can be used to show that Requirement 1 is fulfilled. Again, however, this does not have to be the case—a limit might be well-defined and yet its result might just be in conflict with observations—so we list them separately.

We added the phrase "at least in principle" in Requirement 3 to make clear that no one expects it to be of much practical use to calculate from first principles what the arrangements of elementary particles in a detector are. To find out what a detector is, it is much more practical to just test whether it actually detects the thing we want to detect. However, even though it may be unpractical or even unfeasible to perform an exact calculation, we should be able to identify some general properties for what it takes for a collection of particles to act as a measurement device.

As mentioned previously, Requirement 5 might be resolved by a theory of quantum gravity. However, the currently most well-developed approaches to quantum gravity do not address the measurement process. Similarly, these requirements could feasibly be met by a theory which isn't also a theory of quantum gravity. Therefore, we emphasise a theory of quantum gravity is neither necessary nor sufficient to solve the measurement problem.

Solution Properties

So what does it take to meet these requirements? Requirement 5 requires a local evolution law that connects the initial state of the system with the observed measurement outcome. This means that either

Solution 1. *The wave function itself evolves according to Axiom 3 and is updated according to Axiom 4, but it is is an incomplete description. The physical state of the system is described by something else that evolves locally though not necessarily deterministically.*

or

Solution 2. *The wave function is the complete description of the system, but does not evolve by Axiom 3 and Axiom 4. It evolves instead according to a different, local evolution law, that is necessarily non-deterministic.*

or

Solution 3. *A combination of Solutions 1 and 2.*

Solution 1 is what is commonly called a "hidden variables" theory. We will instead adopt the convention of [Mau95] and refer to it as an "additional variables" theory, to acknowledge that the variables may not, in fact, be hidden—the variables are merely not included in axioms 1-6. It should be noted that according to this definition, details of the detector and environment (as they appear in the decoherence approach) count as additional variables. We will come back to this point later.

For the instrumentalist, all three solutions come down to local, deterministic evolution laws with additional variables. We here use the definition for deterministic from [CW12], which is that measurements have definite outcomes (probability 0 or 1). Using this terminology, the laws may be non-predictable, despite being deterministic.

This can be seen by noting that any local, non-deterministic evolution law can be rewritten into a local, deterministic evolution in a theory with additional variables: any time the evolution law is indeterministic (which could be continuously), we encode the possible time-evolutions with an additional variable.

To give a concrete albeit trivial example: we can make the Collapse Postulate deterministic by just using the eigenstate that is the outcome of the collapse as an additional variable. The collapse is then "determined" by the "additional variable". Of course, this is somewhat pointless, because in quantum mechanics this additional variable is unpredictable (hence would well deserve the name "hidden"), but it serves to show that additional variables *could* be used to describe the process.

The difference between the solutions is then merely what we call the "wave function". Do we reserve the term for that which evolves under Schrödinger equation (Solution 1) which has to come out of the mathematics at least in some limit (Requirement 2), or are we okay with using the term for whatever it is that we use to describe the state of the system (Solutions 2 and 3)? Do we require wave functions to be elements of a Hilbert-space, regardless of their evolution law? Do we want them to factorise for separable systems? Do we want them to be green, or married, or all of the above? In the end, this is just a matter of definition. The instrumentalist doesn't care and concludes that all possible solutions to the problem can be covered with a local and deterministic evolution in a theory with additional variables.

This, together with the undeniable resemblance between the von Neumann-Dirac equation and the Liouville equation, makes it plausible that quantum mechanics is indeed an ensemble description of an underlying statistical theory with the additional variables [HH22a].

Solution Parameterisation

That this new underlying theory must explain just where the Heisenberg Cut is (Requirement 3) means that it has to bring in some new parameter to quantify how good the statistical approximation is. This parameter cannot be derived from quantum mechanics itself; it must be extracted from experiment.

It is clear that the new transition parameter cannot be something as simple as just the number of particles, not least because that quantity is not in general well-defined. (How many particles are inside an atom? How many particles does the vacuum contain?) Seeing the problems of the decoherence approach (see 7.3.4) it is also unlikely to be any measure of decoherence or entanglement, though the frequency or strength of interactions must play some role. Parameters based on the total mass or energy or gravitational self-energy have no *a priori* relation to generic measurements, and parameters based on the classical limit (possibly using the action as a quantifier) create a circular problem, because we would need them to know how the classical limit works to begin with.

However, we have reason to be optimistic that this problem is solvable because we have observational limits on the Heisenberg Cut both from above and below, and current technologies are pushing both these limits: by bringing larger objects into quantum states, while at the same time shrinking the size of detectors. It is only a matter of time until experiment will reach a regime in which deviations from quantum mechanics become noticeable. However, this process could be accelerated if we knew better what to look for, which is a task for theory development.

7.3.4 Solution Attempts

Second Quantization: The problems we discussed in the previous sections do not disappear with second quantisation. All quantum field theories are built on the basic axioms of quantum mechanics which we listed in Section 7.3.1. In quantum field theory we merely have more complicated ways of describing interactions, and calculating the time-evolution of the system and observables related to it. If anything, it has been argued these problems become more complicated in quantum field theory [Gri22].

Understanding Quantum Mechanics: Some readers may wonder if it is possible that these problems will one day be solved within the context of quantum mechanics itself. Maybe the problem with the classical limit is just that no one has found the right limit. Indeed, it is widely known that generalised coherent states make a promising basis for taking the classical limit [Yaf82; Lan06]. But, in light of the problems pointed out by Klein [Kle11], this would at least entail adding further axioms about what to do for obtaining the classical limit. Within the context of this present argument we would therefore have to consider it a new theory, because its set of axioms would not be equivalent to the one we listed earlier.

Wave Functions as Epistemic States: One common strategy to explain away the measurement problem is to argue that the wave function is not an ontic but an epistemic state, and that its collapse is not a physical process. The collapse, so the argument goes, is merely an update of our knowledge about the system, and its non-locality therefore should not worry us. After all, as Bell put it, when the Queen dies, Prince Charles will instantaneously become King, yet no information had to be sent non-locally for this update [Bel04a].

A common strategy to counter this argument is to point out that if the wave function is the complete description of the system, then there is nothing else the wave function could describe knowledge about [Mau95]. Therefore, most wave function-epistemic views require the wavefunction to be an incomplete representation of the system.

While attempts have been made at formalising this view [Spe07; HS10] and using it to come to no-go theorems on the wavefunction being in some way epistemic [PBR12; PPM13; RLE20], these formalisations have fundamental issues [SF12; OL20; HRL22]. Therefore, these no-go theorems contribute little to telling us whether the wavefunction can or cannot be epistemic.

In reality, the two sides of this argument haven't much advanced since the debate between Einstein and Bohr, and at this point it seems unlike they ever will. Let us therefore note that the problems listed in Section 7.3.2 exist regardless of whether one believes the wave function is epistemic or ontic or how one wants to interpret the collapse. Quantum mechanics is unsatisfactory for the instrumentalist simply because we cannot answer questions about physical reality with it: just what properties does a measurement device need to have? Just what happens with space-time curvature when a photon passes through a beam splitter? Saying that the wave function is epistemic doesn't answer these questions.

Decoherence: The virtues of the decoherence program can be briefly summarised as follows. Given any system that includes a prepared state, detector, and environment, a detector is a subsystem that can keep a record of at least one aspect of another subsystem, which is the prepared state one wants to measure. To be able to keep a record of (some property of) the prepared state, the detector itself must have states that are stable under interaction with the rest of the system, which is the environment.

These stable detector states are often called 'pointer states' and we will denote them with $|I\rangle$. They keep a record of the prepared state's projection on the eigenstate corresponding to the pointer state, i.e., one gets a product state $|I\rangle|O_I\rangle$. Any superposition of pointer states would rapidly decohere under interaction of the environment, hence not keep a record of what we are interested in.

To describe the process of decoherence formally, one takes the density matrix of pointer states, prepared states, and environment. One estimates how quickly they become entangled and how much this affects the relative phases between the pointer states. This is the process of decoherence. It must be stressed that this process is fully described by the Schrödinger equation. After that, one traces out the environment, and obtains a mixed state whose probabilities are given by Born's rule.

What one learns from this is that a useful detector, loosely speaking, must be large enough so that superpositions of its pointer states rapidly decohere. Decoherence hence gives us a criterion for identifying detector pointer states by what Zurek termed 'einselection' by the environment [Zur03].

But we do not observe mixed pointer states any more than we observe their superpositions. We only observe detectors in pointer states. In terms of the density matrix, we observe a matrix that has one entry equalling one, somewhere on the diagonal, and all other entries equal zero. The result of decoherence, however, is a density matrix with the Born probabilities on the diagonal.

Thus, while decoherence explains why we do not measure cats that are in a superposition of dead and alive, it does not explain why we do not measure cats that are 50% dead and 50% alive (a classical mixture) either [Sch05]. To agree with observations, the wave function, or its density matrix, respectively, must therefore still be updated upon measurement.

Another way to see that decoherence does not solve the measurement problem is noting that it is based on counterfactual reasoning: the typical initial state of the system will, under the Schrödinger equation, evolve into a final state that is highly entangled with the environment, susceptible to decoherence, and hence not what we observe. According to the decoherence program, the state we observe is instead an (almost) decoherence-free subspace which is exactly what we generically do *not* expect. But the decoherence program gives us no clue as to how we get from evaluating the amount of decoherence in a state we do not observe to the not-decohering state we do observe.

This discrepancy raises whether the notion of entropy we use in quantum mechanics, that increases under an evolution that we do not actually observe, can possibly be correct. Indeed, if the wave function describes an ensemble average, we do not expect a notion of entropy derived from it to be meaningful.

The decoherence program suffers from another problem, as pointed out by Kastner [Kas14]. It requires one to specify a division between the observed system, the detector, and the environment already. Without that division, one does not know what the environment is that one should trace out. For this reason, decoherence does not allow us to define what a detector is. It merely quantifies certain properties that we know detectors do have.

We do not mean to deny the usefulness of studying and quantifying decoherence and entanglement. But they can ultimately not solve the problem of how to define a measurement, because these properties are basis-dependent. They will just reformulate the question into one about the choice of basis or the division into subsystems, respectively.

Many Worlds: The many-worlds interpretation [Eve57; DG15] and similar approaches are often claimed to be simpler than the Copenhagen Interpretation [Fay19] because they do not require the Collapse Postulate. That we only observe detector eigenstates is allegedly explained by the branching of the wave function and is supposedly a consequence of the Schrödinger equation alone.

Alas, this is just not the case. To make a prediction for a measurement outcome in a many-words approach, one has to replace the Collapse Postulate with sufficiently many assumptions that achieve the same. Those are normally stated as assumptions about what constitutes an observer or a detector or a branching event [SC18]; in any case, they are clearly not any simpler than the Collapse Postulate.

The easiest way to see that many worlds does not do away with the Collapse Postulate is to note that if it was possible to make predictions for our observations using the Schrödinger equation alone, then this would be possible in any interpretation of the mathematics. One therefore clearly needs the Collapse Postulate or at least equivalent assumptions in many worlds, regardless of how they are called or interpreted.

This is not to say that many world approaches are wrong. From the instrumentalist perspective, they are as good or as bad as the Copenhagen Interpretation. Anybody who doubts this statement is strongly encouraged to make a prediction with the many worlds interpretation and try to figure out how this differs from one made with the Copenhagen Interpretation.

Bohmian Mechanics: Bohmian Mechanics [Boh52a; Boh52b] comes in two different versions, one in which the equilibrium hypothesis is counted as an axiom, and one in which it is not an axiom, but merely approximately fulfilled in the situations we typically observe in the laboratory.

Bohmian Mechanics with the equilibrium hypothesis is mathematically equivalent to the Copenhagen Interpretation in the sense that one can be derived from the other. They make exactly the same predictions. Bohmian Mechanics is usually formulated in position space, but one can easily extend this definition just by requiring it to respect invariance under basis transformations.

Since Bohmian Mechanics with the equilibrium hypothesis is equivalent to the Copenhagen Interpretation, it cannot solve the problems laid out in Subsection 7.3.2. It adds rather than removes non-locality, does not give us any clue about how to define a detector, and doesn't help us take a classical limit.

The reason Bohmian Mechanics is often said to solve the measurement problem is that the outcome of the time-evolution, interpreted suitably, is a detector eigenstate. In Bohmian Mechanics, one has a distribution of particles but interprets the actual ontic state of the system to be only one of them. Loosely speaking, Bohmian Mechanics combines the Schrödinger evolution and the Collapse Postulate to one local evolution for the particle and a non-local one for the guiding field. Since by assumption there is only one particle in the initial distribution, there is only one final outcome.[1]

This solution however only works if one measures positions, so if one wants this solution to go through one has to argue that the only thing we ever measure are really positions of particles and everything else is derived from that. Given that the Collapse Postulate also brings the system into a detector eigenstate, one thus doesn't gain any advantage from switching to Bohmian Mechanics, one just gets this new headache. Further, since the ontology of Bohmian mechanics is itself non-local, it makes it even more difficult to conceive of a solution to the measurement problem.

Again, this is not to say that Bohmian Mechanics is wrong. Being equivalent to the Copenhagen Interpretation, it is isn't any better or worse: Nikolic proposed comparing Bohmian Mechanics to the Coulomb gauge of electrodynamics [Nik06]. It seems non-local, and though it doesn't give rise to non-local observables, the explicit non-locality makes it difficult to generalise the formulation. It is quite possibly for this reason that quantum field theories based on Bohmian

[1]Though this leaves one with the long-standing problem of the empty waves.

mechanics have been complete non-starters [Dür+04; Dür+14]. See also [Wal22] for more about the difficulty of generalising different interpretations of quantum mechanics to quantum field theory. Bohmian Mechanics may suffer from more severe problems than its failure to solve the measurement problem (see e.g. [Ein53; Ein11; Hel19]), but we will not discuss these here because it is not relevant for our purposes.

Bohmian mechanics without the equilibrium hypothesis [Val91a; Val91b] is distinct from quantum mechanics, and to our best knowledge it has not helped solving the measurement problem. Since it has to reproduce the equilibrium hypothesis in the situations we typically encounter in the laboratory, it is also implausible that it would be of use.

Other Interpretations: At this point, it should be clear that the problem can't be solved by reinterpreting the maths: we actually need new maths. If we cannot derive what a measurement device is, or what the source of gravity is in one interpretation, we can't do it in any interpretation. This means that QBism [Fuc10; FS14b; Fuc17], the modal interpretation [DV98], the previously mentioned statistical interpretation [Bal70], the transactional interpretation [Cra86; Kas13], Rovelli's relational interpretation [Rov96; AR22], or Smolin's ensemble interpretation [Smo12], or any other reinterpretation of the mathematics cannot actually solve the problems we laid out in Section 7.3.2. Those are not just questions whose answers are difficult to calculate; they're questions whose answers can't be calculated in quantum mechanics—regardless of its interpretation.

We want to stress however that we certainly do not mean to say that it is useless to think about different interpretations of quantum mechanics. This is because some interpretations may make it easier to answer certain questions. A good example is the arrival time, a quantity that is notoriously difficult to calculate in the Copenhagen Interpretation, but that was recently successfully calculated using Bohmian Mechanics [DD19].

Collapse Models: Collapse Models (be they gravitational, such as the Penrose-Diosi model, or spontaneous [BG03], such as the GRW [GRW86] and CSL [GPR90] models) have a chance to actually solve the problem, because they are not just reinterpreting the same mathematics. In accordance with what we discussed in 7.3.3, they all bring in new parameters that quantify the deviation from standard quantum mechanics. However, the currently existing collapse models run into a well-known problem: Bell's theorem.

Remember that we can interpret any non-deterministic evolution as a deterministic evolution with additional hidden variables. This means, so long as collapse models fulfil the assumptions for Bell's theorem, they have to violate local causality (or they cannot reproduce observations). The currently used collapse models are therefore either still non-local [Tum06], or the evolution law explicitly contains the basis that the evolution collapses into. In the latter case they either violate statistical independence, or one is forced to assume that measurements can only be made in one particular basis (usually the position basis).

7.3.5 The role of statistical independence

We have argued above that quantum mechanics suffers from several problems, and that any solution to the problem can be expressed as a local, deterministic theory with additional variables.

However, we can measure different observables of the same prepared state, and different observables correspond to different detector pointer states. If the evolution into the detector eigenstate is to be local, then the additional variables which determine the outcome must contain information about the pointer states from the outset. If the prepared state only gets this information by the time it arrives at the detector, then the collapse will generically have to be faster than light—this is exactly what happens in standard quantum mechanics.

If we want to avoid this, then the additional variables, commonly denoted λ, must be correlated with the measurement settings. This is known as a violation of statistical independence, or a violation of measurement independence. If ρ is the probability distribution of the hidden variables, and X are the detector settings (of possibly multiple detectors), then a violation of statistical independence means $\rho(\lambda|X) \neq \rho(\lambda)$. Theories with this property have been dubbed superdeterministic [HP20; Hos20].

To our best knowledge we do not presently have any theory that fulfills requirements 1-5, but from our above arguments we do know that any such theory has to violate statistical independence. Seeing that most of what we argued above (except the reference to Frauchinger-Renner) could have been said 50 years ago, it is curious that progress on this has been so slow. We would like to offer some thoughts on why that may be.

One reason we don't yet have a solution to the problems of Subsection 7.3.2 is almost certainly that so far it just wasn't necessary. Quantum mechanics in its present form has worked well, made a stunning amount of correct predictions, and for the most part saying that we know a detector when we see one is sufficient to make predictions with quantum mechanics' mathematical machine.

Let us not forget that it wasn't until 1982 [AGR81] that violations of Bell's inequality were conclusively measured. Until about a decade ago, interpretations of quantum mechanics were discussed primarily by philosophers, simply because they weren't relevant for physicists. The re-approach between philosophy and physics in quantum foundations is a quite recent development, and it has been driven by technological advances.

Even now, the problems that we outlined in Section 7.3.2 have grown of interest to the instrumentalist, but not yet for the experimentalist. This is about to change though. Soon, experiments will probe into the mesoscopic range, and investigate when a device ceases to be a useful detector. Further, tests of the weak-field limit of quantum gravity are on the way [Bos+17; MV17]—granted, the latter experiments weren't designed to test the Collapse Postulate, but they examine the parameter range where the questions raised above are also relevant.

Another potential reason why we haven't yet managed to solve the measurement problems is that the only viable option—violating statistical independence—was discarded 50 years ago on purely philosophical grounds. For peculiar reasons, statistical independence became referred to as the ‘free will’ or ‘free choice’ assumption, which has discouraged physicists from considering that this assumption may just not be fulfilled, and hence that a local description of the measurement process may still be possible. It has already been explained elsewhere [HP20; Hos20] that this terminology is meaningless; whether statistical independence is violated or not bears no relevance for the existence of free will. The option of violating statistical independence was so strongly discouraged that as of today many physicists do not even know that Bell’s theorem does not generally rule out local and deterministic completions with additional variables [SKZ13].

A third reason that probably added to the lack of interest in the option is that the additional variables are often interpreted as new degrees of freedom that reside inside elementary particles. This possibility is strongly constrained by experiments, and has only become more unappealing the more thoroughly we have tested the Standard Model of particle physics.

This, however, is based on a misunderstanding. The additional variables don’t need to be new variables, and they don’t need to reside on short distance scales; they merely need to be variables that don’t appear in the standard axioms A1-A6. Further, as we mentioned above, the additional variables may merely be a stand-in for a non-deterministic (or non-computable) evolution law [HHP22].

Most importantly though, there is no reason why the additional variables must be located inside particles or located anywhere for that matter. They could for example be the details of the detector or the environment or more generally variables that quantify large-scale properties or correlations. This means the solution we seek for may not to be found on the route of ontological reductionism that has preoccupied thinking in the foundations of physics for the past century (i.e., building bigger particle colliders won’t solve this problem).

Indeed, as we have argued in [HH22a], the additional variables in Bell’s theorem are better interpreted as labels for trajectories (which also explains why they can alternatively be understood as encoding a non-deterministic evolution law). Whatever the reason for the slow progress, we think that the problems we have laid out here are eminently solvable with existing mathematics and will become accessible for experimental test in the near future.

7.3.6 What is it good for?

The brief answer to what solving the measurement problem is good for (besides solving the problem, that is), is that we don’t know. However, we can make some speculations.

For one thing, it might be that the underlying theory which solves the measurement problem turns out to be deterministic again, and explains the seeming indeterminism of quantum mechanics as being epistemic in origin. In this case, it stands to reason that the theory would allow us to overcome limits and bounds set by quantum mechanics on measurement accuracy with suitably configured experiments. This could turn out to be useful for many things, not least for quantum metrology and quantum computing.

However, it could turn out the go the other way. If deviations from the Schrödinger equation become important for, say, quantum computers beyond a certain number of qubits (as they are in IST [Pal20a]), then maybe large quantum computers will be impossible [Sla21; HPR21].

More generally, understanding in which cases a measurement process occurs and just what happens would almost certainly improve our ability to control quantum states.

That is to say, solving the measurement problem is not just a philosophical enterprise. Its solution will quite possibly have technological applications.

We have identified several different aspects of the measurement problem in quantum mechanics, and argued that solving these problems requires a new theory. Deviations from quantum mechanics will likely become experimentally accessible in the near future. However, a theoretical understanding of the measurement process could greatly speed up the discovery. We have listed five requirements that any satisfactory solution of the measurement problem must fulfil.

7.4 Chapter Conclusion

In this Chapter, we went through the necessary background for this second Part of the thesis, which investigates (specifically statistical independence-violating) extensions of quantum mechanics. For this, we first introduced Bell's Theorem, and looked at both the assumptions necessary to formulate a Bell inequality, and the experimental loopholes which can allow models which still meet all these assumptions to violate such an inequality. We then looked specifically at the statistical independence assumption, and investigates proposed extensions of quantum mechanics where this assumption is violated. Secondly, we moved onto looking at the measurement problem in quantum mechanics, illustrating how it arises from the axioms of quantum mechanics, the requirements a solution to the problem must possess, and why currently-proposed interpretations and extensions of quantum mechanics fail to solve this problem.

This Chapter, and Part overall, serve to allow us to survey and illustrate cases where current or developing quantum technologies can be used to test foundational ideas—the second part of the interplay between quantum foundations and quantum technologies. In this Chapter, this is most visible in the discussion of Bell tests, and the loopholes in such tests. However, Chapter 9 will also show this, as we consider ways to experimentally test certain statistical independence-violating models.

CHAPTER 8

EPISTEMIC AND ENSEMBLE INTERPRETATIONS OF THE WAVE FUNCTION

Declaration of contribution: This Chapter is adapted from the published papers *Could wavefunctions simultaneously represent knowledge and reality?* [HRL22], which was conceived by myself and Prof James Ladyman, and written by myself, supervised and edited by Prof Ladyman and Prof John Rarity; and *The wave function as a true ensemble* [HH22a], conceived and written by myself and Dr Sabine Hossenfelder. This second paper was written at a later stage of my PhD, when I had been developing independent collaborations, and my supervisors were happy for this independent work to be included in my thesis.

8.1 Chapter Introduction

In this chapter, we look at a recently-proposed pair of definitions for characterising interpretations and extensions of quantum mechanics. By these definitions, models that meet a specific mathematical criterion have the wavefunction represents the state of the world, and are called ψ-ontic. Models that do not meet this mathematical definition supposedly have the wavefunction represent knowledge, and are called ψ-epistemic. We show that nothing in the informal ideas of the wavefunction representing the state of the world or representing the knowledge of an experimenter require these to be contradictories, and so that this mathematical criterion fails to adequately reflect whether the wavefunction should be described as representing reality, or representing knowledge (or both).

We further argue that the ψ-ontic/ψ-epistemic distinction fails to properly identify ensemble interpretations. In these interpretations, the wavefunction predicts probabilities of possible measurement outcomes, but not which individual outcome is realised in each run of an experiment.

We therefore propose a more useful definition, where the wavefunction describes an ensemble of states with different values of a hidden variable (ψ-ensemble). All local ψ-ensemble interpretations which reproduce quantum mechanics violate Statistical Independence. We then show how these interpretations can help make sense of otherwise puzzling phenomena in quantum mechanics, such as the delayed choice experiment, the Elitzur-Vaidman Bomb Tester, and the Extended Wigner's Friends Scenario.

This chapter is foundational, and so does not show the links between quantum foundations and quantum technologies. However, it reinforces the importance of the next chapter, showing why we should be interested in statistical independence-violating extensions of quantum mechanics. This illustrates why we should be interested in trying to develop experimental tests to distinguish these extensions from standard quantum mechanics.

8.2 Could wavefunctions simultaneously represent knowledge and reality?

In this section, we examine Harrigan and Spekkens's definitions for the wavefunction being ψ-ontic and ψ-epistemic, and argue that their definitions do not reflect the standard usage of the terms ontic and epistemic. Harrigan and Spekkens formally define the wavefunction being ψ-ontic or being ψ-epistemic as contradictory, such that the wavefunction can only be one or the other [HS10]. However, as we show in Subsection 8.2.1, the informal ideas of pertaining to what exists, and pertaining to knowledge do not exclude each other. Subsection 8.2.2 explains Harrigan and Spekkens's formalisation, and shows how their definitions of ψ-ontology and ψ-epistemicism simply presuppose that the wavefunction cannot pertain to both knowledge and reality. If, as we argue, it is right that nothing about the informal ideas of epistemic and ontic interpretations rules out wavefunctions pertaining to both knowledge and reality, then the now standard definitions create a false dichotomy[1]. Subsection 8.2.4 considers the consequences of this for discussions of quantum foundations.

8.2.1 Ontic and Epistemic Interpretations of Terms in Physical Theories

Quantum mechanics describes the behaviour of subatomic particles. Unlike classical mechanics, which attributes particles particular values of position and momentum, quantum mechanics attributes wavefunctions, from which can be derived probabilities for possible values of position, momentum and other observables. In the quantum case the products of the standard deviation of position and the momentum (and other pairs of incompatible observables) have a minimum value given by the uncertainty principle [Ken27], unlike in the classical case. However, in some

[1]Myrvold independently noted that taking ψ-epistemic to be simply the negation of ψ-ontic is 'potentially misleading' [Myr20] and seems to take this to be obvious. However, it is not a prominent point in his paper, and is worth arguing here since some people claim it is not just not obvious but false.

ways the situation is similar to how a probability distribution in classical statistical mechanics can be interpreted as describing what is known about microstate of the particles in a gas given its macrostate, and so as ‘epistemic’. Quantum mechanical wavefunctions behave like classical ensembles in the limit, and so perhaps the wavefunction in quantum mechanics is epistemic in a similar way. On the other hand, Fuch’s Quantum Bayesianism (QBism) treats the wavefunction as representing knowledge of possible results, rather than of the underlying microstate [Fuc02; FS14b] and so as epistemic, but in a different way, while Ben-Menahem’s interpretation holds “quantum probabilities as objective constraints on the information made available by measurement" [Ben17]. All these responses to the peculiarities of quantum mechanics—phenomena such as contextuality, entanglement and collapse—do not posit any kind of novel metaphysics and are exclusively epistemic interpretations.

By contrast, dynamical collapse theories [GRW86; Pen96], Bohmian mechanics [Boh52a; Boh52b] and Everett’s relative-state interpretation [Eve57], take the wavefunction to represent something physical, and not to be epistemic in any way.[2]

The notions of epistemic and ontic have wide application in philosophy. For example, an epistemic interpretation of probability takes it to be a matter of belief, information or knowledge, while an ontic interpretation of probability takes it to be something in the world and nothing to do with the representations of agents. In this and many other cases, epistemic and ontic interpretations are taken to be mutually exclusive. In general however, ‘epistemic’ means pertaining to cognition or knowledge, while ‘ontic’ means pertaining to what is exists (the words are directly based on Greek genitive forms of the words for cognition or knowledge and reality respectively), and there is clearly no contradiction in predicating both of the same thing without the assumption that nothing can be both epistemic and ontic. There is no argument for such an assumption in general of which we are aware, and indeed it is arguably false in particular cases.

For example, the thermodynamic entropy S of a system is a candidate for being both ontic and epistemic. Arguably, the entropy assigned to a system depends on knowledge of it as well as how the system is objectively. For instance, the entropy assigned to an equally divided box of particles is higher if an observer can distinguish these particles than if they cannot [Che09]. David Wallace [Wal14] argues that thermodynamics should be interpreted as a ‘control theory’ in which what agents know about systems is essential to what they can do with them. Arguably, at least, the entropy of a system represents the system as it is independently of us to some extent, while also representing what is known about it [LPS08].

Another example is that of proper mixtures in quantum mechanics, which represent both systems and what is known (or not known) about them. For example, the (proper) mixed state $({}_x|\uparrow\rangle\langle\uparrow|_x + {}_x|\downarrow\rangle\langle\downarrow|_x)/2$, is assigned to a system when it is really either in the state $|\uparrow\rangle_x$ or $|\downarrow\rangle_x$), and when the epistemic probability of state $|\uparrow\rangle_x$ and of state $|\downarrow\rangle_x$ is a half. Clearly, proper mixtures pertain to both what exists and what is known about it. Similarly, one can imagine another

[2] Recently an epistemic interpretation of the wavefunction in Bohmian mechanics has been proposed by [Esf+14].

observer validly assigning the state $(2_x |\uparrow\rangle\langle\uparrow|_x +_x |\downarrow\rangle\langle\downarrow|_x)/3$, illustrating that these states are to some degree epistemic—but given they are also based on the knowledge the system is actually in one of the two eigenstates of spin-x with certainty (we are just unsure which), they are also ontic.

However, the terminology of 'epistemic' and 'ontic' that has become standard among physicists working in quantum foundations is understood so as to make the application of the terms mutually exclusive without an argument being given that they should be. In his influential review of the PBR theorem, Matt Leifer says that an ontic state represents 'something that objectively exists in the world, independently of any observer or agent' [Lei14]. This is stronger than the definition above because it rules out that the state also represents our knowledge of the world. He gives the example of the ontic state of a single classical particle being its position and momentum. His corresponding example of an epistemic state is then a probability distribution over the particle's phase space. Note that in the case Leifer considers there is clearly a many-one map between the epistemic state and the underlying state, in the sense that many epistemic states are compatible with the true state of the system (because there are many probability distributions that give a non-zero probability to a given state of the particle). However, this is not so for the ontic state. Only one mathematical representation is compatible with the underlying state of the system, because the mathematical representation of each possible microstate is a complete specification of all the degrees of freedom of the system over which the probability distribution is defined. As the next section explains, Harrigan and Spekkens make this difference between a many-one and one-one relationship definitive of epistemic and ontic interpretations respectively.

Simon Friedrich says a (ψ-)epistemic interpretation is any view according to which "quantum states do not represent features of physical reality, but reflect, in some way to be specified, the epistemic relations of the agents who assign them to systems they are assigned to" [Fri14]. So he (and Leifer) agree that epistemic and ontic interpretations are exclusive of each other, so that an epistemic interpretation of a term rules out that the term represents the world as well. This is so for Harrigan and Spekkens too, as explained in the next section.

However, not everything in physics is as straightforward as the toy model Leifer considers, and it is not always clear whether terms in physical theories represent real things (for example, component versus resultant forces, gauges and so on). In quantum mechanics, individual particles do not in general have their own pure quantum states. Nothing about the idea of a determinate reality requires that all physical states are decomposable into the states of particles, and the states of quantum field theory do not take individual particles to be the fundamental bearers of physical properties. The probability distributions we get from quantum mechanics are over eigenvalues associated with observables, not over underlying states-of-the-world as in classical statistical mechanics. For all these reasons and more the framework and definitions of the next section that have become orthodox are at least questionable[3].

[3]There is a long history of interpretations of the wavefunction as both ontic and epistemic, which shows precisely

Moreover, it is possible to give definitions of ontic and epistemic, which capture the informal usage of the terms reasonably well, and which do not make them exclusive of each other. Wavefunctions, like particle coordinates or vector fields, are part of the mathematical apparatus of physical theory. Such a term in a physical theory has an ontic interpretation (is an 'ontic state') when it is taken to represent how the world is independently of our knowledge of it *to some extent or other*.[4] For example, the number 9.81 represents the acceleration due to gravity near the surface of the Earth. It is not perfectly accurate and depends on the conventional choice of units, but arguably all representations are like this to some extent. Many mathematical terms used in classical physics, whether scalars, like mass or charge, or vectors, such as magnetic field strength or linear momentum, are apt to be taken as representing real physical properties. On the other hand, a term in physical theory has an epistemic interpretation (is an 'epistemic state') when it is taken to represent knowledge or information about the world *to some extent or other*. (Note that the positive second part of Friederich's definition of a (ψ-)epistemic interpretation is very similar to ours.)

8.2.2 The Ontological Models Framework

The wavefunction ψ gives the relative (Born) probability $p_A^\psi(s)$ that, upon measurement, a certain degree of freedom (described by observable A) will have the specific value s. The wavefunction can be represented as a normalised vector $|\psi\rangle$ in a Hilbert Space $\mathscr{H}$, where there is a complete orthonormal basis of vectors, corresponding to possible values of A (and we write vector $|s\rangle$ with eigenvalue s for A). Harrigan and Spekkens's Ontological Models framework considers how the wavefunction relates to some particular underlying real state of the world λ, from the space Λ of possible such states [Spe05; Rud06; HR07; Spe07; HS10].

In any hidden variable model, each observable A has an associated response function $\mathscr{A}(S|\lambda)$, which gives the probability for state λ that a measurement of variable A would give an outcome s in set S [SF12]. $p(\lambda|\psi)$ is the probability a situation described by wavefunction ψ has the underlying state λ. This probability function is normalised over ψ's support in state space Λ_ψ. [5] The framework assumes these are both probabilities (i.e., non-negative and additive) rather than quasiprobabilities (which need not be either) [HRA07].[6] From these we get the Born rule,

$$p_A^\psi(S) = \sum_S |\langle s|\psi\rangle|^2 = \int_\Lambda \mathscr{A}(S|\lambda)p(\lambda|\psi)d\lambda \tag{8.1}$$

why it is not appropriate to treat these two terms as mutually exclusive [Boh28; Fri14; Hei58; Sch52]. However, for some reason, this has been neglected by Harrigan and Spekkens, who define ψ-ontic and ψ-epistemic as contradictories.

[4]That this is not to say that wavefunctions are physical things, but that they represent something physical, just as the correspondence between your fingers and the numbers one to ten does not make the numbers themselves physical [Her21; SF12; WT10].

[5]The 'support' of a probability distribution is the set of all the values to which it assigns a non-zero probability.

[6]Note, Harrigan and Rudolph extend this framework further by including a treatment of preparation and measurement devices.

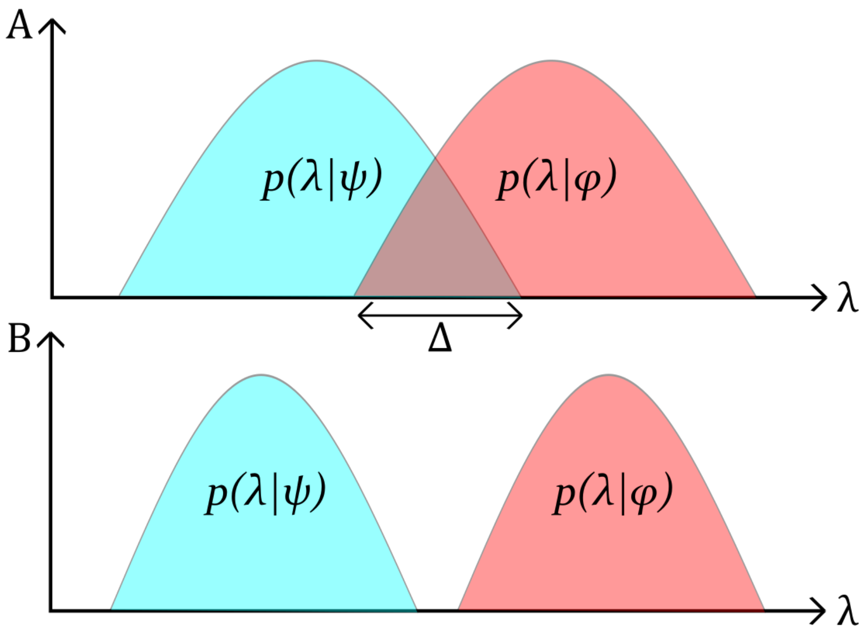

Figure 8.1: Harrigan and Spekkens's ψ-epistemic (A) and ψ-ontic (B) models of reality. Wavefunctions ψ and φ each have probability distributions over state-space $\Lambda = \{\lambda\}$. In a ψ-epistemic model, these can overlap over a subspace Δ, so a state λ within this overlap could be represented by both ψ and φ. However, in their formalism, in ψ-ontic models, each state can only be represented by one wavefunction.

