A Cryptographic Test of Quantumness and Certifiable Randomness from a Single Quantum Device

Zvika Brakerski* Paul Christiano† Urmila Mahadev‡ Umesh Vazirani§
Thomas Vidick¶

Abstract

We give a protocol for producing certifiable randomness from a single untrusted quantum device that is polynomial-time bounded. The randomness is certified to be statistically close to uniform from the point of view of any computationally unbounded quantum adversary, that may share entanglement with the quantum device. The protocol relies on the existence of post-quantum secure trapdoor claw-free functions, and introduces a new primitive for constraining the power of an untrusted quantum device. We show how to construct this primitive based on the hardness of the learning with errors (LWE) problem, and prove that it has a crucial adaptive hardcore bit property. The randomness protocol can be used as the basis for an efficiently verifiable "test of quantumness", thus answering an outstanding challenge in the field.

*Weizmann Institute of Science, Israel. Email: zvika.brakerski@weizmann.ac.il.
†OpenAI, USA. Work performed while at UC Berkeley
‡UC Berkeley, USA. Email: mahadev@berkeley.edu
§UC Berkeley, USA. Email: vazirani@cs.berkeley.edu
¶California Institute of Technology, USA. Email: vidick@cms.caltech.edu

Contents

1 Introduction **3**

2 Preliminaries **10**
2.1 Notation . . . 10
2.2 Distributions . . . 10
2.3 The Learning with Errors problem . . . 11
2.4 Entropies . . . 13

3 Trapdoor claw-free hash functions **14**

4 A Trapdoor Claw-Free family based on LWE **16**
4.1 Efficient Function Generation . . . 17
4.2 Trapdoor Injective Pair . . . 17
4.3 Efficient Range Superposition . . . 18
4.4 Adaptive Hardcore Bit . . . 19
4.4.1 Moderate matrices . . . 19
4.4.2 LWE Hardcore bit . . . 21
4.4.3 Adaptive hardcore lemma . . . 23

5 Protocol description **25**
5.1 The randomness expansion protocol . . . 25
5.2 The simplified protocol . . . 26
5.3 Completeness . . . 26

6 Devices **29**
6.1 General devices . . . 29
6.2 Simplified devices . . . 30

7 Single-round analysis **31**
7.1 A constraint on the measurements of any efficient device . . . 32
7.2 Angles between incompatible measurements . . . 33
7.3 Simulating an efficient device using a simplified device . . . 34

8 Accumulating randomness across multiple rounds **35**
8.1 Reduction to the simplified protocol . . . 36
8.2 Randomness accumulation in the simplified protocol . . . 40
8.3 Randomness accumulation in the general protocol . . . 42

1 Introduction

In this paper we propose solutions to two basic tasks: how to generate *certifiably random* strings from a *single untrusted* quantum device (also referred to as a prover), and how to efficiently verify that an untrusted device is "truly" quantum, as opposed to classical. The setting we consider is one where the quantum device is polynomial-time bounded but untrusted, and the verifier is entirely classical and also polynomial-time bounded. The peculiarity of this setting is that it allows the verifier to leverage post-quantum cryptography, i.e. the existence of cryptographic primitives that can be implemented efficiently on a classical computer but that cannot be broken by any efficient quantum computer.

There has been considerable research into certifiable random number expansion from quantum devices [Col06, PAM+10, VV11, MS16, AFDF+18], including experimental demonstrations [BKG+18]. However, all prior works providing verifiable guarantees have focused on the setting where there are multiple quantum devices that share entanglement, and where the randomness certification relies on the violation of a Bell inequality. More generally, the violation of Bell inequalities provides a powerful technique for the classical testing of quantum devices — the obvious downside being that it is limited to situations with multiple non-communicating quantum devices that share entanglement. Here we propose a new setting for classical testing, where bounds on the computational power of a single quantum device are leveraged by a classical verifier. Specifically, in the context of certifiable random number expansion, the guarantee we seek is that provided the device is unable to break the post-quantum cryptographic assumption during the execution of the protocol, the output of the protocol must be statistically indistinguishable from a uniformly random sequence of bits to any *computationally unbounded* adversary that may share prior entanglement with the computationally bounded device. This information-theoretic guarantee, the same guarantee as that offered in the aforementioned works [VV11, MS16, AFDF+18], is stronger than computational pseudorandomness (that is easily achievable under standard cryptographic assumptions, since the verifier starts with a short uniformly random seed).

The specific cryptographic primitive we rely on is the existence of a post-quantum secure trapdoor claw-free (in short, TCF) family of functions $f : \{0,1\}^n \to \{0,1\}^m$, the post-quantum analogue of a notion introduced by Goldwasser, Micali and Rivest in the context of digital signatures [GMR84]. A TCF is a 2-to-1 function f that satisfies the following properties: $f(x)$ is efficiently computable on a classical computer, and if $f(x) = y$, then there is a unique $x' \neq x$ such that $f(x') = y$. Moreover, with knowledge of a secret trapdoor it is possible to efficiently (classically) compute x and x' from y, but without the trapdoor there is no efficient quantum algorithm that can compute such a claw, a triple (x, x', y), for any y.

By contrast, a quantum algorithm can simultaneously hold an image y, as well as a superposition $\frac{1}{\sqrt{2}}(|x\rangle + |x'\rangle)$ over two preimages of y, simply by evaluating f on a uniform superposition over all inputs and measuring the image y. At this point, measuring the quantum state in the standard basis will yield a random preimage, x or x'. This is not any stronger than a classical device could do, by first sampling a random x and then computing $y = f(x)$. However, the quantum device can do something different from directly measuring a preimage. Instead, the device can perform Fourier sampling (Hadamard transform followed by a measurement), which yields a random n-bit string d such that $d \cdot (x \oplus x') = 0$, thereby revealing some joint information about *both* preimages of y. From the point of view of the classical verifier a device that performs these tasks is modeled as an untrusted black box that outputs y, and then either a string x or a string d. Assuming the verifier knows the secret trapdoor, given y she can efficiently compute both preimages x and x' and verify that indeed $d \cdot (x \oplus x') = 0$. At a high level, the consideration of a cryptographic primitive equipped with a trapdoor restores some symmetry between the quantum prover (the untrusted quantum device) and the classical verifier, by providing a primitive which allows the quantum

capabilities of the prover to play a useful role while at the same time giving the classical verifier a handle, namely the ability to compute both x and x', that the prover does not have access to.

A test of quantumness. We refer to the protocol that consists of first, requesting a value y in the range of f, and then, executing either the preimage test (receive a string x and check that $f(x) = y$), or the equation test (receive a string d and check, using the trapdoor, that $d \cdot (x \oplus x') = 0$, where x, x' are the two preimages of y), each chosen with probability $\frac{1}{2}$, as the *single round test*.

The ability to succeed in the single round test seems to be a unique quantum capability. To argue that no classical procedure could succeed in the test, we describe a TCF construction based on the learning with errors problem (LWE) that has additional security guarantees, most notably the "adaptive hardcore bit" property that is explained below. For clarity, in this paper we refer to the specific kind of TCF that we rely on as an NTCF, or post-quantum secure noisy trapdoor claw-free family. We give a construction of an NTCF that rests on the hardness of the Learning with Errors (LWE) problem, introduced by Regev [Reg05], with slightly super-polynomial noise ratio, against quantum polynomial-time attacks with nonuniform quantum advice (for which the state of the art classical and quantum attacks scale exponentially with the dimension). This construction is similar to the one used in [Mah17a], albeit with some changes in parameters that allows us to prove the following crucial adaptive hardcore bit property: roughly, that given the public parameters of the NTCF it is computationally intractable to sample from any distribution on quadruples (y, x, d, b) such that both conditions $f(x) = y$ and $b = d \cdot (x \oplus x')$ hold with probability non-negligibly larger than $\frac{1}{2}$. Very informally, this says that the condition that no efficient quantum algorithm can exhibit a claw (x, x', y) for the NTCF can be greatly strengthened to assert that no efficient quantum algorithm can even exhibit (x, y) and a nontrivial advantage in guessing any generalized bit of x' of its choice — where a generalized bit is $d \cdot x'$ for any choice of d. Note that this is much stronger than a standard hardcore bit, which would assert intractability only for fixed d.

Assuming such an adaptive hardcore bit property, it is possible to show that passing the single round test constitutes a proof of quantumness of the device. This is because any efficient classical algorithm that can reliably succeed in both the preimage test and the equation test could be "rewound" to simultaneously answer both challenges, thus contradicting the adaptive hardcore bit property.

This result has implications for an important milestone in the experimental realization of quantum computers, namely "quantum supremacy": a proof that an (untrusted) quantum computing device performs some computational task that cannot be solved classically without impractical resources. While this could in principle be achieved by demonstrating quantum factoring, the latter requires quantum resources well beyond the capability of near term experiments. Instead current proposals are based on sampling problems (see e.g. [HM17] for a recent survey). The major challenge for these proposals is verifying that the quantum computer did indeed sample from the desired probability distribution, and all existing proposals rely on exponential time classical algorithms for verification. By contrast, our single round test provides a proof of quantumness that can be verified by a classical verifier in polynomial time. This proposal seems promising from a practical viewpoint — indeed, even using off-the-shelf bounds for LWE-based cryptography suggests that a protocol providing 50 bits of security could be implemented with a quantum device of around 2000 qubits (see e.g. [LP11]). It would be worth exploring whether there are clever implementations of this scheme that can lead to a protocol in the $200 - 500$ qubit range. Our protocol is robust to a device that only successfully answers the verifier's challenges with a sufficiently large, but constant, success probability; it would be interesting to explore whether such a device could be implemented without resorting to general fault-tolerance techniques.

Certifiable randomness. The challenge in achieving certifiable randomness lies in using computational assumptions to establish not pseudorandomness, but rather that the output of the protocol must be (close to) statistically random. The goal of the protocol we sketch below is to leverage the properties of the NTCF to characterize the quantum state and measurements of the untrusted quantum device — essentially showing that it must have a qubit initialized in state $|+\rangle$, which it measured in the standard basis, thus generating statistical randomness. This is the analogue of the use of the violation of Bell inequalities to characterize the state of the device in entanglement-based testing.

We first explain how to show a device that succeeds in the single round test must generate randomness. In the test the device must make one of two measurements: either a "preimage" measurement, or an "equation" measurement. We focus on a single bit of information provided by each measurement. The "preimage" measurement can be treated as a projection into one of two orthogonal subspaces corresponding to the two preimages x, x' for the element y that the device has returned to the verifier. The "equation" measurement can similarly be coarse-grained into a projection on one of two orthogonal subspaces, "valid" or "invalid", i.e. the subspace that corresponds to all measurement outcomes d such that $d \cdot (x \oplus x') = 0$, or the subspace associated with outcomes d such that $d \cdot (x \oplus x') = 1$.

Applying Jordan's lemma, it is possible to decompose the device's Hilbert space into a direct sum of one- and two-dimensional subspaces, such that within each two-dimensional subspace the "preimage" and "equation" measurements each correspond to an orthonormal basis, such that the two bases make a certain angle with each other. We argue that almost all angles must be very close to $\pi/4$. Indeed, whenever the angles are *not* near-maximally unbiased, it is possible to show that by considering the effect of performing the measurements in sequence, one can devise an "attack" on the NTCF of a kind that contradicts the adaptive hardcore bit property of the NTCF — informally, the attack can simultaneously produce a valid preimage and a valid equation, with non-negligible advantage.

As a result it is possible to show that the state and (coarse-grained) measurements of the device are, up to a global change of basis, close to the following: the device starts with a qubit initialized to $|+\rangle$, which it measures in the standard basis for the case of a preimage test and in the Hadamard basis for the case of an equation test. The fact that an efficient quantum device cannot break the cryptographic assumption has thus been translated into a characterization of the state and actions of the quantum device, which further implies that the output of the device in the single round test must contain close to a bit of true (information theoretic) randomness.

One might further conjecture that for a generic TCF (e.g. modeled as a random oracle), if the output of any efficient quantum device passes the single round test with non-negligible advantage over $\frac{1}{2}$, then the pair y, d returned in the equation test must have high min-entropy. Such a strong statement would immediately yield a randomness certification protocol. Among the many difficulties in showing such a statement is that both y and d may be adaptively and adversarially chosen — in the single round protocol above this issue is addressed by the adaptive hardcore bit property of the NTCF.

Outline of randomness generation protocol. Going beyond the analysis of the single round test requires significantly more work. So far we have argued that if an efficient quantum algorithm has the ability to generate a valid equation with probability sufficiently close to 1, then, if instead it is asked for a preimage, this preimage must be close to uniformly distributed over the two possibilities. To leverage this our randomness expansion protocol proceeds in multiple rounds, repeatedly asking for new images y and a preimage of y (to generate randomness) while inserting a few randomly located equation tests to test the device. Each time an "equation" challenge has been answered, we refresh the pseudorandom keys used for the NTCF. This is required to avoid a simple "attack" by the device, which would repeatedly use the same y, preimage x, and

guessed equation d — succeeding in the protocol with probability $\frac{1}{2}$ without generating any randomness.

Let's call the sequence of rounds with a particular set of pseudorandom keys an epoch. Intuitively, we would like to claim that if the device passes all the equation tests, then for most epochs and for most rounds within that epoch, the state of the device and its measurements must be (close to) as characterized above: it starts with a qubit initialized to $|+\rangle$, which it measures in the standard basis for the case of a preimage test, and in the Hadamard basis for the case of an equation test. To show this we would like to claim that if the device passes all the equation tests, for most such tests it must produce a valid equation with probability close to 1. Since each equation test occurs at a random round in the epoch, it should follow from the adaptive hardcore bit property that the sequence of bits that the verifier extracts from the device's answers to preimage tests during that epoch must look statistically random. We give a martingale-based argument to formalize this intuition.

There is however a bigger challenge to analyzing the protocol — we must show that the sequence that the verifier extracts from the device's answers to preimage tests must look statistically random even to an infinitely powerful quantum adversary, who may share an arbitrary entangled state with the quantum device. If we could assert that each round of the protocol is played with a qubit exactly in state $|+\rangle$, and measured in the standard basis basis for the case of a preimage test, then this would lead to an easy proof that the extracted sequence looks random to the adversary. Unfortunately the characterization of the device's qubits leaves plenty of room for entanglement with the adversary. Showing that such entanglement cannot leak too much information about the device's measurements was the major challenge in previous work on certified randomness through Bell inequality violations [VV11, MS14, AFDF$^+$18]. Our cryptographic setting presents a new difficulty, which is that in contrast to the Bell inequality violation scenarios, in our setting it is not *impossible* for a deterministic device to succeed in the test: it is merely *computationally hard* to do so. This prevents us from directly applying the results in [MS14, AFDF$^+$18] and requires us to suitably modify their framework. We describe this part of the argument in more detail below.

In terms of efficiency, for the specific LWE-based NTCF that we construct, our protocol can use as few as $\mathrm{poly}\log(N)$ bits of randomness to generate $O(N)$ bits that are statistically within negligible distance from uniform. However, this requires assuming that the underlying LWE assumption is hard even for sub-exponential size quantum circuits with polynomial-size quantum advice (which is consistent with current knowledge). The more conservative assumption that our variant of LWE is only hard for polynomial size quantum circuits requires $O(N^\epsilon)$ bits of randomness for generating the NTCF, for any constant $\epsilon > 0$. The following is an informal description; see Theorem 8.10 for a more formal statement.

Theorem 1.1 (Informal)**.** *Let $\mathcal{F}$ be an NTCF family and λ a security parameter. Let $N = \Omega(\lambda^2)$ and assume the quantum hardness of solving lattice problems of dimension λ in time $\mathrm{poly}(N)$. There is an N-round protocol for the interaction between a classical polynomial-time verifier and a quantum polynomial-time device such that the protocol can be executed using $\mathrm{poly}(\log(N), \lambda)$ bits of randomness, and for any efficient device and side information E correlated with the device's initial state,*

$$H_\infty^\delta(O|CE)_{\overline{\rho}} \geq (\xi - o(1))N \ .$$

Here ξ is a positive constant, δ is a negligible function of λ, and $\overline{\rho}$ is the final state of the classical output register O, the classical register C containing the verifier's messages to the device, and the side information E, restricted to transcripts that are accepted by the verifier in the protocol.

Sketch of the security analysis. We describe the protocol in slightly more detail (see Section 5 for a formal description). The verifier first uses $\mathrm{poly}(\log(N), \lambda)$ bits of randomness to select a pair of functions

$\{f_{k,b}\}_{b\in\{0,1\}}$ from an NTCF family, and sends the public function key k to the quantum device. This pair of functions can be interpreted as a single 2-to-1 function $f_k : (b, x) \mapsto f_{k,b}(x)$. The verifier keeps private the trapdoor information that allows to invert f_k. The protocol then proceeds for N rounds. In each round the device first outputs a value y in the common range of $f_{k,0}$ and $f_{k,1}$. After having received y, the verifier issues one of two challenges: 0 or 1, preimage or equation. If the challenge is "preimage", then the device must output an x such that $f(x) = y$. If the challenge is "equation" then the device must output a nontrivial binary vector d such that $d \cdot (x_0 \oplus x_1) = 0$, where x_0 and x_1 are the unique preimages of y under $f_{k,0}$ and $f_{k,1}$ respectively. Since the verifier has the secret key, she can efficiently compute x_0 and x_1 from y, and therefore check the correctness of the device's response to each challenge. The verifier chooses $\operatorname{poly}\log(N)$ rounds in which to issue the challenge 1, or "equation", at random. Selecting these rounds requires only $\operatorname{poly}\log(N)$ random bits. At the end of each such round, the verifier samples a new pair of functions from the NTCF family, and communicates the new public key to the device. On each of the remaining $N - \operatorname{poly}\log(N)$ rounds the verifier records a bit according to whether the device returns the preimage x_0, or x_1 (e.g. recording 0 for the lexicographically smaller preimage). At the end of the protocol the verifier uses a strong quantum-proof randomness extractor to extract $\Omega(N)$ bits of randomness from the recorded string (this requires at most an additional $\operatorname{poly}\log(N)$ bits of uniformly random seed).