Imagining our (arbitrary) observable A had the original wavefunction $|\psi\rangle$ as a possible state after measurement (i.e., $|\psi\rangle = |s\rangle$ for some $|s\rangle$ that is an eigenstate of A with some value q), we find

$$\forall \lambda \in \Lambda_\psi, \mathscr{A}(q|\lambda) = 1 \tag{8.2}$$

(as $|\langle\psi|\psi\rangle|^2 = 1$, and $p(\lambda|\psi)$ is normalised over Λ_ψ [Mar12]).

If we now consider a second wavefunction, φ, with probability distribution $p(\lambda|\varphi)$ over its support Λ_φ, obeying these same rules, then, by Eq. 8.1,

$$|\langle\psi|\varphi\rangle|^2 = \int_\Lambda \mathscr{A}(q|\lambda)p(\lambda|\varphi)d\lambda \tag{8.3}$$

By restricting this to the subspace Λ_ψ, we define

$$\begin{aligned}\Delta &\equiv \int_{\Lambda_\psi} \mathscr{A}(q|\lambda)p(\lambda|\varphi)d\lambda \\ &= \int_{\Lambda_\psi} p(\lambda|\varphi)d\lambda \leq |\langle\psi|\varphi\rangle|^2\end{aligned} \tag{8.4}$$

If $\Delta > 0$, Harrigan and Spekkens claim this overlap of $p(\lambda|\varphi)$ and Λ_ψ implies the two wavefunctions in this model are epistemic states, because the same underlying state λ could be represented by either depending on what it known about the system.[7] Harrigan and Spekkens take the existence of two distinct wavefunctions with overlapping support within a model as necessary and sufficient for that model being ψ-epistemic (and not ψ-ontic). This is as, given two different wavefunctions can represent the same underlying state, the wavefunction is not determined completely by the underlying state. Therefore, some other factor must fix the wavefunction, and that factor is what is known about the system.

Harrigan and Spekkens stipulate that all models that do not meet their criterion for being ψ-epistemic are ψ-ontic, based on how, classically, states can be understood as either ontic (points of state space) or epistemic (probability distributions over state space) as discussed above.[8]

Harrigan and Spekkens give three subcategories of ψ-ontic model. In the strongest (ψ-complete), the wavefunction describes the system's state completely and no hidden variables are left out by it (as in, for example, Everettian interpretations). In the other two models, both referred to as ψ-supplemented, the state is supplemented by some hidden variable—in the first as a one-to-one mapping between the wavefunction and the state; and in the second where wavefunctions can map to more than one state (but each state still only maps to one wavefunction). While the wavefunction may not itself be the state, in ψ-ontic models each ontic state corresponds to at most one wavefunction (see Fig. 8.1B), so all wavefunctions have disjoint support in the state-space:

$$\forall \psi, \varphi,\ s.t. \psi \neq \varphi;\ \Lambda_\psi \cap \Lambda_\varphi = \emptyset \tag{8.5}$$

8.2.3 Analysing Harrigan and Spekkens's Definitions

In Harrigan and Spekkens's framework, only ψ-epistemic models allow multiple wavefunctions to have overlapping support from the same underlying real state (see Fig. 8.1A). As noted above, they also take such overlap also to be necessary for a model to interpreted epistemically. However, while we grant that within the ontological models formalism overlap is sufficient for a model to be interpreted epistemically, we believe it is not necessary for it to be interpreted epistemically, because there is no reason to suppose that a one-one map between the wavefunction and the ontic state rules out that the wavefunction represents knowledge. Indeed, if the response functions for individual λ within a wavefunction's support Λ_ψ for S, are not equal to the Born probability for S conditional on the wavefunction, then this surely suggests the wavefunction could be interpreted epistemically:

$$\begin{aligned} &\exists \{\psi, S, A\},\ s.t.\ \neg(\mathscr{A}(S|\lambda) = p^{\psi}_{A}(S)\ \forall \lambda \in \Lambda_\psi) \\ &\qquad\qquad\qquad\qquad \implies \text{epistemic} \end{aligned} \tag{8.6}$$

[7]This doesn't however mean that they necessarily represent the underlying state equally well—for a given λ in this overlap, $p(\lambda|\psi)$ needn't equal $p(\lambda|\varphi)$.

[8]Schlosshauer and Fine apply the terms 'Mixed' and 'Segregated', rather than 'ψ-epistemic' and 'ψ-ontic' [SF12], which better reflects the distinction between wavefunction overlap and not.

On this condition the Born probability is just the weighted average of response functions over Λ_ψ, rather than equalling the response function of the state of the world λ (suggesting there is a real difference between the λs in the support of ψ, but that the experimenter does not know enough to distinguish between these different λs)[9]. All cases of overlap obey this condition, as within Λ_ψ, $\mathscr{A}(\psi|\lambda) = 1$, but $p_A^{\varphi}(\psi) = |\langle\psi|\varphi\rangle|^2$ which isn't equal to 1 if $\psi \neq \varphi$. However, cases without overlap could also obey this condition, as it applies to the response function of any observable, rather than the one represented by the operator whose eigenstates include ψ whatever that is. Hence, Harrigan and Spekkens's criterion for a model being ψ-ontic—that wavefunctions never have overlapping support in the space of underlying states—is wrong to make it not ψ-epistemic by definition.

ψ-ontic models either take the wavefunction to represent the exact underlying state, or to represent some part of the state and so to be incomplete. What Harrigan and Spekkens ignore is the possibility that the wavefunction represents both the state of reality and the epistemic state of an observer, as with the case of a proper mixture or entropy discussed above. Hence, even though the wavefunction represents reality, wavefunctions could still have overlapping support on state space, because they could simultaneously represent different states of knowledge.

Because of this, there is nothing in Harrigan and Spekkens' formal definition for a model being ψ-ontic, nor a criterion we can build from their model, which is necessary or sufficient for a model to have wavefunctions representing reality. The closest we can come to is that, if the wavefunction doesn't represent knowledge, it must represent something else (and so represent reality)—but, given our condition for a model representing knowledge above is sufficient rather than necessary, there is no reason models without overlap need represent reality.

Harrigan and Spekkens admit that all models that are not ψ-complete could be said to have an epistemic character, especially where multiple states map onto one wavefunction, given this associates the wavefunction with a probability distribution over state space. However, they say they are instead interested only in pure quantum states—hence their definitions being based on overlap, given they claim, with pure states, "ψ has an ontic character if and only if a variation of ψ implies a variation of reality and an epistemic character if and only if a variation of ψ does not necessarily imply a variation of reality" [HS10]. Hardy agrees, saying, given no overlap between the support of wavefunctions on the underlying state, we could deduce the wavefunction from λ, which he takes to mean the wavefunction would be written into the underlying reality of the world [Har13]. However, this is questionable, given that, as noted above, correlation between λ and a wavefunction does not necessarily make that wavefunction physical, any more than the correspondence between your fingers and the numbers 1-10 makes those numbers physical [SF12] (which is why we defined wavefunction-ontic as the wavefunction representing the real state, rather than being real itself).

[9]The Colbeck-Renner-Leegwater Theorem claims to prohibit cases like this, where the quantum state does not provide a full description for the prediction of future measurement outcomes [CR12; Lee16]. However, Hermens has shown this rests on a faulty assumption, and so these cases are not prohibited [Her20].

Further, considering the wavefunction to be ontic if and only if variation in it implies variation in the underlying state ignores that, while ontic models require that the wavefunction represent an element of reality, they need not only represent that. Even for pure states, so long as we are not considering a ψ-complete model, there is always some factor supplementing the wavefunction in the full underlying state of the world—so, conversely, knowledge could supplement that representation of reality to give us the wavefunction. Therefore, there is no reason for change in wavefunction to imply change in reality in an ontic model, even for pure states.

This is demonstrated by ψ-dependent models [SF12]. In a ψ-dependent model, the response functions depends on ψ—individual overlap possibilities for each possible state are given by the relevant measurement probabilities, rather than being uniform. The simplest case of this is when p_ψ is uniform across Λ_ψ, as, to retrieve the Born probabilities, $\mathscr{A}^\psi(S,\lambda)$ must then depend on ψ. As the wavefunction still to some degree reflects the underlying state, these models are ontic in our sense. However, they also allow overlap of wavefunctions on state space, making them ψ-epistemic. Schlosshauer and Fine give an ontic model which is also ψ-epistemic, also showing models can be both.

8.2.4 Consequences

The Pusey-Barrett-Rudolph (PBR) argument shows that wavefunctions cannot overlap on state space, but it says nothing about whether they are epistemic in the broader sense of Section 8.2.1. The same applies for other no-go theorems based on these definitions, such as Colbeck and Renner's [CR12], Patra et al's [PPM13], Hardy's [Har13], and Ruebeck et al's [RLE20]. Pusey et al showed, given certain assumptions, that any model that is ψ-epistemic by Harrigan and Spekkens's criterion (i.e. allows wavefunction-overlap on state space) always contradicts quantum theory [PBR12]. However, above we showed that wavefunctions being able to overlap on state space is a sufficient, rather than necessary condition for a model being epistemic. This means, even if those assumptions are valid, and so wavefunction overlap is incompatible with quantum physics, that does not rule out all epistemic models.

Even in the ontological models framework, models can be simultaneously ontic and epistemic—the wavefunction can represent both elements of reality, and knowledge about that reality. Harrigan's and Spekkens's terms, ψ-ontic and ψ-epistemic, do not formalise these informal ideas.

In light of our analysis, it is important that people do not conflate the ideas of ontic and epistemic in general with Harrigan and Spekkens' specific definitions, as to do so confuses the debate about quantum foundations.

8.2.5 Issues with an Assumed Underlying State λ

As mentioned above, there are a number of issues with the idea of an underlying state assumed by the Ontological Models Framework. Classically, points in state space define exact values for all observables (e.g. a point in (3+3)-dimensional position-momentum state space gives the exact position and momentum of an object), so associating a real possible state-of-the-world to each point ensures all values are simultaneously single-valued. However, if we try and say a single point in a state space corresponds to a real quantum state-of-the-world, it enforces locality of hidden variables, so violates Bell's theorem. (The only major interpretation of quantum physics in which this is not the case is in Bohmian mechanics, where the state-of-the-world would pin down the quantum potential—but existence as real values would reinforce the nonlocality-vs-special-relativity issue (as information sent nonlocally would definitively be writ onto the state-of-the-world, so this nonlocal communication would be, to use Shimony's terms, action, rather than passion, at a distance [Shi93]).)

Further, there is an issue with representing a projective measurement in this form, if one assumes all values are simultaneously single-valued—given, if we assume all system evolution except measurement is deterministic (as quantum mechanics does), it implies measurement must physically disturb the system in a random way to replicate our observations. Spekkens tries to say this random disturbance could come from the initial state of the measurement device, and that if we knew this too, the whole interaction would be deterministic, but this sounds like superdeterminism-by-stealth.

Another issue, specifically with the idea of the real underlying state of the world assumed by the ontological models formalism, is that both the response function for a variable at an initial state, and the probability that a situation represented by a given wavefunction being in a given state, are real probabilities, rather than quasiprobabilities—they are always non-negative, and sum additively. This is unlike quantum 'probability amplitudes', which are really quasiprobabilities—they can be negative or even imaginary, and so can interfere destructively with one another. This negative probability interference is a key part of quantum mechanics, explaining peculiar phenomena such as the Elitzur-Vaidman Bomb Tester [EV93], and counterfactual communication [Sal+13b; HLR21] and imaging [HR21b]. It is unclear how the real non-negative probabilities of this real underlying state-of-the-world could replicate these effects—casting doubt on this framework.

See Oldofredi et al for additional issues with this underlying real state assumption in [OL20], and below (and [HH22a]) for an alternative terminology for underlying-state models of the wavefunction.

8.3 The wavefunction as a true ensemble

The wavefunction is weird. Its most salient feature — that it merely predicts probabilities for measurement outcomes, rather than the outcomes themselves — suggests that quantum mechanics is an emergent, average description of an underlying, more complicated dynamics. In this underlying theory, the time-evolution of the system would be determined and measurement outcomes could be predicted. It is because we lack information about the details of the initial state that we can only make a probabilistic prediction.

That the wavefunction might be an emergent description of yet-to-be-discovered underlying physics is often called a hidden variable interpretation. The hidden variables are the information that is missing in quantum mechanics. This straightforward explanation for the strange properties of the wavefunction, however, has seemingly been disfavoured by Bell's and related theorems [Bel64; Cla+69], and, more recently, by the Pusey-Barrett-Rudolph theorem [PBR12; Lei14] (hereafter PBR-theorem). We here want to look closer at what these theorems actually say about the wavefunction as an ensemble of different values of the hidden variables.

The PBR-theorem in particular builds on a classification of models put forward by Harrigan and Spekkens [HS10]. Their framework attempts to mathematically formalise the distinction between a wavefunction that represents reality (ψ-ontic) and one that represents knowledge (ψ-epistemic). That this is a false dichotomy was previously pointed out in [HRL22]. We here want to spell out another problem with the ψ-ontic/epistemic framework — it doesn't identify the hidden variables theories that Bell was trying to rule out.

We begin with defining in what sense we take the wavefunction to be an ensemble in Section 8.3.1, and explain the relation to the ψ-ontic/epistemic framework in Section 8.3.2. In Section 8.3.3 we will discuss the most important consequence, that is the different definition of what the hidden variables describe, and how to interpret it. In Section 8.3.4, we will revisit the PBR-theorem in this light. In Section 8.3.5 we discuss Statistical Independence, locality, and the sometimes used classification of superdeterminism and retrocausality for violations of Statistical Independence. Finally, in Section 8.3.6, we will explain how the ψ-ensemble interpretation helps making sense of quantum mechanics.

8.3.1 The ψ-ensemble interpretation

We will begin by explaining what exactly we mean by the wavefunction being an ensemble. We do not mean the Ensemble, or Statistical, Interpretation of quantum mechanics [Bal70], according to which the wavefunction describes an ensemble of identically prepared states in identical experiments. It arguably does, but if the states are indeed identical, that's not much of an ensemble. Similarly, we do not mean Smolin's ensemble interpretation [Smo12], where the ensemble considered is the ensemble of all the systems in the same quantum state in the universe.

We mean instead that the supposedly identically prepared states in supposedly identical experiments are in fact different: they have different hidden variables and this is the reason why the measurement outcomes can be different for identical wavefunctions. We will refer to this interpretation of the wavefunction as the ψ-ensemble interpretation.

In our ψ-ensemble interpretation, we have an underlying theory with variables that we will collectively name κ and we will denote the state-space of all κ with K. The κ are generically time-dependent, $\kappa(t)$. These $\kappa(t)$s are not the same as Bell's hidden variables, as will become clear in a moment.

In the ψ-ensemble interpretation, the wavefunction in quantum mechanics emerges as an average description from the hidden variables theory, much like thermodynamics emerges from statistical mechanics. This interpretation suggests itself because of the apparent similarity of the von Neumann-Dirac equation with the Liouville equation, and also, as was pointed out in [Kle11], because the naive $\hbar \to 0$ limit of quantum mechanics results in a statistical theory.

We use the term 'hidden variables' because it has become common terminology. However, we want to stress that these variables are not a priori unmeasurable. They are 'hidden' merely in the sense that they do not appear in quantum mechanics. If the underlying theory was better understood, they could well become measurable one day. That is to say, the ψ-ensemble is an interpretation of the wavefunction in quantum mechanics, but it implies the existence of an underlying hidden variables theory. This underlying theory is better referred to as a completion or modification. This is as opposed to interpretations, such as the Transactional Interpretation [Cra86; Kas13] or Decoherent Histories [GH96], which have the same ontological basis as standard quantum mechanics, or Modal interpretations [LD21] that do not rely on an underlying completion (and do not violate Statistical Independence — we will get to this point below).

We assume that the underlying theory is deterministic, that is, if one specifies an initial state, one can use the theory to uniquely calculate the outcome of a measurement.

$|\psi\rangle$ is, as usual, an element of a projected Hilbert space, $\mathscr{H}$. It fulfils the Schrödinger equation. We are considering a generic quantum mechanical measurement, in which $|\psi\rangle$ is prepared at time t_{p} and measured at time t_{m}. The measurement is described by an orthonormal basis of pointer states that we will denote $|I\rangle$, where $I \in \{0...N-1\}$ and N is the dimension of the Hilbert-space. $|I\rangle$ could be describing multiple different detectors, and might be a product-state or an entangled state when expressed in a basis of the individual detectors. The basis $|I\rangle$ implicitly contains the measurement settings at the time of measurement.

For simplicity we will in the following assume that the basis $|I\rangle$ is time-independent. This does not mean that the detector setting cannot change. It merely means that if the detector setting before the measurement was different from the setting at the actual measurement, then it wasn't described by $|I\rangle$.

In quantum mechanics now, the Schrödinger evolution of $|\psi\rangle$ will generically result in a state that is not one of the detector eigenstates at the time of measurement. In this case, we assume that $|\psi(t_{\mathrm{m}})\rangle$ is an ensemble of different $\kappa(t_{\mathrm{m}})$, that is, a distribution $\mu(\kappa(t))$ over K. This means in the ψ-ensemble theory we have a map

$$\mathscr{P}(K) \ni \mu(\kappa(t)) \to |\psi(t)\rangle \in \mathscr{H} \,, \tag{8.7}$$

where $\mathscr{P}(K)$ is the space of all normalisable probability distributions on K. That is to say, we interpret $|\psi(t_{\mathrm{m}})\rangle$ as an ensemble because we empirically know it corresponds to different measurement outcomes.

For each pointer state $|I\rangle$, there is a subset of hidden variables, $\{\kappa\}_I$ that will lead to this outcome. We will refer to these subsets $\{\kappa\}_I$ as "clusters". The measurement reveals which cluster the hidden variable of the actual state belonged to. Hence, the measurement reveals some information about the hidden variable.

The reason we have for considering this interpretation is that the wavefunction update is non-local when interpreted as a physical process. This makes it difficult to combine quantum theory with general relativity. If we hence introduce a ψ-ensemble interpretation, our motivation is to develop an underlying theory which restores locality, causality, and determinism. Of course one can conceive of ψ-ensembles that are not local, but these (unsurprisingly) are not useful for restoring locality. Therefore, for the rest of this Section we will only consider ψ-ensembles that are locally causal in Bell's sense [Bel04b; CW12; Nor11].

8.3.2 Difference between ψ-ensemble and ψ-epistemic

The introduction of the ψ-ensemble in the previous section sounds superficially very similar to the definition of a ψ-epistemic model introduced by Harrigan and Spekkens we discuss above [HS10].

As a simple example, consider you have a two-state system with detector eigenstates $|0\rangle$ and $|1\rangle$. You could prepare the state so that, under the Schrödinger-evolution it goes to $|0\rangle$ with certainty. Or you could prepare it so that it goes to $(|0\rangle + |1\rangle)/\sqrt{2}$. If the wavefunction was ψ-epistemic, then the latter case must have some overlap with the former in the underlying hidden-variables theory because it can result in the same measurement outcome. At least that is the idea.

There are several problems with the Harrigan and Spekkens ψ-ontic/epistemic distinction. One of them is purely linguistic. According to the Harrigan and Spekkens definition, for ψ-ontic theories, the wavefunction in quantum mechanics itself is ontic. It seems odd to declare something ontic that can't be observed, especially given that Bohr already argued ψ is far better interpreted as knowledge. Settling a hundred-year old debate by definition is not insightful.

One might put this aside as historical nitpicking, but this linguistic issue reflects a deeper problem. If there's no observational requirement tied to calling something ontic, then every model can be made ontic just by declaring the wavefunction to be part of the hidden variables [SF12; Lei14]. Such an easily malleable definition of 'ontic' is not what one wants to base theorems on.

It is also highly confusing that, according to the Harrigan and Spekkens-definition, most hidden variables models which have been proposed so far are actually ψ-ontic. Think back to the example with the two state system that evolves into $(|0\rangle + |1\rangle)/\sqrt{2}$ but is measured in either $|0\rangle$ or $|1\rangle$. While the $(|0\rangle + |1\rangle)/\sqrt{2}$ state must have been made up of different underlying states (because otherwise it couldn't give rise to different outcomes) there is no reason why any one of those underlying states must have been *the same* as those of a system prepared to evolve into $|0\rangle$

or $|1\rangle$. Hence, according to the Harrigan and Spekkens-definition, a ψ-ontic model can solve the measurement problem as well as a ψ-epistemic one.

We believe our definition of a ψ-ensemble to be more descriptive of hidden variables models, and less ambiguous than the ψ-epistemology proposed by Harrigan and Spekkens. Whether or not the state prepared to evolve into $(|0\rangle+|1\rangle)/\sqrt{2}$ had overlap with any other wavefunction, it will always be an ensemble if the underlying hidden variables theory solves the measurement problem. That is to say a ψ-ensemble could be either ψ-ontic or ψ-epistemic according to the Harrigan and Spekkens-definition.

8.3.3 Hidden variables label trajectories

In our definition of a ψ-ensemble, we have been careful to point out that naturally the underlying hidden variables theory will be dynamic, hence the hidden variables will be functions of time. This is of utmost importance for locality considerations, as those usually pertain to arguments about the distribution of the hidden variables. (We will discuss this in Section 8.3.5.)

If we evaluate the distribution of hidden variables at the time of measurement, then we can infer the measurement outcome directly from that. However, for Bell's theorem, and also for the PBR-theorem, one normally considers the distribution of variables at the time of preparation instead. These two distributions will in general not be identical. The problem is, if one merely takes the distribution at the time of preparation, this will *not* constitute the hidden variables of Bell's theorem.

To see this, note that since we assume the underlying theory is deterministic, we can always take the distribution at the time of preparation, $t_{\rm p}$, and evolve it forward with whatever is the transport function of the theory. Assuming time-translation invariance, let us denote the transport function as $T(t_1-t_0,\cdot)$ where the free slot is for the hidden variables. It has the property that

$$\kappa(t_1)=T(t_1-t_0,\kappa(t_0))\ ,\ T(0,\cdot)\equiv \mathrm{I}d\ . \tag{8.8}$$

Now arguably the entire information that is necessary to determine the outcome at $t_{\rm m}$ is *both* the initial distribution at preparation $\kappa(t_{\rm p})$ and the transport function $T(t_{\rm m}-t_{\rm p},\cdot)$.

Bell, therefore, in the derivation of his inequality, correctly defines the hidden variables as the 'full specification' of the information necessary to predict the outcome [Bel04b]. In his own words, the "values of these variables together with the state vector determine precisely the results of individual measurements" [Bel66]. However, an initial state without an evolution law does not determine a final state.

Therefore, if we call Bell's hidden variables (as usual) λ, then $\lambda(t_0)=(\kappa(t_0),T(t_{\rm m}-t_0,\kappa(t_0))$ and one can now treat the hidden variables as time-independent. The additional information in the transport function becomes redundant in case one defines the distribution of the hidden variables at the time of measurement already ($t_0=t_{\rm m}$) but normally one chooses $t_{\rm p}<t_0<t_{\rm m}$.

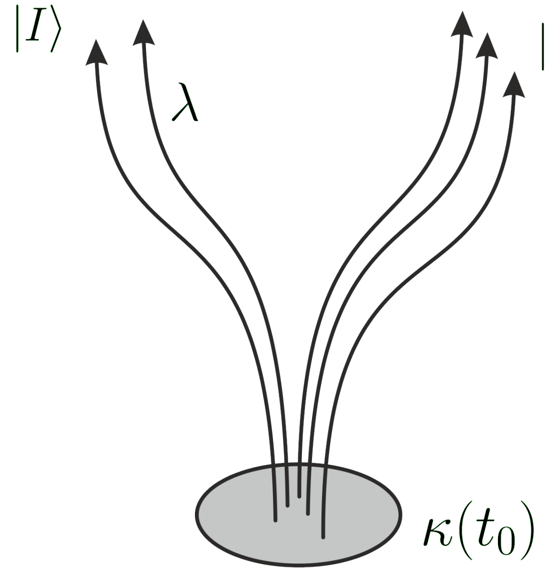

Figure 8.2: Illustration of the relation between the hidden variables $\kappa(t)$ which define the state, and the hidden variables λ which define the entire trajectory. $|I\rangle$ and $|J\rangle$ depict two different detector eigenstates. The trajectories that go to the eigenstate $|I\rangle$ belong to cluster $\{\kappa\}_I$, and those which go to $|J\rangle$ belong to cluster $\{\kappa\}_J$, respectively.

Pictorially this means that Bell's hidden variables (the λ) can be interpreted as labels for histories (in state space), whereas the Harrigan and Spekkens hidden variables (that we called κ) define a state (see Fig. 8.2). For this reason, assuming that Bell's hidden variables belong to a particular moment in the history of the (entire) state makes no sense. We will comment in Section 8.3.5 on what this means for locality requirements that one normally uses for the distribution of the hidden variables.

A particularly illustrative example for hidden variables that define trajectories is the model proposed in [Pal09; Pal20a]. In this case, the hidden variables simply are the detector eigenstates towards which the prepared state evolves. Given that we cannot measure the state before we measure the state, this is the most minimal assumption that one can make.

Following the terminology introduced in [CW12], we note that a deterministic model may not be predictable, in the sense that a model may contain variables from which one *could* calculate measurement outcomes when one had them, and yet one cannot find out the values of the hidden variables to actually do that. Bohmian mechanics is an example of such a deterministic, yet unpredictable mode. The ψ-ensembles we consider here are all deterministic. They may or may not also be predictable. In the model laid out in [Pal09; Pal20a] the hidden variables are incalculable with also makes the model unpredictable.

A few words are in order here about the sense in which the ψ-ensemble is local. In the ψ-ensemble, the hidden variables are *in general* not localised in the sense that they do not belong to some compact region of space because they actually describe trajectories. The initial values $\kappa(t_0)$ are in general not localised either because they are not uniquely defined. Suppose, for example, you have a distribution of variables in space and instead replace them with their Fourier-transformation. That would contain the same information but one may be localised, the other not. However, the initial states are localised if one considers experiments in which the prepared state is localised and is detected in localised detectors, which is for all practical purposes always the case, and certainly in Bell-type experiments. It is for this reason that the ψ-ensembles can fulfil the definition of local causality, as Bell intended.

We should add here that it is quite possible that the type of model we consider, in which the hidden variables actually describe trajectories, is not what Bell meant when he conceived of his definition of local causality. But regardless of what Bell may have meant, such models can formally fulfil the condition of local causality.

8.3.4 The PBR theorem revisited

Let us then look at what happens when one confuses these two sets of hidden variables. The PBR-theorem [PBR12], in a nutshell, shows that models which are ψ-epistemic according to the Harrigan and Spekkens-definition are incompatible with observations. This superficially sounds similar to Bell's theorem which takes on hidden variables theories, but the PBR-theorem works entirely differently. For this reason, one cannot draw conclusions from it about ψ-ensemble theories. It is not our intention here to criticise the PBR theorem which has instilled much fruitful discussion and helped sharpen terminology. We merely want to clarify what conclusions can be drawn from it.

The above discussion on ψ-ensembles is relevant as the hidden variables in the Harrigan and Spekkens-definition merely define the state. They do not also specify the time-evolution of the underlying hidden variable system. If one hence does not define the state at the time of measurement, then the hidden variables in the Harrigan and Spekkens-definition do not determine the measurement outcome. How could they if they do not contain information about the evolution law? Now one may say that a definition is just that, a definition. However, for the PBR-theorem one treats the hidden variables *as if* they did determine the measurement outcome, which implicitly assumes that the transport function contains no further information.

If however the transport function contains necessary information to determine the outcome, then one runs into the following problem. For the PBR-theorem one considers the distribution of hidden variables at preparation in the Harrigan and Spekkens-definition. However, those don't tell us what the outcome is, so no conclusions about whether the theory is viable or not can be derived. If one instead replaces the variables in the PBR-theorem with the variables which actually determine the outcome — i.e., those which contain information about the transport function — then one needs an additional assumption. This assumption is that the complete variables are not correlated with the detector settings. That is, one needs the assumption that Bell called Statistical Independence. This is because if Statistical Independence is not fulfilled, then one of the assumptions for the PBR-theorem, the Preparation Independence Postulate (PIP) is not fulfilled.

Bell defined Statistical Independence (SI) as the absence of correlations between the hidden variables and the two detector settings which he considered in a particular experiment. However, of course one can define Statistical Independence in general as

$$\mu(\lambda|S) = \mu(\lambda) \ , \tag{8.9}$$

where S are the detector settings (of whatever experiment), and μ is the probability distribution of the hidden variables. As laid out in [HHP22], violating this condition does not necessarily require correlations between the settings and the distribution of the hidden variables over the state space; the correlations can instead be induced by the structure of the state-space.

Statistical Independence sneaks into the PBR-theorem because the Harrigan and Spekkens-definition for a ψ-epistemic model assumes that the hidden variables which define the state at preparation have nothing to do with the measurement. And that sounds reasonable, until one realises that if one wants to say anything about measurement outcomes one needs to know the time-evolution of the hidden variables too.

The PIP now says that if we prepare two independent (spatially separate) experiments, then the distributions of the hidden variables should factorise. In the PBR argument one first considers two experiments in isolation. Then one combines the two so prepared states — using PIP — and measures them instead with a smartly chosen basis of distinguishable states. From this, one derives a contradiction: If the distribution of the hidden variables was actually the product of the distributions for the separate measurements, then one cannot reproduce the predictions of quantum mechanics.

Needless to say, the theorem is correct for what the mathematics is concerned, but the relevant physical assumption was that the distributions of the hidden variables for the individual experiments already didn't depend on the measurement settings. So of course if one multiplies them, they still do not depend on the measurement settings and one cannot reproduce the predictions of quantum mechanics. That is, in the ψ-epistemic framework, we have SI $\Rightarrow$ PIP[10].

However, the relevant assumption was the independence of the complete hidden variables (the λs which determine the outcome) from the measurement settings, not the factorisation. In particular, the initial distribution of the variables, $\kappa(t_\mathrm{p})$, might well factorise at preparation.

An example for a model in which this happens is [DH22]. This model uses one (complex valued) hidden variable for each dimension of the Hilbert-space. These hidden variables, which correspond to κ, are uniformly distributed in the unit disk. Hence, the model fulfils the requirements for the PBR-theorem and yet reproduces the predictions of quantum mechanics, seemingly defying the theorem. The reason is that the dynamical law of the model depends on the detector settings, and so do the full hidden variables λ. When one uses these variables, the model violates PIP, hence

[10]Note that it's not in general the case that $\neg$ PIP $\Rightarrow \neg$ SI because it could be that other assumptions of the ψ-epistemic framework are violated rather than SI.

no contradiction with the theorem arises. However, since the variables then no longer describe the initial state at preparation, it's unclear why they should fulfil PIP in the first place.

We may note that this issue does not appear in Bell's theorem, simply because Bell considers a situation where the state is prepared at one place but measured in two, whereas PBR consider a situation where the state is prepared in two places but measured in one. In Bell's case one has no reason to assume that the distribution of the hidden variables factorises.

For completeness, let us mention that the approach presented in [Bar+14] to rule out (Harrigan and Spekkens) ψ-epistemic models. This approach works by constructing sets of incompatible states, and deriving properties about the overlap of probability distributions from measurements on those. This proof implicitly assumes that the probability distribution of the hidden variables does not depend on the measurement settings.

8.3.5 Properties of ψ-ensemble theories

Statistical Independence

We know from Bell's theorem that any ψ-ensemble theory which is locally causal (in Bell's sense) and reproduces the predictions of quantum mechanics must violate Statistical Independence. Indeed, that one can reproduce quantum mechanical correlations while respecting local causality provided Statistical Independence is violated is quite possibly the reason why Bell (and others after him) sought other ways to rule out Statistical Independence, notably by arguing that it requires finetuning (discussed below).

However, one does not need Bell's theorem to see that local and causal ψ-ensemble models will violate Statistical Independence. We just need to note that if we want states in the hidden variables theory to evolve locally and within the light-cone into detector eigenstates, then we need information about the detector before the prepared state reaches the detector — otherwise how do we know what state to evolve into?

For the purposes of this Section, we shall say that a hidden variables theory solves the measurement problem if it respects Bell's local causality and gives rise to detector eigenstates in each single run of an experiment with a probability distribution that agrees (to within measurement accuracy) with that of quantum mechanics. There are other ways to look at the measurement problem [Mau95], but for the purposes of this Section, this simplified notion that specifically applies to hidden variables theories will suffice.

To illustrate what solutions to the measurement problem have to do with Statistical Independence, consider the generic setting of an experiment with two detectors (Fig. 8.3). The state is prepared at $\mathbf{P}$ and measured at detectors $\mathbf{D}_1$ and $\mathbf{D}_2$. One may have in mind for example a single photon that goes through a beam splitter. The state does not have to be entangled. We want to know how it evolves in the underlying hidden variables theory.

When the prepared state arrives at $\mathbf{D}_1$ it needs to know the setting at $\mathbf{D}_2$. Yet, it is clear that if that information wasn't available already at $\mathbf{P}$, then it can't have gotten there within the light

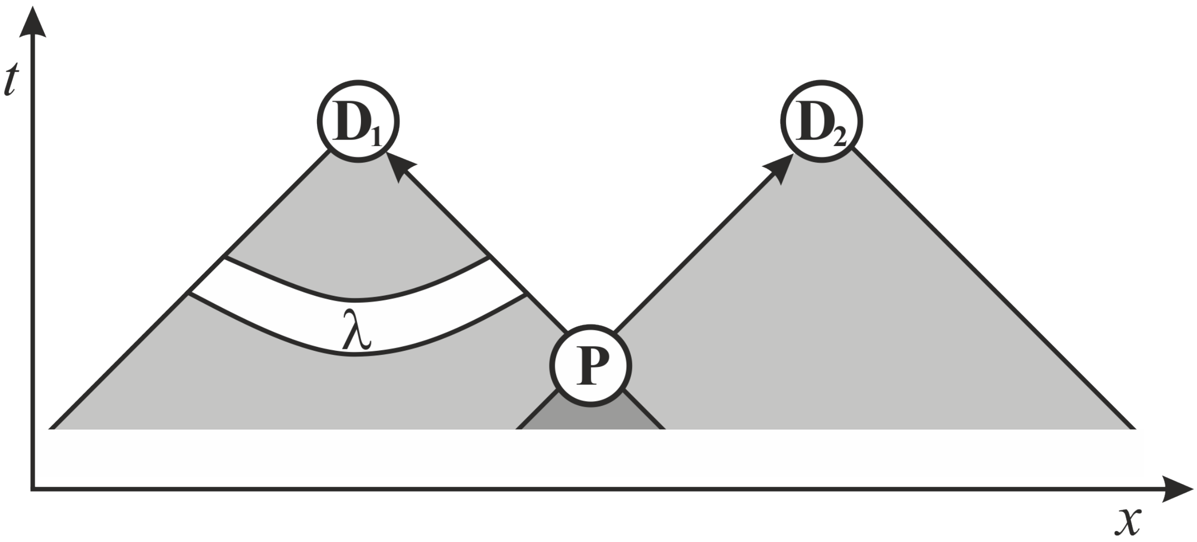

Figure 8.3: A diagram showing the two backwards light cones from detectors $\mathbf{D}_1$ and $\mathbf{D}_2$, and their overlap at point of preparation $\mathbf{P}$.

cone. Hence, both detector settings $\mathbf{D}_1$ and $\mathbf{D}_2$ must have been either in the initial distribution of hidden variables or the evolution law. In either case, Statistical Independence is violated.

To be sure, we have drawn an extreme case where the prepared state actually moves with the speed of light. If one took a massive particle, the last moment in which information about $\mathbf{D}_2$ must be available for the underlying state going to $\mathbf{D}_1$ could be after preparation. But that doesn't change the fact that Statistical Independence must have been violated because information about the detector setting must have been in the full specification of the hidden variables.

It is a corollary of this — and not the starting point — that any ψ-ensemble theory that solves the measurement problem and reproduces quantum mechanics will violate the assumptions of Bell's theorem and the PBR-theorem. It follows that experiments which show the violation of a Bell inequality or the CHSH inequality [AGR81], or the agreement of the measurement proposed by PBR with the predictions of quantum mechanics [Rin+15], cannot rule out ψ-ensemble theories. A different type of experiment is necessary to test such theories.