To guarantee that the extractor produces bits that are statistically close to uniform, we would like to prove that the $N - \operatorname{poly}\log(N)$ random bits recorded by the verifier must have $\Omega(N)$ bits of (smoothed) min-entropy,[1] even conditioned on the side information available to an infinitely powerful quantum adversary, who may share an arbitrary entangled state with the quantum device.

The analysis proceeds as follows. First we assume without loss of generality that the entire protocol is run coherently, i.e. we may assume that the initial state of the quantum device (holding quantum register D) and the adversary (holding quantum register E) is a pure state $|\phi\rangle_{\mathsf{DE}}$, since the adversary may as well start with a purification of their joint state. We may also assume that the verifier starts with a cat state on $\operatorname{poly}\log(N)$ qubits, and uses one of the registers of the state, C, to provide the random bits used to select the type of test being performed in each round. (This is for the sake of analysis only, the actual verifier is of course completely classical.) We can similarly arrange that the state remains pure throughout the protocol by using the principle of deferred measurement. Our goal is to show a lower bound on the smooth min-entropy of the output register O in which the verifier has recorded the device's outputs, conditioned on the state E of the adversary, and on the register C of the cat state (conditioning on the latter represents the fact that the verifier's choice of challenges may be leaked to the adversary, and we would like security even in this scenario). Intuitively, this amounts to bounding the information accessible to the most powerful adversary quantum mechanics allows, conditioned on the joint state of the verifier and device.

In order to bound the entropy of the final state we need to show that the entropy "accumulates" at each round of the protocol. A general framework to establish entropy accumulation in quantum protocols such as the one considered here was introduced in [AFDF+18]. At a high level, the approach consists in reducing the goal of a min-entropy bound to a bound on the appropriate notion of $(1+\varepsilon)$ quantum conditional Rényi entropy, and then arguing that, under suitable conditions on the process that generates the outcomes recorded in the protocol, entropy accumulates sequentially throughout the protocol.

In a little more detail, the first step on getting a handle on the smooth min-entropy is to use the quantum asymptotic equipartion property (QAEP) [TCR09] to relate it to the $(1+\varepsilon)$ conditional Rényi entropy, for suitably small ε. The second step uses a duality relation for the conditional Rényi entropy to relate the $(1+\varepsilon)$ conditional Rényi entropy of the output register O, conditioned on the adversary side information in R and the register C of the cat state, to a quantity analogous to the $(1-\varepsilon')$ conditional Rényi entropy

[1] We refer to Section 2 for definitions of entropic quantities.

of the output register, conditioned on the register E for the device, and a purifying copy of the register C of the cat state. The latter quantity, a suitable conditional entropy of the output register conditioned on the challenge register and the state of the device, is the quantity that we ultimately aim to bound. Note what these transformations have achieved for us: it is now sufficient to consider as side information only "known" quantities in the protocol, the verifier's choice of challenges and the device's state; the information held by the adversary plays no other role than that of a purifying register.

As mentioned earlier, our cryptographic setting presents the additional difficulty that our guarantee is only that it is computationally hard for a deterministic device to succeed in the protocol. The results in [AFDF$^+$18, MS14] crucially rely on the fact that the process that generates the randomness does so irrespective of the quantum state in which it is initialized (as long as the output of the process satisfies the test's success criterion). This requirement comes from the conditioning that is performed in order to show that entropy accumulates; in our setting, conditioning is more delicate as it can in principle induce non-computationally efficient states for the device.

Recall that we argued that for a single round of the protocol, we can decompose the device's Hilbert space into a direct sum of one- or two-dimensional subspaces, such that within most two-dimensional subspace the "preimage" and "equation" measurements correspond to orthonormal bases that make an angle close to $\pi/4$ with each other. Showing that the Rényi entropy accumulates in each round requires a device in which *all* angles are close to $\pi/4$, not "almost all". To accommodate for this we "split" the state of the device into its component on the good subspace, where the angles are unbiased, and the bad subspace, where the measurements may be aligned. The fact that the distinction between good and bad subspace is not measured in the protocol, but is only a distinction made for the analysis, requires us to apply a fairly delicate martingale based argument that takes into account possible interference effects and bounds those "branches" where the state has gone through the bad subspace an improbably large number of times. Whenever the state lies in the good susbpace, we can appeal to an uncertainty principle from [MS14] to show that the device's measurement increases the conditional Rényi entropy of the output register by a small additive constant. Pursuing this approach across all N rounds, we obtain a linear lower bound on the conditional Rényi entropy of the output register, conditioned on the state of the device. As argued above this in turn translates into a linear lower bound on the smooth conditional min-entropy of the output, conditioned on the state of the adversary and the verifier's choice of challenges. It only remains to apply a quantum-proof randomness extractor to the output, using a poly-logarithmic number of additional bits of randomness, to obtain the final result.

Our NTCF Family. Our goal is to construct a family of pairs of injective functions f_0, f_1 with the same image such that it is hard to find a collision x_0, x_1 with $f_0(x_0) = f_1(x_1)$, but so that given a suitable trapdoor it is possible to recover, for any y, values x_0, x_1 such that $f_0(x_0) = f_1(x_1) = y$. We do this by relying on the hardness of the Learning with Errors (LWE) problem [Reg05]. LWE states that given a public uniformly random matrix $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ for $m \gg n$, it is intractable to distinguish between $\mathbf{u} = \mathbf{A}\mathbf{s} + \mathbf{e} \pmod q$ and a uniform vector, for a uniform vector $\mathbf{s}$ and small discrete Gaussian vector $\mathbf{e}$ (all arithmetic from here on is performed modulo q; we use $\oplus$ to denote binary XOR). Inspired by [Mah17a], our function pair will be characterized by $(\mathbf{A}, \mathbf{u} = \mathbf{A}\mathbf{s} + \mathbf{e})$, but for a *binary* vector $\mathbf{s}$.[2] The trapdoor for our function is a lattice trapdoor for $\mathbf{A}$ that allows to recover $\mathbf{s}, \mathbf{e}$ given a vector of the form $\mathbf{A}\mathbf{s} + \mathbf{e}$ (it is possible to generate $\mathbf{A}$ together with a trapdoor such that $\mathbf{A}$ is indistinguishable from uniform, as originally shown by Ajtai [Ajt99]).

[2] It is known that LWE is hard even with binary secrets. We do not use this property explicitly but rather employ the respective techniques in our proof.

The structure of the LWE problem motivates us to consider functions f_0, f_1 that range over probability distributions. Specifically, we define the distribution $f_b(\mathbf{x})$ as $f_b(\mathbf{x}) = \mathbf{A}\mathbf{x} + b \cdot (\mathbf{A}\mathbf{s}) + \mathbf{e}'$ where $\mathbf{e}'$ is a discrete Gaussian random variable with a sufficiently wide Gaussian parameter. Observe that the functions have overlapping images in the following sense: $f_0(\mathbf{x}) = f_1(\mathbf{x} - \mathbf{s})$. Moreover, since $\mathbf{A}\mathbf{s} + \mathbf{e}'$ is statistically indistinguishable from $\mathbf{u} + \mathbf{e}'$, we can efficiently sample from the distribution $f_b(\mathbf{x})$ up to negligible statistical distance.

This probabilistic notion complicates the definition and use of the family, but the principles are similar to the deterministic version.[3]

Adaptive Hardcore Bit. Finally, we need to show the adaptive hardcore bit property. Formally, letting $\mathbf{x}_0, \mathbf{x}_1 \in \mathbb{Z}_q^n$ be a collision, and letting $z_0, z_1 \in \{0,1\}^{n \log q}$ be their binary representations, respectively, we need to show that it is intractable to come up with a pair (b, z_b), for some $b \in \{0,1\}$, together with a nontrivial vector d and with the value $d \cdot (z_0 \oplus z_1)$, with probability noticeably better than $\frac{1}{2}$. "Nontrivial" here means belonging to a well defined and efficiently recognizable set D with density ≈ 1 in $\{0,1\}^{n \log q}$ (e.g. the zero vector is obviously excluded). Assume for the sake of this overview that we get a tuple (z_0, d, c). We first notice that since $\mathbf{x}_0, \mathbf{x}_1$ is a collision, then $\mathbf{x}_1 = \mathbf{x}_0 - \mathbf{s} \pmod q$. We now use the fact that $\mathbf{s}$ is a binary vector to show, using simple arithmetic, that $z_0 \oplus z_1$ can be expressed as a linear function of the bits of $\mathbf{s}$, so that $d \cdot (z_0 \oplus z_1) = \hat{d} \cdot \mathbf{s} \pmod 2$, for some $\hat{d} \in \{0,1\}^n$. (the description of this transformation will effect our choice of the set D). We thus need to show that it is intractable, given the instance $(\mathbf{A}, \mathbf{u} = \mathbf{A}\mathbf{s} + \mathbf{e})$, to come up with $\hat{d}, \hat{d} \cdot \mathbf{s} \pmod 2$. To prove this we use the lossiness technique used in [GKPV10] and show that this is equivalent to coming up with $\hat{d}, \hat{d} \cdot \mathbf{s} \pmod 2$ given $\mathbf{B}, \mathbf{B}\mathbf{s}$ where $\mathbf{B} \in \mathbb{Z}_q^{k \times n}$ is now a highly shrinking function, even for binary inputs, i.e. $k \log q \ll n$. This seems like an easy task since the adversary now doesn't have the complete information about $\mathbf{s}$ so it shouldn't be able to compute $\hat{d} \cdot \mathbf{s} \pmod 2$ for any reasonable $\hat{d}$, except $\hat{d}$ might depend on $\mathbf{B}$ itself (recall that $\hat{d}$ is chosen adversarially). We prove via Fourier analysis that if $\mathbf{B}$ is sufficiently shrinking, then there is no $\hat{d}$ that can take advantage of the dependence on $\mathbf{B}$, which completes the proof.