Understanding that Statistical Independence is violated *because* that is required to solve the measurement problem also explains why one need not expect such violations to occur in non-quantum experiments, like the often quoted tobacco trials to infer a correlation between smoking and cancer (see [HP20] and references therein). No preparation of such an experiment will ever result in a superposition of measurement outcomes. Even if one used a quantum experiment to, say, sample mice into two groups (one exposed to tobacco smoke, the other not), then the quantum superposition would be destroyed before one even wrote down the outcome of the quantum experiment. But Statistical Independence is only violated between the preparation of an experiment that, under the Schrödinger-evolution, evolves into a superposition of detector eigenstates, and the measurement of that state.

It must be understood here that we use the term 'measurement' in the same sense as in quantum mechanics. That is, it does not necessarily require an observer or even a detector, it just requires a sufficient amount of interactions with the environment. We don't need to be specific about this because this aspect of the measurement problem was already solved by decoherence and einselection [Zur03]: A detector is any system that decoheres the prepared state quickly enough, and the detector eigenstates are defined as those which are stable in that process. We merely need violations of Statistical Independence to explain why we always measure a detector eigenstate and not, as decoherence predicts, a mixed state.

It is worth stressing that the statements made in this section are not suggested interpretations of mathematical expressions, they are conclusions drawn from empirical facts. We know, as a matter of fact, that measurements yield definite results. If the results come about locally, causally, and deterministically from the prepared state, then the underlying theory must violate Statistical Independence. We also know empirically that we never observe superpositions of detector eigenstates, hence violations of Statistical Independence are negligible for macroscopic objects.

Theories which violate Statistical Independent are often further broken down into those which are superdeterministic and those that are retrocausal. The superdeterministic ones were recently further broken down into those that are and are not supermeasured (see [HHP22] and Chapter 9). However, the terms superdeterministic and retrocausal have been used to mean different things by different authors which has made the situation very confusing. For example, some authors assume superdeterminism is deterministic [HP20], while others do not [SV20a]. Some equate retrocausality with future input dependence [WA20], others do not [Pri94; Pri12]. Some claim that retrocausality is just mislabelled superdeterminism [NV22]. Yet other authors have proposed to use the terms all-at-once or atemporal [Adl22]. None seem to assign much importance to the ensemble that the wavefunction should describe, which is at the centre of our definition.

We believe that the ψ-ensemble is a more useful classification because it captures those approaches which are candidates to solve the measurement problem(s).

Locality

It is quite common to get confused here about how such a requirement — that the evolution law depends on the measurement settings at the time of measurement — can possibly be local. However, keep in mind that the hidden variables λ label trajectories, not initial states, of the underlying theory. The trajectories necessarily connect the prepared state with the detectors locally and causally (if they didn't, we couldn't detect the prepared state). Hence, the requirement that the hidden variables are correlated with the detector settings at measurement can come about entirely by local propagation and by local interactions. (Again, we do not claim that such a theory necessarily has to be local. It's just that this is the case we are interested in.)

Such ψ-ensemble theories that violate Statistical Independence fulfil Bell's criterion of local causality because if the information about both detector settings was available already at preparation, then it is unnecessary to later add it.

Fine-tuning

It is frequently argued that theories which violate Statistical Independence would need a lot of detail to specify the initial state for the detector setting at the time of measurement, and that they are thence fine-tuned "conspiracies" [SV20b]. This initial state is often believed to be required long before the experiment even began, possibly as far back as the Big Bang.

However, the idea that we need to define an initial state for the detector setting is patently absurd. In quantum mechanics, we also never define the initial condition at the Big Bang that will give rise to a specific detector setting. Indeed, we never do this in any theory. It is in practice unnecessary to ever write down this initial state, regardless of whether it is feasible in the first place. We simply make predictions for any possible measurement setting. Hence, violating Statistical Independence does not require a lot of detail; it merely requires the detector settings at the time of measurement — an input we also use in ordinary quantum mechanics.

A second finetuning argument [WS15] has it that it would require specifying a lot of details for the distribution of hidden variables in order to prevent violations of Born's rule that would allow for superluminal signalling. However, leaving aside that the absolute impossibility of superluminal signalling (rather than its observed rarity) is a postulate, not an empirical fact, we can arguably always achieve that Born's rule is fulfilled simply by postulating that it is fulfilled, which one also does in quantum mechanics.

That is, these two arguments use a double-standard: Assumptions are not called fine-tuned when used in normal quantum mechanics (detector settings at measurement and Born's rule are required to calculate probabilities) but are called fine-tuned when used in other frameworks. As already argued in [Hos20], such a notion of fine-tuning is scientifically meaningless and can at best be considered metaphysical. (Finetuning in these contexts has been discussed in more detail by Adlam [Adl21a].)

8.3.6 Using the ψ-ensemble to make sense of quantum experiments

Of course using a ψ-ensemble, even if deterministic, local, and causal doesn't remove the weirdness of quantum mechanics. It remains in the property that for a spatially distributed detector, a state in the underlying theory that goes to one particular eigenstate must have had information about that part of the detector to which it *didn't* go. Again, this requirement can be formulated purely in local terms (if the state would have gone to the other part of the detector, it would have interacted with it), but this is where the famed quantum weirdness goes. If the underlying hidden variables theory is local and causal and deterministic, then the evolution which the actual state takes — and which we ultimately observe — takes into account what would have happened if the state had gone elsewhere. We will now explain how this nicely explains some otherwise very unintuitive quantum behaviour.

The Double-Slit Experiment

To warm up, we consider the familiar double-slit experiment. The odd thing about this experiment is that the behaviour of the quantum particle changes depending on whether we know which slit the particle goes through. If we don't know, the particle seems to go through both slits and interferes with itself, like we expect for a wave. But if we measure which slit the particle goes through, the interference pattern vanishes.

This seemingly strange behaviour makes total sense from the ψ-ensemble perspective because the trajectory of the state in the underlying hidden variables theory depends on what is being measured.

Wheeler's Delayed-Choice Experiment

The delayed choice-experiment [Whe78] comes in a large number of variations, but in the simplest version, one changes the variable that is being measured after the prepared state has begun its propagation. In normal quantum mechanics, this is even more difficult to understand than the double slit experiment because it seems that the prepared state must somehow 'change its mind' as the setting is changed.

However, from the ψ-ensemble perspective this is clearly not what happened. This is because the trajectories depend on the detector setting *at the time of measurement*. What the settings were before that or how often they were changed is irrelevant.

The Elitzur-Vaidman Bomb Tester

The Elitzur-Vaidman Bomb Tester allows one to detect whether a bomb is live (would blow up when it detects a photon) without triggering the bomb [EV93]. In this experiment, the bomb acts as a detector. Whether the bomb is live or a dud therefore constitutes two different detector settings. From the ψ-ensemble perspective, it is hence rather unsurprising that the case in which the bomb does not blow up contains information about whether the bomb is live, because this is just a type of detector setting.

Counterfactual Computation/Communication

In counterfactual computation [Hos+06] one infers the result of a computer query without interacting with the computer. This experiment is a variant of the Elitzur-Vaidman experiment and can be explained the same way. The computer acts as a detector depending on what the outcome of the calculation is. In the ψ-ensemble interpretation, the particle which one measures contains information about the setting of the detectors which it wasn't detected in, which allows one to infer the result of the computation.

The same logic also applies to Salih et al's counterfactual communication device and related protocols [Sal+13b; Sal+22; HLR21], and so protocols for imaging [HR21b] and for quantum information transfer derived from it [Sal+20; Sal+21].

The Extended Wigner's Friend Scenario

The previous examples all work more or less the same way. The Extended Wigner's Friend experiment [FR18] is a more recent proposal that beautifully highlights the problems one runs into when rejecting hidden variables theories, but works entirely differently.

The Extended Wigner's Friend Scenario is a Bell-type test with two entangled particles, one measured by Charlie and one by Debbie (the friends) in spatially separated laboratories. Each of the two observers is observed by a super-observer, Alice and Bob respectively. Charlie and Debbie measure an entangled state and Alice and Bob then further measure correlations between Alice's and Bob's measurement outcomes. It turns out that, given suitable measurement settings, the observers cannot all agree on what the measurement results are.

The issue with this experiment is a fuzzy definition of what constitutes a measurement. As already pointed out in [Rel18; ŻM21], one can interpret the friends' measurements as either having resulted in a definite outcome, or as still being in an entangled quantum state. The resulting contradiction does not occur if one has a theory in which a measurement outcome is an unambiguous result of the time-evolution.

In a ψ-ensemble theory, the friends either make a measurement and violations of Statistical Independence are for all practical purposes destroyed. Then the super-observers can no longer measure quantum correlations between the friends. Or the friends really just make a transformation on a quantum state, rather than a measurement, in which case the super-observers can well observe further quantum effects.

We have argued here that the ψ-ontic/epistemic distinction is not descriptive of hidden variables models that can restore locality and causality by giving a physical meaning to the wavefunction update. Instead, they can better be described as ψ-ensemble theories. If they are local, causal, deterministic and reproduce the predictions of quantum mechanics, then such theories violate the assumptions of both Bell's theorem and the PBR-theorem. We have shown that taking this possibility seriously can help making sense of some otherwise strange quantum phenomena.

8.4 Chapter Conclusion

In this chapter, we undertook foundational work—evaluating whether Harrigan and Spekkens's criteria for models being ψ-ontic and ψ-epistemic actually corresponded to the wavefunction representing the state of the world (the wavefunction being ontic) and the wavefunction representing some experimenter's knowledge about the state of the world (the wavefunction being epistemic). While Harrigan and Spekkens's terms ψ-ontic and ψ-epistemic are necessarily defined as contradictories, we showed the wavefunction being ontic doesn't necessarily exclude it from also being epistemic. Harrigan and Spekkens's mathematical definitions therefore fail to capture the underlying intuitions they are supposed to represent. We then highlight a category of model specifically excluded from Harrigan and Spekkens's definitions, where the wavefunction represents an ensemble of identically-prepared states with differing hidden variables (ψ-ensemble models). Those ψ-ensemble models which are locally causal in Bell's sense can be used to solve the measurement problem, at the expense of violating statistical independence. Combined with the review of the measurement problem in Chapter 7, this shows why we should be interested in models that violate statistical independence—while they have been attacked as unpalatable, they present the only viable path towards resolving the measurement problem.

While this chapter does not directly link to the interplay between quantum foundations and quantum technologies, it shows why we should be interested in statistical independence-violating models. This reinforces why we need ways to experimentally distinguish between these models and standard quantum mechanics.

CHAPTER 9

Statistical Independence-Violating Models

Declaration of contribution: This Chapter is adapted from one published paper and one preprint. The published paper is *Supermeasured: Violating Bell-Statistical Independence without violating physical statistical independence* [HHP22], which was conceived and written by myself, Dr Sabine Hossenfelder, and Prof Tim Palmer. The preprint is *Experimental Tests of Invariant Set Theory* [HPR21], which was written by myself, edited by Prof Tim Palmer, and supervised and edited by Prof John Rarity. These papers were written at a later stage of my PhD, when I had been developing independent collaborations, and my supervisors were happy for this independent work to be included in my thesis.

9.1 Chapter Introduction

In this chapter, we look in more depth at statistical independence-violating models of quantum mechanics. First, we show that, while statistical independence-violating models are typically split into either superdeterministic or retrocausal models, there is actually a third type of model which violates statistical independence: we name this "supermeasured". We show how this type of model differs from superdeterministic and retrocausal models: these differences allow proponents of such models to counter arguments posed against statistical independence-violating models. We then focus on an example of a supermeasured model—Invariant Set Theory (IST)—and investigate how to experimentally test predictions it makes which differ from those of standard quantum mechanics.

The work in this chapter, especially towards the latter end, forms a key aspect of Part II's case study into the interrelation between foundations and quantum technology. By showing how quantum technology can be used to test foundational models, it illustrates how new and developing/hypothetical quantum technologies can feed back into quantum foundations. This demonstrates that quantum technological research serves not only to develop spin-off applications of quantum foundations, but also provides key tools for experimentally validating foundational research.

9.2 Supermeasured: Violating Bell-Statistical Independence without violating physical statistical independence

Bell's theorem [Bel64] has been a milestone in our understanding of quantum mechanics, detailing just what correlations are necessary to reproduce observations. Unfortunately, many physicists have jumped to incorrect conclusions from it. A 2016 survey among professional physicists [SN16] found that 34% believe Bell's theorem shows that "Hidden variables are impossible": that is, they think Bell's theorem rules out theories in which measurement outcomes are determined by variables that are not accounted for in standard quantum mechanics. Similarly, in a survey conducted among professional quantum physicists at a conference in 2012 [SKZ13], 64% claimed that Bell's theorem rules out hidden variable theories (and said local realism is untenable). This is of course not so. Bell's theorem just shows that a hidden variables theory *which fulfils all the assumptions of the theorem* is ruled out by observation.

Bell's theorem however contains a questionable assumption: (Bell-)Statistical Independence, sometimes called the "Free Will" or "Free Choice" assumption (here labelled Bell/capitalised to distinguish it from the intuitive ideas of physical statistical independence, free will and free choice). Indeed, one can interpret *all* experiments that have found violations Bell inequalities as simply demonstrating that if quantum mechanics is underpinned by a local, causal, and deterministic theory, then that underlying theory must violate Bell-Statistical Independence. Clearly the conclusion to draw from this is that we should look for a hidden-variables theory that violates Bell-Statistical Independence, not least to develop a quantum formalism that is compatible with General Relativity. Of course this is not historically what has happened. Instead, physicists have collectively discarded the possibility that Bell-Statistical Independence might be violated because they misunderstood what it means. For example, referring to Bell-Statistical Independence as "free will" or "free choice" seems to have created a strong cognitive bias for accepting the assumption unthinkingly.

That this "free will" nomenclature is highly misleading has already been clarified elsewhere [HP20; Hos20]; we don't want to repeat this entire discussion here (though we will briefly comment on the relation between statistical independence and free will in Subsection 9.2.3). Our aim in this Section is to investigate the physical interpretation of Bell-Statistical Independence, and explain why it is widely misunderstood.

This misunderstanding is well-illustrated by a quote from a recent paper by Sen [Sen21]:

> "The [Statistical Independence] assumption states that the hidden variables that

determine the measurement outcomes are uncorrelated with the measurement settings."

Similar interpretations can be found in [SV20a; SV20b]. This indeed is the standard way of interpreting the mathematical statement of Bell-Statistical Independence. However, we will show below that physically this interpretation is ambiguous at best. Subsection 9.2.1 gives a general argument for this. Subsection 9.2.2 looks at the CHSH inequality in particular. Subsection 9.2.3 discusses in more detail IST [Pal20a], which is to our knowledge the only example of a theory that 'violates Bell-Statistical Independence without violating physical statistical independence', before discussing misconceptions regarding free will, fine tuning and conspiracy.

9.2.1 Understanding Statistical Independence

As per Subsection 7.2.5, Statistical Independence is often said to be the assumption that

$$\rho(\lambda|X) = \rho(\lambda)\ , \tag{9.1}$$

where λ is a set of hidden variables, X are the detector settings, and ρ is a probability distribution of the hidden variables. In Bell's theorem one normally uses two separate detectors and their settings. We comment on that specifically in Subsection 9.2.2, but first look at the general interpretation. It is possible in principle that ρ depends on further variables, but this will not matter in the following.

That $\rho(\lambda, X)$ is a probability distribution means it is normalised over a space, which we will denote $\mathscr{S}_{math}$ for the mathematical state space: it comprises all mathematically possible states of the hidden-variables theory. By "mathematically possible" we literally just mean that we can write them down mathematically. We might however later discard some of the mathematically possible states as not physically possible or meaningful. This isn't so uncommon. For example, some mathematically possible solutions to the Schrödinger equation are not normalisable, and hence not physically possible.

Our key point in this section is that any space we integrate over must have a measure, $\mu(\lambda, X)$, and generically this measure is non-trivial, i.e., isn't just identical to some normalisation constant. A measure roughly speaking quantifies the volume of the space. The probability distribution ρ can only be normalised by help of the appropriate measure:

$$\int_{\mathscr{S}_{math}} d\lambda dX\ \rho(\lambda, X)\mu(\lambda, X) = 1\ . \tag{9.2}$$

Measure theory [CD15] is not usually discussed in physics textbooks. However, a variety of measures make their appearance in physics nevertheless. The most widely used one is the Lebesgue measure on $\mathbb{R}^n$ and (pseudo-)Riemannian manifolds. On fractals, it can be generalised to the Hausdorff measure. In the context of Hamiltonian dynamical systems, a non-trivial measure on state space arises in the theory of symplectic manifolds (leading, for example, to the

Gromov non-squeezing theorem). In Subsection 9.2.3, we discuss non-trivial invariant measures associated with chaotic attractors.

The measure of $\mathscr{S}_{math}$ appears in the calculation of any expectation value; therefore it should enter the derivation of Bell's theorem together with the probability distribution ρ. Since these two functions always appear together, it is tempting to simply combine them into one $\rho_{Bell}(\lambda, X) := \rho(\lambda, X)\mu(\lambda, X)$, where we use the index "Bell" to emphasise that this is the quantity that really enters Bell's theorem. The assumption of Statistical Independence in Bell's theorem is therefore actually the assumption that

$$\rho_{Bell}(\lambda|X) = \rho_{Bell}(\lambda) \ . \tag{9.3}$$

We call this Bell-Statistical Independence (with the prefix Bell- and capital letters), and distinguish it from Eq. 9.1 which we now call physical statistical independence in lower-case letters. (Note, in the most general formulation of the assumptions of Bell's theorem, one considers a dependence of ρ_{Bell} on both the measurement settings as well as the preparation procedure. However, the dependence on the preparation procedure is irrelevant to our point, and we have therefore omitted it for simplicity. The reader may consider the preparation details to be part of the hidden variables.)

To avoid confusion with the standard interpretation of superdeterminism, we propose calling a theory which violates Eq. 9.3 but does not violate Eq. 9.1 a "supermeasured" theory, with μ being the supermeasure. (Note, despite the similarity in terminology, this has nothing to do with measurement.)

Since Eq. 9.3 is mathematically indistinguishable from Eq. 9.1 given a suitable redefinition of the probability density, one may wonder why we should bother introducing the two distributions ρ and ρ_{Bell}? It is important to distinguish them because physically they mean something different. ρ is the distribution of states on $\mathscr{S}_{math}$. It can be affected by factors under the control of the experimenter, such as the preparation of the state. ρ_{Bell}, by contrast, is the distribution weighted by the measure $\mu(\lambda, X)$. This measure is not under the control of the experimenter—it's just a property of the laws of physics. As such ρ_{Bell} contains information *both* about the intrinsic properties of the space *and* the distribution over the space.

The problem with the common interpretation of Bell-Statistical Independence is that typically the measure μ is not explicitly defined in the assumptions for Bell's theorem. This means that one implicitly assumes that the measure μ is identical to the uniform measure μ_0 on $\mathscr{S}_{math}$. The consequence is that interpretations of Bell's theorem run afoul of physics whenever one is dealing with a theory in which $\mu(\lambda, X) \neq \mu_0$.

To see why this distinction matters, let us look at a simple idealised example. The following example is not meant to describe a realistic physical theory. We merely present it to elucidate that it is always possible to replace a correlation on one space with a non-correlated distribution on a subset of the first space—without changing any of the probabilities. This shows that the common definition of statistical independence is ambiguous for what the physical interpretation is concerned, and so claims of “fine-tuning” based on this definition are also ambiguous. We will come to a more physically relevant example later.

Let $\mathscr{S}_{math}$ be a compact continuous space with uniform measure $\mu \equiv \mu_0 =$ constant, and ρ a probability distribution over it. This probability distribution may violate Eq. 9.1. Our task here will be to show that we can remove this correlation entirely without changing any probabilities.

To see this, we use ρ to randomly choose a set $\mathscr{S}_N = \{(\lambda_1, X_1), (\lambda_2, X_2)...(\lambda_N, X_N)\}$ of N points in $\mathscr{S}_{math}$. For illustration, see Fig. 9.1. From this we define the discrete measure

$$\tilde{\mu}(\lambda, X) := \sum_{i=1}^{N} \delta(\lambda - \lambda_i)\delta(X - X_i) \ . \tag{9.4}$$

Note that the probability that two points are equal to another is zero. We then define a uniform probability distribution $\tilde{\rho}(\lambda, X) \equiv 1/N$ on $\mathscr{S}_{math}$ which is normalized with respect to $\tilde{\mu}$ and from that $\tilde{\rho}_{Bell}(\lambda, X) := \tilde{\rho}(\lambda, X)\tilde{\mu}(\lambda, X)$.

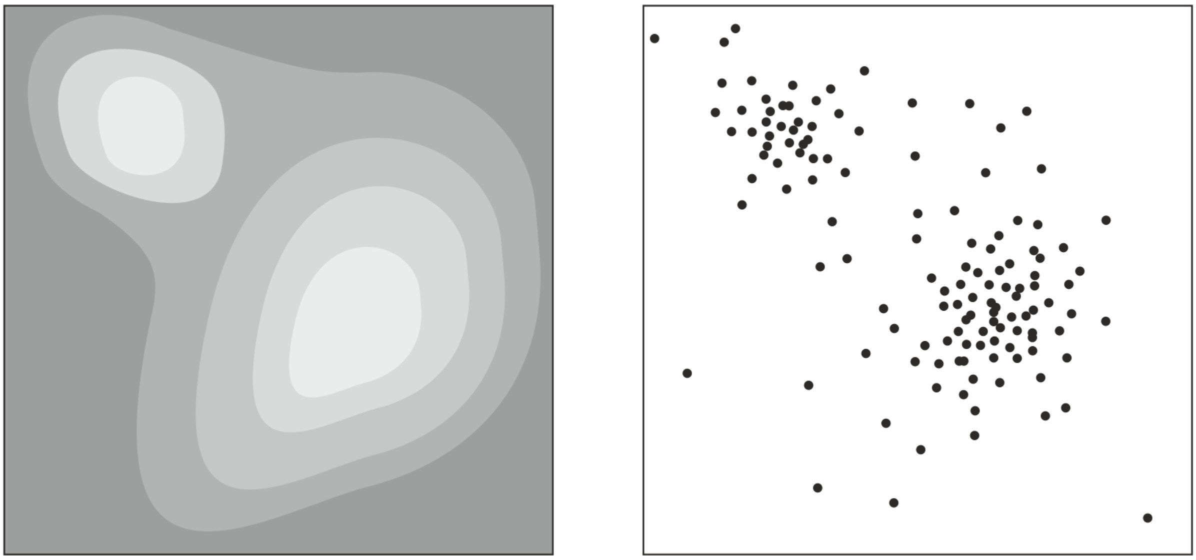

Figure 9.1: Illustration of sampling procedure. Left: The square represents the space $\mathscr{S}_{math}$ over hidden variables and measurement settings, and the shading is the probability distribution ρ over it. The brighter the shading, the higher the probability. Right: We randomly distribute a set of N points using ρ. In the limit $N \to \infty$ a uniform distribution $\tilde{\rho}$ on the points will reproduce the probabilities defined by ρ on $\mathscr{S}_{math}$ with a uniform measure. The set of points defines the new space $\mathscr{S}_{phys}$, also over hidden variables and measurement settings. It has a non-trivial measure $\tilde{\mu}$ in $\mathscr{S}_{math}$. Any correlations that were present in ρ are thereby moved into the structure of $\mathscr{S}_{phys}$.

Let us now take a subset A of $\mathscr{S}_{math}$ with non-zero volume (according to μ), $A \subset \mathscr{S}_{math}$. In the limit $N \to \infty$, the probability $P(A)$ of finding the system in that subset is

$$P(A) = \int_A d\lambda dX \ \tilde{\rho}_{Bell}(\lambda, X) = \int_A d\lambda dX \ \rho_{Bell}(\lambda, X) \ , \tag{9.5}$$

for any A. This means that all probabilities calculated from ρ on $\mathscr{S}_{math}$ with uniform measure $\mu \equiv \mu_0$ are by construction identical to those of the uniform distribution $\tilde{\rho}$ with measure $\tilde{\mu}$.

Finally, we define the new, physical state-space $\mathscr{S}_{phys} := \lim_{N\to\infty} \mathscr{S}_N$. Since $\tilde{\mu} \equiv 0$ on $\mathscr{S}_{math} \setminus \mathscr{S}_{phys}$, we discard the complement, only keep $\mathscr{S}_{phys}$, and restrict the probability $\tilde{\rho}$ to $\mathscr{S}_{phys}$.

Once we have done that, the entire information that was previously in ρ has moved into the definition of the physical state-space $\mathscr{S}_{phys}$. $\rho_{Bell} = \rho\mu$ and $\tilde{\rho}_{Bell} = \tilde{\rho}\tilde{\mu}$ give exactly the same probabilities. Since we never experimentally measure probability densities but only probabilities, these two theories are physically indistinguishable. Both violate Bell-Statistical Independence as defined in Eq. 9.3. But $\tilde{\rho}(\lambda|X) = \tilde{\rho}(\lambda)$: the hidden variables are by construction uncorrelated on the new space. On $\mathscr{S}_{phys}$, the theory violates Eq. 9.3, but not Eq. 9.1.

We here used a uniform distribution on the physically possible states. This corresponds to what is commonly called the principle of indifference [Key21]. We just use this as the simplest example of a distribution on state space. Regardless of the distribution, however, one cannot be indifferent about the state-space itself because this state space is a property of the laws of physics. It is whatever it is. This can induce violations of Bell-Statistical Independence even if the distribution over the space is uniform.

It is in this sense that the common interpretation of Bell-Statistical Independence is wrong: on a state space with non-trivial measure, the hidden variables may not be correlated with the detector settings, yet the Bell-Statistical Independence assumption will be violated. The common interpretation neglects the possibility that the theory is supermeasured rather than superdeterministic.

Moving the correlation into the definition of the physical state space is a simple way to avoid fine-tuning (the claim that the experimentally confirmed correlations are sensitive to small changes in the distribution, raised in e.g., [SV20b; WS15])). If the correlations are created by the intrinsic properties of the space itself, rather than the distribution over the space, then small changes just can't happen. This is the idea of IST [Pal20a] which we will discuss in more detail in Subsection 9.2.3. Other reasons why the fine-tuning argument can fail were previously discussed in [Hos20; WA20; DH22].

Another example of a theory which uses a non-trivial measure is Spekkens toy model [Spe07]. This model relies on an "epistemic restriction," that requires certain combinations of phase-space distributions to have measure zero. The Spekkens model is not supermeasured, however, because this measure does not depend on the measurement setting. For this reason the Spekkens model cannot reproduce Bell-inequality violations, whereas IST can.

One could apply the principle of indifference to say the measure should be uniform (or trivial). This occurs in fields such as statistical mechanics (e.g. with the Gibbs measure [Geo11]). However, empirical evidence/observed physics gives us reason to believe the measure is (or at least to consider theories where the measure is) non-trivial, and so the principle of indifference does not apply here. Specifically, we can take clues from the non-commutativity of certain variables in quantum mechanics that we might have gotten the measure of the space wrong, and hence try a different measure to see if it allows us to explain more.

We want to stress, however, that just because a measure *can* remove fine-tuning does not mean it necessarily does. The measure itself may be fine-tuned, in the sense that it requires a large number of details to be specified. Whether that is so must be evaluated for each model on a case by case basis. But since we know already that the measurement settings are sufficient to obtain the correct predictions of quantum mechanics as average values (because that is what we do in quantum mechanics), it is reasonable to think that the measure need not be fine-tuned. We will now show this with a simple example.

9.2.2 Statistical Independence in the CHSH inequality

Before turning to IST as a natural example for a non-trivial measure, we will go through a simple example, the CHSH inequality [Cla+69]. The CHSH setting describes a measurement of two entangled particles with two different detectors, commonly assigned to two observers, Alice (A) and Bob (B). The inequality states that any locally causal hidden variable theory which respects Bell-Statistical Independence fulfils

$$\left|E(X_0,Y_0)-E(X_1,Y_0)+E(X_0,Y_1)+E(X_1,Y_1)\right| \leq 2 \,, \tag{9.6}$$

where $X_{0/1}$ are detector settings on Alice's side and $Y_{0/1}$ are the detector settings on Bob's side. The quantum correlations $E(X_{0/1},Y_{0/1})$ are the expectation values of the results, $A(X_{0/1})$ and $B(Y_{0/1})$. $A(X_{0/1})$ is Alice's result given Alice's setting $X_{0/1}$, and $B(Y_{0/1})$ is Bob's result given Bob's setting $Y_{0/1}$. We will consider the usual case in which there are only two possible measurement outcomes relative to those settings, $A(\cdot),B(\cdot) \in \{-1,+1\}$.

In experimental tests of the CHSH inequality, the correlations for four different combinations of settings are estimated from four separate sub-ensembles of particles. Physical statistical independence is then the assumption that

$$\tilde{\rho}(\lambda|X_0Y_0) = \tilde{\rho}(\lambda|X_0Y_1) = \tilde{\rho}(\lambda|X_1Y_0) = \tilde{\rho}(\lambda|X_1Y_1) = \tilde{\rho}(\lambda) \,, \tag{9.7}$$

that is, the distribution of hidden variables does not depend on the (combination of) detector settings. We will assume that our hidden variables theory fulfils this assumption.

We now explain how the CHSH inequality can be violated in a hidden variables model by violating Bell-Statistical Independence but not physical statistical independence. For this, we first assume that the entangled state is represented by a hidden variable, λ, which has the probability distribution $\tilde{\rho}(\lambda,XY)$. Secondly, we assume the measurement outcome is determined by the hidden variable and the settings: $A(\lambda,X),B(\lambda,Y)$.

We denote the space of all the hidden variables with Λ, and divide it up into subsets for each possible combination of detector settings and outcomes Λ^{AB}_{XY}. That is, the subset Λ^{++}_{00} contains all

λs for setting X_0Y_0 that will give the result $A = +1, B = +1$, the subset Λ_{00}^{+-} contains all λs for setting X_0Y_0 that will give the result $A = +1, B = -1$, and so on.

However, we will next assume that the hidden variable cannot occur for two different combinations of measurement settings. If the variable λ described the case with setting X_0Y_0, then the combination (λ, X_0Y_1) is in $\mathscr{S}_{math}$ but not in $\mathscr{S}_{phys}$. This means

$$\Lambda = \bigcup_{i,j,k,l} \Lambda_{kl}^{ij} \quad \text{for} \quad i,j \in \{+,-\} \wedge k,l \in \{0,1\} , \tag{9.8}$$

but that these spaces are mutually disjoint

$$\Lambda_{kl}^{ij} \cap \Lambda_{cd}^{ab} = \emptyset \quad \text{for} \quad ijkl \neq abcd . \tag{9.9}$$

This does not contradict Eq. 9.7, because physical statistical independence is a statement about the probability distribution. The probability distributions for the four different combinations of settings can be made similar to arbitrary precision, even though no value of λ appears twice.

Imagine for example that λ is a real number from the interval $[0,1] \in \mathbb{Q}$. We randomly sample N points from this interval using a uniform distribution and assign each to one combination of settings and outcomes, i.e. one of the Λ_{kl}^{ij}. The so-generated probability distributions will be statistically indistinguishable for $N \to \infty$, even though the probability that λ appears for two different settings is zero. This is true regardless of how many of the λs we assign to each detector, provided the cardinality of all subsets is the same. In the limit $N \to \infty$ the distribution for λ conditioned on one of the detectors will just be uniform on each subspace Λ_{kl}^{ij}.

However, importantly, the measures of the subspaces Λ_{kl}^{ij} don't have to be the same. All we need to do now is choose $\tilde{\rho}$ to be constant, and the measure of the space Λ_{kl}^{ij} to be proportional to the quantum mechanical probability $P\Big(A(X_k)B(Y_l)|X_iY_j\Big)$ (the constant of proportionality will cancel with the normalisation of $\tilde{\rho}$) for each possible combination of outcomes. Since λ together with the detector setting determines the outcome, the outcome isn't an independent variable. If we want to calculate the expectation value for a certain combination of measurement settings in the hidden variables model, we therefore have

$$\begin{aligned} E(X_k, Y_l) = & \sum_{AB} \int_{\Lambda_{kl}^{AB}} d\lambda \; A(\lambda, X_k) B(\lambda, Y_l) \tilde{\rho}(\lambda) \tilde{\mu}(\lambda | X_k Y_l) \\ = & \sum_{AB} A(X_k) B(Y_l) P\Big(A(X_k)B(Y_l)|X_kY_l\Big) , \end{aligned} \tag{9.10}$$

which reproduces the correlations of quantum mechanics.

This might seem trivial: we just pushed the quantum mechanical correlation into the definition of the physical state space, then uniformly sampled the hidden variables over this space. This way, physical statistical independence is respected because the correlation comes from the definition of the space rather than from the distribution over it.

The above example can be generalised straight-forwardly to any quantum mechanical measurement, regardless of what variables are measured in what order or how many detectors there are. The above construction will always give the exact same result as quantum mechanics. In particular, it obeys the same bounds and violate the same inequalities as quantum mechanics.

This example is admittedly somewhat pointless, because we did not specify the model sufficiently to even know whether it was locally causal. However, any deterministic hidden variable theory that violates local causality can be made locally causal at the expense of violating Bell-Statistical Independence. We have further shown—contrary to what is often stated—violating Bell-Statistical Independence does not necessarily require correlations between the hidden variables and measurement settings.

This is not the only way to reproduce quantum mechanics with a locally causal and deterministic model without fine-tuning [DH22], but it is a nice example to see just why the distribution of the hidden variables is not fine-tuned. It is uniform on the sample-space, and the sample-space just describes what happens in reality. If the detector setting is one thing, it is not also another thing. The correlations come from the sample-space itself, and the information that goes into the construction of the sample-space is just the same as in quantum mechanics. Hence, this model is exactly as fine-tuned or not fine-tuned as quantum mechanics. That is to say, a rational reader who has no quarrels with quantum mechanics should have no quarrels with this model either.

9.2.3 Invariant Set Theory

The Mathematical Basis of IST We next consider supermeasures in the context of IST. In the following we do not directly use the fractal structure of an invariant set. That invariant sets are generically fractals merely serves as a motivation to consider a finite discretisation of Hilbert space in which certain combinations of states do not exist. As we explain below, this discretisation naturally acts as a supermeasure.

IST [Pal20a; Pal95; Pal09] is a putative theory of quantum physics based on the assumption that the universe is a causal deterministic dynamical system whose state-space is a fractal set, I_U. This fractal set corresponds to $\mathscr{S}_{phys}$ of the previous section. It is invariant under the action of dynamical equations: if a point lies on I_U, its time evolution always lies on I_U; if a point does not lie on I_U, its time evolution never has and never will. The nontrivial measure $\tilde{\mu}$ is the measure of this invariant set. It is sometimes called the invariant measure. Each trajectory of I_U actually comprises a Cantor Set's worth of trajectories. As such $\tilde{\mu}$ is a Hausdorff measure [Rog98]—a generalisation of Lebesgue measure for spaces with non-integer dimension.

Notably, if the state space is a fractal, it has gaps. (Indeed one could say it is mostly gaps, since the set has measure zero in the continuum embedding space.) In IST, the states in the gaps are not ontic, they are counterfactual states that are mathematically possible (and hence lie in $\mathscr{S}_{math}$), but are inconsistent with the assumed laws of physics (and hence do not lie in $\mathscr{S}_{phys}$). With respect to a Euclidean metric on $\mathscr{S}_{math}$, the restricted states are arbitrarily close to the allowed states. That is, perturbations (which are tiny with respect to the Euclidean metric) will generically take an allowed state to one that is inconsistent with the assumed laws of physics. This does not make the theory fine-tuned as perturbations which take ontic states off the invariant set are necessarily p-large with respect to a p-adic norm. Such a norm, and associated metric, is the natural one to use on a fractal geometry.

The fractal structure that underpins I_U is further assumed to be isomorphic to the p-adic integers, for some very large p. The theory of dynamical systems defined on p-adic numbers is an established part of arithmetic dynamics (see [WS98] and references therein). Indeed the famous Lorenz model based on three simple ordinary differential equations

$$\begin{aligned}\frac{dX}{dt} &= \sigma(Y-X)\\ \frac{dY}{dt} &= X(\rho-Z)-Y\\ \frac{dZ}{dt} &= XY-\beta Z\end{aligned} \tag{9.11}$$

provides a motivational example of what we have in mind. No matter where in state space these equations are initialised, the solutions of the Lorenz equations define trajectories which, after an infinite length of time, fall onto the fractal Lorenz attractor. IST proposes that laws of physics are not based on differential equations like Eq. 9.11, but on geometric equations which describe the attractor. With such laws, a point in state space which does not lie on the attractor is not physically consistent with such laws. Such a point is assigned a prior probability of zero. On the other hand, it is impossible to know *a priori* whether a point lies on the attractor: the geometric properties of fractal structures like the Lorenz attractor are formally non-computable [Blu+98; Dub93].

IST in its current form does not have a dynamical law. However, this is commonplace in the quantum foundational literature as in many cases one only cares about transition amplitudes between initial and final times. Those amplitudes are in addition often between spin states, so that one does not need to consider a space-time evolution. An example of this is Spekkens's Toy Model [Spe07], which does not have a dynamical evolution equation but has still proved useful. Like in Spekkens's Toy Model, we study here what insights we can extract directly from the structure of state space.

As a consequence of this fractal structure, an ensemble-based probabilistic state of the system in IST can be expanded in the basis of detector eigenstates $|A_j\rangle$ in the form

$$|\psi\rangle = a_1|A_1\rangle + a_2|A_2\rangle \ldots + a_J|A_J\rangle\ , \tag{9.12}$$

where the modulus-squared of the coefficients sun to 1 (i.e., $\sum_j a_j a_j^* = 1$). The crucial difference to standard quantum theory is that in IST the complex amplitudes a_j belong to a subset of the complex numbers $a_j \in \mathbb{C}_p, \mathbb{C}_p \subset \mathbb{C}$. The elements of $\mathbb{C}_p$ obey rationality restrictions on the coefficients (and so do not form a field). Specifically, if we write a_j in polar form $a_j = R_j e^{i\phi_j}$ then the fractal structure of I_U demands

$$R_j^2 = m_j/p; \quad \phi_j = 2\pi n_j/p \ , \tag{9.13}$$

where $m_j, n_j, p \in \mathbb{N}_0$, $m_j, n_j < p$.

This discretisation is effectively a nontrivial measure, allowing the violation of Bell-Statistical Independence—it gives measure zero to any states whose coefficients in Eq. 9.12 do not obey the conditions in Eq. 9.13. This means distributions over the set of allowed states will violate the assumption of Bell-Statistical Independence, even if the distributions themselves contain no information about the detector setting.

In the limit $p \to \infty$, the set of such "rational" Hilbert states is dense in the projection of the quantum mechanical Hilbert-space. From this point of view, for large enough p, IST can be made as experimentally indistinguishable from quantum theory as one likes. This makes it difficult to devise experimental tests for this idea. Such tests must ultimately be based on the fact that, at the end of the day, p is some finite number. We discuss this further in Section 9.3 and [HPR21].

However, no matter how large is p, the state-space of this theory will continue to have gaps. That is to say, the limit $p \to \infty$ is singular: quantum mechanics does not correspond to IST in the large p limit. Importantly, no matter how large is p, ensembles of trajectories described by Hilbert States where the complex amplitudes do not belong to $\mathbb{C}_p$, do not lie on I_U. As such, these trajectories have the measure $\tilde{\mu} = 0$ in the continuous embedding space.

With this IST explains why it is impossible to simultaneously measure conjugated variables in quantum mechanics with certainty. In quantum theory, this is a consequence of having non-commuting operators acting on a Hilbert-space, but is otherwise unexplained. In IST this arises in a deterministic framework because of the geometric structure of the invariant set and associated fractal measure. The incomplete algebraic structure of $\mathbb{C}_p$ reflects the "gappy" geometric structure of I_U. For example, superpositions of two states which are in $\mathbb{C}_p$ are generically not also in $\mathbb{C}_p$.

Of course, thinking about rational numbers and constraints among which states can mutually exist does not explain all results of quantum mechanics—that would require, amongst other things, a dynamical law. It does however provide a mathematical basis for the impossibility of measuring certain combinations of variables at the same time.

The CHSH Inequality in IST [Pal20a] previously demonstrated that IST correctly reproduces the observed violations of Bell's inequality and the results of sequential Stern-Gerlach experiments. We here use the CHSH inequality to explain how the measure of state space is relevant to obtain the correct probabilities.

In IST, if we keep the setting of the first detector fixed (say, at X_0), then the settings Y_0 and Y_1 of the second detector cannot both be on the set if $Y_0 \neq Y_1$. If Y_0 was on the set together with X_0, then Y_1 won't be on it, and vice versa. One of them is not physically possible. This property is a consequence of the rationality requirement on the amplitudes and phases, and a number-theoretic result known as Niven's theorem [Niv56; Jah10]:

Niven's Theorem: *Let* $\phi/2\pi \in \mathbb{Q}$. *Then* $\cos\phi \notin \mathbb{Q}$ *except when* $\cos\phi = 0, \pm\frac{1}{2}, \pm 1$.

If the first combination of settings, $(X_0 Y_0)$ for given λ, fulfils the rationality condition, then the second one, $(X_0 Y_1)$ for the same λ, can't fulfil it; it is associated with a state in $\mathscr{S}_{math}$ where $\tilde{\mu} = 0$.

This means

$$\tilde{\rho}_{\mathrm{B}ell}(\lambda|X_0Y_0) \neq 0 \implies \tilde{\rho}_{\mathrm{B}ell}(\lambda|X_0Y_1) = \tilde{\rho}_{\mathrm{B}ell}(\lambda|X_1Y_0) = 0 \tag{9.14}$$

$$\tilde{\rho}_{\mathrm{B}ell}(\lambda|X_1Y_0) \neq 0 \implies \tilde{\rho}_{\mathrm{B}ell}(\lambda|X_0Y_0) = \tilde{\rho}_{\mathrm{B}ell}(\lambda|X_1Y_1) = 0 \tag{9.15}$$

etc. However, this does not imply

$$\tilde{\rho}(\lambda|X_0Y_0) \neq 0 \implies \tilde{\rho}(\lambda|X_0Y_1) = \tilde{\rho}(\lambda|X_1Y_0) = 0 \tag{9.16}$$

$$\tilde{\rho}(\lambda|X_1Y_0) \neq 0 \implies \tilde{\rho}(\lambda|X_0Y_0) = \tilde{\rho}(\lambda|X_1Y_1) = 0 \tag{9.17}$$

etc.

As the statistics of trajectories in IST can be represented by complex Hilbert vectors over $\mathbb{C}_p$ for large p, the measures of the spaces Λ^{ij}_{kl} introduced in Subsection 9.2.2 are proportional to the quantum mechanical probabilities $P\big(A(X_k)B(Y_l)|X_kY_l\big)$. Hence IST reproduces the correlations of quantum mechanics.

Free Will As mentioned at the beginning of this Section, Bell-Statistical Independence is sometimes called the Free Will and/or Free Choice assumption. This refers to the notion that if, for example, Alice actually chose X_0 and Bob Y_0, then, keeping λ and Alice's choice fixed, Bob could have chosen Y_1 instead—he had the freedom to have done otherwise. But as we have also discussed, such a counterfactual state is incompatible with IST: if Bob actually chose Y_0, then the state corresponding to the triple (λ, X_0, Y_1) is not in the physical state-space; it corresponds to a point in $\mathscr{S}_{math}$ where $\tilde{\mu} = 0$.

But does this constraint actually have anything to do with free choice? In a deterministic framework, like IST, free will can at best be interpreted as the absence of constraints that could prevent an agent from doing as they please. But there is nothing in IST that would prevent an agent from doing as he or she pleases any more than is always the case in any deterministic theory—the laws of nature always constrain what we can do. And just as it is possible to violate Bell-Statistical Independence without violating physical statistical independence, it is possible to violate Free Choice (the assumption in Bell's Theorem) without violating free choice—after all, Free Choice is just a fancy name for Bell-Statistical Independence. That is, in IST Bob can freely choose *among the physically possible options* in the sense that there is no constraint on them.

While the difference between the correlations being on state space or in the probability distributions makes little difference to our observable world, it has relevance to the difference between Free Will as an assumption in Bell's Theorem, and "free will" as debated in metaphysics. Were the "free will" debate based on what we observe or experience, we would be tempted to say that moving correlations to a supermeasure makes no difference; however, most of the free will debate is metaphysical—it has nothing to do with what we observe or experience. It concerns the question of what we even mean by being "free" from something. And for that part of the debate, it matters whether an event or process is even physically possible (in the physical state space) or merely mathematically possible.

For instance, no one seems to ever be worried that we are not free to move around with complex-valued velocities or momenta. Mathematically, this is totally possible. It just does not happen in reality—observables are Hermitian operators. However, no one has ever argued that this restricts our "free will". Why not? Because that velocities are real-valued is just a property of the universe. Obtaining a complex-valued velocity is not a physically possible change, so we do not think it restricts our "free will". (It is probably something most people do not think about in the first place.) It is for this reason that the distinction between mathematical and physical possibilities matters, even though this does not affect what actually happens experimentally.

In some parts of the literature (as we sketch in Subsection 7.2.5), authors have tried to distinguish two types of theories which violate Bell-Statistical Independence: those which are superdetermined, and those which are retrocausal. The most naive form of this (e.g. [Sen21]) seems to ignore the prior existence of the measurement settings, and confuses a correlation with a causation. More generally, we are not aware of an unambiguous definition of the term "retrocausal" and therefore do not want to use it.

In the supermeasured models that we consider, the distribution of hidden variables is correlated with the detector settings at the time of measurement. The settings do not cause the distribution. We find Adlam's terms—that superdeterministic/supermeasured theories apply an "atemporal" or "all-at-once" constraint—more apt and more useful [Adl22].

While Bell's theorem is often said to imply that local causality (which is violated by standard quantum mechanics) cannot be restored with a deterministic hidden variables theory, this is only correct if the hidden-variables theory respects Bell-Statistical Independence. Violations of Bell-Statistical Independence are commonly interpreted as implying a correlation between the measurement settings and the hidden variables which determine the measurement outcomes. However, as we have shown here, one can violate the (Bell-)Statistical Independence assumption in Bell's theorem without any correlations between the measurement outcomes and the hidden variables. The violations of Bell-Statistical Independence can instead come about by the geometry of the underlying state space. We have argued that this is a simple way to see that violating Bell-Statistical Independence does not require fine tuning. In the next Section we look again at the example we gave above of a model whereby violations of Bell-Statistical Independence come from the geometry of state space—IST.

9.3 Experimental Tests of Invariant Set Theory

For all the successes of modern physics over the last century-and-a-half, we have been left with two apparently incompatible branches—the nonlinear and deterministic general relativity, and the linear but indeterminate quantum theory. For us to have a Theory of Everything, that describes all observed physical phenomena, we need a way to unite these, so we can describe physical phenomena at any scale. However, due to their differing takes on the determinacy of the universe, this has so far proved difficult.

IST attempts to unify these three disparate branches by using insight from chaos theory to create a fully local and determinate model of quantum phenomena [Pal95; Pal09; Pal11; Pal12b; Pal12a; Pal14; Pal16; Pal17; Pal18; Pal19; Pal20a; Pal20b]. It does by assuming that the universe is a determinate dynamical system evolving precisely on a fractal invariant set in state space. The natural metric to describe distances on a fractal set is the p-adic metric. This replaces the standard Euclidean metric of distance between states in state space. A consequence of this is that putative counterfactual states which lie in the fractal gaps of the invariant set are considered distant from states which do lie on the invariant set, even though from a Euclidean perspective such distances may appear small. Given the uncomputability of the possible states on any given fractal attractor, we cannot in advance distinguish states allowed and disallowed by this metric—hence, in IST, quantum-scale phenomena appear random despite being deterministic.

p-adic numbers form a back-bone of modern number theory and as such provide a framework to describe quantum physics within a finite number-theoretic framework. An example of this how IST provide an explanation for complementarity, a concept underpinning the uncertainty principle in quantum mechanics. In IST, complementarity is an emergent phenomenon arising from a number-theoretic property of trigonometric functions known as Niven's theorem—that $\cos\phi$ is not a rational number when $\exp i\phi$ is a primitive pth root of unity. The complex Hilbert Space of standard quantum mechanics arises as a singular limit of IST when p is set equal to infinity.

However, despite showing how key examples of quantum phenomena can be described deterministically, the theory deviates from standard quantum physics in some of its predictions—mainly in ways which stem from the p-adic metric being finite. In this Section, we give these key points of deviation, and investigate the extent to which these could be used to experimentally test the theory.

9.3.1 Entanglement Limits

In standard quantum theory, there is no limit to the number of quantum objects n which can be n-partite maximally entangled (i.e. saturate the Coffman-Kundu-Wootters inequality for an n-partite state [CKW00; OV06]). However, in IST, there is. Here, we codify this limit, and design experiments using optics/NISQ devices to probe it.

For this, we use the M-qubit W state [DVC00; RA09],

$$\left|W^M\right\rangle = \frac{1}{\sqrt{M}} \sum_{i=0}^{M-1} |0\rangle^{\otimes i} |1\rangle |0\rangle^{\otimes(M-1-i)} \tag{9.18}$$

(where $|\psi\rangle^{\otimes M}$ is the tensor product of $|\psi\rangle$ with itself M times). For instance, the $M = 3$ W state is

$$\left|W^3\right\rangle = \frac{|100\rangle + |010\rangle + |001\rangle}{\sqrt{3}} \tag{9.19}$$

The W state is a maximally entangled state of M qubits—and in standard quantum theory, there is no limit to how high M can be.

However, in IST, the finiteness of the p-adic metric provides a limit to the number of qudits that can be maximally entangled. For multiple-qubit entanglement, this limit is codified in [Pal20a] as a maximum of $\log_2 N$ qubits being able to be maximally entangled, in a p-adic system where the equatorial great circle of the Bloch sphere consists of N equally-spaced discrete points.

A system of maximally-entangled photon-vacuum qubits can be created using a single photon and a number of mirrors and 50:50 beamsplitters, as shown in Fig. 9.2. This naturally forms a W state across M qubits, and, by standard quantum theory, should potentially be able to be extended to $M \to \infty$. However, this disagrees with IST, which limits to a maximum of $M = \log_2 N$ entangled qubits, where the two orthogonal spherical dimensions of the Bloch sphere (θ and ϕ) are each discrete in N divisions. While p is expected to be very large, each qubit will only have been affected by $I = \log_2 \log_2 N$ beamsplitters, so, for realistic experimental beamsplitter loss of 0.1%, the chance of losing a given qubit to decoherence only reaches 1% once the system has entangled over 1000 qubits, which is only possible by IST if $N \geq 10^{250}$ (as we show in Fig. 9.3). Further, an advantage of the W state is, even if decoherence effectively measures one of the qubits, so long as the result is 0 (the photon isn't in that mode), this collapse leaves the remaining qubits still maximally entangled in the $(M-1)^{th}$ W state.

Even if we obtain this state, we need to prove it is entangled. Gräfe et al [Grä+14] and Heilmann et al [Hei+15] have done this for 8 and 16 qubit W states respectively, confirming that they generated an entangled W state of that size (assuming they inputted a single photon), and NISQ devices such as Wang et al's integrated silicon photonics chip could be used to do this for a 32-qubit W state [Wan+18]. It is an ongoing problem to specifically discern an entanglement-confirming optical layout for an arbitrarily-large W state, but Lougovski et al give the quantum-information-theoretical groundwork for doing so [Lou+09]. This involves using beamsplitters to shift the optical-path modes to instead each represent one possible permutation of phase combinations for the sub-components (ignoring the global phase of the state). For instance, for the 4-qubit W state, combining beamsplitters after the state creation so as to have each final

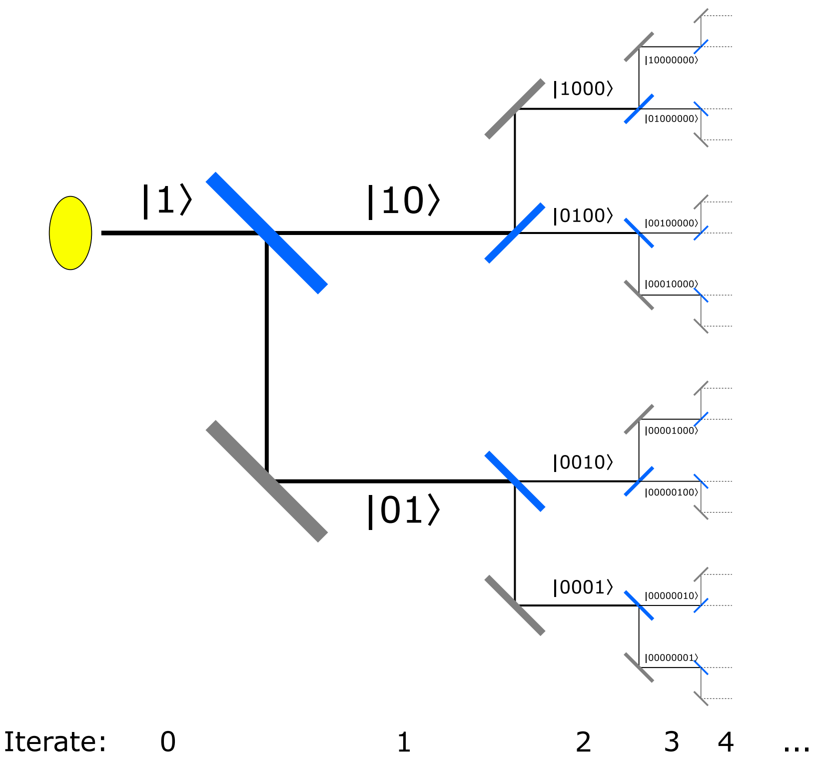

Figure 9.2: The first 4 iterates of a set-up to create the entangled state $|W^M\rangle$ (as given in Eq. 9.18), where $M = 2^I$ at the I^{th} iterate. The diagonal blue lines are 50:50 beamsplitters, the diagonal grey lines mirrors, the yellow oval a single-photon source, and the black lines the possible paths of the photon. Given this maximally entangles 2^I qubits, IST predicts entanglement generated by an experiment like this should begin to fail after $I = \log_2 \log_2 N$ iterations, where the two spherical dimensions of the Bloch sphere are each N-discrete. We can test whether this entanglement holds or fails by putting mirrors at the ends of each path—if the photon returns with 100% probability to the input port, it was maximally entangled; if each beamsplitter splits it evenly, such that it only returns 2^{-I} of the time, the entanglement has decayed completely. Return probabilities between show various levels of entanglement decay.

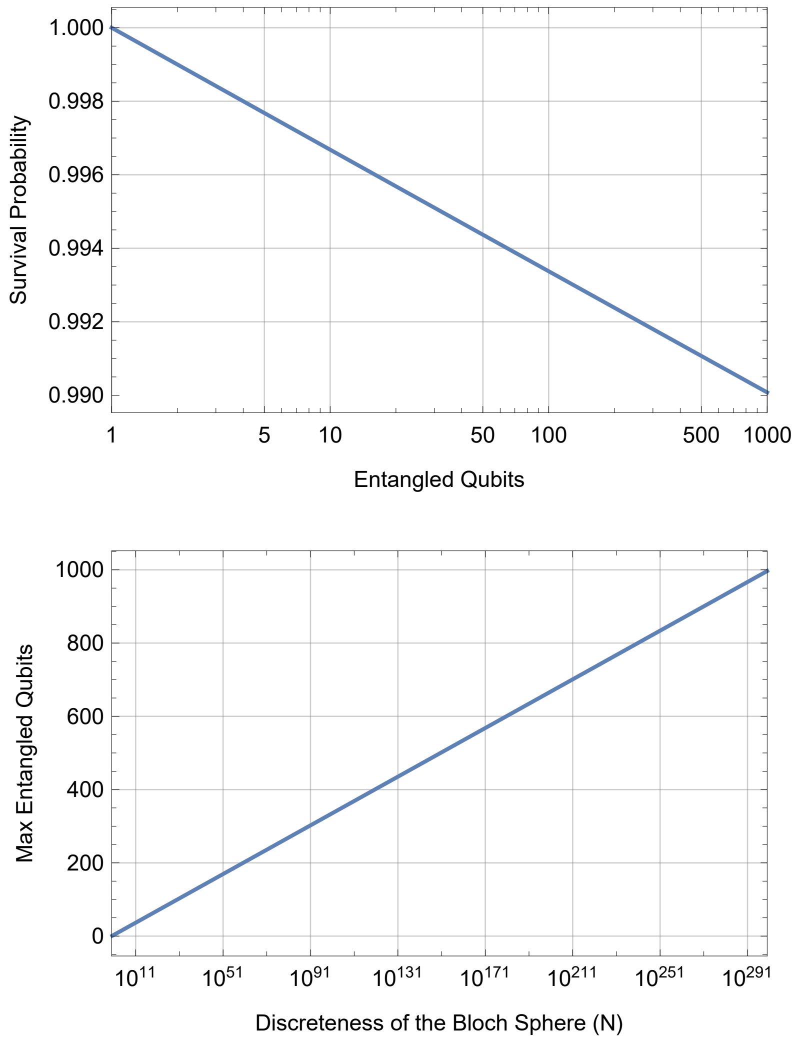

Figure 9.3: The survival probability of each qubit for a given set of entangled qubits created using the experiment in Fig. 9.2, and the maximum number of entangled qubits that can be created in a version of IST where the p-adicity causes the Bloch sphere to be split into N divisions in each angular direction. This shows how this beamsplitter experiment allows us to test this entanglement limit for very high-p versions of IST, due to the comparative lack of loss-induced decoherence on the W state created.

path act to project on one of the 4 states

$$\begin{aligned}
\left|W_1^4\right\rangle &= (|1000\rangle + |0100\rangle + |0010\rangle + |0001\rangle)/2 \\
\left|W_2^4\right\rangle &= (|1000\rangle - |0100\rangle - |0010\rangle + |0001\rangle)/2 \\
\left|W_3^4\right\rangle &= (|1000\rangle + |0100\rangle - |0010\rangle - |0001\rangle)/2 \\
\left|W_4^4\right\rangle &= (|1000\rangle - |0100\rangle + |0010\rangle - |0001\rangle)/2
\end{aligned} \tag{9.20}$$

Doing this means a consistent detection on just one of the paths over many runs (e.g. the one corresponding to just $\left|W_1^4\right\rangle$) indicates a pure entangled state is consistently being created (specifically here the state $\left|W_1^4\right\rangle$). Were the entanglement to break, the detections would begin to spread between the targeted state $\left|W_1^4\right\rangle$ and the other three states, until, for a maximally mixed state, each detector would click 25% of the time.

In the same way, for the I^{th} iterate, consisting of $M = 2^I$ qubits, using linear optical components one can project the eventual state into one of the 2^I phase permutations of $\left|W^M\right\rangle$, and so detect with certainty that a pure entangled state of M qubits was created. Interestingly, preparing these states to certify entanglement requires each optical mode to again only interact with I beamsplitters, to allow us to certify $M = 2^I$-qubit entanglement, which simply squares the survival probability. This means for 1000 qubits, it becomes 98% rather than 99%. Given the resilience of the overall state to loss-induced decoherence, and the fact that Lougovski et al show this certification method also allows us to detect any entangled states of fewer than M qubits, this loss probability poses very little issue to our test of IST. Further, despite the loss, the total number of surviving (maximally-entangled) qubits tends to infinity as I tends to infinity, rather than peaking at a certain value.

This W state-based experimental analysis of IST can be extended by looking at an experiment such as Rarity and Tapster's, where a pair of photons are generated in a cone of possible positions. Here, the angular position of one photon is anti-correlated with the position of the other [RT90]. We show this in Fig. 9.4. Considering just one photon in the cone, this is equivalent to a W state where M is the number of sectors into which you subdivide the cone. Adding a second photon, position-entangled with the first, doubles the number of entangled qubits in the system.

Rarity and Tapster also give a way to prove these photons are entangled—by interfering them to violate a Bell inequality. However, as this is done assuming their position is a continuous variable, we need to adjust it to prove just how large it holds for as discrete variables.

This can be done by making a set of $2M$ apertures on the circumference of the cone, and splitting the ring into two half-circumferences. After this, similarly to what we do in Fig. 9.2, we can iteratively combine adjacent apertures to get position-momentum entanglement between adjacent apertures. Once this projects to equal superpositions across all M apertures on each half-circle, we can record detected position for each half-circle's photon. By comparing the final detected position between upper half-circumference and lower half-circumference, and seeing if they still correlate, we can confirm this double-W_M state.

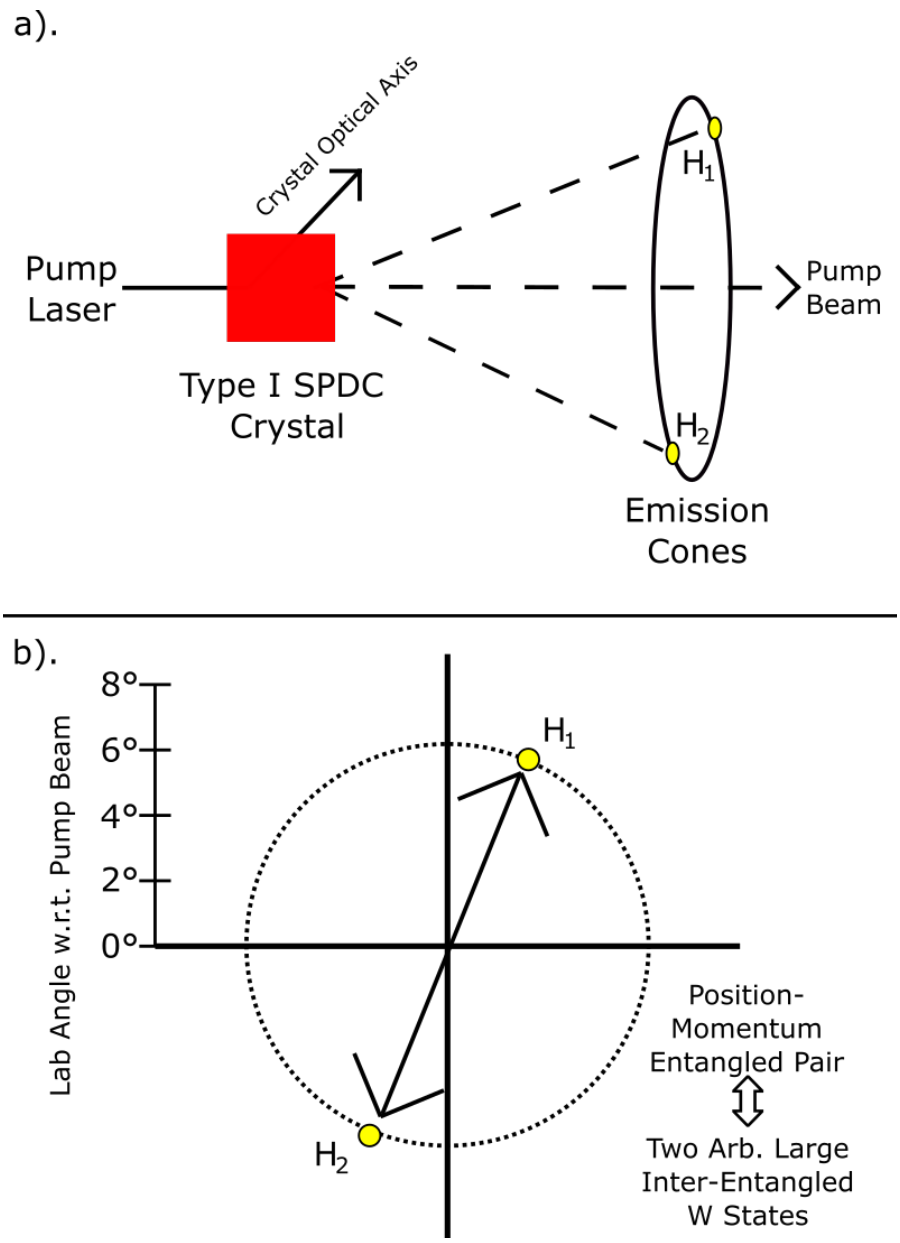

Figure 9.4: Type I SPDC source for the generation of pairs of position-and-momentum-entangled photons, as given by Rarity and Tapster [RT90]. The generated position of each photon on the cone can be viewed as a W state of arbitrary number of qubits N, and so the system of the two photons is a double-W state of $2N$ qubits. This arbitrary number of qubits N can be lower-bounded as the resolution of a circular single-photon position detector array used to detect where on the circle each photon is emitted.

While the phase between the upper photon and some other discrete division in the upper half will be random, it will be the same as the phase between the lower photon and some discrete division in the lower half. The correlation is always the same, but specific phases at different points on the circumference are not. This is why, using the two photons (and two split half-circles), we can prove the correlations still exist—a similar (continuous-variable) method was used by Rarity and Tapster to provably violate a Bell inequality.

9.3.2 No Continuous Variables

A second, related implication of IST is that it permits no continuous quantum variables.As the p-adic metric used in IST is necessarily finite-dimensional, the space of states allowed must also be finite. Since we can lower bound the number of states allowed as the dimension of the Hilbert space we use (to replicate classical information theory), we can say that, the existence of a qudit of dimension d implies a state space of at least dimension d (e.g. a qubit requires at least two distinct states: 0 and 1; a qutrit requires 3 states: 0, 1 and 2, etc...). Hardy extends this argument, saying that, to satisfy his axioms for quantum theory, between any two pure states in a system, there needs to be a continuous reversible transformation available on a system that goes from one to the other. To allow this, Hardy argues a qudit of dimension d requires a state space of dimension d^2 [Har01].

This means for continuous variables to exist, given they have an infinite-dimensional Hilbert space [BV05], there must be an infinite number of states allowed. This violates IST. Therefore, in IST, there can be no quantum continuous variables.

In standard quantum physics, many variables are continuous (e.g., position, momentum, electric field strength, and time) [Wee+12]. Therefore, for IST to hold true, all of these variables would actually need to be discrete: of finite (but very high) dimension. While some models hold one or another of these variables to be discrete (e.g., space-time in Loop Quantum Gravity [Rov98] and certain toy models of the Universe [FS14a]), the idea that all 'continuous' variables are actually discrete would be controversial.

9.3.3 Gravitational Decoherence

Palmer describes IST as not so much a quantum theory of gravity (like String Theory and Loop Quantum Gravity), but a gravitational theory of the quantum [Pal16]. Aside from its determinate nature, nowhere is this more true than in how IST models regimes where gravitational and quantum effects are both present. The paper *Invariant Set Theory* describes the theory as positing no gravitons and so no supersymmetry (spin-2 gravitons typically being seen as hinting at supersymmetry) [Pal16]. Instead, the paper holds that gravity is inherently decoherent, turning gravitationally-affected superpositions into maximally mixed states. This paper also claims that effects typically considered signs of either dark matter or dark energy could instead be in some way due to various manifestations of the "smearing" of energy-momentum on space-times

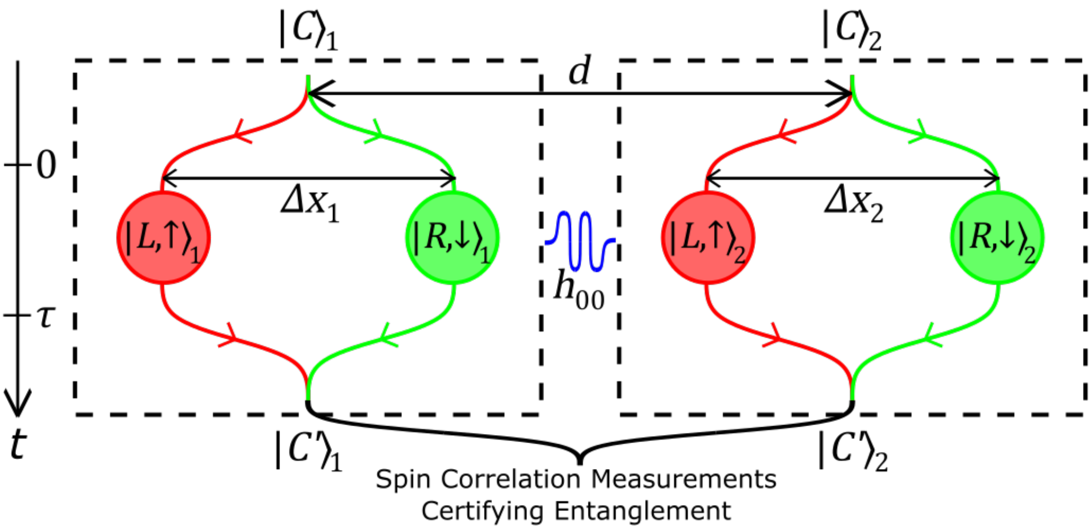

Figure 9.5: The experiment described by Bose et al [Bos+17] and Marletto and Vedral [MV17], for testing gravity's ability to entangle two masses. Two masses, m_i for $i \in \{1,2\}$ are separated from each other by distance d. Both are initially in state $|C\rangle_i$, with embedded spin $(|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$. They are then both admitted into Stern-Gerlach devices, which put them both into the spin-dependent superposition $(|L,\uparrow\rangle_i + |R,\downarrow\rangle_i)/\sqrt{2}$, where $|L\rangle_i$ and $|R\rangle_i$ are separated from each other by distance Δx_i. They are left in these superpositions for time τ. During this time, if gravity is quantum-coherent, evolution under mutual gravitational attraction h_{00} would entangle the two particles, adding relevant phases to both. After time τ, an inverse Stern-Gerlach device is applied to return each mass to their initial state (potentially modulo the phases applied by h_{00}). By applying this process, and measuring spin correlations between the two particles after each run, we can detect if relative phases have been applied to each, and so if gravity is coherent. For IST to hold, gravity must be decoherent, and so cannot entangle two masses. This means IST predicts no alteration of phases will be detected.

neighbouring our universe $\mathscr{M}_U$ on the invariant fractal set I_U influencing curvature of $\mathscr{M}_U$. It claims this smearing avoids precise singularities in $\mathscr{M}_U$: avoiding singularities being a key goal of many previous attempts to quantise General Relativity.

Palmer suggests an alteration of the Einstein Field Equations (EFEs) based on the presence and effects of possible universes $\mathscr{M}'_U$ on our universe $\mathscr{M}_U$, leading to the EFEs instead being

$$G_{\mu\nu}(\mathscr{M}_U) = \frac{8\pi G}{c^4}\int_{\mathscr{N}(\mathscr{M}_U)} T_{\mu\nu}(\mathscr{M}'_U)F(\mathscr{M}_U,\mathscr{M}'_U)d\mu \tag{9.21}$$

where $F(\mathscr{M}_U,\mathscr{M}'_U)$ is some propagator to be determined and $d\mu$ is a suitably normalised Haar measure in some neighbourhood $\mathscr{N}(\mathscr{M}_U)$ on I_U [Pal16]. Note, in this altered form of the EFEs, the cosmological constant Λ is set to zero, given Palmer claims the alteration would separately resolve the issue of dark matter and the acceleration of the expansion of the universe.

This gravitational decoherence could be tested by experiments that involve putting heavy objects in spatial superpositions. This would involve allowing them to gravitationally interact, then returning the spatial superposition components back to a single position, then seeing if there are any signs of entanglement between the objects from the resulting interference pattern (see Fig. 9.5) [Bos+17; MV17].