Concurrent and related work. The idea of using a TCF as a basic primitive in interactions between an efficient quantum prover and a classical verifier has been further developed in recent work by Mahadev [Mah17b], giving the first construction of a quantum fully homomorphic encryption scheme with classical keys. In further follow-up work, Mahadev [Mah18] shows a remarkable use of a NTCF family with adaptive hardcore bit. Namely, that the NTCF can be used to certify that a prover measures a qubit in a prescribed basis (standard or Hadamard). This allows to achieve single prover *verifiability* for quantum computations using a purely classical verifier (but relying on computational assumptions).

Independently of this work, a construction of trapdoor one-way functions with second preimage resistance based on LWE was recently introduced in [CCKW18], where it is used to achieve delegated computation in the weaker honest-but-curious model for the adversary (i.e. without soundness against provers not following the protocol). The family of functions considered in [CCKW18] is not sufficient for our purposes, as it lacks the adaptive hardcore bit property.

We believe that the technique of constraining the power of a quantum device using NTCFs promises to be a powerful tool for the field of untrusted quantum devices.

[3] Another possible variant is to define $f_b(\mathbf{x}) = \lfloor \mathbf{A}\mathbf{x} + b \cdot (\mathbf{A}\mathbf{s}) \rceil$ where $\lfloor \cdot \rceil$ is a rounding function that truncates "many" of the least significant bits of its operand. However, we remain with the Gaussian variant which is easier to analyze.

Organization. We start with some notation and preliminaries in Section 2. Section 3 contains the definition of a noisy trapdoor claw-free family (NTCF). Our construction for such a family is given in Section 4 (with Appendix 2.3 containing relevant preliminaries on the learning with errors problem). The randomness generation protocol is described in Section 5. In Section 6 we introduce our formalism for modeling the actions of an arbitrary prover, or device, in the protocol. In Section 7 we analyze a single round of the protocol, and in Section 8 we show that randomness accumulates across multiple rounds.

Acknowledgments. Zvika Brakerski is supported by the Israel Science Foundation (Grant No. 468/14), Binational Science Foundation (Grants No. 2016726, 2014276), and by the European Union Horizon 2020 Research and Innovation Program via ERC Project REACT (Grant 756482) and via Project PROMETHEUS (Grant 780701). Paul Christiano and Urmila Mahadev are supported by a Templeton Foundation Grant 52536, ARO Grant W911NF-12-1-0541, and NSF Grant CCF-1410022. Umesh Vazirani is supported by Templeton Foundation Grant 52536, ARO Grant W911NF-12-1-0541, NSF Grant CCF-1410022, MURI Grant FA9550-18-1-0161 and a Vannevar Bush Faculty Fellowship. Thomas Vidick is supported by NSF CAREER Grant CCF-1553477, AFOSR YIP award number FA9550-16-1-0495, MURI Grant FA9550-18-1-0161, a CIFAR Azrieli Global Scholar award, and the IQIM, an NSF Physics Frontiers Center (NSF Grant PHY-1125565) with support of the Gordon and Betty Moore Foundation (GBMF-12500028).

2 Preliminaries

2.1 Notation

$\mathbb{Z}$ is the set of integers, and $\mathbb{N}$ the set of natural numbers. For any $q \in \mathbb{N}$ such that $q \geq 2$ we let $\mathbb{Z}_q$ denote the ring of integers modulo q. We generally identify an element $x \in \mathbb{Z}_q$ with its unique representative $[x]_q \in (-\frac{q}{2}, \frac{q}{2}] \cap \mathbb{Z}$. For $x \in \mathbb{Z}_q$ we define $|x| = |[x]_q|$. When considering an $s \in \{0,1\}^n$ we sometimes also think of s as an element of $\mathbb{Z}_q^n$, in which case we write it as $\mathbf{s}$.

We use the terminology of polynomially bounded and negligible functions. A function $n : \mathbb{N} \to \mathbb{R}_+$ is *polynomially bounded* if there exists a polynomial p such that $n(\lambda) \leq p(\lambda)$ for all $\lambda \in \mathbb{N}$. A function $n : \mathbb{N} \to \mathbb{R}_+$ is *negligible* if for every polynomial p, $p(\lambda)n(\lambda) \to_{\lambda\to\infty} 0$. We write $\mathrm{negl}(\lambda)$ to denote an arbitrary negligible function of λ. For two parameters κ, λ we write $\kappa \ll \lambda$ to express the constraint that κ should be "sufficiently smaller than" λ, meaning that there exists a small universal constant $c > 0$ such that $\kappa \leq c\lambda$, where c is usually implicit for context.

$\mathcal{H}$ always denotes a finite-dimensional Hilbert space. We use indices $\mathcal{H}_\mathsf{A}$, $\mathcal{H}_\mathsf{B}$, etc., to refer to distinct spaces. $\mathrm{Pos}(\mathcal{H})$ is the set of positive semidefinite operators on $\mathcal{H}$, and $\mathrm{D}(\mathcal{H})$ the set of density matrices, i.e. the positive semidefinite operators with trace 1. For an operator X on $\mathcal{H}$, we use $\|X\|$ to denote the operator norm (largest singular value) of X, and $\|X\|_{tr} = \frac{1}{2}\|X\|_1 = \frac{1}{2}\mathrm{Tr}\sqrt{XX^\dagger}$ for the trace norm.

2.2 Distributions

We generally use the letter D to denote a distribution over a finite domain X, and f for a density on X, i.e. a function $f : X \to [0,1]$ such that $\sum_{x\in X} f(x) = 1$. We often use the distribution and its density interchangeably. We write U for the uniform distribution. We write $x \leftarrow D$ to indicate that x is sampled from distribution D, and $x \leftarrow_U X$ to indicate that x is sampled uniformly from the set X. We write $\mathcal{D}_X$ for the set of all densities on X. For any $f \in \mathcal{D}_X$, $\textsc{Supp}(f)$ denotes the support of f,

$$\textsc{Supp}(f) = \{x \in X \mid f(x) > 0\}.$$

For two densities f_1 and f_2 over the same finite domain X, the Hellinger distance between f_1 and f_2 is

$$H^2(f_1, f_2) \;=\; 1 - \sum_{x \in X} \sqrt{f_1(x) f_2(x)}\;. \tag{1}$$

The total variation distance between f_1 and f_2 is

$$\|f_1 - f_2\|_{TV} \;=\; \frac{1}{2} \sum_{x \in X} |f_1(x) - f_2(x)| \;\leq\; \sqrt{2H^2(f_1, f_2)}\;. \tag{2}$$

The following immediate lemma relates the Hellinger distance and the trace distance of superpositions.

Lemma 2.1. *Let X be a finite set and $f_1, f_2 \in \mathcal{D}_X$. Let*

$$|\psi_1\rangle = \sum_{x \in X} \sqrt{f_1(x)}|x\rangle \qquad \textit{and} \qquad |\psi_2\rangle = \sum_{x \in X} \sqrt{f_2(x)}|x\rangle\;.$$

Then

$$\||\psi_1\rangle - |\psi_2\rangle\|_{tr} \;=\; \sqrt{1 - (1 - H^2(f_1, f_2))^2}\;.$$

We say that a family of quantum circuits $\{C_\lambda\}_{\lambda \in \mathbb{N}}$ is *polynomial-time generated* if there exists a polynomial-time deterministic Turing machine that, on every input $\lambda \in \mathbb{N}$, returns a gate-by-gate encoding of the circuit C_λ. We introduce a notion of efficient distinguishability between distributions.

Definition 2.2. We say that two families of distributions $D_0 = \{D_{0,\lambda}\}_{\lambda \in \mathbb{N}}$ and $D_1 = \{D_{1,\lambda}\}_{\lambda \in \mathbb{N}}$ on the same finite set $\{X_\lambda\}$ are *computationally indistinguishable* if for every polynomial-time generated family of quantum circuits $\mathcal{A} = \{A_\lambda : X_\lambda \to \{0,1\}\}$ it holds that

$$\Big| \Pr_{x \leftarrow D_{0,\lambda}}[A_\lambda(x) = 0] - \Pr_{x \leftarrow D_{1,\lambda}}[A_\lambda(x) = 0] \Big| \;=\; \mathrm{negl}(\lambda)\;. \tag{3}$$

The next definition generalizes the previous one to the case of quantum states.