In such an experiment, assuming gravity is coherent, the combined state of the two masses initially is

$$|\Psi_{Init}\rangle_{12} = (|\uparrow\rangle_1 + |\downarrow\rangle_1)(|\uparrow\rangle_2 + |\downarrow\rangle_2)|C\rangle_1|C\rangle_2/2 \tag{9.22}$$

Passing both masses through a Stern-Gerlach apparatus, this combined state then evolves at $t = 0$ to

$$|\Psi(t=0)\rangle_{12} = (|L,\uparrow\rangle_1 + |R,\downarrow\rangle_1)(|L,\uparrow\rangle_2 + |R,\downarrow\rangle_2)/2 \tag{9.23}$$

After allowing the two masses to gravitationally interact for time $t = \tau$, the overall state has become

$$\begin{aligned}|\Psi(t=\tau)\rangle_{12} = \frac{e^{i\phi}}{2}\Big(&|L,\uparrow\rangle_1(|L,\uparrow\rangle_2 + e^{i\Delta\phi_{LR}}|R,\downarrow\rangle_2)\\ &+ |R,\downarrow\rangle_1(e^{i\Delta\phi_{RL}}|L,\uparrow\rangle_2 + |R,\downarrow\rangle_2)\Big)\end{aligned} \tag{9.24}$$

where

$$\begin{aligned}&\phi \approx \frac{Gm_1m_2\tau}{\hbar d}\\ &\phi_{RL} \approx \frac{Gm_1m_2\tau}{\hbar(d-\Delta x)},\ \phi_{LR} \approx \frac{Gm_1m_2\tau}{\hbar(d+\Delta x)}\\ &\Delta\phi_{LR} = \phi_{LR} - \phi,\ \Delta\phi_{RL} = \phi_{RL} - \phi\end{aligned} \tag{9.25}$$

After applying the opposite of the initial Stern-Gerlach interaction, the final state is

$$\begin{aligned}|\Psi_{End}\rangle_{12} = |C'\rangle_1|C'\rangle_2 = \Big(&|\uparrow\rangle_1(|\uparrow\rangle_2 + e^{i\Delta\phi_{LR}}|\downarrow\rangle_2)\\ &+ |\downarrow\rangle_1(e^{i\Delta\phi_{RL}}|\uparrow\rangle_2 + |\downarrow\rangle_2)\Big)|C\rangle_1|C\rangle_2/2\end{aligned} \tag{9.26}$$

However, if gravity isn't coherent, there are two possible final states. If gravity doesn't also collapse the state, the final state will be equivalent to the initial one ($|\Psi_{Init}\rangle_{12} = |\Psi_{End}\rangle_{12}$). However, if gravity does collapse the superposition, each particle will be forced into the (spin) maximally mixed state

$$|\Psi_{MM}\rangle_i = |C\rangle\langle C|_i(|\uparrow\rangle\langle\uparrow|_i + |\downarrow\rangle\langle\downarrow|_i)/2,\ i \in \{1,2\} \tag{9.27}$$

By measuring spin correlations to estimate the entanglement witness

$$\mathscr{W} = |\langle\sigma_x^{(1)} \otimes \sigma_z^{(2)}\rangle - \langle\sigma_y^{(1)} \otimes \sigma_z^{(2)}\rangle| \tag{9.28}$$

we can distinguish the entangled state from the two other possible final states (if $\mathscr{W} > \mathbb{1}$, the state is entangled), and so see if gravity is coherent. For IST to hold, $\mathscr{W}$ needs to be less than or equal to $\mathbb{1}$.

While such a test may sound challenging to implement, recent experimental work indicates it should be possible within the next few years [Del+20; Mar+21]. This test would also have the benefit of either ruling for or ruling out other theories which hold gravity to be decoherent (e.g. Penrose's gravitational collapse interpretation [Pen96]).

We have identified points of difference between IST and standard quantum mechanics. While not fatal to IST, they provide potential avenues to experimentally test the theory, to see whether its deterministic, fractal-attractor-based structure is compatible with observed reality.

9.4 Chapter Conclusion

In this chapter, we looked further at statistical independence-violating models of quantum mechanics. We showed that, while statistical independence-violating models are typically categorised into either superdeterministic or retrocausal models, there is actually a third type of model which violates statistical independence, which we introduced, those which are "supermeasured". We showed how this type of model differs from superdeterministic and retrocausal models, and how these differences can help proponents of models in this category respond to arguments typically levied against statistical independence-violating models. Finally, we focused specifically on an example of a supermeasured model—IST—and presented potential methods for how to experimentally test it against standard quantum mechanics.

The work in this Chapter illustrates the second part of the interplay between quantum foundations and quantum technologies—the use of current and developing/hypothetical technologies to test ideas in quantum foundations. In this Chapter, we first specified a subtype of model in which statistical independence is violated—supermeasured models—and described in detail a model which meets fits this categorisation: IST. We then went on to describe differences between the empirical predictions of IST and standard quantum mechanics. We then proposed experiments, based on current or near-future quantum technologies (e.g. NISQ devices), which would test which set of predictions most closely matches reality. This serves as an example of how these near-term quantum technologies allow us to probe into the foundational mysteries of quantum mechanics.

Chapter 10

Conclusion and Outlook

This final chapter first summarises the work presented in this thesis, before looking at related potential future research directions, with specific focus on directions which illustrate the links between quantum foundations and quantum technologies. We then discuss broader themes within this work.

10.1 Summary of Work

This thesis focused on two areas within quantum foundations—counterfactual communication, and statistical independence-violating extensions of quantum mechanics. We here summarise work presented in each of these two areas.

10.1.1 Counterfactual Communication

Part I of this thesis examined counterfactual communication, and the related effects of interaction-free measurement and counterfactual computation. Chapter 2 reviewed both proposed criteria to determine where a quantum particle has been, and the history of interaction-free and counterfactual effects. Chapter 3 analysed one of these proposed criteria for determining the path of a quantum particle—where a non-zero weak value of the spatial projection operator corresponding to a given path (or some operator comprising the product of the spatial projection operator with any other operator) implies a particle was on that path. We showed issues with these weak values approaches (such as Vaidman's weak trace approach), and discussed whether we should associate these weak values with the intuitive idea of presence.

Chapter 4 considered foundational issues related to counterfactual communication: whether protocols are counterfactual by the definitions in Chapter 2; and whether protocols are necessarily quantum. Chapter 5 developed protocols using counterfactual communication as a primitive, for the transfer of quantum rather than classical information. We first gave a protocol for enacting an arbitrary single-qubit unitary onto a qubit across a channel counterfactually. We then gave a protocol by which a controlled-Z gate operation can be applied across a channel onto a given qubit counterfactually. This second protocol allowed us to build on this counterfactual controlled-Z gate, for counterfactual quantum teleportation and telecloning. Chapter 6 turned counterfactual communication towards metrology, giving a protocol for counterfactual ghost imaging. This allows delicate samples to be imaged optically with far reduced (asymptotically zero) light flux through the object, and no light flux through the object on runs of the protocol where the object is successfully imaged.

Part I shows how an initially foundational idea, the Elitzur-Vaidman Bomb Tester, can still be adapted to develop novel and interesting quantum technological tools, even thirty years after it was originally described.

10.1.2 Extensions of Quantum Mechanics

Part II of this thesis looked at extensions of quantum mechanics which violate statistical independence.

Chapter 7 first reviewed Bell's Theorem. We examined the assumptions necessary to construct a Bell inequality. We then reviewed the experiments performed that show that Bell inequalities can be violated, and so that one of these assumptions must be false. We then looked more closely at models where one of these assumptions, the statistical independence assumption, are violated. These models are often referred to as superdeterministic or retrocausal. Finally, we introduced the measurement problem: even though all other time evolution in quantum mechanics is linear and unitary as governed by (forms of) the Schrödinger equation, the update of a quantum state after measurement is explicitly non-unitary. This is despite measurement not being a rigorously-defined part of quantum mechanics. We showed that the measurement problem motivates us to look for theories that extend quantum mechanics, and that developing models which violate statistical independence may give extensions which could resolve this problem.

Chapter 8 dived further into foundational research—evaluating the extent to which Harrigan and Spekkens's definitions ψ-ontic and ψ-epistemic correspond respectively to the wavefunction representing the state of the world (the wavefunction being ontic) and the wavefunction representing some experimenter's knowledge about the state of the world (the wavefunction being epistemic). We showed that, while Harrigan and Spekkens's definitions for ψ-ontic and ψ-epistemic are necessarily contradictory, the wavefunction being ontic doesn't necessarily exclude it from also being epistemic. This shows Harrigan and Spekkens's definitions fail to adequately represent these underlying ideas. We then identified a specific category of model excluded by Harrigan and Spekkens's definitions—those where the wavefunction represents an ensemble of identically-prepared states, with differing hidden variables (ψ-ensemble models). We showed

ψ-ensemble models which obey Bell's definition of local causality can be used to solve the measurement problem, so long as they violate statistical independence.

Finally, Chapter 9 specifically examined models which violate statistical independence. We showed that there is a third type of model which violates statistical independence, alongside superdeterministic and retrocausal—those which are "supermeasured". We illustrated how this type of model differs from superdeterministic and retrocausal models, and how these differences can help proponents of these models respond to typical arguments against models which violate statistical independence. Finally, we focused on a specific example of a supermeasured model—Invariant Set Theory (IST)—and presented potential ways to experimentally test it against standard quantum mechanics.

10.2 Future Work

This Section gives possible future directions for work. These are either related to the two case-study areas presented in this thesis, or more generally show the interplay between quantum foundations and quantum technologies. These new directions comprise: interaction-free polarimetry (Subsection 10.2.1), the interpretation and use of weak values (Subsection 10.2.2), the broader experimental testing of extensions of quantum mechanics (Subsection 10.2.3), the examination of data from quantum computing research for interesting patterns (Subsection 10.2.4), and philosophy of science in practice (Subsection 10.2.5).

10.2.1 Interaction-Free Polarimetry

Chapter 6 proposed counterfactual ghost imaging [HR21b]: a form of interaction-free metrology where pixels of a sample could be discriminated as optically transmissive or absorptive without any light flux through the sample. However, this and other interaction-free imaging protocols (such as those described in Chapter 2) suggest interaction-free imaging could be used for other metrological tasks.

A task this tool could be applied to is polarimetry—the determination of the effect of transmission through a certain sample on the polarisation of any arbitrary ray of light. (See [Tyo+06; Gol20] for more information on polarimetry.) Samples tested using current polarimetric protocols are often optically sensitive. An example is where polarimetry is used to identify the relative concentrations of chiral isomers of complex biological molecules [Boe+82; MPB01; Due+21]. We want to reduce the light flux through these samples when undertaking polarimetry on them. Interaction-free methods are ideally suited to this task, given they reduce light flux through a sample, potentially to vanishing levels.

Our preprint *Interaction-Free Polarimetry* does preliminary work on combining interaction-free methods with polarimetry [HR22]. However, this is still a work-in-progress, and needs validation and verification—hence we did not include it in this thesis. However, we hope this work can be developed further in the near future, and this tool deployed to aid in polarimetry—a potential future example of a useful quantum technology derived from quantum foundations.

10.2.2 Interpretation and Use of Weak Values

Chapters 2 and 3 mention that weak values can have practical applications in metrology, despite originally being a foundational curiosity (e.g., in Aharonov, Albert and Vaidman's paper, *How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100* [AAV88]). From weak value amplification, to protocols for the "direct measurement" of a quantum state (weak value state tomography), weak values can obtain more information in a given scenario than either classical or quantum projective measurement allow [Dre+14]. Therefore, weak values present another potential case study in how quantum foundations can aid the development of new/hypothetical quantum technologies. Similarly, quantum protocols/devices have been used to try and show that foundational inferences made using weak values hold true. Examples of this include Danan et al's *Asking photons where they have been* [Dan+13], or the experiment performed in Chapter 4 and the paper *Do the laws of physics prohibit counterfactual communication* [Sal+22]. Weak value metrology has even been used to observe the Bohmian trajectories of quantum particles [Koc+11; Mah+16]. This illustrates how an idea from quantum foundations can be developed into a quantum technological tool (in the broadest sense), then used to aid in quantum foundational research.

Due to this, an avenue for future research within the scope of this thesis is into weak values, both foundationally and for potential future applications. We performed some initial research in this area, both in Chapter 3 and the Appendix to this thesis. However, as this was not yet as fully developed as the research we performed into counterfactual communication or statistical independence-violating extensions of quantum mechanics, we decided not to include this work in its own Part.

A key direction for this research, both in the Appendix and beyond, is examining the difference between how we treat some quantities as "negligible in practice" (as being able to take some infinitesimally small limit of, and treat as zero), and how the necessarily non-zero size of these quantities affects actual experimental practice. This difference suggests that looking at how protocols and experiments are actually implemented is important for foundational research. For weak values, this specifically involves looking at the perturbation to a system caused by weakly measuring the system. In weak value theory, this perturbation is typically taken as being reducible to zero. However, initial results in Appendix B suggest different results in situations when there is strictly zero perturbation.

10.2.3 Broader Experimental Testing of Extensions of Quantum Mechanics

Chapter 9 presented initial ideas as to how we could experimentally test IST [Pal16], a supermeasured extension of quantum mechanics [HHP22]. However, this is one of a plethora of proposed extensions to quantum mechanics, to resolve our current issues with measurement and the relationship between quantum phenomena and gravity. A key aspect of these extensions is that they provide differing predictions to standard quantum mechanics in some parameter space. If we can access that parameter space experimentally, we can test whether that extension reflects observations better than standard quantum mechanics. This foundational phenomenology is a necessary if underappreciated endeavour—evaluating how the predictions of newly-proposed models differ from standard quantum mechanics, proposing experiments that use cutting-edge quantum devices and technologies to test those differing predictions, and properly identifying the set of models given results rule out. The best current examples of foundational quantum phenomenology are the tabletop quantum gravity testing programme [MV17; Bos+17; Del+20], and collapse-model testing programme [BG03; Bas+13]. Given developments in recent years on quantum technologies, it is surprising so few groups leverage this to investigate foundational ideas, even just to demonstrate their technological research.

Specifically testing models which violate statistical independence, proposed experiments such as Hossenfelder's [Hos11] are slowly becoming more feasible with modern quantum technologies (such as NISQ devices). These could bound which statistical-independence-violating models are compatible with observed reality. Therefore, an ideal direction for future work is collaborating with experimentalists to perform such experiments, to provide new insights into the foundational problems discussed throughout this thesis.

10.2.4 Examining Data from Quantum Computing Research

Also worth examining is data being generated in the current races between companies and research groups to be the first to create a working quantum computer either of a given size, or which can perform a given quantum algorithm [Wit22]. Achieving these goals will require large numbers of interconnected qubits, protected in some way from environmental noise and decoherence [CTV17]. (This protection is typically through application of quantum error correction protocols over even larger numbers of either qubits or Gaussian states [DMN13].) These quantum computers will comprise the largest entangled states ever controlled by man.

As discussed previously, some extensions of quantum mechanics differ from standard quantum mechanics in how they treat large entangled states. For instance, IST sets stronger limits on the size of maximally entangled states than the standard monogamy of entanglement [Pal16], and the Diosi-Penrose model involves gravitationally-induced collapse [Pen96]. Therefore it seems sensible for the foundations community to examine data from these experiments and NISQ devices: compare observed decoherence/error rates to those expected from standard quantum mechanics, and consider the foundational implications of practical "short-cuts" taken in quantum error-correction research to mitigate unknown-yet-quantifiable error-inducing effects. This will allow us to leverage the work done in developing quantum computers to investigate and potentially resolve foundational questions—or even formulate new questions entirely.

10.2.5 Philosophy of Science in Practice

As mentioned in Section 1.1, the work in this thesis acts as an example of philosophy of science in practice—bringing philosophers and (especially experimental) quantum physicists together to do collaborative research, both to consider philosophical problems within the field (philosophy-of-science in practice) and to present case studies of scientific practice to develop philosophical conclusions (philosophy of science-in-practice) [BL13].

A future direction would be to cement this interdisciplinary link between applied quantum physics/quantum technologies and philosophy of science, to expand these two research strands. Given how little interaction philosophers of science have had with quantum experimentalists (especially compared to their interaction with quantum theorists), this presents an underexplored research direction which could provide insights into major problems in quantum foundations. We expect the key insights of such a research programme to come from cases where abstractions and idealisations used by theorists fall apart when compared with actual experimental practice.

An example of this we mentioned in Subsection 10.2.2 is the idealisation with weak values where the effect of weak measurement on the system under consideration can be treated as negligible. As we discuss in Appendix B, experimentally this may not be necessarily true. While still under evaluation, this presents an example of the sort of insight philosophical analysis of real experimental practice could bring to the foundations of quantum mechanics.

10.3 Broad Themes

As discussed in Section 1.1, the overarching aim for this thesis was to show how quantum foundations and quantum technologies interrelate: how quantum foundations can inform the development of new/hypothetical quantum technologies, and how quantum devices and technologies can allow us to probe quantum foundations. This thesis examined two case studies for this interrelation: the development of counterfactual communication from the foundational effect of interaction-free measurement into a tool for both quantum information transfer and practical metrology (Part I); and the use of quantum technologies to test extensions of quantum mechanics where statistical independence is violated (Part II).

While some work in this thesis relates more specifically to either pure quantum foundations or pure quantum technologies (due to open-ended nature of academic research), the goal of this work was to illustrate the interconnectedness of these two areas. Between the two case studies, we have elucidated this interplay, as well as showing how certain examples may not neatly fit into either the schema of "quantum foundations inspiring quantum technologies", or "quantum technologies testing quantum foundations". For instance, this thesis showed cases where foundational criteria have been used to analyse and evaluate developing quantum technologies, that were themselves originally based on quantum foundations (as in Chapter 4); or where quantum protocols which were proposed to test one foundational question can instead be used to test another model entirely (as in Chapter 9). This shows the rich and rewarding area of potential developments at the intersection of research in quantum technologies and quantum foundations, and the potential for these research areas to re-seed new research directions in one another. Hopefully, by illustrating the progress which can be made even in two small sub-areas of this intersection, this thesis has advocated for greater interest in, and investigation of, the interplay between quantum foundations and quantum technologies.

APPENDIX

A

MISAPPLICATION OF WEAK MEASUREMENT

Declaration of contribution: This Appendix is adapted from the published article *Comment on "Scheme of the arrangement for attack on the protocol BB84"* [HR21a], which was conceived, simulated, and written by myself, supervised and edited by Prof John Rarity.

In his recent paper, Khoklhov claims to have developed a protocol that can be used to distinguish between the four basis states used in BB84 to encode information—the H-, V-, D- and A-polarised states of a single photon [Kho16b]. He claims this is through weak measurement—where weak coupling of a quantum variable to an ancilla allows data about a quantum state to be obtained without collapsing the state. However, weak measurement typically only obtains a small amount of information per measurement, so a large number of identically-prepared quantum objects are needed to obtain this fully. An alternative proposal given by Aharonov et al discusses the possibility of performing a weak measurement on a single particle [ACE14]—however, this is still subject to Busch's limit on information gained for a given disturbance [Bus09].

Khoklhov previously gave an interferometric device that he claims allows a weak measurement to tell the path a photon travelled via, without disturbing the state of that photon (see Fig. A.1a) [Kho16a]. For this, a photon of state $\alpha|H\rangle + \beta|V\rangle$ goes through a PBS, which transmits H- and reflects V-polarised components. The two components each travel down a respective arm, where they have a momentum-kick applied to them, such that their eventual arrival position on the second PBS isn't affected—here, the H-component gets a downwards kick, and the V-component an upwards kick. The two components meet at the second PBS, recombine, and then exit, but the difference in momentum allows, at a far distance from the second PBS, the respective

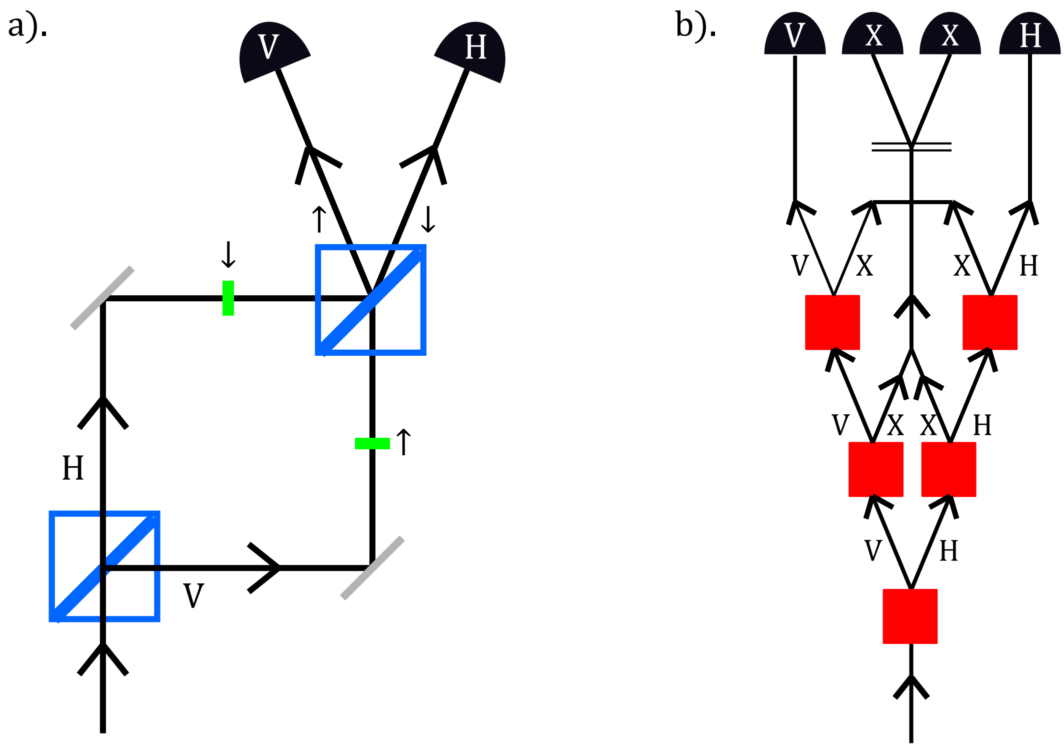

Figure A.1: Diagrams of a) the single-interferometer device, which forms the building blocks of the attack protocol, and b) the layout formed of these, where each red square is an interferometer of the sort given in a. As can be seen, once initially split by the first interferometer, the H and V components remain these separate components, and so never travel the paths marked X. Therefore, only two of the four detectors ever receive photons, making the protocol useless for determining the full initial state for a single photon.

components to be identified. Khoklhov claims, by using many photons, this allows the probability amplitudes of the two polarisation components to be determined, and so the preparation state (in truth, we only get the moduli-squared of those components, and so the classical balance of probabilities). Further, Khoklhov implies that the photon emitted into one of the two distinguishable far-field paths would still be in its original state, rather than collapsed to either H or V. This is also incorrect—as the polarisation becomes entangled with the momentum degree-of-freedom, the collapse to either upwards- or downwards-momentum causes the simultaneous collapse to either H- or V-polarisation.

In this paper, Khoklov then makes the bold claim that, using these single-interferometer units as building blocks, he can make a device which can perfectly distinguish between the four basis states used in BB84—H, V, D (or $\frac{H+V}{\sqrt{2}}$) and A (or $\frac{H-V}{\sqrt{2}}$). Given the security of BB84 rests on the quantum assumption that, even with an optimal choice of measuring basis, one cannot distinguish between all four bases perfectly (thus bounding an eavesdropper's potential knowledge), this claim threatens the security of one of the most well-known QKD protocols.

Khoklhov claims this is possible by taking the separated outputs of one of his single-interferometer devices, then putting each output through another device. Those outputs which disagree with their original polarisation-determination (due to his assumption that the photon exits in its original state) are then combined—the components which agree with their original polarisation are put through another device, and again have any further outputs which disagree combined. The combined beam is then put through a polarising beamsplitter in the $A-D$ basis, with its two outputs sent to detectors, alongside the other two outputs (that always agreed with their initially-determined polarisation). We present this in Fig. A.1b. Khoklhov claims that if the input photon's initial state is D (A) it will end up in the inner left (right) detector. This claim is false.

Let us examine the path of the input photon using standard quantum mechanics (as represented by Bra-Ket notation). We can describe the action of one of Khoklhov's single-interferometer devices (pre-detection), as given in Fig. A.1a, on a single polarisation-encoded photon qubit by

$$\alpha|H\rangle + \beta|V\rangle \to \alpha|H,\downarrow\rangle + \beta|V,\uparrow\rangle \tag{A.1}$$

This effectively entangles the polarisation and path degrees of freedom. Applying this to the larger set-up (as given in Fig. A.1b), we then see the two components (H and V) then act as separate for the remainder of the chain, obeying

$$\alpha|H\rangle + \beta|V\rangle \to \alpha|H,R,R,R\rangle + \beta|V,L,L,L\rangle \tag{A.2}$$

where L and R describe the paths on the figure after each interferometer. This means the inner paths, $|R,L\rangle$, $|L,R\rangle$, $|L,L,R\rangle$ and $|R,R,L\rangle$ are never explored, so the two inner detectors never click. Given the determination of the single photon's polarisation qubit in the attack is predicated on these two detectors being able to click (being taken to represent $\frac{H+V}{\sqrt{2}}$ and $\frac{H-V}{\sqrt{2}}$), this shows the attack does not work—as expected, standard quantum mechanics preserves the security of BB84 from this attack.

We finally give a more proper account of the result of using true weak measurements to attempt Khoklhov's scheme, using the description of weak measurement from [TC13].

To do this, in each of Khoklhov's apparatuses, we couple our photon's polarisation

$$\begin{aligned}|\psi\rangle &\in \{|H\rangle;|V\rangle;\frac{|H\rangle+|V\rangle}{\sqrt{2}};\frac{|H\rangle-|V\rangle}{\sqrt{2}}\} \\ &= \alpha|H\rangle + \beta|V\rangle\end{aligned} \tag{A.3}$$

with the pointer (the photon's momentum)

$$|\phi\rangle = |\phi_d\rangle = \int_p \phi(p)|p\rangle\, dp \tag{A.4}$$

where p is the vertical momentum of the photon. $\hat{P}_d$ is the momentum operator such that $\hat{P}_d\,|p\rangle = p\,|p\rangle$.

We assume $\phi(p)$ has a Gaussian distribution around 0 (input vertical momentum), such that

$$\phi(p) = e^{-p^2/4\sigma^2}/\sqrt{2\pi\sigma^2} \tag{A.5}$$

If we define the polarisation-distinguishing operator

$$\hat{A} = |H\rangle\langle H| - |V\rangle\langle V| \tag{A.6}$$

we can consider an interaction Hamiltonian between the two

$$\hat{H}_{int} = g(t)\hat{A} \otimes \hat{X}_d \tag{A.7}$$

where $\hat{X}_d$ is the operator conjugate to $\hat{P}_d$ such that $[\hat{P}_d, \hat{X}_d] = i\hbar$, and $g(t)$ is the coupling function such that

$$\int_0^T g(t)dt = 1 \tag{A.8}$$

for coupling time T.

This means, applying this Hamiltonian

$$e^{i\hat{H}t/\hbar}|\psi\rangle \otimes |\phi\rangle \tag{A.9}$$

we see for each of $|H\rangle \otimes |\phi(p)\rangle$, $|V\rangle \otimes |\phi(p)\rangle$, the Hamiltonian takes $\hat{P}_d$ to $\hat{P}_d + 1$, $\hat{P}_d - 1$ respectively, as

$$\hat{P}_d(T) - \hat{P}_d(0) = \int_0^T \frac{i}{\hbar}[\hat{H}, \hat{P}_d]dt \in \{+1, -1\} \tag{A.10}$$

Therefore, the corresponding transformation is

$$\begin{aligned}
&e^{i\hat{H}t/\hbar}|\psi\rangle \otimes |\phi(p)\rangle \\
&= \alpha\,|H\rangle \otimes |\phi(p-1)\rangle + \beta\,|V\rangle \otimes |\phi(p+1)\rangle \\
&= \int_p \big(\alpha\,|H\rangle \otimes \phi(p-1) + \beta\,|V\rangle \otimes \phi(p+1)\big)\,|p\rangle\, dp
\end{aligned} \tag{A.11}$$

The above wavefunctions $\phi(p-1)$ and $\phi(p+1)$ need to overlap each other for the measurement to be weak—and so need to have high variance, σ. The higher the variance, the weaker the measurement—if these Gaussian wavefunctions don't overlap, them the measurement is strong. Given σ is initially defined from the vertical momentum of the photon, this means the photon input into the system must also have high σ.

The effect of the strength of the measurement can be most readily seen when observe and collapse the pointer to a specific momentum-value, p_0, to read out our weak measurement, which gives

$$\left(e^{-\frac{(p_0-1)^2}{4\sigma^2}}\alpha|H\rangle + e^{-\frac{(p_0+1)^2}{4\sigma^2}}\beta|V\rangle\right)\otimes|p_0\rangle \tag{A.12}$$

where both the coefficients on $|H\rangle$ and $|V\rangle$ are biased slightly depending on where they are in relation to p_0. This makes sense, analogous to how measuring an eigenvalue for an observable collapses the measured state to the relevant eigenstate—the only difference here is the variance σ providing some uncertainty in that measurement.

The far-field vertical position of the photon will depend on the vertical momentum of the photon, as Khoklhov rightly says. However, the variance in this momentum (which must be large enough to allow overlap between the momenta for up and for down in order for the measurement to be weak, and the final polarisation state to not have changed too far) means that these positions must overlap heavily too. While, in the limit of many identically-prepared photons, we could obtain information about whether the polarisation-state was $|H\rangle$, $|V\rangle$, or a superposition of the two, we cannot gain this for a single run without inducing collapse. Therefore, a protocol built up of several of these devices, as Khoklhov's attack protocol is, either doesn't work due to collapse (as we show with the standard quantum-mechanical approach above), or gains effectively no information about the polarisation state of the photon, making it useless as an attack.

Appendix B

The difference between the dynamical quantum Cheshire cat and counterfactual communication

Declaration of contribution: This Appendix is adapted from the comment *The difference between the dynamical quantum Cheshire cat and counterfactual communication*, which was conceived and written by myself, supervised and edited by Prof John Rarity and Prof James Ladyman.

B.1 Introduction

In 2013 Aharonov et al first described the quantum Cheshire cat effect: where a property of a particle (for example, its polarisation) is disembodied from said particle, and travels through regions the particle cannot traverse [Aha+13]. (See Subsection 2.2.3 for a description of the quantum Cheshire Cat.) Aharonov et al further developed a protocol which they argue allows a disembodied property to be altered while separated from the particle that possesses it — a dynamical Cheshire cat [ACP21]. They say it can be detected only indirectly using weak values. They further claim that this dynamical quantum Cheshire cat effect is behind counterfactual communication [Sal+13b; Sal+22; AV19; AR20; HLR21] (and related counterfactual (quantum) information transfer/metrology [Sal16; Sal18c; Sal+20; Sal+21; HR21b]).

In this Appendix, we show how to obtain these weak values, and so to detect a dynamical Cheshire cat effect experimentally. We note however this involves adding a field orthogonal to the state claimed, travelling with the property, which can measure the evolution of the state. The measured weak value remains constant as we reduce the value of this orthogonal probe field, but its uncertainty increases and becomes infinite when the probe field is zero. Since the weak value is completely indeterminate when the probe field is zero, the effect cannot be observed experimentally unless the probe field travels with it. This is different from counterfactual communication as presented by Salih et al [Sal+13b], where the transmitted information is provably not associated with a probe field passing between communicating parties when information is sent.

B.2 Theoretically calculating the weak value of σ_x

As Aharonov et al themselves say, the dynamical quantum Cheshire cat effect cannot be observed directly—a particle in the input state they give (spin $|\uparrow_z\rangle$) returns after $2N$ cycles with its spin-state unchanged, regardless of whether the spin-dependent mirror is present or absent [ACP21, p3]. Therefore, they suggest observing it using weak values, by weakly measuring σ_x at the start of the experiment, then strongly measuring it after time $2NT$, to get the weak value of σ_x.

To do this, we first replace spin (the degree of freedom Aharonov et al use) with polarisation. We replace $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$ with $|H\rangle$ (horizontal) and $|V\rangle$ (vertical) polarisation respectively, and so $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$ with $|D\rangle$ (diagonal) and $|A\rangle$ (anti-diagonal) polarisation. Therefore, operator $\sigma_x = |D\rangle\langle D| - |A\rangle\langle A|$. Similarly, we replace the spin-dependent mirror with a polarising mirror, which transmits H-polarised and reflects V-polarised light, and replace the leaky mirror with a beamsplitter with reflection:transmission ratio $\cos(\epsilon) : \sin(\epsilon)$.

Aharonov et al use the weak value formula

$$\hat{A}^w = \frac{\langle \psi_f | \hat{A} | \psi_i \rangle}{\langle \psi_f | \psi_i \rangle} \tag{B.1}$$

where $\hat{A} = \sigma_x$, $|\psi_i\rangle = |H\rangle |L\rangle$, and

$$\begin{aligned} |\psi_f\rangle &= \hat{U}^{2N} |D\rangle |L\rangle \\ &= \frac{\cos^{2N}(\pi/2N)|H\rangle - |V\rangle}{\sqrt{2}} |L\rangle + \text{ Loss} \\ &\approx |A\rangle |L\rangle \end{aligned} \tag{B.2}$$

Note, $\hat{U}^{2N}$ is the evolution in the device, which applies a phase of -1 to the $|V\rangle$ state, but no phase to the $|H\rangle$ state, which is equivalent to flipping $|D\rangle$ to $|A\rangle$ and vice-versa; i.e.,

$$\hat{U}^{\dagger,2N} |L\rangle\langle L| \sigma_x \hat{U}^{2N} \approx |L\rangle\langle L| (-\sigma_x) \tag{B.3}$$

Therefore,

$$\begin{aligned} \sigma_x^w &= \frac{\langle L| \langle A| \Big(|D\rangle\langle D| - |A\rangle\langle A| \Big) |H\rangle |L\rangle}{\langle L| \langle A|H\rangle |L\rangle} \\ &= \frac{-1/\sqrt{2}}{1/\sqrt{2}} = -1 \end{aligned} \tag{B.4}$$

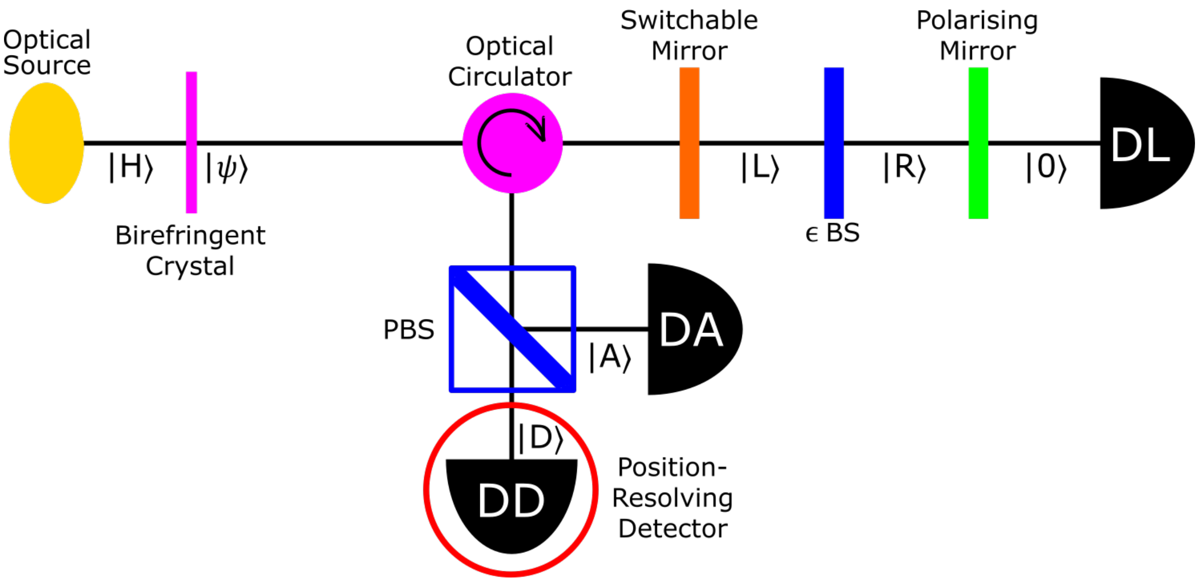

Figure B.1: Adapted form of Aharonov et al's dynamical quantum Cheshire cat apparatus, allowing us to obtain the weak value of σ_x. We postselect on the wavepacket ending up D-polarised, and going to detector DD.

More generally, for postselected final state

(B.5)
$$\left|\psi_f^\beta\right\rangle = \left(\cos(\beta)\cos^{2N}(\pi/2N)\left|H\right\rangle - \sin(\beta)\left|V\right\rangle\right)\left|L\right\rangle,$$

(B.6)
$$\begin{aligned}\sigma_x^w(\beta) &= \frac{\left\langle L\right|\left(\cos(\beta)\cos^{2N}(\pi/2N)\left\langle H\right| - \sin(\beta)\left\langle V\right|\right)\left(\left|D\right\rangle\!\left\langle D\right| - \left|A\right\rangle\!\left\langle A\right|\right)\left|H\right\rangle\left|L\right\rangle}{\left\langle L\right|\left(\cos(\beta)\cos^{2N}(\pi/2N)\left\langle H\right| - \sin(\beta)\left\langle V\right|\right)\left|H\right\rangle\left|L\right\rangle}\\ &= \frac{-\sin(\beta)}{\cos(\beta)\cos^{2N}(\pi/2N)} \approx -\tan(\beta)\end{aligned}$$

which goes to 0 as $\beta \to 0$ (i.e., we postselect on $|H\rangle$), and becomes infinite as $\beta \to \pi/2$ (i.e., we postselect on $|V\rangle$).

Compare this to the weak value of σ_x when the polarising mirror is removed: $\left|\psi_f\right\rangle$ remains $|D\rangle|L\rangle$, so both the numerator and the denominator of σ_x are $1/\sqrt{2}$, and $\sigma_x^w = 1$. Similarly, $\sigma_x^w(\beta) = \tan(\beta)$—therefore, the two cases become indistinguishable as $\beta \to 0$ (i.e., we postselect on $|H\rangle$), and become infinitely distinguishable as $\beta \to \pi/2$ (i.e., we postselect on $|V\rangle$).

B.3 Experimentally measuring the weak value of σ_x

We now show, unlike in the theoretical treatment above, actually obtaining a weak value experimentally causes a coupling disturbance, due to the initial weak measurement required to entangle the system with the pointer [Kas17; Ips22]. We see this by showing how to experimentally obtain the weak value Aharonov et al give for the σ_x operator, using a variation of

Ritchie et al's protocol for obtaining weak values for polarisation by weakly coupling to the beam centroid position as a pointer state (see Fig. B.1) [RSH91]. Putting an H-polariser between the weak measurement and Aharonov et al's protocol shows the polarisation-position entanglement caused by the weak measurement has the same effect as a perturbation of the polarisation — so we cannot see the dynamical quantum Cheshire cat experimentally if we allow only H-polarised light to enter the protocol.

While we interpret the Gaussian pointer below as the position of the wavepacket centroid, these calculations, and so these conclusions, apply to any case where we use weak measurement to entangle a qubit with a Gaussian continuous-variable pointer. While other protocols exist for obtaining weak values experimentally (see [CP18; Wag+23]), these make use of strong measurements, so also do not necessarily give us information on/require assumptions to allow us to infer about about the unperturbed system.

To obtain a weak value for σ_x, we describe our wavepacket's position as a Gaussian distribution:

$$|E_i\rangle = \sqrt{\frac{2\sqrt{2}}{w_0\sqrt{\pi}}} e^{-\frac{x^2+y^2}{w_0^2}} |H\rangle = \sqrt{\frac{\sqrt{2}}{w_0\sqrt{\pi}}} e^{-\frac{x^2+y^2}{w_0^2}} \Big(|D\rangle + |A\rangle\Big) \tag{B.7}$$

We then weakly measure polarisation, using a birefringent-crystalline quartz plate to separate the polarisation-components by a distance a much smaller than the beam width w_0:

$$|E_w\rangle = \sqrt{\frac{\sqrt{2}}{w_0\sqrt{\pi}}} e^{-\frac{x^2}{w_0^2}} \Big(e^{-\frac{(y-a/2)^2}{w_0^2}} |D\rangle + e^{-\frac{(y+a/2)^2}{w_0^2}} |A\rangle\Big) \tag{B.8}$$

If we now did not apply Aharonov et al's protocol, and instead send the beam straight to the detector DA, this would give intensity

$$I_w(y) = |\langle D|E_w\rangle|^2 = \frac{\sqrt{2}}{w_0\sqrt{\pi}} e^{-\frac{2x^2}{w_0^2}} \Big(e^{-\frac{2(y-a/2)^2}{w_0^2}}\Big) \tag{B.9}$$

and so average beam position at $x = 0$

$$\langle y_w\rangle = \int_{-\infty}^{\infty} y I_w(y) dy = a/2 \tag{B.10}$$

Instead applying system evolution $\hat{U}^{2N}$ to $|E_w\rangle$, we get state

(B.11)

$$\begin{aligned}\big|E_f\big\rangle = \hat{U}^{2N}|E_w\rangle =& \sqrt{\frac{\sqrt{2}}{w_0\sqrt{\pi}}} e^{-\frac{x^2}{w_0^2}} \Big(e^{-\frac{(y-a/2)^2}{w_0^2}} \frac{\cos^{2N}(\pi/2N)|H\rangle - |V\rangle}{\sqrt{2}} + e^{-\frac{(y+a/2)^2}{w_0^2}} \frac{\cos^{2N}(\pi/2N)|H\rangle + |V\rangle}{\sqrt{2}}\Big) \\ \approx& \sqrt{\frac{\sqrt{2}}{w_0\sqrt{\pi}}} e^{-\frac{x^2}{w_0^2}} \Big(e^{-\frac{(y-a/2)^2}{w_0^2}} |A\rangle + e^{-\frac{(y+a/2)^2}{w_0^2}} |D\rangle\Big)\end{aligned}$$

Postselecting on polarisation state

$$\big\langle\beta\big| = \cos(\beta)\langle H| + \sin(\beta)\langle V| = \frac{\cos(\beta)+\sin(\beta)}{\sqrt{2}}|D\rangle + \frac{\cos(\beta)-\sin(\beta)}{\sqrt{2}}|A\rangle \tag{B.12}$$

gives us amplitude

$$\langle\beta|E_f\rangle = \sqrt{\frac{\sqrt{2}}{w_0\sqrt{\pi}}}e^{-\frac{x^2}{w_0^2}}\Big(\frac{\cos(\beta)-\sin(\beta)}{\sqrt{2}}e^{-\frac{(y-a/2)^2}{w_0^2}}+\frac{\cos(\beta)+\sin(\beta)}{\sqrt{2}}e^{-\frac{(y+a/2)^2}{w_0^2}}\Big) \tag{B.13}$$

and so normalised intensity (at $x=0$)

$$I_f(y) = |\langle\beta|E_f\rangle|^2_{x=0} = \frac{\sqrt{2}}{w_0\sqrt{\pi}}\Big(\big(\frac{1-\sin(2\beta)}{2}\big)e^{-\frac{2(y-a/2)^2}{w_0^2}}+\frac{1+\sin(2\beta)}{2}e^{-\frac{2(y+a/2)^2}{w_0^2}}+\cos(2\beta)e^{-\frac{2y^2+a^2/2}{w_0^2}}\Big) \tag{B.14}$$

Therefore, the average position $\langle y_f\rangle$ of the centre of the beam on detector DA after the protocol is

$$\langle y_f\rangle = \int_{-\infty}^{\infty} yI_f(y)dy = \frac{a}{2}\big(\big(\frac{\cos(\beta)-\sin(\beta)}{\sqrt{2}}\big)^{-}\big(\frac{\cos(\beta)+\sin(\beta)}{\sqrt{2}}\big)^2\big) \tag{B.15}$$

Aharonov et al postselect on state $|D\rangle$ ($\beta=\pi/4$), so the average position is $\langle y_f\rangle = -a/2$, and the weak value $\sigma_x^w = \langle y_f\rangle/\langle y_w\rangle = -\frac{a}{2}/\frac{a}{2} = -1$. This agrees with Aharonov et al's theoretical weak value for σ_x. However, by changing postselection, we can change the weak value—as $\beta \to 0, \pi/2$ (i.e., we postselect on either $|H\rangle$ or $|V\rangle$), this value of σ_x goes to 0.

Were there no polarising mirror in the apparatus, any light travelling to the right-hand cavity of the apparatus would be lost, regardless of polarisation. Therefore, both H and V components would be attenuated by $\cos^{2N}(\pi/2N)$, rather than the V-polarised component's sign flipping. The D-polarised component remain D-polarised rather than flipping to A, and the A-polarised component would remain A-polarised rather than flipping to D, so the beam centroid would stay at $\langle y\rangle = a/2$, and the weak value would be 1, not -1. This shows a clear experimentally-detectable difference between when there is and when there isn't a polarising mirror in the right-hand cavity, when we perturb the system to detect this.

We now show entangling the wavepacket's polarisation with a pointer variable is equivalent to perturbing the polarisation from state $|H\rangle$ by small perturbation, as, when we put a H-polariser between the entangling birefringent crystal and the protocol, we cannot observe the weak value experimentally.

We first take the state in Eq. B.8, and apply a $|H\rangle\langle H|$ projector

$$|H\rangle\langle H|\,|E_w\rangle = \sqrt{\frac{\sqrt{2}}{2w_0\sqrt{\pi}}}e^{-\frac{x^2}{w_0^2}}\Big(e^{-\frac{(y-a/2)^2}{w_0^2}}+e^{-\frac{(y+a/2)^2}{w_0^2}}\Big)|H\rangle \tag{B.16}$$

We then apply the evolution $\hat{U}^{2N}$

$$\begin{aligned}\hat{U}^{2N}|H\rangle\langle H|\,|E_w\rangle &= \sqrt{\frac{\sqrt{2}}{2w_0\sqrt{\pi}}}e^{-\frac{x^2}{w_0^2}}\cos^{2N}(\pi/2N)\Big(e^{-\frac{(y-a/2)^2}{w_0^2}}+e^{-\frac{(y+a)^2}{w_0^2}}\Big)|H\rangle + \text{Loss}\\ &\approx \sqrt{\frac{\sqrt{2}}{2w_0\sqrt{\pi}}}e^{-\frac{x^2}{w_0^2}}\Big(e^{-\frac{(y-a/2)^2}{w_0^2}}+e^{-\frac{(y+a)^2}{w_0^2}}\Big)|H\rangle\end{aligned} \tag{B.17}$$

Finally, we postselect on the wavepacket being D-polarised (arriving at detector DD), to get intensity

$$I(y) = \frac{\sqrt{2}}{4w_0\sqrt{\pi}} e^{-\frac{2x^2}{w_0^2}} \left(e^{-\frac{2(y-a/2)^2}{w_0^2}} + e^{-\frac{(y+a/2)^2}{w_0^2}} + e^{-\frac{2y^2+a^2/2}{w_0^2}} \right) \tag{B.18}$$

and so, at $x=0$, y-expectation value

$$\langle y \rangle = \int_{-\infty}^{\infty} y I(y) dy = \frac{a}{8} - \frac{a}{8} = 0 \tag{B.19}$$

Now, when there is no polarising mirror in the apparatus, the protocol just attenuates the H-polarised state by $\cos^{2N}(\pi/2N)$, so the state after the protocol would be identical to Eq. B.17, and the y-expectation value would also be 0. There is now no experimentally-detectable difference between the protocol having or missing the polarising mirror.

This shows entangling some pointer variable with the polarisation when obtaining the weak value is equivalent to putting some indeterminate perturbation on the state, taking it from $|H\rangle$ to $\cos\delta|H\rangle + \sin\delta|V\rangle$ for some indeterminate tiny perturbation δ. By postselecting on some non-zero β, we force this δ to be non-zero too, and so observe the effect of Aharonov et al's protocol on some non-zero $|V\rangle$ component. This is different from the $|H\rangle$ input state Aharonov et al claim to evaluate—so we cannot experimentally detect the dynamical quantum Cheshire cat with an unperturbed state.

B.4 What does this have to do with counterfactual communication?

Aharonov et al claim this dynamical quantum Cheshire cat effect underlies counterfactual communication (an interesting development of exchange-free measurement [Sal+13b]). Given recent interest in counterfactual communication, both practically for quantum computation [Sal16; Sal18c; Sal+20; Sal+21] and metrology [HR21b], and foundationally [Sal+22; AV19; AR20; HLR21], it is important to note the implications of our result for this claim.

The term 'counterfactual' in the name 'counterfactual communication' comes from Alice (the receiver)'s photon never having travelled to/via Bob (the sender), when she receives a bit of information from Bob in the protocol (i.e. by her postselection on photon either going to her detector $D0$ for a 0-bit, or her detector $D1$ for a 1-bit). Rather than factually interacting with anything on Bob's side, the photon 'counterfactually' tells Alice whether Bob is blocking his channel or not (and so whether his bit-value is 1 or 0 respectively), without ever having been to his channel. Therefore, counterfactual communication seems superficially similar to the dynamical quantum Cheshire cat, where there is a difference in the (theoretically-calculated) weak value depending on whether or not a polarisation-dependent mirror is placed at the end of the right-hand cavity, despite the H-polarised photon not being able to travel to-and-back-from this mirror.

However, in the counterfactual communication protocol of Salih et al (see [Sal+13b; Sal+22], we do not require any small amount of field to travel to-and-back-from Bob's channel to allow Alice to receive information from Bob—anything which goes to Bob is lost (like the H-polarised light which travels to the right-hand cavity and is then lost in the unperturbed dynamical quantum Cheshire cat setup). On the other hand, as shown above, experimentally determining whether or not the spin-dependent mirror is present requires some small amount of V-polarised light to travel to, and back from, the mirror. This V-polarised light is obviously not counterfactual. Therefore, the dynamical quantum Cheshire cat effect is not how Alice is able to detect Bob's blocker setting (and so bit-value) in Salih et al's counterfactual communication.

B.5 Weak value of $|R\rangle\langle R|$

Looking at the perturbed state we give above, when we want to find the weak value of the projector $\hat{P}_R = |R\rangle\langle R|$ at some time during the protocol (again, with polarising mirror), we follow Aharonov et al by saying

$$P_R^w = \frac{\langle L|\langle D|\hat{U}^{2N-n-1}\hat{U}(T-\tau)\hat{P}_R\hat{U}(\tau)\hat{U}^n|H\rangle|L\rangle}{\langle L|\langle D|\hat{U}^{2N}|H\rangle|L\rangle} \tag{B.20}$$

but inserting our weakly perturbed state $|E_w(\alpha=\pi/4)\rangle$ instead of the initial $|H\rangle$. For ease, we can say $\hat{U}(\tau) = \hat{\mathbb{1}}$, and so $\hat{U}(T-\tau) = \hat{U}$.

To see the result we get from this, we must first evaluate the effect of both the $|D\rangle$ and $|A\rangle$ components of this, getting

$$\begin{aligned}\langle L|\langle D|\hat{U}^{2N-n}\hat{P}_R\hat{U}^n|D\rangle|L\rangle &= -\sin(\frac{(2N-n)\pi}{2N})\sin(\frac{n\pi}{2N})/2\\ \langle L|\langle D|\hat{U}^{2N-n}\hat{P}_R\hat{U}^n|A\rangle|L\rangle &= \sin(\frac{(2N-n)\pi}{2N})\sin(\frac{n\pi}{2N})/2\end{aligned} \tag{B.21}$$

and so

$$\langle L|\langle D|\hat{U}^{2N-n}\hat{P}_R\hat{U}^n|E_w\rangle|L\rangle = \sqrt{\frac{\sqrt{2}}{w_0\sqrt{\pi}}}e^{-\frac{x^2}{w_0^2}}\sin(\frac{(2N-n)\pi}{2N})\sin(\frac{n\pi}{2N})\Big(e^{-\frac{(y+a/2)^2}{w_0^2}} - e^{-\frac{(y-a/2)^2}{w_0^2}}\Big)/2 \tag{B.22}$$

From above, we get

$$\langle L|\langle D|\hat{U}^{2N}|E_w\rangle|L\rangle = \sqrt{\frac{\sqrt{2}}{w_0\sqrt{\pi}}}e^{-\frac{x^2}{w_0^2}}e^{-\frac{(y+a/2)^2}{w_0^2}} \tag{B.23}$$

Therefore

$$P_R^w = \frac{\langle L|\langle D|\hat{U}^{2N-n}\hat{P}_R\hat{U}^n|E_w\rangle|L\rangle}{\langle L|\langle D|\hat{U}^{2N}|E_w\rangle|L\rangle} = \sin(\frac{(2N-n)\pi}{2N})\sin(\frac{n\pi}{2N})\Big(1-e^{\frac{2ay}{w_0^2}}\Big)/2 \tag{B.24}$$

While this is 0 at $y=0$, away from this point, it has non-zero weak values. For instance, if we set $y=-a/2$, it reaches its greatest magnitude at $n=N$ of

$$P_R^{w,\max} = \frac{1-e^{-\frac{a^2}{w_0^2}}}{2} \tag{B.25}$$

If we interpret this non-zero weak value of the projection operator $\hat{P}_R$ as showing the presence of some field at a given y, this shows the measurable perturbation of the beam centroid by a gives us a non-zero weak value for both polarisation change in the protocol and presence in the right-hand half of the apparatus. This makes sense, given the polarisation-dependent perturbation caused by the weak value effectively creates positive and negative V-polarised components, each in a Gaussian distribution, one shifted by $a/2$, the other by $-a/2$. Therefore, while these cancel at $y = 0$, they do not away from here, and where they do not cancel, some of the V-polarised wavepacket travels into the protocol. This is further proof that the dynamical quantum Cheshire cat involves elements of the wavepacket "factually" travelling with the disembodied property, and so further illustrates the difference between this and counterfactual communication.

B.6 Conclusion

We have shown Aharonov et al's dynamical quantum Cheshire cat effect can be detected by obtaining weak values through weak coupling and postselection. However, this necessarily requires some perturbation away from the initial state, and so requires something to travel through the second cavity, interact with the probed barrier, and return to the first cavity. This is different from Salih et al's counterfactual communication, which does not require any field to travel between communicants for information to be transferred.

Bibliography

[AAV88] Yakir Aharonov, David Z. Albert, and Lev Vaidman. "How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100". In: *Phys. Rev. Lett.* 60 (14 Apr. 1988), pp. 1351–1354. DOI: 10.1103/PhysRevLett.60.1351.

[AB16] D. R. M. Arvidsson-Shukur and C. H. W. Barnes. "Quantum counterfactual communication without a weak trace". In: *Phys. Rev. A* 94 (6 Dec. 2016), p. 062303. DOI: 10.1103/PhysRevA.94.062303.

[AB19] D. R. M. Arvidsson-Shukur and C. H. W. Barnes. "Postselection and counterfactual communication". In: *Phys. Rev. A* 99 (6 June 2019), 060102(R). DOI: 10.1103/PhysRevA.99.060102.

[ABL64] Yakir Aharonov, Peter G Bergmann, and Joel L Lebowitz. "Time symmetry in the quantum process of measurement". In: *Physical Review* 134.6B (1964), B1410. DOI: 10.1103/PhysRev.134.B1410.

[Abo+04] Ayman F. Abouraddy et al. "Entangled-Photon Imaging of a Pure Phase Object". In: *Phys. Rev. Lett.* 93 (21 Nov. 2004), p. 213903. DOI: 10.1103/PhysRevLett.93.213903.

[ACE14] Yakir Aharonov, Eliahu Cohen, and Avshalom C Elitzur. "Foundations and applications of weak quantum measurements". In: *Physical Review A* 89.5 (2014), p. 052105. DOI: 10.1103/PhysRevA.89.052105.

[ACP21] Yakir Aharonov, Eliahu Cohen, and Sandu Popescu. "A dynamical quantum Cheshire Cat effect and implications for counterfactual communication". In: *Nature Communications* 12.1 (2021), pp. 1–8. DOI: 10.1038/s41467-021-24933-9.

[ACT18] Yakir Aharonov, Eliahu Cohen, and Jeff Tollaksen. "Completely top–down hierarchical structure in quantum mechanics". In: *Proceedings of the National Academy of Sciences* 115.46 (2018), pp. 11730–11735. DOI: 10.1073/pnas.1807554115.

[Adl21a] Emily Adlam. "Contextuality, Fine-Tuning and Teleological Explanation". In: *Foundations of Physics* 51.6 (Nov. 2021), p. 106. ISSN: 1572-9516. DOI: 10.1007/s10701-021-00516-y.

[Adl21b] Emily Adlam. *Foundations of Quantum Mechanics*. Elements in the Philosophy of Physics. Cambridge University Press, 2021. DOI: 10.1017/9781108885515.

[Adl22] Emily Adlam. “Two Roads to Retrocausality”. In: *arXiv preprint arXiv:2201.12934* (2022). DOI: 10.48550/arXiv.2201.12934.

[AGB17] D. R. M. Arvidsson-Shukur, A. N. O. Gottfries, and C. H. W. Barnes. “Evaluation of counterfactuality in counterfactual communication protocols”. In: *Phys. Rev. A* 96 (6 2017), p. 062316. DOI: 10.1103/PhysRevA.96.062316.

[AGR81] Alain Aspect, Philippe Grangier, and Gérard Roger. “Experimental Tests of Realistic Local Theories via Bell’s Theorem”. In: *Phys. Rev. Lett.* 47 (7 Aug. 1981), pp. 460–463. DOI: 10.1103/PhysRevLett.47.460.

[AGR82] Alain Aspect, Philippe Grangier, and Gérard Roger. “Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell’s Inequalities”. In: *Phys. Rev. Lett.* 49 (2 July 1982), pp. 91–94. DOI: 10.1103/PhysRevLett.49.91.

[Aha+13] Yakir Aharonov et al. “Quantum cheshire cats”. In: *New Journal of Physics* 15.11 (2013), p. 113015. DOI: 10.1088/1367-2630/15/11/113015.

[Aha+17] Yakir Aharonov et al. “The case of the disappearing (and re-appearing) particle”. In: *Scientific reports* 7.1 (2017), p. 531. DOI: 10.1038/s41598-017-00274-w.

[Aha+18] Yakir Aharonov et al. “The weak reality that makes quantum phenomena more natural: novel insights and experiments”. In: *Entropy* 20.11 (2018), p. 854. DOI: 10.3390/e20110854.

[AJ15] Miguel A Alonso and Andrew N Jordan. “Can a Dove prism change the past of a single photon?” In: *Quantum Studies: Mathematics and Foundations* 2.3 (2015), pp. 255–261. DOI: 10.1007/s40509-015-0044-8.

[ALZ16] M Al Amri, Zheng-Hong Li, and Muhammad Suhail Zubairy. “Quantum teleportation without classical channel”. In: *Second International Seminar on Photonics, Optics, and Its Applications (ISPhOA 2016)*. Vol. 10150. International Society for Optics and Photonics. 2016, 101501N. DOI: 10.1117/12.2267245.

[Ank+11] Rachel Ankeny et al. “Introduction: philosophy of science in practice”. In: *European Journal for Philosophy of Science* 1.3 (Sept. 2011), p. 303. ISSN: 1879-4920. DOI: 10.1007/s13194-011-0036-4.

[AP04] S. E. Ahnert and M. C. Payne. “Weak measurement of the arrival times of single photons and pairs of entangled photons”. In: *Phys. Rev. A* 69 (4 Apr. 2004), p. 042103. DOI: 10.1103/PhysRevA.69.042103.

[AR05] Yakir Aharonov and Daniel Rohrlich. *Quantum paradoxes: quantum theory for the perplexed*. John Wiley & Sons, 2005. ISBN: 978-3-527-40391-2. DOI: 10.1002/9783527619115.

[AR20] Yakir Aharonov and Daniel Rohrlich. "What Is Nonlocal in Counterfactual Quantum Communication?" In: *Phys. Rev. Lett.* 125 (26 Dec. 2020), p. 260401. DOI: 10.1103/PhysRevLett.125.260401.

[AR22] Emily Adlam and Carlo Rovelli. "Information is physical: Cross-perspective links in relational quantum mechanics". In: *arXiv preprint arXiv:2203.13342* (2022). URL: https://arxiv.org/abs/2203.13342.

[Asp+13] Reuben S Aspden et al. "EPR-based ghost imaging using a single-photon-sensitive camera". In: *New J. Phys.* 15.7 (2013), p. 073032. DOI: 10.1088/1367-2630/15/7/073032.

[Asp+15] Reuben S. Aspden et al. "Photon-sparse microscopy: visible light imaging using infrared illumination". In: *Optica* 2.12 (2015), pp. 1049–1052. DOI: 10.1364/OPTICA.2.001049.

[Asp76] Alain Aspect. "Proposed experiment to test the nonseparability of quantum mechanics". In: *Phys. Rev. D* 14 (8 Oct. 1976), pp. 1944–1951. DOI: 10.1103/PhysRevD.14.1944.

[AV19] Yakir Aharonov and Lev Vaidman. "Modification of counterfactual communication protocols that eliminates weak particle traces". In: *Phys. Rev. A* 99 (1 Jan. 2019), 010103(R). DOI: 10.1103/PhysRevA.99.010103.

[Bal70] Leslie E Ballentine. "The statistical interpretation of quantum mechanics". In: *Reviews of Modern Physics* 42.4 (1970), p. 358. DOI: 10.1103/RevModPhys.42.358.

[Bar+14] Jonathan Barrett et al. "No ψ-epistemic model can fully explain the indistinguishability of quantum states". In: *Physical review letters* 112.25 (2014), p. 250403. DOI: 10.1103/PhysRevLett.112.250403.

[Bar+95] Adriano Barenco et al. "Elementary gates for quantum computation". In: *Physical review A* 52.5 (1995), p. 3457. DOI: 10.1103/PhysRevA.52.3457.

[Bas+13] Angelo Bassi et al. "Models of wave-function collapse, underlying theories, and experimental tests". In: *Reviews of Modern Physics* 85.2 (2013), p. 471. DOI: 10.1103/RevModPhys.85.471.

[BB84] Charles H Bennett and Gilles Brassard. "Quantum cryptography: Public key distribution and coin tossing". In: *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*. 1984, pp. 175–179.

[BBB02] Ryan S. Bennink, Sean J. Bentley, and Robert W. Boyd. ""Two-Photon" Coincidence Imaging with a Classical Source". In: *Phys. Rev. Lett.* 89 (11 Aug. 2002), p. 113601. DOI: 10.1103/PhysRevLett.89.113601.

[Bel04a] John S Bell. “La nouvelle cuisine”. In: *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*. Ed. by Alain Aspect. 2nd ed. Cambridge University Press, 2004, pp. 232–248. DOI: 10.1017/CBO9780511815676.026.

[Bel04b] John S Bell. *Speakable and unspeakable in quantum mechanics: Collected papers on quantum philosophy*. 2nd ed. Cambridge University Press, 2004. ISBN: 9780511815676. DOI: 10.1017/CBO9780511815676.

[Bel64] John S Bell. “On the Einstein Podolsky Rosen paradox”. In: *Physics Physique Fizika* 1.3 (1964), p. 195. DOI: 10.1103/PhysicsPhysiqueFizika.1.195.

[Bel66] John S Bell. “On the Problem of Hidden Variables in Quantum Mechanics”. In: *Rev. Mod. Phys.* 38 (3 July 1966), pp. 447–452. DOI: 10.1103/RevModPhys.38.447.

[Bel71] John S Bell. *Introduction to the hidden-variable question*. Tech. rep. CM-P00058691, 1971.

[Bel81] John S Bell. “Quantum mechanics for cosmologists”. In: *Quantum Gravity*. Ed. by Christopher Isham, Roger Penrose, and Dennis Sciama. Vol. 2. Oxford University Press, 1981, pp. 611–637.

[Ben+17] Alon Ben-Israel et al. “An improved experiment to determine the ‘past of a particle’in the nested Mach-Zehnder interferometer”. In: *Chinese Physics Letters* 34.2 (2017), p. 020301. DOI: 10.1088/0256-307X/34/2/020301.

[Ben+93] Charles H Bennett et al. “Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels”. In: *Physical review letters* 70.13 (1993), p. 1895. DOI: 10.1103/PhysRevLett.70.1895.

[Ben17] Yemima Ben-Menahem. “The PBR theorem: Whose side is it on?” In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 57 (2017), pp. 80–88. ISSN: 1355-2198. DOI: 10.1016/j.shpsb.2016.11.004.

[Ber01] Michael V Berry. “Chaos and the semiclassical limit of quantum mechanics (is the moon there when somebody looks?)” In: *Quantum Mechanics: Scientific perspectives on divine action* 41 (2001), p. 56.

[BG03] Angelo Bassi and GianCarlo Ghirardi. “Dynamical reduction models”. In: *Physics Reports* 379.5-6 (2003), pp. 257–426. DOI: 10.1016/S0370-1573(03)00103-0.

[BGP99] Dario Bambusi, Sandro Graffi, and Thierry Paul. “Long time semiclassical approximation of quantum flows: a proof of the Ehrenfest time”. In: *Asymptotic analysis* 21.2 (1999), pp. 149–160. DOI: 10.48550/arXiv.math-ph/9805018.

[BL13] Marcel Boumans and Sabina Leonelli. "Introduction: On the Philosophy of Science in Practice". In: *Journal for General Philosophy of Science* 44.2 (Dec. 2013), pp. 259–261. ISSN: 1572-8587. DOI: 10.1007/s10838-013-9232-6.

[Blu+98] Lenore Blum et al. *Complexity and real computation*. Springer Science & Business Media, 1998. ISBN: 0387982817.

[Boe+82] W Boehme et al. "Spectrophotometric and polarimetric detectors in liquid chromatography for the determination of enantiomer ratios in complex mixtures". In: *Analytical Chemistry* 54.4 (1982), pp. 709–711.

[Boh28] N. Bohr. "The Quantum Postulate and the Recent Development of Atomic Theory". In: *Nature* 121.3050 (Apr. 1928), pp. 580–590. ISSN: 1476-4687. DOI: 10.1038/121580a0.

[Boh51] David Bohm. *Quantum Theory*. Dover Books on Physics Series. Dover Publications, 1951. ISBN: 9780486659695. URL: https://books.google.co.uk/books?id=9DWim3RhymsC.

[Boh52a] David Bohm. "A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I". In: *Phys. Rev.* 85 (2 Jan. 1952), pp. 166–179. DOI: 10.1103/PhysRev.85.166.

[Boh52b] David Bohm. "A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. II". In: *Phys. Rev.* 85 (2 Jan. 1952), pp. 180–193. DOI: 10.1103/PhysRev.85.180.

[Bos+17] Sougato Bose et al. "Spin Entanglement Witness for Quantum Gravity". In: *Phys. Rev. Lett.* 119 (24 Dec. 2017), p. 240401. DOI: 10.1103/PhysRevLett.119.240401.

[Bra88] Carl H. Brans. "Bell's theorem does not eliminate fully causal hidden variables". In: *International Journal of Theoretical Physics* 27.2 (Feb. 1988), pp. 219–226. ISSN: 1572-9575. DOI: 10.1007/BF00670750.

[Bri+11] G. Brida et al. "Systematic analysis of signal-to-noise ratio in bipartite ghost imaging with classical and quantum light". In: *Phys. Rev. A* 83 (6 June 2011), p. 063807. DOI: 10.1103/PhysRevA.83.063807.

[Bri+12] Giorgio Brida et al. "Experimental realization of counterfactual quantum cryptography". In: *Laser Physics Letters* 9.3 (2012), p. 247. DOI: 10.1002/lapl.201110120.

[Bru+14] Nicolas Brunner et al. "Bell nonlocality". In: *Reviews of Modern Physics* 86.2 (2014), p. 419. DOI: 10.1103/RevModPhys.86.419.

[Bus09] Paul Busch. ""No Information Without Disturbance": Quantum Limitations of Measurement". In: *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle*. Springer, 2009, pp. 229–256.

[BV05] Samuel L Braunstein and Peter Van Loock. “Quantum information with continuous variables”. In: *Reviews of modern physics* 77.2 (2005), p. 513. DOI: 10.1103/RevModPhys.77.513.

[BV17] Alon Ben-Israel and Lev Vaidman. “Comment on “Non-representative Quantum Mechanical Weak Values””. In: *Foundations of Physics* 47.4 (2017), pp. 467–470. DOI: 10.1007/s10701-017-0071-x.

[BZ78] Gennady P Berman and George M Zaslavsky. “Condition of stochasticity in quantum nonlinear systems”. In: *Physica A: Statistical Mechanics and its Applications* 91.3-4 (1978), pp. 450–460. DOI: 10.1016/0378-4371(78)90190-5.

[Cal+19] I Alonso Calafell et al. “Trace-free counterfactual communication with a nanophotonic processor”. In: *npj Quantum Information* 5.1 (2019), p. 61. DOI: 10.1038/s41534-019-0179-2.

[Cao+14] Yuan Cao et al. “Direct counterfactual communication with single photons”. In: *CLEO: QELS_Fundamental Science*. Optical Society of America. 2014, FM4A–6.

[Cao+17] Yuan Cao et al. “Direct counterfactual communication via quantum Zeno effect”. In: *Proceedings of the National Academy of Sciences* (2017). ISSN: 0027-8424. DOI: 10.1073/pnas.1614560114.

[CD15] Piermarco Cannarsa and Teresa D’Aprile. *Introduction to Measure Theory and Functional Analysis*. Springer International Publishing, 2015. DOI: 10.1007/978-3-319-17019-0.

[Che+15a] Yuanyuan Chen et al. “Counterfactual quantum cryptography network with untrusted relay”. In: *International Journal of Modern Physics B* 29.20 (2015), p. 1550134. DOI: 10.1142/S0217979215501349.

[Che+15b] Yuanyuan Chen et al. “Tripartite counterfactual entanglement distribution”. In: *Optics express* 23.16 (2015), pp. 21193–21203. DOI: 10.1364/OE.23.021193.

[Che+16] Yuanyuan Chen et al. “Counterfactual entanglement distribution using quantum dot spins”. In: *J. Opt. Soc. Am. B* 33.4 (Apr. 2016), pp. 663–669. DOI: 10.1364/JOSAB.33.000663.

[Che09] Chi-Ho Cheng. “Thermodynamics of the system of distinguishable particles”. In: *Entropy* 11.3 (2009), pp. 326–333. DOI: 10.3390/e11030326.

[Che21] Eddy Keming Chen. “Bell’s Theorem, Quantum Probabilities, and Superdeterminism”. In: *The Routledge Companion to Philosophy of Physics*. Routledge, 2021, pp. 184–199. ISBN: 9781138653078.

[Cir80] Boris S Cirel’son. “Quantum generalizations of Bell’s inequality”. In: *Letters in Mathematical Physics* 4.2 (1980), pp. 93–100. DOI: 10.1007/BF00417500.

[CKW00] Valerie Coffman, Joydip Kundu, and William K Wootters. “Distributed entanglement”. In: *Physical Review A* 61.5 (2000), p. 052306. DOI: 10.1103/PhysRevA.61.052306.

[Cla+69] John F Clauser et al. “Proposed experiment to test local hidden-variable theories”. In: *Physical review letters* 23.15 (1969), p. 880. DOI: 10.1103/PhysRevLett.23.880.

[COS20] Gerardo Sanjuán Ciepielewski, Elias Okon, and Daniel Sudarsky. “On superdeterministic rejections of settings independence”. In: *British Journal for the Philosophy of Science* (2020). DOI: 10.1086/714819.

[CP18] Eliahu Cohen and Eli Pollak. “Determination of weak values of quantum operators using only strong measurements”. In: *Phys. Rev. A* 98 (4 Oct. 2018), p. 042112. DOI: 10.1103/PhysRevA.98.042112.

[CR12] Roger Colbeck and Renato Renner. “Is a System’s Wave Function in One-to-One Correspondence with Its Elements of Reality?” In: *Phys. Rev. Lett.* 108 (15 Apr. 2012), p. 150402. DOI: 10.1103/PhysRevLett.108.150402.

[CR97] Monique Combescure and Didier Robert. “Semiclassical spreading of quantum wave packets and applications near unstable fixed points of the classical flow”. In: *Asymptotic Analysis* 14.4 (1997), pp. 377–404. DOI: 10.3233/ASY-1997-14405.

[Cra86] John G. Cramer. “The transactional interpretation of quantum mechanics”. In: *Rev. Mod. Phys.* 58 (3 July 1986), pp. 647–687. DOI: 10.1103/RevModPhys.58.647.

[CTV17] Earl T Campbell, Barbara M Terhal, and Christophe Vuillot. “Roads towards fault-tolerant universal quantum computation”. In: *Nature* 549.7671 (2017), pp. 172–179. DOI: 10.1038/nature23460.

[CW12] Eric G Cavalcanti and Howard M Wiseman. “Bell nonlocality, signal locality and unpredictability (or what Bohr could have told Einstein at Solvay had he known about Bell experiments)”. In: *Foundations of Physics* 42.10 (2012), pp. 1329–1338. DOI: 10.1007/s10701-012-9669-1.

[CZ05] Yangjian Cai and Shi-Yao Zhu. “Ghost imaging with incoherent and partially coherent light radiation”. In: *Phys. Rev. E* 71 (5 May 2005), p. 056607. DOI: 10.1103/PhysRevE.71.056607.

[Dan+13] Ariel Danan et al. “Asking photons where they have been”. In: *Physical review letters* 111.24 (2013), p. 240402. DOI: 10.1103/PhysRevLett.111.240402.

[Dan+15] Ariel Danan et al. “Response: Commentary: “Asking photons where they have been” - without telling them what to say”. In: *Frontiers in Physics* 3 (2015), p. 48. ISSN: 2296-424X. DOI: 10.3389/fphy.2015.00048.

[DD19] Siddhant Das and Detlef Dürr. “Arrival time distributions of spin-1/2 particles”. In: *Scientific reports* 9.1 (2019), pp. 1–8. DOI: 10.1038/s41598-018-38261-4.

[DDX13] Deyang Duan, Shaojiang Du, and Yunjie Xia. “Multiwavelength ghost imaging”. In: *Phys. Rev. A* 88 (5 Nov. 2013), p. 053842. DOI: 10.1103/PhysRevA.88.053842.