Definition 2.3. We say that two families of sub-normalized densities $\sigma_0 = \{\sigma_{0,\lambda}\}_{\lambda \in \mathbb{N}}$ and $\sigma_1 = \{\sigma_{1,\lambda}\}_{\lambda \in \mathbb{N}}$ on the same Hilbert space $\{\mathcal{H}_\lambda\}$ are *computationally indistinguishable* if for every polynomial-time generated family of observables $O = \{O_\lambda\}_{\lambda \in \mathbb{N}}$ it holds that

$$\big|\mathrm{Tr}\big(O_\lambda(\sigma_{0,\lambda} - \sigma_{1,\lambda})\big)\big| \;=\; \mathrm{negl}(\lambda)\;.$$

2.3 The Learning with Errors problem

We give some background on the Learning with Errors problem (LWE). For a positive real B and a positive integer q, the truncated discrete Gaussian distribution over $\mathbb{Z}_q$ with parameter B is the distribution supported on $\{x \in \mathbb{Z}_q : \|x\| \leq B\}$ with density

$$D_{\mathbb{Z}_q,B}(x) \;=\; \frac{e^{\frac{-\pi\|x\|^2}{B^2}}}{\sum\limits_{x \in \mathbb{Z}_q,\, \|x\| \leq B} e^{\frac{-\pi\|x\|^2}{B^2}}}\;. \tag{4}$$

More generally, for a positive integer m the truncated discrete Gaussian distribution over $\mathbb{Z}_q^m$ with parameter B is the distribution supported on $\{x \in \mathbb{Z}_q^m : \|x\| \leq B\sqrt{m}\}$ with density

$$\forall x = (x_1, \ldots, x_m) \in \mathbb{Z}_q^m\;, \qquad D_{\mathbb{Z}_q^m,B}(x) \;=\; D_{\mathbb{Z}_q,B}(x_1) \cdots D_{\mathbb{Z}_q,B}(x_m)\;. \tag{5}$$

Lemma 2.4. *Let B be a positive real and q, m positive integers. Let $\mathbf{e} \in \mathbb{Z}_q^m$. The Hellinger distance between the distribution $D = D_{\mathbb{Z}_q^m,B}$ and the shifted distribution $D + \mathbf{e}$, with density $(D+\mathbf{e})(x) = D(x-\mathbf{e})$, satisfies*

$$H^2(D, D+\mathbf{e}) \leq 1 - e^{\frac{-2\pi\sqrt{m}\|\mathbf{e}\|}{B}}, \tag{6}$$

and the statistical distance between the two distributions satisfies

$$\left\|D - (D+\mathbf{e})\right\|_{TV}^2 \leq 2\left(1 - e^{\frac{-2\pi\sqrt{m}\|\mathbf{e}\|}{B}}\right). \tag{7}$$

Proof. Let $\tau = \sum_{x \in \mathbb{Z}_q,\, \|x\| \leq B} e^{\frac{-\pi\|x\|^2}{B^2}}$. We can compute

$$\begin{aligned}
\sum_{\mathbf{e}_0 \in \mathbb{Z}_q^m} \sqrt{D_{\mathbb{Z}_q^m,B}(\mathbf{e}_0) D_{\mathbb{Z}_q^m,B}(\mathbf{e}_0 - \mathbf{e})} &= \sum_{\mathbf{e}_0 \in \mathbb{Z}_q^m} \sqrt{D_B(\mathbf{e}) D_B(\mathbf{e}_0 - \mathbf{e})} \\
&= \frac{1}{\tau^m} \sum_{\mathbf{e} \in \mathbb{Z}_q^m} e^{\frac{-\pi(\|\mathbf{e}_0\|^2 + \|\mathbf{e}_0 - \mathbf{e}\|^2)}{2B^2}} \\
&\geq \frac{1}{\tau^m} \sum_{\mathbf{e}_0 \in \mathbb{Z}_q^m} e^{\frac{-\pi(\|\mathbf{e}_0\|^2 + (\|\mathbf{e}_0\| + \|\mathbf{e}\|)^2)}{2B^2}} \\
&= \frac{1}{\tau^m} \sum_{\mathbf{e}_0 \in \mathbb{Z}_q^m} e^{\frac{-\pi(\|\mathbf{e}_0\|^2)}{B^2}} e^{\frac{-\pi(2\|\mathbf{e}_0\|\|\mathbf{e}\|)}{2B^2}} e^{\frac{-\pi(\|\mathbf{e}\|^2)}{2B^2}} \\
&\geq e^{\frac{-\pi(\|\mathbf{e}\|^2 + 2\|\mathbf{e}_0\|\|\mathbf{e}\|)}{2B^2}} \frac{1}{\tau^m} \sum_{\mathbf{e}_0 \in \mathbb{Z}_q^m} e^{\frac{-\pi(\|\mathbf{e}_0\|)^2}{B^2}} \\
&= e^{\frac{-\pi(\|\mathbf{e}\|^2 + 2\|\mathbf{e}_0\|\|\mathbf{e}\|)}{2B^2}} \\
&\geq e^{\frac{-2\pi\|\mathbf{e}_0\|\|\mathbf{e}\|}{B^2}}.
\end{aligned}$$

Using the fact that for any $\mathbf{e}_0$ in the support of $D_{\mathbb{Z}_q^m,B}$, $\|\mathbf{e}_0\| \leq B\sqrt{m}$, gives the claimed bound. The bound on the statistical distance follows from the bound on the Hellinger distance using the inequality in (2). □

We define the main assumption that underlies all computational hardness claims made in the paper.

Definition 2.5. For a security parameter λ, let $n, m, q \in \mathbb{N}$ be integer functions of λ. Let $\chi = \chi(\lambda)$ be a distribution over $\mathbb{Z}$. The $\mathrm{LWE}_{n,m,q,\chi}$ problem is to distinguish between the distributions $(\mathbf{A}, \mathbf{As} + \mathbf{e} \pmod q)$ and $(\mathbf{A}, \mathbf{u})$, where $\mathbf{A} \leftarrow_U \mathbb{Z}_q^{n \times m}$, $\mathbf{s} \leftarrow_U \mathbb{Z}_q^n$, $\mathbf{e} \leftarrow \chi^m$, and $\mathbf{u} \leftarrow_U \mathbb{Z}_q^m$. Often we consider the hardness of solving LWE for any function m such that m is at most a polynomial in $n \log q$. This problem is denoted $\mathrm{LWE}_{n,q,\chi}$.

In this paper we make the assumption that no quantum polynomial-time procedure can solve the $\mathrm{LWE}_{n,q,\chi}$ problem with more than a negligible advantage in λ, even when given access to a quantum polynomial-size advice state depending on the parameters n, m, q and χ of the problem. We refer to this assumption as "the $\mathrm{LWE}_{n,q,\chi}$ assumption".

As shown in [Reg05, PRS17], for any $\alpha > 0$ such that $\sigma = \alpha q \geq 2\sqrt{n}$ the $\mathrm{LWE}_{n,q,D_{\mathbb{Z}_q,\sigma}}$ problem, where $D_{\mathbb{Z}_q,\sigma}$ is the discrete Gaussian distribution, is at least as hard as approximating the shortest independent vector problem (SIVP) to within a factor of $\gamma = \widetilde{O}(n/\alpha)$, where $\tilde{O}$ hides factors logarithmic in the argument,

in *worst case* dimension n lattices. This is proven using a quantum reduction. Classical reductions (to a slightly different problem) exist as well [Pei09, BLP+13] but with somewhat worse parameters. The best known (classical or quantum) algorithm for these problems run in time $2^{\widetilde{O}(n/\log\gamma)}$. For our construction, given in Section 4, we assume hardness of the problem against a quantum polynomial-time adversary in the case that γ is a super polynomial function in n. This is a commonly used assumption in cryptography (for e.g. homomorphic encryption schemes such as [GSW13]).

We use two additional properties of the LWE problem. The first is that it is possible to generate LWE samples $(\mathbf{A}, \mathbf{As}+\mathbf{e})$ such that there is a trapdoor allowing recovery of $\mathbf{s}$ from the samples.

Theorem 2.6 (Theorem 5.1 in [MP12])**.** *Let $n, m \geq 1$ and $q \geq 2$ be such that $m = \Omega(n \log q)$. There is an efficient randomized algorithm* $\text{GENTRAP}(1^n, 1^m, q)$ *that returns a matrix $\mathbf{A} \in \mathbb{Z}_q^{m\times n}$ and a trapdoor $t_{\mathbf{A}}$ such that the distribution of $\mathbf{A}$ is negligibly (in n) close to the uniform distribution. Moreover, there is an efficient algorithm* INVERT *that, on input $\mathbf{A}, t_{\mathbf{A}}$ and $\mathbf{As}+\mathbf{e}$ where $\|\mathbf{e}\| \leq q/(C_T\sqrt{n\log q})$ and C_T is a universal constant, returns $\mathbf{s}$ and $\mathbf{e}$ with overwhelming probability over $(\mathbf{A}, t_{\mathbf{A}}) \leftarrow \text{GENTRAP}(1^n, 1^m, q)$.*

The second property is the existence of a "lossy mode" for LWE. The following definition is Definition 3.1 in [AKPW13].

Definition 2.7. Let $\chi = \chi(\lambda)$ be an efficiently sampleable distribution over $\mathbb{Z}_q$. Define a lossy sampler $\tilde{\mathbf{A}} \leftarrow \text{LOSSY}(1^n, 1^m, 1^\ell, q, \chi)$ by $\tilde{\mathbf{A}} = \mathbf{BC} + \mathbf{F}$, where $\mathbf{B} \leftarrow_U \mathbb{Z}_q^{m\times\ell}$, $\mathbf{C} \leftarrow_U \mathbb{Z}_q^{\ell\times n}$, $\mathbf{F} \leftarrow \chi^{m\times n}$.

Theorem 2.8 (Lemma 3.2 in [AKPW13])**.** *Under the* $\text{LWE}_{\ell,q,\chi}$ *assumption, the distribution of a random $\tilde{\mathbf{A}} \leftarrow \text{LOSSY}(1^n, 1^m, 1^\ell, q, \chi)$ is computationally indistinguishable from $\mathbf{A} \leftarrow_U \mathbb{Z}_q^{m\times n}$.*

2.4 Entropies

For $p \in [0,1]$ we write $H(p) = -p\log p - (1-p)\log(1-p)$ for the binary Shannon entropy. We measure randomness using Rényi conditional entropies. For a positive semidefinite matrix $\sigma \in \text{Pos}(\mathcal{H})$ and $\varepsilon \geq 0$, let

$$\langle\sigma\rangle_{1+\varepsilon} = \text{Tr}\big(\sigma^{1+\varepsilon}\big) .$$

This quantity satisfies the following approximate linearity relations:

$$\forall\varepsilon \in [0,1] , \qquad \langle\sigma\rangle_{1+\varepsilon} + \langle\tau\rangle_{1+\varepsilon} \leq \langle\sigma+\tau\rangle_{1+\varepsilon} \leq \big(1+O(\varepsilon)\big)\big(\langle\sigma\rangle_{1+\varepsilon} + \langle\tau\rangle_{1+\varepsilon}\big) . \tag{8}$$

In addition, for positive semidefinite $\sigma, \rho \in \text{Pos}(\mathcal{H})$ such that the support of ρ is included in the support of σ, and $\varepsilon \geq 0$, let

$$\tilde{Q}_{1+\varepsilon}(\rho\|\sigma) = \langle\sigma^{-\frac{\varepsilon}{2(1+\varepsilon)}}\rho\sigma^{-\frac{\varepsilon}{2(1+\varepsilon)}}\rangle_{1+\varepsilon} . \tag{9}$$

Quantum analogues of the conditional Rényi entropies can be defined as follows.

Definition 2.9. Let $\rho_{\mathsf{AB}} \in \text{Pos}(\mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}})$ be positive semidefinite. Given $\varepsilon > 0$, the $(1+\varepsilon)$ *Rényi entropy* of A conditioned on B is defined as

$$H_{1+\varepsilon}(A|B)_\rho = \sup_{\sigma\in\text{D}(\mathcal{H}_{\mathsf{B}})} H_{1+\varepsilon}(A|B)_{\rho|\sigma} ,$$

where for any $\sigma_{\mathsf{B}} \in \text{D}(\mathcal{H}_{\mathsf{B}})$,

$$H_{1+\varepsilon}(A|B)_{\rho|\sigma} = -\frac{1}{\varepsilon}\log\tilde{Q}_{1+\varepsilon}(\rho\|\sigma) .$$

.

Rényi entropies are used in the proofs because they have better "chain-rule-like" properties than the min-entropy, which is the most appropriate measure for randomness quantification.

Definition 2.10. Let $\rho_{\mathsf{AB}} \in \mathrm{Pos}(\mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}})$ be positive semidefinite. Given a density matrix the *min-entropy* of A conditioned on B is defined as

$$H_\infty(A|B)_\rho \;=\; \sup_{\sigma \in \mathrm{D}(\mathcal{H}_{\mathsf{B}})} H_\infty(A|B)_{\rho|\sigma}\,,$$

where for any $\sigma_{\mathsf{B}} \in \mathrm{D}(\mathcal{H}_{\mathsf{B}})$,

$$H_\infty(A|B)_{\rho|\sigma} \;=\; \max\left\{\lambda \geq 0 \,|\, 2^{-\lambda}\,\mathrm{Id}_A \otimes \sigma_B \geq \rho_{AB}\right\}\,.$$

It is often convenient to consider the *smooth* min-entropy, which is obtained by maximizing the min-entropy over all positive semidefinite operators matrices in an ε-neighborhood of ρ_{AB}. The definition of neighborhood depends on a choice of metric; the canonical choice is the "purified distance". Since this choice will not matter for us we defer to [Tom15] for a precise definition.

Definition 2.11. Let $\varepsilon \geq 0$ and $\rho_{\mathsf{AB}} \in \mathrm{Pos}(\mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}})$ positive semidefinite. The *ε-smooth min-entropy* of A conditioned on B is defined as

$$H^\varepsilon_\infty(A|B)_\rho \;=\; \sup_{\sigma_{\mathsf{AB}} \in \mathcal{B}(\rho_{\mathsf{AB}},\varepsilon)} H_\infty(A|B)_\sigma\,,$$

where $\mathcal{B}(\rho_{\mathsf{AB}},\varepsilon)$ is the ball of radius ε around ρ_{AB}, taken with respect to the purified distance.

The following theorem relates the min-entropy to the the Rényi entropies introduced earlier. The theorem expresses the fact that, up to a small amount of "smoothing" (the parameter δ in the theorem), all these entropies are of similar order.

Theorem 2.12 (Theorem 4.1 [MS14])**.** *Let $\rho_{\mathsf{XE}} \in \mathrm{Pos}(\mathcal{H}_{\mathsf{X}} \otimes \mathcal{H}_{\mathsf{E}})$ be positive semidefinite of the form $\rho_{\mathsf{XE}} = \sum_{x\in\mathcal{X}} |x\rangle\langle x| \otimes \rho^x_{\mathsf{E}}$, where $\mathcal{X}$ is a finite alphabet. Let $\sigma_{\mathsf{E}} \in \mathrm{D}(\mathcal{H}_{\mathsf{E}})$ be an arbitrary density matrix. Then for any $\delta > 0$ and $0 < \varepsilon \leq 1$,*

$$H^\delta_\infty(X|E)_\rho \;\geq\; -\frac{1}{\varepsilon}\log\Big(\sum_x \tilde{Q}_{1+\varepsilon}(\rho^x_{\mathsf{E}}\|\sigma_{\mathsf{E}})\Big) - \frac{1+2\log(1/\delta)}{\varepsilon}\,.$$

3 Trapdoor claw-free hash functions

Let λ be a security parameter, and $\mathcal{X}$ and $\mathcal{Y}$ finite sets (depending on λ). For our purposes an ideal family of functions $\mathcal{F}$ would have the following properties. For each public key k, there are two functions $\{f_{k,b} : \mathcal{X} \to \mathcal{Y}\}_{b\in\{0,1\}}$ that are both injective and have the same range, and are invertible given a suitable trapdoor t_k (i.e. t_k can be used to compute x given b and $y = f_{k,b}(x)$). Furthermore, the pair of functions should be claw-free: it must be hard for an attacker to find two pre-images $x_0, x_1 \in \mathcal{X}$ such that $f_{k,0}(x_0) = f_{k,1}(x_1)$. Finally, the functions should satisfy an adaptive hardcore bit property, which is a stronger form of the claw-free property: assuming for convenience that $\mathcal{X} = \{0,1\}^w$, we would like that it is computationally infeasible to simultaneously generate a pair $(b, x_b) \in \{0,1\} \times \mathcal{X}$ and a $d \in \{0,1\}^w \setminus \{0^w\}$ such that with non-negligible advantage over $\frac{1}{2}$ the equation $d \cdot (x_0 \oplus x_1) = 0$, where x_{1-b} is defined as the unique element such that $f_{k,1-b}(x_{1-b}) = f_{k,b}(x_b)$, holds.

Unfortunately, we do not know how to construct a function family that exactly satisfies all these requirements under standard cryptographic assumptions. Instead, we construct a family that satisfies slightly relaxed requirements, that we will show still suffice for our purposes, based on the hardness of the learning with errors problem introduced in Section 2.3. The requirements are relaxed as follows. First, the range of the functions is no longer a set $\mathcal{Y}$; instead, it is $\mathcal{D}_\mathcal{Y}$, the set of probability densities over $\mathcal{Y}$. That is, each function returns a density, rather than a point. The trapdoor injective pair property is then described in terms of the support of the output densities: these supports should either be identical, for a colliding pair, or be disjoint, in all other cases.