[Del+20] Uroš Delić et al. “Cooling of a levitated nanoparticle to the motional quantum ground state”. In: *Science* 367.6480 (2020), pp. 892–895. ISSN: 0036-8075. DOI: 10.1126/science.aba3993.

[Deu85] David Deutsch. “Quantum theory, the Church–Turing principle and the universal quantum computer”. In: *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 400.1818 (1985), pp. 97–117. DOI: 10.1098/rspa.1985.0070.

[DG15] Bryce Seligman DeWitt and Neill Graham. *The many-worlds interpretation of quantum mechanics*. Vol. 61. Princeton University Press, 2015. ISBN: 140086805X.

[DH22] Sandro Donadi and Sabine Hossenfelder. “Toy model for local and deterministic wave-function collapse”. In: *Phys. Rev. A* 106 (2 Aug. 2022), p. 022212. DOI: 10.1103/PhysRevA.106.022212.

[Dic81] Robert H Dicke. “Interaction-free quantum measurements: A paradox?” In: *American Journal of Physics* 49.10 (1981), pp. 925–930. DOI: 10.1119/1.12592.

[Dix+09] P Ben Dixon et al. “Ultrasensitive beam deflection measurement via interferometric weak value amplification”. In: *Physical review letters* 102.17 (2009), p. 173601. DOI: 10.1103/PhysRevLett.102.173601.

[DM03] Jonathan P Dowling and Gerard J Milburn. “Quantum technology: the second quantum revolution”. In: *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 361.1809 (2003), pp. 1655–1674. DOI: 10.1098/rsta.2003.1227.

[DMN13] Simon J Devitt, William J Munro, and Kae Nemoto. “Quantum error correction for beginners”. In: *Reports on Progress in Physics* 76.7 (2013), p. 076001. DOI: 10.1088/0034-4885/76/7/076001.

[Doy94] Arthur Conan Doyle. *The Memoirs of Sherlock Holmes*. George Newnes, 1894. Chap. Adventure 1: “Silver Blaze.

[Dre+14] Justin Dressel et al. “Colloquium: Understanding quantum weak values: Basics and applications”. In: *Reviews of Modern Physics* 86.1 (2014), p. 307. DOI: 10.1103/RevModPhys.86.307.

[DRS22] Patrick J. Daley, Kevin J. Resch, and Robert W. Spekkens. “Experimentally adjudicating between different causal accounts of Bell-inequality violations via statistical model selection”. In: *Phys. Rev. A* 105 (4 Apr. 2022), p. 042220. DOI: 10.1103/PhysRevA.105.042220.

[Dub93] S. Dube. “Undecidable problems in fractal geometry”. In: *Complex Systems* 7 (1993), pp. 423–444.

[Due+21] Anna K Duell et al. “Determination of (R)-(+)-and (S)-(-)-Nicotine Chirality in Puff Bar E-Liquids by 1H NMR Spectroscopy, Polarimetry, and Gas Chromatography–Mass Spectrometry”. In: *Chemical Research in Toxicology* 34.7 (2021), pp. 1718–1720.

[Dür+04] Detlef Dürr et al. “Bohmian mechanics and quantum field theory”. In: *Physical Review Letters* 93.9 (2004), p. 090402. DOI: 10.1103/PhysRevLett.93.090402.

[Dür+14] Detlef Dürr et al. “Can Bohmian mechanics be made relativistic?” In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 470.2162 (2014), p. 20130699. DOI: 10.1098/rspa.2013.0699.

[DV98] Dennis Dieks and Pieter E Vermaas. *The modal interpretation of quantum mechanics*. Vol. 60. Springer Science & Business Media, 1998. ISBN: 0792352076.

[DVC00] Wolfgang Dür, Guifre Vidal, and J Ignacio Cirac. “Three qubits can be entangled in two inequivalent ways”. In: *Physical Review A* 62.6 (2000), p. 062314. DOI: 10.1103/PhysRevA.62.062314.

[Ein05] Albert Einstein. “Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt”. In: *Annalen der Physik* 322 (1905), pp. 132–148. DOI: 10.1002/andp.19053220607.

[Ein11] Albert Einstein. “Elementary considerations on the interpretation of the foundations of quantum mechanics”. In: *arXiv preprint arXiv:1107.3701* (2011).

[Ein15] Albert Einstein. “Die feldgleichungen der gravitation”. In: *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)* (1915), pp. 844–847.

[Ein53] Albert Einstein. “Elementare Überlegungen zur Interpretation der Grundlagen der Quanten-Mechanik”. In: *Scientific Papers Presented to Max Born on his retirement from the Tait Chair of Natural Philosophy in the University of Edinburgh*. London: Oliver and Boyd, 1953, pp. 33–40.

[Ein71] Albert Einstein. “Letter from Einstein to Max Born, 3 March 1947”. In: *The Born-Einstein Letters; Correspondence between Albert Einstein and Max and Hedwig Born from 1916 to 1955*. Ed. by Max Born and Irene Born. London: Macmillan, 1971, pp. 157–158.

[Eng+17] Berthold-Georg Englert et al. “Past of a quantum particle revisited”. In: *Phys. Rev. A* 96 (2 Aug. 2017), p. 022126. DOI: 10.1103/PhysRevA.96.022126.

[Eng96] Berthold-Georg Englert. “Fringe visibility and which-way information: An inequality”. In: *Physical review letters* 77.11 (1996), p. 2154. DOI: 10.1103/PhysRevLett.77.2154.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" In: *Phys. Rev.* 47 (10 May 1935), pp. 777–780. DOI: 10.1103/PhysRev.47.777.

[Esf+14] Michael Esfeld et al. "The Ontology of Bohmian Mechanics". In: *The British Journal for the Philosophy of Science* 65.4 (2014), pp. 773–796. ISSN: 00070882, 14643537. DOI: 10.1093/bjps/axt019.

[EV93] Avshalom C. Elitzur and Lev Vaidman. "Quantum mechanical interaction-free measurements". In: *Foundations of Physics* 23.7 (July 1993), pp. 987–997. ISSN: 1572-9516. DOI: 10.1007/BF00736012.

[Eve57] Hugh Everett III. ""Relative state" formulation of quantum mechanics". In: *Reviews of modern physics* 29.3 (1957), p. 454. DOI: 10.1103/RevModPhys.29.454.

[Fay19] Jan Faye. "Copenhagen Interpretation of Quantum Mechanics". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2019. Metaphysics Research Lab, Stanford University, 2019.

[FC72] Stuart J Freedman and John F Clauser. "Experimental test of local hidden-variable theories". In: *Physical Review Letters* 28.14 (1972), p. 938. DOI: 10.1103/PhysRevLett.28.938.

[Fer+05] F. Ferri et al. "High-Resolution Ghost Image and Ghost Diffraction Experiments with Thermal Light". In: *Phys. Rev. Lett.* 94 (18 May 2005), p. 183602. DOI: 10.1103/PhysRevLett.94.183602.

[For82] LH Ford. "Gravitational radiation by quantum systems". In: *Annals of Physics* 144.2 (1982), pp. 238–248. DOI: 10.1016/0003-4916(82)90115-4.

[FR18] Daniela Frauchiger and Renato Renner. "Quantum theory cannot consistently describe the use of itself". In: *Nature Communications* 9.1 (Sept. 2018), p. 3711. ISSN: 2041-1723. DOI: 10.1038/s41467-018-05739-8.

[Fri14] Simon Friederich. "Interpreting quantum theory: A therapeutic approach". In: First. New Directions in the Philosophy of Science. Basingstoke: Palgrave Macmillan UK, Nov. 2014, p. 55. DOI: 10.1057/9781137447159.

[FS14a] Terence C Farrelly and Anthony J Short. "Discrete spacetime and relativistic quantum particles". In: *Physical Review A* 89.6 (2014), p. 062109. DOI: 10.1103/PhysRevA.89.062109.

[FS14b] Christopher A Fuchs and Rüdiger Schack. "QBism and the Greeks: why a quantum state does not represent an element of physical reality". In: *Physica Scripta* 90.1 (2014), p. 015104. DOI: 10.1088/0031-8949/90/1/015104.

[Fuc02] Christopher A Fuchs. “Quantum mechanics as quantum information (and only a little more)”. In: *arXiv preprint quant-ph/0205039* (2002). URL: `https://arxiv.org/pdf/quant-ph/0205039.pdf`.

[Fuc10] Christopher A Fuchs. “QBism, the perimeter of quantum Bayesianism”. In: *arXiv preprint arXiv:1003.5209* (2010). URL: `https://arxiv.org/abs/1003.5209`.

[Fuc17] Christopher A Fuchs. “Notwithstanding Bohr, the reasons for QBism”. In: *Mind and Matter* 15.2 (2017), pp. 245–300.

[GC18] Danko Georgiev and Eliahu Cohen. “Probing finite coarse-grained virtual Feynman histories with sequential weak values”. In: *Phys. Rev. A* 97 (5 May 2018), p. 052102. DOI: `10.1103/PhysRevA.97.052102`.

[Gen16] Marco Genovese. “Real applications of quantum imaging”. In: *J. Opt. UK* 18.7 (June 2016), p. 073002. DOI: `10.1088/2040-8978/18/7/073002`.

[Geo11] Hans-Otto Georgii. *Gibbs measures and phase transitions*. de Gruyter, 2011. ISBN: 9783110250299.

[GH10] Wenlin Gong and Shensheng Han. “Phase-retrieval ghost imaging of complex-valued objects”. In: *Phys. Rev. A* 82 (2 Aug. 2010), p. 023828.

[GH96] Murray Gell-Mann and James B Hartle. “Quantum mechanics in the light of quantum cosmology”. In: *Foundations Of Quantum Mechanics In The Light Of New Technology: Selected Papers from the Proceedings of the First through Fourth International Symposia on Foundations of Quantum Mechanics*. World Scientific. 1996, pp. 347–369. DOI: `10.1142/9789812819895_0036`.

[Gil+19] Marta Gilaberte Basset et al. “Perspectives for applications of quantum imaging”. In: *Laser Photonics Rev.* 13.10 (2019), p. 1900097.

[Gis13] Nicolas Gisin. “Optical communication without photons”. In: *Phys. Rev. A* 88 (3 Sept. 2013), 030301(R). DOI: `10.1103/PhysRevA.88.030301`.

[Giu+15] Marissa Giustina et al. “Significant-Loophole-Free Test of Bell’s Theorem with Entangled Photons”. In: *Phys. Rev. Lett.* 115 (25 Dec. 2015), p. 250401. DOI: `10.1103/PhysRevLett.115.250401`.

[GLM04] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. “Quantum-Enhanced Measurements: Beating the Standard Quantum Limit”. In: *Science* 306.5700 (2004), pp. 1330–1336. DOI: `10.1126/science.1104149`.

[Gol20] Aaron Z Goldberg. “Quantum theory of polarimetry: From quantum operations to Mueller matrices”. In: *Physical Review Research* 2.2 (2020), p. 023038. DOI: `10.1103/PhysRevResearch.2.03038`.

[GPR90] Gian Carlo Ghirardi, Philip Pearle, and Alberto Rimini. "Markov processes in Hilbert space and continuous spontaneous localization of systems of identical particles". In: *Physical Review A* 42.1 (1990), p. 78. DOI: 10.1103/PhysRevA.42.78.

[Grä+14] Markus Gräfe et al. "On-chip generation of high-order single-photon W-states". In: *Nature Photonics* 8.10 (2014), pp. 791–795. DOI: 10.1038/nphoton.2014.204.

[Gri05] David J. Griffiths. *Introduction to Electrodynamics*. 2005.

[Gri16] Robert B. Griffiths. "Particle path through a nested Mach-Zehnder interferometer". In: *Phys. Rev. A* 94 (3 Sept. 2016), p. 032115. DOI: 10.1103/PhysRevA.94.032115.

[Gri17] Robert B. Griffiths. "Reply to "Comment on 'Particle path through a nested Mach-Zehnder interferometer' "". In: *Phys. Rev. A* 95 (6 June 2017), p. 066102. DOI: 10.1103/PhysRevA.95.066102.

[Gri18] Robert B. Griffiths. "Reply to "Comment on 'Particle path through a nested Mach-Zehnder interferometer' "". In: *Phys. Rev. A* 97 (2 Feb. 2018), p. 026102. DOI: 10.1103/PhysRevA.97.026102.

[Gri19] Robert B. Griffiths. "The Consistent Histories Approach to Quantum Mechanics". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2019. Metaphysics Research Lab, Stanford University, 2019.

[Gri22] Daniel Grimmer. "The Pragmatic QFT Measurement Problem and the need for a Heisenberg-like Cut in QFT". In: *arXiv preprint arXiv:2205.09608* (2022). URL: https://arxiv.org/abs/2205.09608.

[Gri84] Robert B. Griffiths. "Consistent histories and the interpretation of quantum mechanics". In: *Journal of Statistical Physics* 36.1-2 (1984), pp. 219–272. DOI: 10.1007/BF01015734.

[GRW80] Gian-Carlo Ghirardi, Alberto Rimini, and Tullio Weber. "A general argument against superluminal transmission through the quantum mechanical measurement process". In: *Lettere al Nuovo Cimento* 27.10 (1980), pp. 293–298. DOI: 10.1007/BF02817189.

[GRW86] Gian-Carlo Ghirardi, Alberto Rimini, and Tullio Weber. "Unified dynamics for microscopic and macroscopic systems". In: *Physical Review D* 34.2 (1986), p. 470. DOI: 10.1103/PhysRevD.34.470.

[GS99] Guang-Can Guo and Bao-Sen Shi. "Quantum cryptography based on interaction-free measurement". In: *Physics Letters A* 256.2-3 (1999), pp. 109–112. DOI: 10.1016/S0375-9601(99)00235-2.

[Guo+14a] Qi Guo et al. "Counterfactual distributed controlled-phase gate for quantum-dot spin qubits in double-sided optical microcavities". In: *Physical Review A* 90.4 (2014), p. 042327. DOI: 10.1103/PhysRevA.90.042327.

[Guo+14b] Qi Guo et al. "Counterfactual entanglement distribution without transmitting any particles". In: *Optics express* 22.8 (2014), pp. 8970–8984. DOI: 10.1364/OE.22.008970.

[Guo+15] Qi Guo et al. "Counterfactual quantum-information transfer without transmitting any physical particles". In: *Scientific reports* 5 (2015), p. 8416. DOI: 10.1038/srep08416.

[Guo+17] Qi Guo et al. "Counterfactual quantum cloning without transmitting any physical particles". In: *Phys. Rev. A* 96 (5 Nov. 2017), p. 052335. DOI: 10.1103/PhysRevA.96.052335.

[Guo+18] Qi Guo et al. "Counterfactual entanglement swapping enables high-efficiency entanglement distribution". In: *Optics express* 26.21 (2018), pp. 27314–27325. DOI: 10.1364/OE.26.027314.

[Han+19] Jonte Hance et al. "Modal, Truly Counterfactual Communication with On-Chip Demonstration Proposal". In: *Quantum Information and Measurement (QIM) V: Quantum Technologies*. Optical Society of America, 2019, T5A.50. DOI: 10.1364/QIM.2019.T5A.50.

[Har01] Lucien Hardy. "Quantum theory from five reasonable axioms". In: *arXiv preprint quant-ph/0101012* (2001). URL: https://arxiv.org/abs/quant-ph/0101012.

[Har13] Lucien Hardy. "Are quantum states real?" In: *International Journal of Modern Physics B* 27.01n03 (2013), p. 1345012. DOI: 10.1142/S0217979213450124.

[Has+16] FA Hashmi et al. "Two-state vector formalism and quantum interference". In: *Journal of Physics A: Mathematical and Theoretical* 49.34 (2016), p. 345302. DOI: 10.1088/1751-8113/49/34/345302.

[HBL17] Jérémie Harris, Robert W Boyd, and Jeff S Lundeen. "Weak value amplification can outperform conventional measurement in the presence of detector saturation". In: *Physical review letters* 118.7 (2017), p. 070802. DOI: 10.1103/PhysRevLett.118.070802.

[Hei+15] René Heilmann et al. "A novel integrated quantum circuit for high-order W-state generation and its highly precise characterization". In: *Science Bulletin* 60.1 (2015), pp. 96–100. ISSN: 2095-9273. DOI: 10.1007/s11434-014-0688-5.

[Hei52] Werner Heisenberg. "Questions of principle in modern physics". In: *Philosophic Problems in Nuclear Science*. Ed. by F. C. Hayes. Greenwich, Conneticut: Fawcett Publishing, 1952, pp. 41–52.

[Hei58] Werner Heisenberg. *Physics and philosophy: The revolution in modern science*. Harper & Row, 1958. ISBN: 1989970893.

[Hel19] Robert C Helling. "No Signalling and Unknowable Bohmian Particle Positions". In: *arXiv preprint arXiv:1902.03752* (2019). DOI: 10.48550/arXiv.1902.03752.

[Hen+15] B. Hensen et al. "Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres". In: *Nature* 526.7575 (Oct. 2015), pp. 682–686. ISSN: 1476-4687. DOI: 10.1038/nature15759.

[Her20] Ronnie Hermens. "Completely real? A critical note on the claims by Colbeck and Renner". In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 72 (2020), pp. 121–137. ISSN: 1355-2198. DOI: 10.1016/j.shpsb.2020.08.005.

[Her21] Ronnie Hermens. "How Real are Quantum States in ψ-ontic Models?" In: *arXiv preprint arXiv:2101.01972* (2021). URL: https://arxiv.org/abs/2101.01972.

[HH22a] Jonte R Hance and Sabine Hossenfelder. "The wave function as a true ensemble". In: *Proceedings of the Royal Society A* 478 (2022), p. 20210705. DOI: 10.1098/rspa.2021.0705.

[HH22b] Jonte R Hance and Sabine Hossenfelder. "What does it take to solve the measurement problem?" In: *arXiv preprint arXiv:2206.10445* (2022). DOI: 10.48550/arXiv.2206.10445.

[HH22c] Jonte R. Hance and Sabine Hossenfelder. "Bell's theorem allows local theories of quantum mechanics". In: *Nature Physics* 18.12 (Dec. 2022), pp. 1382–1382. ISSN: 1745-2481. DOI: 10.1038/s41567-022-01831-5.

[HHP22] Jonte R Hance, Sabine Hossenfelder, and Tim N. Palmer. "Supermeasured: Violating Bell-Statistical Independence Without Violating Physical Statistical Independence". In: *Foundations of Physics* 52.4 (July 2022), p. 81. ISSN: 1572-9516. DOI: 10.1007/s10701-022-00602-9.

[HLR21] Jonte R Hance, James Ladyman, and John Rarity. "How Quantum is Quantum Counterfactual Communication?" In: *Foundations of Physics* 51.1 (Feb. 2021), p. 12. DOI: 10.1007/s10701-021-00412-5.

[HLR22] Jonte R Hance, James Ladyman, and John Rarity. "Is the dynamical quantum Cheshire cat detectable?" In: *arXiv preprint arXiv:2204.03374* (2022). DOI: 10.48550/ARXIV.2204.03374.

[Hoo16] Gerard t Hooft. *The cellular automaton interpretation of quantum mechanics*. Springer Nature, 2016. ISBN: 978-3-319-41284-9. DOI: 10.1007/978-3-319-41285-6.

[Hos+06] Onur Hosten et al. "Counterfactual quantum computation through quantum interrogation". In: *Nature* 439.7079 (2006), p. 949. DOI: 10.1038/nature04523.

[Hos11] Sabine Hossenfelder. “Testing super-deterministic hidden variables theories”. In: *Foundations of Physics* 41.9 (2011), pp. 1521–1531. DOI: 10.1007/s10701-011-9565-0.

[Hos18] Sabine Hossenfelder. *Lost in math: How beauty leads physics astray*. Basic Civitas Books, 2018. ISBN: 0465094252.

[Hos20] Sabine Hossenfelder. “Superdeterminism: A Guide for the Perplexed”. In: *arXiv preprint arXiv:2010.01324* (2020). URL: https://arxiv.org/abs/2010.01324.

[HP20] Sabine Hossenfelder and Tim Palmer. “Rethinking superdeterminism”. In: *Frontiers in Physics* 8 (2020), p. 139. DOI: 10.3389/fphy.2020.00139.

[HPR21] Jonte R Hance, Tim N Palmer, and John Rarity. “Experimental Tests of Invariant Set Theory”. In: *arXiv preprint arXiv:2102.07795* (2021). URL: https://arxiv.org/abs/2102.07795.

[HR07] Nicholas Harrigan and Terry Rudolph. “Ontological models and the interpretation of contextuality”. In: *arXiv preprint arXiv:0709.4266* (2007). URL: https://arxiv.org/pdf/0709.4266.pdf.

[HR21a] Jonte Hance and John Rarity. “Comment on “Scheme of the arrangement for attack on the protocol BB84””. In: *Optik* (2021), p. 167451. DOI: 10.1016/j.ijleo.2021.167451.

[HR21b] Jonte R Hance and John Rarity. “Counterfactual ghost imaging”. In: *npj Quantum Information* 7.1 (2021), pp. 1–7. DOI: 10.1038/s41534-021-00411-4.

[HR22] Jonte R Hance and John Rarity. “Interaction-Free Polarimetry”. In: *arXiv preprint arXiv:2207.13539* (2022). DOI: 10.48550/arXiv.2207.13539.

[HRA07] Nicholas Harrigan, Terry Rudolph, and Scott Aaronson. “Representing probabilistic data via ontological models”. In: *arXiv preprint arXiv:0709.1149* (2007). URL: https://arxiv.org/pdf/0709.1149.pdf.

[HRL21] Jonte R Hance, John Rarity, and James Ladyman. “Do weak values show the past of a quantum particle?” In: *arXiv preprint arXiv:2109.14060* (2021). URL: https://arxiv.org/abs/2109.14060.

[HRL22] Jonte R Hance, John Rarity, and James Ladyman. “Could wavefunctions simultaneously represent knowledge and reality?” In: *Quantum Studies: Mathematics and Foundations* (Apr. 2022). ISSN: 2196-5617. DOI: 10.1007/s40509-022-00271-3.

[HS10] Nicholas Harrigan and Robert W Spekkens. “Einstein, incompleteness, and the epistemic view of quantum states”. In: *Foundations of Physics* 40.2 (2010), pp. 125–157. DOI: 10.1007/s10701-009-9347-0.

[Hu+09] CY Hu et al. "Proposed entanglement beam splitter using a quantum-dot spin in a double-sided optical microcavity". In: *Physical Review B* 80.20 (2009), p. 205326. DOI: 10.1103/PhysRevB.80.205326.

[Ips22] Asger C. Ipsen. "Anomalous Weak Values are Caused by Disturbance". In: *Foundations of Physics* 52.1 (Jan. 2022), p. 20. ISSN: 1572-9516. DOI: 10.1007/s10701-021-00534-w.

[Jah10] Jörg Jahnel. "When is the (co) sine of a rational angle equal to a rational number?" In: *arXiv preprint arXiv:1006.2938* (2010). URL: https://arxiv.org/abs/1006.2938.

[JMH14] Andrew N Jordan, Julián Martínez-Rincón, and John C Howell. "Technical advantages for weak-value amplification: when less is more". In: *Physical Review X* 4.1 (2014), p. 011031. DOI: 10.1103/PhysRevX.4.011031.

[Joz99] Richard Jozsa. "Quantum Effects in Algorithms". In: *Quantum Computing and Quantum Communications*. Ed. by Colin P. Williams. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp. 103–112. ISBN: 978-3-540-49208-5. DOI: 10.1007/3-540-49208-9_7.

[JSL11] Musheng Jiang, Shihai Sun, and Linmei Liang. "Practical Stabilization of Counterfactual Quantum Cryptography". In: *Journal of Quantum Information Science* 1.03 (2011), p. 116.

[JSV95] Gregg Jaeger, Abner Shimony, and Lev Vaidman. "Two interferometric complementarities". In: *Physical Review A* 51.1 (1995), p. 54. DOI: 10.1103/PhysRevA.51.54.

[Kas13] Ruth E Kastner. *The transactional interpretation of quantum mechanics: the reality of possibility*. Cambridge University Press, 2013.

[Kas14] Ruth E Kastner. "'Einselection' of pointer observables: The new H-theorem?" In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 48 (2014), pp. 56–58. ISSN: 1355-2198. DOI: 10.1016/j.shpsb.2014.06.004.

[Kas17] Ruth E Kastner. "Demystifying Weak Measurements". In: *Foundations of Physics* 47.5 (May 2017), pp. 697–707. ISSN: 1572-9516. DOI: 10.1007/s10701-017-0085-4.

[KB18] Andreĭ Nikolaevich Kolmogorov and Albert T Bharucha-Reid. *Foundations of the theory of probability: Second English Edition*. Courier Dover Publications, 2018.

[Ken27] Earle H Kennard. "Zur quantenmechanik einfacher bewegungstypen". In: *Zeitschrift für Physik* 44.4-5 (1927), pp. 326–352. DOI: 10.1007/BF01391200.

[Key21] John Maynard Keynes. *A treatise on probability*. Macmillan and Company, limited, 1921. ISBN: 978-0-486-49580-4.

[KF93] Chung-I Kuo and L. H. Ford. "Semiclassical gravity theory and quantum fluctuations". In: *Phys. Rev. D* 47 (1993), pp. 4510–4519. DOI: 10.1103/PhysRevD.47.4510.

[Kho16a] DL Khokhlov. “Scheme of the weak measurement of the polarization state of the photon”. In: *Optik* 127.8 (2016), pp. 4089–4091. DOI: 10.1016/j.ijleo.2016.01.049.

[Kho16b] Dmitri L Khokhlov. “Scheme of the arrangement for attack on the protocol BB84”. In: *Optik* 127.18 (2016), pp. 7083–7087. ISSN: 0030-4026. DOI: https://doi.org/10.1016/j.ijleo.2016.05.023.

[Kle11] U Klein. “What is the limit $\hbar \to 0$ of quantum theory?” In: *arXiv preprint arXiv:1201.0150* (2011).

[Koc+11] Sacha Kocsis et al. “Observing the Average Trajectories of Single Photons in a Two-Slit Interferometer”. In: *Science* 332.6034 (2011), pp. 1170–1173. DOI: 10.1126/science.1202218.

[KR11] V. Alan Kostelecký and Neil Russell. “Data tables for Lorentz and CPT violation”. In: *Reviews of Modern Physics* 83.1 (Mar. 2011), pp. 11–31. DOI: 10.1103/revmodphys.83.11.

[Kwi+95] Paul Kwiat et al. “Interaction-Free Measurement”. In: *Phys. Rev. Lett.* 74 (24 June 1995), pp. 4763–4766. DOI: 10.1103/PhysRevLett.74.4763.

[Kwi+99] Paul G Kwiat et al. “High-efficiency quantum interrogation measurements via the quantum Zeno effect”. In: *Physical Review Letters* 83.23 (1999), p. 4725. DOI: 10.1103/PhysRevLett.83.4725.

[KWZ96] Paul Kwiat, Harald Weinfurter, and Anton Zeilinger. “Quantum seeing in the dark”. In: *Scientific American* 275.5 (1996), pp. 72–78.

[Lan06] Nicolaas P Landsman. “Between classical and quantum”. In: *Handbook of the Philosophy of Science* 2 (2006), pp. 417–553.

[Lar+14] Jan-Åke Larsson et al. “Bell-inequality violation with entangled photons, free of the coincidence-time loophole”. In: *Phys. Rev. A* 90 (3 Sept. 2014), p. 032107. DOI: 10.1103/PhysRevA.90.032107.

[Lar14] Jan-Åke Larsson. “Loopholes in Bell inequality tests of local realism”. In: *Journal of Physics A: Mathematical and Theoretical* 47.42 (2014), p. 424003. DOI: 10.1088/1751-8113/47/42/424003.

[LAZ13] Zheng-Hong Li, Mohammad Al-Amri, and Muhammad Suhail Zubairy. “Comment on “Past of a quantum particle””. In: *Phys. Rev. A* 88 (4 Oct. 2013), p. 046102. DOI: 10.1103/PhysRevA.88.046102.

[LAZ15] Zheng-Hong Li, Mohammad Al-Amri, and Muhammad Suhail Zubairy. “Direct counterfactual transmission of a quantum state”. In: *Phys. Rev. A* 92 (5 Nov. 2015), p. 052315. DOI: 10.1103/PhysRevA.92.052315.

[LD21] Olimpia Lombardi and Dennis Dieks. "Modal Interpretations of Quantum Mechanics". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2021. Metaphysics Research Lab, Stanford University, 2021. URL: https://plato.stanford.edu/archives/win2021/entries/qm-modal/.

[Lee16] Gijs Leegwater. "An impossibility theorem for parameter independent hidden variable theories". In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 54 (2016), pp. 18–34. DOI: 10.1016/j.shpsb.2016.01.003.

[Leg05] AJ Leggett. "The quantum measurement problem". In: *science* 307.5711 (2005), pp. 871–872. DOI: 10.1126/science.1109541.

[Lei14] Matthew Leifer. "Is the Quantum State Real? An Extended Review of ψ-ontology Theorems". In: *Quanta* 3.1 (2014), pp. 67–155. ISSN: 1314-7374. DOI: 10.12743/quanta.v3i1.22.

[LG04] J-Å Larsson and Richard David Gill. "Bell's inequality and the coincidence-time loophole". In: *EPL (Europhysics Letters)* 67.5 (2004), p. 707. DOI: 10.1209/epl/i2004-10124-7.

[Li+19] Zheng-Hong Li et al. "Counterfactual exchange of unknown quantum states". In: *Physical Review A* 100.2 (2019), p. 022110. DOI: 10.1103/PhysRevA.100.022110.

[Li14] Yan-Bing Li. "Analysis of counterfactual quantum key distribution using error-correcting theory". In: *Quantum information processing* 13.10 (2014), pp. 2325–2342. DOI: 10.1007/s11128-014-0786-y.

[Liu+12] Yang Liu et al. "Experimental Demonstration of Counterfactual Quantum Communication". In: *Phys. Rev. Lett.* 109 (3 July 2012), p. 030501. DOI: 10.1103/PhysRevLett.109.030501.

[Liu+14] Xingtong Liu et al. "Eavesdropping on counterfactual quantum key distribution with finite resources". In: *Physical Review A* 90.2 (2014), p. 022318. URL: 10.1103/PhysRevA.90.022318.

[Liu+18] Chao Liu et al. "Improvement of reliability in multi-interferometer-based counterfactual deterministic communication with dissipation compensation". In: *Optics express* 26.3 (2018), pp. 2261–2269. DOI: 10.1364/OE.26.002261.

[Lou+09] Pavel Lougovski et al. "Verifying multipartite mode entanglement of W states". In: *New Journal of Physics* 11.6 (2009), p. 063029. DOI: 10.1088/1367-2630/11/6/063029.

[LPS08] James Ladyman, Stuart Presnell, and Anthony J Short. "The use of the information-theoretic entropy in thermodynamics". In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 39.2 (2008), pp. 315–324. DOI: 10.1016/j.shpsb.2007.11.004.

[LWL13] Yan-Bing Li, Qiao-Yan Wen, and Zi-Chen Li. "Security flaw of counterfactual quantum cryptography in practical setting". In: *arXiv preprint arXiv:1312.1436* (2013). URL: https://arxiv.org/abs/1312.1436.

[Mah+16] Dylan H. Mahler et al. "Experimental nonlocal and surreal Bohmian trajectories". In: *Science Advances* 2.2 (2016), e1501466. DOI: 10.1126/sciadv.1501466.

[Mal+10] Mehul Malik et al. "Quantum ghost image identification with correlated photon pairs". In: *Phys. Rev. Lett.* 104.16 (2010), p. 163602. DOI: 10.1103/PhysRevLett.104.163602.

[Mar+21] Yair Margalit et al. "Realization of a complete Stern-Gerlach interferometer: Toward a test of quantum gravity". In: *Science Advances* 7.22 (2021). DOI: 10.1126/sciadv.abg2879.

[Mar12] Owen JE Maroney. "How statistical are quantum states?" In: *arXiv preprint arXiv:1207.6906* (2012). URL: https://arxiv.org/pdf/1207.6906.pdf.

[Mat05] David Mattingly. "Modern Tests of Lorentz Invariance". In: *Living Reviews in Relativity* 8.1 (Dec. 2005), p. 5. ISSN: 1433-8351. DOI: 10.12942/lrr-2005-5.

[Mau02] Tim Maudlin. *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics*. Blackwell, 2002.

[Mau95] Tim Maudlin. "Three measurement problems". In: *topoi* 14.1 (1995), pp. 7–15. DOI: 10.1007/BF00763473.

[Mer22] N. David Mermin. "There is no quantum measurement problem". In: *Physics Today* 75.6 (2022), pp. 62–63. DOI: 10.1063/PT.3.5027.

[MJ01] Graeme Mitchison and Richard Jozsa. "Counterfactual computation". In: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 457.2009 (2001), pp. 1175–1193. DOI: 10.1098/rspa.2000.0714.

[Mor+15] Peter A Morris et al. "Imaging with a small number of photons". In: *Nat. Commun.* 6.1 (2015), pp. 1–6. DOI: 10.1038/ncomms6913.

[Mor+18] Paul-Antoine Moreau et al. "Resolution limits of quantum ghost imaging". In: *Opt. Express* 26.6 (Mar. 2018), pp. 7528–7536. DOI: 10.1364/OE.26.007528.

[Mot29] Nevill Francis Mott. "The wave mechanics of α-Ray tracks". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 126.800 (1929), pp. 79–84. DOI: 10.1098/rspa.1929.0205.

[MPB01] Markus D Müller, Thomas Poiger, and Hans-Rudolf Buser. "Isolation and identification of the metolachlor stereoisomers using high-performance liquid chromatography, polarimetric measurements, and enantioselective gas chromatography". In: *Journal of agricultural and food chemistry* 49.1 (2001), pp. 42–49. DOI: 10.1021/jf000857f.

[MR02] Donald Marolf and Carlo Rovelli. "Relativistic quantum measurement". In: *Physical Review D* 66.2 (2002), p. 023510. DOI: 10.1103/PhysRevD.66.023510.

[MS77] B. Misra and E. C. G. Sudarshan. "The Zeno's paradox in quantum theory". In: *Journal of Mathematical Physics* 18.4 (1977), pp. 756–763. DOI: 10.1063/1.523304.

[MTW73] Charles W Misner, Kip S Thorne, and John Archibald Wheeler. *Gravitation*. San Francisco: W.H. Freeman, 1973. ISBN: 0716703440.

[Müc+10] Martin Mücke et al. "Electromagnetically induced transparency with single atoms in a cavity". In: *Nature* 465 (2010), pp. 755–758. ISSN: 00280836. DOI: 10.1038/nature09093.

[Mur+99] M Murao et al. "Quantum telecloning and multiparticle entanglement". In: *Physical Review A* 59.1 (1999), p. 156. DOI: 10.1103/PhysRevA.59.156.

[MV17] Chiara Marletto and Vlatko Vedral. "Gravitationally induced entanglement between two massive particles is sufficient evidence of quantum effects in gravity". In: *Physical review letters* 119.24 (2017), p. 240402. DOI: 10.1103/PhysRevLett.119.240402.

[Myr02] Wayne C Myrvold. "On peaceful coexistence: is the collapse postulate incompatible with relativity?" In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 33.3 (2002), pp. 435–466. DOI: 10.1016/S1369-8486(02)00004-3.

[Myr20] Wayne Myrvold. "On the status of quantum state realism". In: *Scientific realism and the quantum*. Ed. by Steven French and Juha Saatsi. Oxford University Press, 2020. DOI: 10.1093/oso/9780198814979.003.0012.

[NC10] Michael A Nielsen and Isaac Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2010. ISBN: 978-1-107-00217-3.

[Nik06] Hrvoje Nikolic. "Many-fingered time Bohmian mechanics". In: *arXiv preprint quant-ph/0603207* (2006). URL: https://arxiv.org/abs/quant-ph/0603207.

[Niv56] Ivan Niven. "Irrational numbers". In: The Carus Mathematical Monographs 11. The Mathematical Association of America, 1956, p. 41. DOI: 10.5948/9781614440116.

[Noh09] Tae-Gon Noh. "Counterfactual Quantum Cryptography". In: *Phys. Rev. Lett.* 103 (23 Dec. 2009), p. 230501. DOI: 10.1103/PhysRevLett.103.230501.

[Nor11] Travis Norsen. "John S. Bell's concept of local causality". In: *American Journal of Physics* 79.12 (2011), pp. 1261–1275. DOI: 10.1119/1.3630940.