The consideration of functions that return densities gives rise to an additional requirement of efficiency: there should exist a quantum polynomial-time procedure that efficiently prepares a superposition over the range of the function, i.e. for any key k and $b \in \{0,1\}$, the procedure can prepare the state

$$\frac{1}{\sqrt{\mathcal{X}}} \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \sqrt{\big(f_{k,b}(x)\big)(y)}\,|x\rangle|y\rangle \ . \tag{10}$$

In our instantiation based on LWE, it is not possible to prepare (10) perfectly, but it is possible to create a superposition with coefficients $\sqrt{(f'_{k,b}(x))(y)}$, such that the resulting state is within negligible trace distance of (10). The density $f'_{k,b}(x)$ is required to satisfy two properties used in our protocol. First, it must be easy to check, without the trapdoor, if an $y \in \mathcal{Y}$ lies in the support of $f'_{k,b}(x)$. Second, the inversion algorithm should operate correctly on all y in the support of $f'_{k,b}(x)$.

We slightly modify the adaptive hardcore bit requirement as well. Since the set $\mathcal{X}$ may not be a subset of binary strings, we first assume the existence of an injective, efficiently invertible map $J : \mathcal{X} \to \{0,1\}^w$. Next, we only require the adaptive hardcore bit property to hold for a subset of all nonzero strings, instead of the set $\{0,1\}^w \setminus \{0^w\}$. Finally, membership in the appropriate set should be efficiently checkable, given access to the trapdoor.

A formal definition follows.

Definition 3.1 (NTCF family)**.** Let λ be a security parameter. Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets. Let $\mathcal{K}_\mathcal{F}$ be a finite set of keys. A family of functions

$$\mathcal{F} \ = \ \big\{f_{k,b} : \mathcal{X} \to \mathcal{D}_\mathcal{Y}\big\}_{k \in \mathcal{K}_\mathcal{F}, b \in \{0,1\}}$$

is called a *noisy trapdoor claw free (NTCF) family* if the following conditions hold:

1. **Efficient Function Generation.** There exists an efficient probabilistic algorithm $\text{GEN}_\mathcal{F}$ which generates a key $k \in \mathcal{K}_\mathcal{F}$ together with a trapdoor t_k:

$$(k, t_k) \leftarrow \text{GEN}_\mathcal{F}(1^\lambda) \ .$$

2. **Trapdoor Injective Pair.** For all keys $k \in \mathcal{K}_\mathcal{F}$ the following conditions hold.

   (a) *Trapdoor*: For all $b \in \{0,1\}$ and $x \neq x' \in \mathcal{X}$, $\text{SUPP}(f_{k,b}(x)) \cap \text{SUPP}(f_{k,b}(x')) = \emptyset$. Moreover, there exists an efficient deterministic algorithm $\text{INV}_\mathcal{F}$ such that for all $b \in \{0,1\}$, $x \in \mathcal{X}$ and $y \in \text{SUPP}(f_{k,b}(x))$, $\text{INV}_\mathcal{F}(t_k, b, y) = x$.

   (b) *Injective pair*: There exists a perfect matching $\mathcal{R}_k \subseteq \mathcal{X} \times \mathcal{X}$ such that $f_{k,0}(x_0) = f_{k,1}(x_1)$ if and only if $(x_0, x_1) \in \mathcal{R}_k$.

3. **Efficient Range Superposition.** For all keys $k \in \mathcal{K}_{\mathcal{F}}$ and $b \in \{0,1\}$ there exists a function $f'_{k,b} : \mathcal{X} \to \mathcal{D}_{\mathcal{Y}}$ such that the following hold.

   (a) For all $(x_0, x_1) \in \mathcal{R}_k$ and $y \in \text{SUPP}(f'_{k,b}(x_b))$, $\text{INV}_{\mathcal{F}}(t_k, b, y) = x_b$ and $\text{INV}_{\mathcal{F}}(t_k, b \oplus 1, y) = x_{b\oplus 1}$.

   (b) There exists an efficient deterministic procedure $\text{CHK}_{\mathcal{F}}$ that, on input k, $b \in \{0,1\}$, $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, returns 1 if $y \in \text{SUPP}(f'_{k,b}(x))$ and 0 otherwise. Note that $\text{CHK}_{\mathcal{F}}$ is not provided the trapdoor t_k.

   (c) For every k and $b \in \{0,1\}$,

   $$\text{E}_{x \leftarrow_U \mathcal{X}} \left[H^2(f_{k,b}(x),\, f'_{k,b}(x)) \right] \;\leq\; \mu(\lambda) \;,$$

   for some negligible function $\mu(\cdot)$. Here H^2 is the Hellinger distance; see (1). Moreover, there exists an efficient procedure $\text{SAMP}_{\mathcal{F}}$ that on input k and $b \in \{0,1\}$ prepares the state

   $$\frac{1}{\sqrt{|\mathcal{X}|}} \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \sqrt{(f'_{k,b}(x))(y)}\, |x\rangle |y\rangle \;. \tag{11}$$

4. **Adaptive Hardcore Bit.** For all keys $k \in \mathcal{K}_{\mathcal{F}}$ the following conditions hold, for some integer w that is a polynomially bounded function of λ.

   (a) For all $b \in \{0,1\}$ and $x \in \mathcal{X}$, there exists a set $G_{k,b,x} \subseteq \{0,1\}^w$ such that $\Pr_{d \leftarrow_U \{0,1\}^w}[d \notin G_{k,b,x}]$ is negligible, and moreover there exists an efficient algorithm that checks for membership in $G_{k,b,x}$ given k, b, x and the trapdoor t_k.

   (b) There is an efficiently computable injection $J : \mathcal{X} \to \{0,1\}^w$, such that J can be inverted efficiently on its range, and such that the following holds. If

   $$\begin{aligned} H_k &= \left\{ (b, x_b, d, d \cdot (J(x_0) \oplus J(x_1))) \mid b \in \{0,1\},\ (x_0, x_1) \in \mathcal{R}_k,\ d \in G_{k,0,x_0} \cap G_{k,1,x_1} \right\}, \text{[4]} \\ \overline{H}_k &= \left\{ (b, x_b, d, c) \mid (b, x, d, c \oplus 1) \in H_k \right\}, \end{aligned}$$

   then for any quantum polynomial-time procedure $\mathcal{A}$ there exists a negligible function $\mu(\cdot)$ such that

   $$\left| \Pr_{(k,t_k) \leftarrow \text{GEN}_{\mathcal{F}}(1^\lambda)} [\mathcal{A}(k) \in H_k] - \Pr_{(k,t_k) \leftarrow \text{GEN}_{\mathcal{F}}(1^\lambda)} [\mathcal{A}(k) \in \overline{H}_k] \right| \;\leq\; \mu(\lambda) \;. \tag{12}$$

4 A Trapdoor Claw-Free family based on LWE

In this section we present our LWE-based construction of an NTCF. For LWE-related preliminaries and definitions see Section 2.3. Let λ be a security parameter. All other parameters are functions of λ. Let $q \geq 2$ be a prime. Let $\ell, n, m, w \geq 1$ be polynomially bounded functions of λ and B_L, B_V, B_P be positive

[4] Note that although both x_0 and x_1 are referred to to define the set H_k, only one of them, x_b, is explicitly specified in any 4-tuple that lies in H_k.

integers such that the following conditions hold:

$$\begin{aligned}
&1.\ n = \Omega(\ell \log q) \text{ and } m = \Omega(n \log q),\\
&2.\ w = n\lceil \log q \rceil,\\
&3.\ B_P = \frac{q}{2C_T\sqrt{mn\log q}}, \text{ for } C_T \text{ the universal constant in Theorem 2.6,}\\
&4.\ 2\sqrt{n} \le B_L < B_V < B_P,\\
&5.\ \text{The ratio } \tfrac{B_P}{B_V} \text{ and } \tfrac{B_V}{B_L} \text{ are both super-polynomial in } \lambda.
\end{aligned} \tag{13}$$

Given a choice of parameters satisfying all conditions in (13), we describe the function family $\mathcal{F}_{\mathrm{LWE}}$. Let $\mathcal{X} = \mathbb{Z}_q^n$ and $\mathcal{Y} = \mathbb{Z}_q^m$. The key space $\mathcal{K}_{\mathcal{F}_{\mathrm{LWE}}}$ is a subset of $\mathbb{Z}_q^{m\times n} \times \mathbb{Z}_q^m$ defined in Section 4.1. For $b \in \{0,1\}$, $x \in \mathcal{X}$ and key $k = (\mathbf{A}, \mathbf{As}+\mathbf{e})$, the density $f_{k,b}(x)$ is defined as

$$\forall y \in \mathcal{Y}, \quad (f_{k,b}(x))(y) = D_{\mathbb{Z}_q^m, B_P}(y - \mathbf{A}x - b\cdot \mathbf{As}) \ , \tag{14}$$

where the density $D_{\mathbb{Z}_q^m,B_P}$ is defined in (4). It follows from the definition of the key generation procedure $\mathrm{GEN}_{\mathcal{F}_{\mathrm{LWE}}}$ given in Section 4.1 that $f_{k,b}$ is well-defined given $k = (\mathbf{A}, \mathbf{As}+\mathbf{e})$, as for our choice of parameters k uniquely identifies s.

The four properties required for a noisy trapdoor claw-free family, as specified in Definition 3.1, are verified in the following subsections, providing a proof of the following theorem. Recall the definition of the hardness assumption $\mathrm{LWE}_{n,q,\chi}$ given in Definition 2.5.