[NV22] Vitaly Nikolaev and Louis Vervoort. “Aspects of Superdeterminism Made Intuitive”. In: *arXiv preprint arXiv:2205.10616* (2022). DOI: `10.48550/arXiv.2205.10616`.

[OCB10] Malcolm N. O’Sullivan, Kam Wai Clifford Chan, and Robert W. Boyd. “Comparison of the signal-to-noise characteristics of quantum versus thermal ghost imaging”. In: *Phys. Rev. A* 82 (5 Nov. 2010), p. 053803. DOI: `10.1103/PhysRevA.82.053803`.

[OL20] Andrea Oldofredi and Cristian López. “On the Classification Between ψ-Ontic and ψ-Epistemic Ontological Models”. In: *Foundations of Physics* 50.11 (Nov. 2020), pp. 1315–1345. ISSN: 1572-9516. DOI: `10.1007/s10701-020-00377-x`.

[OLL02] Sangchul Oh, Soonchil Lee, and Hai-woong Lee. “Fidelity of quantum teleportation through noisy channels”. In: *Physical Review A* 66.2 (Aug. 2002). ISSN: 1094-1622. DOI: `10.1103/physreva.66.022316`.

[OV06] Tobias J Osborne and Frank Verstraete. “General monogamy inequality for bipartite qubit entanglement”. In: *Physical review letters* 96.22 (2006), p. 220503. DOI: `10.1103/PhysRevLett.96.220503`.

[Pal09] Tim N Palmer. “The Invariant Set Postulate: a new geometric framework for the foundations of quantum theory and the role played by gravity”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 465.2110 (2009), pp. 3165–3185. DOI: `10.1098/rspa.2009.0080`.

[Pal11] Tim N Palmer. “The invariant set hypothesis: a new geometric framework for the foundations of quantum theory and the role played by gravity”. In: *Electronic Notes in Theoretical Computer Science* 270.2 (2011), pp. 115–119. DOI: `10.1016/j.entcs.2011.01.026`.

[Pal12a] Tim N Palmer. “Quantum theory and the symbolic dynamics of invariant sets: towards a gravitational theory of the quantum”. In: *arXiv preprint arXiv:1210.3940* (2012). URL: `https://arxiv.org/abs/1210.3940`.

[Pal12b] Tim N Palmer. “THE BUTTERFLY AND THE PHOTON: NEW PERSPECTIVES ON UNPREDICTABILITY, AND THE NOTION OF CASUAL REALITY, IN QUANTUM PHYSICS”. In: *Science: Image In Action*. World Scientific, 2012, pp. 129–139. DOI: `10.1142/9789814383295_0011`.

[Pal14] Tim N Palmer. “Lorenz, Gödel and Penrose: new perspectives on determinism and causality in fundamental physics”. In: *Contemporary Physics* 55.3 (2014), pp. 157–178. DOI: `10.1080/00107514.2014.908624`.

[Pal16] Tim N Palmer. “Invariant set theory”. In: *arXiv preprint arXiv:1605.01051* (2016). URL: `https://arxiv.org/abs/1605.01051`.

[Pal17] Tim N Palmer. “A Gravitational Theory of the Quantum”. In: *arXiv preprint arXiv:1709.00329* (2017). URL: `https://arxiv.org/abs/1709.00329`.

[Pal18] Tim N Palmer. "Experimental non-violation of the Bell inequality". In: *Entropy* 20.5 (2018), p. 356. DOI: 10.3390/e20050356.

[Pal19] Tim N Palmer. "Bell Inequality Violation with Free Choice and Local Causality on the Invariant Set". In: *arXiv preprint arXiv:1903.10537* (2019). URL: https://arxiv.org/abs/1903.10537.

[Pal20a] Tim N Palmer. "Discretization of the Bloch sphere, fractal invariant sets and Bell's theorem". In: *Proceedings of the Royal Society A* 476.2236 (2020), p. 20190350. DOI: 10.1098/rspa.2019.0350.

[Pal20b] Tim N Palmer. "Undecidability, Fractal Geometry and the Unity of Physics". In: *FQXi Undecidability, Uncomputability, and Unpredictability Essay Contest (2019-2020)* (2020). URL: https://fqxi.org/data/essay-contest-files/Palmer_FXQi_Palmer_1.pdf.

[Pal95] Tim N Palmer. "A local deterministic model of quantum spin measurement". In: *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences* 451.1943 (1995), pp. 585–608. DOI: 10.1098/rspa.1995.0145.

[PB17] Miles J. Padgett and Robert W. Boyd. "An introduction to ghost imaging: quantum and classical". In: *Philos. T. R. Soc. A* 375.2099 (2017), p. 20160233. DOI: 10.1098/rsta.2016.0233.

[PBR12] Matthew F Pusey, Jonathan Barrett, and Terry Rudolph. "On the reality of the quantum state". In: *Nature Physics* 8.6 (2012), p. 475. DOI: 10.1038/nphys2309.

[PC17] Dilip Paneru and Eliahu Cohen. "Past of a particle in an entangled state". In: *International Journal of Quantum Information* 15.08 (2017), p. 1740019. DOI: 10.1142/S0219749917400196.

[Pen94] Roger Penrose. *Shadows of the Mind*. Vol. 4. Oxford University Press Oxford, 1994.

[Pen96] Roger Penrose. "On gravity's role in quantum state reduction". In: *General relativity and gravitation* 28.5 (1996), pp. 581–600. DOI: 10.1007/BF02105068.

[Per98] Ian C. Percival. "Quantum transfer functions, weak nonlocality and relativity". In: *Physics Letters A* 244.6 (1998), pp. 495–501. ISSN: 0375-9601. DOI: 10.1016/S0375-9601(98)00353-3.

[PF15] Václav Potoček and Gergely Ferenczi. "Which-way information in a nested Mach-Zehnder interferometer". In: *Physical Review A* 92.2 (2015), p. 023829. DOI: 10.1103/PhysRevA.92.023829.

[Pir+15] Stefano Pirandola et al. "Advances in quantum teleportation". In: *Nature photonics* 9.10 (2015), pp. 641–652. DOI: 10.1038/nphoton.2015.154.

[Pit+95] TB Pittman et al. "Optical imaging by means of two-photon quantum entanglement". In: *Phys. Rev. A* 52.5 (1995), R3429. DOI: 10.1103/PhysRevA.52.R3429.

[Pop14] Sandu Popescu. “Nonlocality beyond quantum mechanics”. In: *Nature Physics* 10.4 (Apr. 2014), pp. 264–270. ISSN: 1745-2481. DOI: 10.1038/nphys2916.

[PPM13] M. K. Patra, S. Pironio, and S. Massar. “No-Go Theorems for ψ-Epistemic Models Based on a Continuity Assumption”. In: *Phys. Rev. Lett.* 111 (9 Aug. 2013), p. 090402. DOI: 10.1103/PhysRevLett.111.090402.

[PR94] Sandu Popescu and Daniel Rohrlich. “Quantum nonlocality as an axiom”. In: *Foundations of Physics* 24.3 (Mar. 1994), pp. 379–385. ISSN: 1572-9516. DOI: 10.1007/BF02058098.

[Pra+21] Thomas Pradeu et al. “Philosophy in Science: Can philosophers of science permeate through science and produce scientific knowledge?” In: *The British Journal for the Philosophy of Science* (2021), to appear. DOI: 10.1086/715518.

[Pri12] Huw Price. “Does time-symmetry imply retrocausality? How the quantum world says “Maybe”?” In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 43.2 (2012), pp. 75–83. DOI: 10.1016/j.shpsb.2011.12.003.

[Pri94] Huw Price. “A neglected route to realism about quantum mechanics”. In: *Mind* 103.411 (1994), pp. 303–336. DOI: 10.1093/mind/103.411.303.

[Pro+19] Massimiliano Proietti et al. “Experimental test of local observer independence”. In: *Science advances* 5.9 (2019), eaaw9832. DOI: 10.1126/sciadv.aaw9832.

[PT04] Asher Peres and Daniel R. Terno. “Quantum information and relativity theory”. In: *Rev. Mod. Phys.* 76 (1 Jan. 2004), pp. 93–123. DOI: 10.1103/RevModPhys.76.93.

[PV18] Uri Peleg and Lev Vaidman. “Comment on" Past of a quantum particle revisited"”. In: *arXiv preprint arXiv:1805.12171* (2018). URL: https://arxiv.org/abs/1805.12171.

[RA02] Daniel Rohrlich and Yakir Aharonov. “Cherenkov radiation of superluminal particles”. In: *Phys. Rev. A* 66 (4 Oct. 2002), p. 042102. DOI: 10.1103/PhysRevA.66.042102.

[RA09] Amit Rai and GS Agarwal. “Possibility of coherent phenomena such as Bloch oscillations with single photons via W states”. In: *Physical Review A* 79.5 (2009), p. 053849. DOI: 10.1103/PhysRevA.79.053849.

[Reb+21] Enrico Rebufello et al. “Anomalous weak values via a single photon detection”. In: *Light: Science & Applications* 10.1 (May 2021), p. 106. ISSN: 2047-7538. DOI: 10.1038/s41377-021-00539-0.

[Red87] Michael Redhead. *Incompleteness, nonlocality, and realism: a prolegomenon to the philosophy of quantum mechanics*. Oxford University Press, 1987.

[Rei+14] Andreas Reiserer et al. “A quantum gate between a flying optical photon and a single trapped atom”. In: *Nature* 508 (2014), pp. 237–240. ISSN: 14764687. DOI: 10.1038/nature13177.

[Rei14] Andreas Reiserer. “A controlled phase gate between a single atom and an optical photon”. PhD thesis. Technische Universität München, 2014.

[Rel18] Armando Relaño. “Decoherence allows quantum theory to describe the use of itself”. In: *arXiv preprint arXiv:1810.07065* (2018). URL: https://arxiv.org/abs/1810.07065.

[Ren+11] M. Ren et al. “Experimental demonstration of counterfactual quantum key distribution”. In: *Laser Physics* 21.4 (Apr. 2011), pp. 755–760. ISSN: 1555-6611. DOI: 10.1134/S1054660X11070267.

[Ren60] M. Renninger. “Messungen ohne Störung des Meßobjekts”. In: *Zeitschrift für Physik* 158.4 (Aug. 1960), pp. 417–421. ISSN: 0044-3328. DOI: 10.1007/BF01327019.

[Rin+15] M. Ringbauer et al. “Measurements on the reality of the wavefunction”. In: *Nature Physics* 11.3 (Mar. 2015), pp. 249–254. ISSN: 1745-2481. DOI: 10.1038/nphys3233.

[RLE20] Joshua B. Ruebeck, Piers Lillystone, and Joseph Emerson. “ψ-epistemic interpretations of quantum theory have a measurement problem”. In: *Quantum* 4 (Mar. 2020), p. 242. ISSN: 2521-327X. DOI: 10.22331/q-2020-03-16-242.

[Rog98] Claude Ambrose Rogers. *Hausdorff measures*. Cambridge University Press, 1998. DOI: 10.1017/S001309150000969X.

[Rov96] Carlo Rovelli. “Relational quantum mechanics”. In: *International Journal of Theoretical Physics* 35.8 (1996), pp. 1637–1678. DOI: 10.1007/BF02302261.

[Rov98] Carlo Rovelli. “Loop Quantum Gravity”. In: *Living Reviews in Relativity* 1.1 (Jan. 1998), p. 1. ISSN: 1433-8351. DOI: 10.12942/lrr-1998-1.

[Row+01] M. A. Rowe et al. “Experimental violation of a Bell’s inequality with efficient detection”. In: *Nature* 409.6822 (Feb. 2001), pp. 791–794. ISSN: 1476-4687. DOI: 10.1038/35057215.

[RR15] Andreas Reiserer and Gerhard Rempe. “Cavity-based quantum networks with single atoms and optical photons”. In: *Reviews of Modern Physics* 87.4 (Dec. 2015), pp. 1379–1418. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.87.1379.

[RSH91] NWM Ritchie, J Greg Story, and Randall G Hulet. “Realization of a measurement of a “weak value””. In: *Physical review letters* 66.9 (1991), p. 1107. DOI: 10.1103/PhysRevLett.66.1107.

[RT90] JG Rarity and PR Tapster. “Experimental violation of Bell’s inequality based on phase and momentum”. In: *Physical Review Letters* 64.21 (1990), p. 2495. DOI: 10.1103/PhysRevLett.64.2495.

[Rud06] Terry Rudolph. "Ontological models for quantum mechanics and the Kochen-Specker theorem". In: *arXiv preprint arXiv:quant-ph/0608120* (2006). URL: `https://arxiv.org/abs/quant-ph/0608120`.

[Sal+13a] Hatim Salih et al. "Optical communication with invisible photons". In: *The Rochester Conferences on Coherence and Quantum Optics and the Quantum Information and Measurement meeting*. Optical Society of America, 2013, T4B.1. DOI: `10.1364/CQO.2013.T4B.1`.

[Sal+13b] Hatim Salih et al. "Protocol for Direct Counterfactual Quantum Communication". In: *Phys. Rev. Lett.* 110 (17 Apr. 2013), p. 170502. DOI: `10.1103/PhysRevLett.110.170502`.

[Sal+14] Hatim Salih et al. "Salih et al. Reply:" in: *Phys. Rev. Lett.* 112 (20 May 2014), p. 208902. DOI: `10.1103/PhysRevLett.112.208902`.

[Sal+20] Hatim Salih et al. "Deterministic Teleportation and Universal Computation Without Particle Exchange". In: *arXiv preprint arXiv:2009.05564* (2020). URL: `https://arxiv.org/abs/2009.05564`.

[Sal+21] Hatim Salih et al. "Exchange-free computation on an unknown qubit at a distance". In: *New Journal of Physics* 23.1 (Jan. 2021), p. 013004. DOI: `10.1088/1367-2630/abd3c4`.

[Sal+22] Hatim Salih et al. "The laws of physics do not prohibit counterfactual communication". In: *npj Quantum Information* 8.1 (May 2022), p. 60. ISSN: 2056-6387. DOI: `10.1038/s41534-022-00564-w`.

[Sal14a] Pablo L Saldanha. "Interpreting a nested Mach-Zehnder interferometer with classical optics". In: *Physical Review A* 89.3 (2014), p. 033825. DOI: `10.1103/PhysRevA.89.033825`.

[Sal14b] Hatim Salih. "Tripartite counterfactual quantum cryptography". In: *Physical Review A* 90.1 (2014), p. 012333. DOI: `10.1103/PhysRevA.90.012333`.

[Sal15] Hatim Salih. "Commentary: "Asking photons where they have been" - without telling them what to say". In: *Frontiers in Physics* 3 (2015), p. 47. ISSN: 2296-424X. DOI: `10.3389/fphy.2015.00047`.

[Sal16] Hatim Salih. "Protocol for counterfactually transporting an unknown qubit". In: *Frontiers in Physics* 3 (2016), p. 94. DOI: `10.3389/fphy.2015.00094`.

[Sal18a] Hatim Salih. "Comment on "Particle path through a nested Mach-Zehnder interferometer"". In: *Phys. Rev. A* 97 (2 Feb. 2018), p. 026101. DOI: `10.1103/PhysRevA.97.026101`.

[Sal18b] Hatim Salih. “Counterfactual quantum erasure: spooky action without entanglement”. In: *Royal Society open science* 5.2 (2018), p. 171250. DOI: 10.1098/rsos.171250.

[Sal18c] Hatim Salih. “From a Quantum Paradox to Counterportation”. In: *arXiv preprint arXiv:1807.06586* (2018). URL: https://arxiv.org/abs/1807.06586.

[SB12] Jeffrey H. Shapiro and Robert W. Boyd. “The physics of ghost imaging”. In: *Quantum Information Processing* 11.4 (2012), pp. 949–993. DOI: 10.1007/s11128-011-0356-5.

[SC18] Charles T. Sebens and Sean M. Carroll. “Self-locating Uncertainty and the Origin of Probability in Everettian Quantum Mechanics”. In: *The British Journal for the Philosophy of Science* 69.1 (2018), pp. 25–74. DOI: 10.1093/bjps/axw004.

[Sch05] Maximilian Schlosshauer. “Decoherence, the measurement problem, and interpretations of quantum mechanics”. In: *Rev. Mod. Phys.* 76 (4 Feb. 2005), pp. 1267–1305. DOI: 10.1103/RevModPhys.76.1267.

[Sch52] Erwin Schrödinger. “Are there quantum jumps? Part I”. In: *The British Journal for the Philosophy of science* 3.10 (1952), pp. 109–123.

[Sen21] Indrajit Sen. “Analysis of the superdeterministic Invariant-set theory in a hidden-variable setting”. In: *arXiv preprint arXiv:2107.04761* (2021).

[SF12] Maximilian Schlosshauer and Arthur Fine. “Implications of the Pusey-Barrett-Rudolph Quantum No-Go Theorem”. In: *Phys. Rev. Lett.* 108 (26 June 2012), p. 260404. DOI: 10.1103/PhysRevLett.108.260404.

[Shi93] Abner Shimony. “Controllable and uncontrollable non-locality”. In: *The Search for a Naturalistic World View*. Vol. 2. Cambridge University Press, 1993, pp. 130–139. DOI: 10.1017/CBO9781139172196.010.

[SKZ13] Maximilian Schlosshauer, Johannes Kofler, and Anton Zeilinger. “A snapshot of foundational attitudes toward quantum mechanics”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 44.3 (2013), pp. 222–230. ISSN: 1355-2198. DOI: 10.1016/j.shpsb.2013.04.004.

[Sla21] Kevin Slagle. “Testing quantum mechanics using noisy quantum computers”. In: *arXiv preprint arXiv:2108.02201* (2021).

[Smo12] Lee Smolin. “A Real Ensemble Interpretation of Quantum Mechanics”. In: *Foundations of Physics* 42.10 (Oct. 2012), pp. 1239–1261. ISSN: 1572-9516. DOI: 10.1007/s10701-012-9666-4.

[SN16] Sujeevan Sivasundaram and Kristian Hvidtfelt Nielsen. “Surveying the attitudes of physicists concerning foundational issues of Surveym mechanics”. In: *arXiv preprint arXiv:1612.00676* (2016). URL: https://arxiv.org/abs/1612.00676.

[Sol+04] D. R. Solli et al. “Fast Light, Slow Light, and Phase Singularities: A Connection to Generalized Weak Values”. In: *Phys. Rev. Lett.* 92 (4 Jan. 2004), p. 043601. DOI: 10.1103/PhysRevLett.92.043601.

[SP14] Chitra Shukla and Anirban Pathak. “Orthogonal-state-based deterministic secure quantum communication without actual transmission of the message qubits”. In: *Quantum Information Processing* 13.9 (Sept. 2014), pp. 2099–2113. ISSN: 1573-1332. DOI: 10.1007/s11128-014-0792-0.

[Spe05] Robert W Spekkens. “Contextuality for preparations, transformations, and unsharp measurements”. In: *Phys. Rev. A* 71 (5 May 2005), p. 052108. DOI: 10.1103/PhysRevA.71.052108.

[Spe07] Robert W Spekkens. “Evidence for the epistemic view of quantum states: A toy theory”. In: *Phys. Rev. A* 75 (3 Mar. 2007), p. 032110. DOI: 10.1103/PhysRevA.75.032110.

[SS13] Akshata Shenoy-Hejamadi and R. Srikanth. “The wave-function is real but nonphysical: A view from counterfactual quantum cryptography”. In: *arXiv preprint arXiv:1311.7127* (2013). URL: https://arxiv.org/abs/1311.7127.

[SS15] Akshata Shenoy-Hejamadi and R. Srikanth. “Counterfactual distribution of Schrödinger cat states”. In: *Phys. Rev. A* 92 (6 Dec. 2015), p. 062308. DOI: 10.1103/PhysRevA.92.062308.

[SSS13a] Akshata Shenoy-Hejamadi, R. Srikanth, and T. Srinivas. “Counterfactual quantum key distribution without polarization encoding”. In: *AQIS 2013 Proceedings* (2013).

[SSS13b] Akshata Shenoy-Hejamadi, R. Srikanth, and T. Srinivas. “Semi-counterfactual cryptography”. In: *EPL (Europhysics Letters)* 103.6 (Sept. 2013), p. 60008. DOI: 10.1209/0295-5075/103/60008.

[SSS14] Akshata Shenoy-Hejamadi, R. Srikanth, and T. Srinivas. “Counterfactual quantum certificate authorization”. In: *Phys. Rev. A* 89 (5 May 2014), p. 052307. DOI: 10.1103/PhysRevA.89.052307.

[SV20a] Indrajit Sen and Antony Valentini. “Superdeterministic hidden-variables models I: non-equilibrium and signalling”. In: *Proceedings of the Royal Society A* 476.2243 (2020), p. 20200212. DOI: 10.1098/rspa.2020.0212.

[SV20b] Indrajit Sen and Antony Valentini. “Superdeterministic hidden-variables models II: conspiracy”. In: *Proceedings of the Royal Society A* 476.2243 (2020), p. 20200214. DOI: 10.1098/rspa.2020.0214.

[Sve13] Bengt E Y Svensson. “What is a quantum-mechanical “weak value” the value of?” In: *Foundations of Physics* 43.10 (2013), pp. 1193–1205. DOI: 10.1007/s10701-013-9740-6.

[Sve15] Bengt E Y Svensson. "Non-representative quantum mechanical weak values". In: *Foundations of Physics* 45.12 (2015), pp. 1645–1656. DOI: 10.1007/s10701-015-9951-0.

[Sve17] Bengt E Y Svensson. "Response to Comment on 'Non-representative Quantum Mechanical Weak Values' by Ben-Israel and Vaidman". In: *Foundations of Physics* 47.9 (2017), pp. 1258–1260. DOI: 10.1007/s10701-017-0108-1.

[SW10] Ying Sun and Qiao-Yan Wen. "Counterfactual quantum key distribution with high efficiency". In: *Physical Review A* 82.5 (2010), p. 052318. DOI: 10.1103/PhysRevA.82.052318.

[SY17] Ya-Qi Song and Li Yang. "Quantum bit commitment protocol based on counterfactual quantum cryptography". In: *arXiv preprint arXiv:1709.08490* (2017). URL: https://arxiv.org/abs/1709.08490.

[TC13] Boaz Tamir and Eliahu Cohen. "Introduction to weak measurements and weak values". In: *Quanta* 2.1 (2013), pp. 7–17. DOI: 10.12743/quanta.v2i1.14.

[Tea22] Nature Physics Editorial Team. "Survey the foundations". In: *Nature Physics* 18.9 (Sept. 2022), pp. 961–961. ISSN: 1745-2481. DOI: 10.1038/s41567-022-01766-x.

[Tie+14] T. G. Tiecke et al. "Nanophotonic quantum phase switch with a single atom". In: *Nature* 508.7495 (Apr. 2014), pp. 241–244. ISSN: 0028-0836. DOI: 10.1038/nature13188.

[Tum06] Roderich Tumulka. "A Relativistic Version of the Ghirardi-Rimini-Weber Model". In: *Journal of Statistical Physics* 125.4 (Nov. 2006), pp. 821–840. ISSN: 1572-9613. DOI: 10.1007/s10955-006-9227-3.

[Tyo+06] J Scott Tyo et al. "Review of passive imaging polarimetry for remote sensing applications". In: *Applied optics* 45.22 (2006), pp. 5453–5469. DOI: 10.1364/AO.45.005453.

[Vai13a] Lev Vaidman. "Past of a quantum particle". In: *Phys. Rev. A* 87 (5 May 2013), p. 052104. DOI: 10.1103/PhysRevA.87.052104.

[Vai13b] Lev Vaidman. "Reply to "Comment on 'Past of a quantum particle' "". In: *Phys. Rev. A* 88 (4 Oct. 2013), p. 046103. DOI: 10.1103/PhysRevA.88.046103.

[Vai14a] Lev Vaidman. "Comment on "Protocol for Direct Counterfactual Quantum Communication"". In: *Phys. Rev. Lett.* 112 (20 May 2014), p. 208901. DOI: 10.1103/PhysRevLett.112.208901.

[Vai14b] Lev Vaidman. "Tracing the past of a quantum particle". In: *Phys. Rev. A* 89 (2 Feb. 2014), p. 024102. DOI: 10.1103/PhysRevA.89.024102.

[Vai15] Lev Vaidman. "Counterfactuality of 'counterfactual' communication". In: *Journal of Physics A: Mathematical and Theoretical* 48.46 (Oct. 2015), p. 465303. DOI: 10.1088/1751-8113/48/46/465303.

[Vai16a] L. Vaidman. "Comment on "Which-way information in a nested Mach-Zehnder interferometer"". In: *Phys. Rev. A* 93 (1 Jan. 2016), p. 017801. DOI: 10.1103/PhysRevA.93.017801.

[Vai16b] Lev Vaidman. ""Counterfactual" quantum protocols". In: *International Journal of Quantum Information* 14.04 (2016), p. 1640012. DOI: 10.1142/S0219749916400128.

[Vai16c] Lev Vaidman. "Comment on "Direct counterfactual transmission of a quantum state"". In: *Phys. Rev. A* 93 (6 June 2016), p. 066301. DOI: 10.1103/PhysRevA.93.066301.

[Vai16d] Lev Vaidman. "Past of a particle inside an interferometer: the two-state vector approach". In: *arXiv preprint arXiv:1610.04781* (2016). URL: https://arxiv.org/abs/1610.04781.

[Vai17] Lev Vaidman. "Comment on "Particle path through a nested Mach-Zehnder interferometer"". In: *Phys. Rev. A* 95 (6 June 2017), p. 066101. DOI: 10.1103/PhysRevA.95.066101.

[Vai19] Lev Vaidman. "Analysis of counterfactuality of counterfactual communication protocols". In: *Physical Review A* 99.5 (2019), p. 052127. DOI: 10.1103/PhysRevA.99.052127.

[Vai96] Lev Vaidman. "Weak-measurement elements of reality". In: *Foundations of Physics* 26.7 (1996), pp. 895–906. DOI: 10.1007/BF02148832.

[Val91a] Antony Valentini. "Signal-locality, uncertainty, and the subquantum H-theorem. I". In: *Physics Letters A* 156.1 (1991), pp. 5–11. ISSN: 0375-9601. DOI: 10.1016/0375-9601(91)90116-P.

[Val91b] Antony Valentini. "Signal-locality, uncertainty, and the subquantum H-theorem. II". In: *Physics Letters A* 158.1 (1991), pp. 1–8. ISSN: 0375-9601. DOI: 10.1016/0375-9601(91)90330-B.

[VT18] Lev Vaidman and Izumi Tsutsui. "When photons are lying about where they have been". In: *Entropy* 20.7 (2018), p. 538. DOI: 10.3390/e20070538.

[WA20] K. B. Wharton and N. Argaman. "Colloquium: Bell's theorem and locally mediated reformulations of quantum mechanics". In: *Rev. Mod. Phys.* 92 (2 May 2020), p. 021002. DOI: 10.1103/RevModPhys.92.021002.

[Wae+22] Mordecai Waegell et al. "Quantum reality with negative-mass particles". In: *arXiv preprint arXiv:2201.09510* (2022).

[Wag+23] Rafael Wagner et al. "Quantum circuits measuring weak values and Kirkwood-Dirac quasiprobability distributions, with applications". In: *arXiv preprint arXiv:2302.00705* (2023).

[Wal14] David Wallace. “Thermodynamics as Control Theory”. In: *Entropy* 16.2 (2014), pp. 699–725. ISSN: 1099-4300. DOI: 10.3390/e16020699.

[Wal22] David Wallace. “The sky is blue, and other reasons quantum mechanics is not underdetermined by evidence”. In: *arXiv preprint arXiv:2205.00568* (2022). DOI: 10.48550/arXiv.2205.00568.

[Wan+18] Jianwei Wang et al. “Multidimensional quantum entanglement with large-scale integrated optics”. In: *Science* 360.6386 (2018), pp. 285–291. ISSN: 0036-8075. DOI: 10.1126/science.aar7053.

[Wat55] Satosi Watanabe. “Symmetry of Physical Laws. Part III. Prediction and Retrodiction”. In: *Rev. Mod. Phys.* 27 (2 Apr. 1955), pp. 179–186. DOI: 10.1103/RevModPhys.27.179.

[WC17] Howard M. Wiseman and Eric G. Cavalcanti. “Causarum Investigatio and the Two Bell’s Theorems of John Bell”. In: *Quantum [Un]Speakables II: Half a Century of Bell’s Theorem*. Ed. by Reinhold Bertlmann and Anton Zeilinger. Cham: Springer International Publishing, 2017, pp. 119–142. ISBN: 978-3-319-38987-5. DOI: 10.1007/978-3-319-38987-5_6.

[WCV21] Alon Wander, Eliahu Cohen, and Lev Vaidman. “Three approaches for analyzing the counterfactuality of counterfactual protocols”. In: *Physical Review A* 104.1 (2021), p. 012610. DOI: 10.1103/PhysRevA.104.012610.

[Wee+12] Christian Weedbrook et al. “Gaussian quantum information”. In: *Reviews of Modern Physics* 84.2 (2012), p. 621. DOI: 10.1103/RevModPhys.84.621.

[Wei+98] Gregor Weihs et al. “Violation of Bell’s inequality under strict Einstein locality conditions”. In: *Physical Review Letters* 81.23 (1998), p. 5039. DOI: 10.1103/PhysRevLett.81.5039.

[Wei16] Steven Weinberg. *What’s the matter with quantum mechanics*. https://www.youtube.com/watch?v=3 CASW, Patrusky Lecture. 2016.

[Wei19] Zachary Alexander Weinersmith. “Duuude”. In: *Saturday Morning Breakfast Cereal (SMBC) Comics* (Sept. 2019). URL: https://www.smbc-comics.com/comic/duuude.

[Wei72] Steven Weinberg. *Gravitation and cosmology: principles and applications of the general theory of relativity*. New York: Wiley, 1972. ISBN: 0471925675.

[Whe78] John Archibald Wheeler. “The “Past” and the “Delayed-Choice” Double-Slit Experiment”. In: *Mathematical foundations of quantum theory*. Elsevier, 1978, pp. 9–48. DOI: 10.1016/B978-0-12-473250-6.50006-6.

[Whi+98] Andrew G. White et al. ““Interaction-free” imaging”. In: *Phys. Rev. A* 58 (1 July 1998), pp. 605–613. DOI: 10.1103/PhysRevA.58.605.

[Wie14] Marcin Wieśniak. "Criticism of" Asking Photons Where They Have Been"". In: *arXiv preprint arXiv:1407.1739* (2014). URL: https://arxiv.org/abs/1407.1739.

[Wig70] Eugene P. Wigner. "On Hidden Variables and Quantum Mechanical Probabilities". In: *American Journal of Physics* 38.8 (1970), pp. 1005–1009. DOI: 10.1119/1.1976526.

[Wit22] Stephen Witt. "The World-Changing Race to Develop the Quantum Computer". In: *The New Yorker* 2022 (19 Dec 2022). URL: https://www.newyorker.com/magazine/2022/12/19/the-world-changing-race-to-develop-the-quantum-computer.

[WLZ16] Tian-Yin Wang, Yan-Ping Li, and Rui-Ling Zhang. "Analysis of counterfactual quantum certificate authorization". In: *International Journal of Theoretical Physics* 55.12 (2016), pp. 5331–5335. DOI: 10.1007/s10773-016-3152-2.

[WS15] Christopher J Wood and Robert W Spekkens. "The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning". In: *New Journal of Physics* 17.3 (2015), p. 033002. DOI: 10.1088/1367-2630/17/3/033002.

[WS98] Christopher F Woodcock and Nigel P Smart. "p-adic chaos and random number generation". In: *Experimental Mathematics* 7.4 (1998), pp. 333–342. DOI: 10.1080/10586458.1998.10504379.

[WT10] David Wallace and Christopher G Timpson. "Quantum mechanics on spacetime I: Spacetime state realism". In: *The British journal for the philosophy of science* 61.4 (2010), pp. 697–727. DOI: 10.1093/bjps/axq010.

[WZ82] William K Wootters and Wojciech H Zurek. "A single quantum cannot be cloned". In: *Nature* 299.5886 (1982), pp. 802–803. DOI: 10.1038/299802a0.

[Yaf82] Laurence G Yaffe. "Large N limits as classical mechanics". In: *Reviews of Modern Physics* 54.2 (1982), p. 407. DOI: 10.1103/RevModPhys.54.407.

[Yan+16] Xiuqing Yang et al. "Trojan horse attacks on counterfactual quantum key distribution". In: *Physics Letters A* 380.18-19 (2016), pp. 1589–1592. DOI: 10.1016/j.physleta.2015.09.027.

[Yin+10] Zhen-Qiang Yin et al. "Security of counterfactual quantum cryptography". In: *Physical Review A* 82.4 (2010), p. 042335. DOI: 10.1103/PhysRevA.82.042335.

[Yin+12] Zhen-Qiang Yin et al. "Counterfactual quantum cryptography based on weak coherent states". In: *Physical Review A* 86.2 (2012), p. 022313. DOI: 10.1103/PhysRevA.86.022313.

[You04] Thomas Young. "The Bakerian Lecture: Experiments and Calculations Relative to Physical Optics". In: *Philosophical Transactions of the Royal Society of London Series I* 94 (1804), pp. 1–16. DOI: 10.1098/rstl.1804.0001.

[Zas81] George M Zaslavsky. "Stochasticity in quantum systems". In: *Physics Reports* 80.3 (1981), pp. 157–250. DOI: 10.1016/0370-1573(81)90127-7.

[Zha+13] Jia-Li Zhang et al. "Private database queries based on counterfactual quantum key distribution". In: *Phys. Rev. A* 88 (2 Aug. 2013), p. 022334. DOI: 10.1103/PhysRevA.88.022334.

[Zha+14] De-Jian Zhang et al. "Lensless ghost imaging of a phase object with pseudo-thermal light". In: *Appl. Phys. Lett.* 104.12 (2014), p. 121113. DOI: 10.1063/1.4869959.

[Zha+19] Yingwen Zhang et al. "Interaction-free ghost-imaging of structured objects". In: *Optics Express* 27.3 (2019), pp. 2212–2224. DOI: 10.1364/OE.27.002212.

[Zha17] Sheng Zhang. "Probabilistic direct counterfactual quantum communication". In: *Chinese Physics B* 26.2 (Feb. 2017), p. 020304. DOI: 10.1088/1674-1056/26/2/020304.

[Zho+17] Zong-Quan Zhou et al. "Experimental observation of anomalous trajectories of single photons". In: *Physical Review A* 95.4 (2017), p. 042121. DOI: 10.1103/PhysRevA.95.042121.

[ZJS18] Fakhar Zaman, Youngmin Jeong, and Hyundong Shin. "Counterfactual Bell-state analysis". In: *Scientific reports* 8.1 (2018), p. 14641. DOI: 10.1038/s41598-018-32928-8.

[ŻM21] Marek Żukowski and Marcin Markiewicz. "Physics and Metaphysics of Wigner's Friends: Even Performed Premeasurements Have No Results". In: *Physical Review Letters* 126.13 (2021), p. 130402. DOI: 10.1103/PhysRevLett.126.130402.

[Zub+14] Muhammad Suhail Zubairy et al. *Method and apparatus for direct counterfactual quantum communication*. US Patent 8,891,767. Nov. 2014.

[Zur03] Wojciech Hubert Zurek. "Decoherence, einselection, and the quantum origins of the classical". In: *Reviews of modern physics* 75.3 (2003), p. 715. DOI: 10.1103/RevModPhys.75.715.

[Zur18] Wojciech Hubert Zurek. "Quantum theory of the classical: quantum jumps, Born's Rule and objective classical reality via quantum Darwinism". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376.2123 (2018), p. 20180107. DOI: 10.1098/rsta.2018.0107.

[ZWT12a] Sheng Zhang, Jian Wang, and Chao-Jing Tang. "Counterfactual attack on counterfactual quantum key distribution". In: *EPL (Europhysics Letters)* 98.3 (May 2012), p. 30012. DOI: 10.1209/0295-5075/98/30012.

[ZWT12b] Sheng Zhang, Jian Wang, and Chao-Jing Tang. “Security proof of counterfactual quantum cryptography against general intercept-resend attacks and its vulnerability”. In: *Chinese Physics B* 21.6 (June 2012), p. 060303. DOI: 10.1088/1674-1056/21/6/060303.

[ZZL14] Sheng Zhang, Bo Zhang, and Xing-tong Liu. “Improved Direct Counterfactual Quantum Communication”. In: *arXiv preprint arXiv:1410.2769* (2014). URL: https://arxiv.org/abs/1410.2769.