Theorem 4.1. *For any choice of parameters satisfying the conditions* (13)*, the function family* $\mathcal{F}_{\mathrm{LWE}}$ *is a noisy trapdoor claw free family under the hardness assumption* $\mathrm{LWE}_{\ell,q,D_{\mathbb{Z}_q,B_L}}$.

4.1 Efficient Function Generation

$\mathrm{GEN}_{\mathcal{F}_{\mathrm{LWE}}}$ is defined as follows. First, the procedure samples a random $\mathbf{A} \in \mathbb{Z}_q^{m\times n}$, together with trapdoor information $t_{\mathbf{A}}$. This is done using the procedure $\textsc{GenTrap}(1^n, 1^m, q)$ from Theorem 2.6. The trapdoor allows the evaluation of an inversion algorithm $\textsc{Invert}$ that, on input $\mathbf{A}$, $t_{\mathbf{A}}$ and $b = \mathbf{As}+\mathbf{e}$ returns $\mathbf{s}$ and $\mathbf{e}$ as long as $\|\mathbf{e}\| \le \frac{q}{C_T\sqrt{n\log q}}$. Moreover, the distribution on matrices $\mathbf{A}$ returned by $\textsc{GenTrap}$ is negligibly close to the uniform distribution on $\mathbb{Z}_q^{m\times n}$.

Next, the sampling procedure selects $s \in \{0,1\}^n$ uniformly at random, and a vector $\mathbf{e} \in \mathbb{Z}_q^m$ by sampling each coordinate independently according to the distribution $D_{\mathbb{Z}_q,B_V}$ defined in (4). $\mathrm{GEN}_{\mathcal{F}_{\mathrm{LWE}}}$ returns $k = (\mathbf{A}, \mathbf{As}+\mathbf{e})$ and $t_k = t_{\mathbf{A}}$.

4.2 Trapdoor Injective Pair

(a) *Trapdoor.* It follows from (14) and the definition of the distribution $D_{\mathbb{Z}_q^m,B_P}$ in (4) that for any key $k = (\mathbf{A}, \mathbf{As}+\mathbf{e}) \in \mathcal{K}_{\mathcal{F}_{\mathrm{LWE}}}$ and for all $x \in \mathcal{X}$,

$$\textsc{Supp}(f_{k,0}(x)) = \left\{\mathbf{A}x + \mathbf{e}_0 \mid \|\mathbf{e}_0\| \le B_P\sqrt{m}\right\} , \tag{15}$$

$$\textsc{Supp}(f_{k,1}(x)) = \left\{\mathbf{A}x + \mathbf{As} + \mathbf{e}_0 \mid \|\mathbf{e}_0\| \le B_P\sqrt{m}\right\} . \tag{16}$$

The procedure $\mathrm{INV}_{\mathcal{F}_{\mathrm{LWE}}}$ takes as input the trapdoor $t_{\mathbf{A}}$, $b \in \{0,1\}$, and $y \in \mathcal{Y}$. It uses the algorithm INVERT to determine $\mathbf{s}_0, \mathbf{e}_0$ such that $y = \mathbf{A}\mathbf{s}_0 + \mathbf{e}_0$, and returns the element $\mathbf{s}_0 - b \cdot \mathbf{s} \in \mathcal{X}$. Using Theorem 2.6, this procedure returns the unique correct outcome provided $y = \mathbf{A}\mathbf{s}_0 + \mathbf{e}_0$ for some $\mathbf{e}_0$ such that $\|\mathbf{e}_0\| \leq \frac{q}{C_T\sqrt{n \log q}}$. This condition is satisfied for all $y \in \mathrm{SUPP}(f_{k,b}(x))$ provided B_P is chosen so that

$$B_P \leq \frac{q}{C_T\sqrt{mn \log q}} \,. \tag{17}$$

(b) *Injective Pair.* We let $\mathcal{R}_k$ be the set of all pairs (x_0, x_1) such that $f_{k,0}(x_0) = f_{k,1}(x_1)$. By definition this occurs if and only if $x_1 = x_0 - \mathbf{s}$, and so $\mathcal{R}_k$ is a perfect matching.

4.3 Efficient Range Superposition

For $k = (\mathbf{A}, \mathbf{A}\mathbf{s} + \mathbf{e}) \in \mathcal{K}_{\mathcal{F}_{\mathrm{LWE}}}, b \in \{0,1\}$ and $x \in \mathcal{X}$, let

$$(f'_{k,b}(x))(y) \;=\; D_{\mathbb{Z}_q^m, B_P}(y - \mathbf{A}x - b \cdot (\mathbf{A}\mathbf{s} + \mathbf{e})) \,. \tag{18}$$

Note that $f'_{k,0}(x) = f_{k,0}(x)$ for all $x \in \mathcal{X}$. The distributions $f'_{k,1}(x)$ and $f_{k,1}(x)$ are shifted by $\mathbf{e}$. Given the key k and $x \in \mathcal{X}$, the densities $f'_{k,0}(x)$ and $f'_{k,1}(x)$ are efficiently computable. For all $x \in \mathcal{X}$,

$$\mathrm{SUPP}(f'_{k,0}(x)) \;=\; \mathrm{SUPP}(f_{k,0}(x)) \,, \tag{19}$$

$$\mathrm{SUPP}(f'_{k,1}(x)) \;=\; \left\{\mathbf{A}x + \mathbf{e}_0 + \mathbf{A}\mathbf{s} + \mathbf{e} \;\middle|\; \|\mathbf{e}_0\| \leq B_P\sqrt{m}\right\} . \tag{20}$$

(a) Using that $B_V < B_P$, it follows that the norm of the term $\mathbf{e}_0 + \mathbf{e}$ in (20) is always at most $2B_P\sqrt{m}$. Therefore, the inversion procedure $\mathrm{INV}_{\mathcal{F}_{\mathrm{LWE}}}$ can be guaranteed to return x on input $t_{\mathbf{A}}$, $b \in \{0,1\}$, $y \in \mathrm{SUPP}(f'_{k,b}(x))$ if we strengthen the requirement on B_P given in (17) to

$$B_P \;\leq\; \frac{q}{2C_T\sqrt{mn \log q}} \,. \tag{21}$$

This strengthened trapdoor requirement also implies that for all $b \in \{0,1\}$, $(x_0, x_1) \in \mathcal{R}_k$, and $y \in \mathrm{SUPP}(f'_{k,b}(x_b))$, $\mathrm{INV}_{\mathcal{F}_{\mathrm{LWE}}}(t_{\mathbf{A}}, b \oplus 1, y) = x_{b\oplus 1}$.

(b) On input $k = (\mathbf{A}, \mathbf{A}\mathbf{s} + \mathbf{e})$, $b \in \{0,1\}$, $x \in \mathcal{X}$, and $y \in \mathcal{Y}$, the procedure $\mathrm{CHK}_{\mathcal{F}_{\mathrm{LWE}}}$ operates as follows. If $b = 0$, it computes $\mathbf{e}' = y - \mathbf{A}x$. If $\|\mathbf{e}'\| \leq B_P\sqrt{m}$, the procedure returns 1, and 0 otherwise. If $b = 1$, it computes $\mathbf{e}' = y - \mathbf{A}x - (\mathbf{A}\mathbf{s} + \mathbf{e})$. If $\|\mathbf{e}'\| \leq B_P\sqrt{m}$, it returns 1, and 0 otherwise.

(c) We bound the Hellinger distance between the densities $f_{k,b}(x)$ and $f'_{k,b}(x)$. If $b = 0$ they are identical. If $b = 1$, both densities are shifts of $D_{\mathbb{Z}_q^m, B_P}$, where the shifts differ by $\mathbf{e}$. Each coordinate of $\mathbf{e}$ is drawn independently from $D_{\mathbb{Z}_q, B_V}$, so $\|\mathbf{e}\| \leq \sqrt{m}B_V$. Applying Lemma 2.4, we get that

$$H^2(f_{k,1}(x), f'_{k,1}(x)) \;\leq\; 1 - e^{\frac{-2\pi m B_V}{B_P}} \,.$$

Using the assumption that B_P/B_V is super-polynomial, this is negligible, as desired. It remains to describe the procedure $\mathrm{SAMP}_{\mathcal{F}_{\mathrm{LWE}}}$. At the first step, the procedure creates the following superposition

$$\sum_{\mathbf{e}_0 \in \mathbb{Z}_q^m} \sqrt{D_{\mathbb{Z}_q^m, B_P}(\mathbf{e}_0)}|\mathbf{e}_0\rangle \,. \tag{22}$$

This state can be prepared efficiently as described in [Reg05, Lemma 3.12].[5]

At the second step, the procedure creates a uniform superposition over $x \in \mathcal{X}$, yielding the state

$$q^{-\frac{n}{2}} \sum_{\substack{x \in \mathcal{X} \\ \mathbf{e}_0 \in \mathbb{Z}_q^m}} \sqrt{D_{\mathbb{Z}_q^m, B_P}(\mathbf{e}_0)} |x\rangle |\mathbf{e}_0\rangle \ . \tag{23}$$

At the third step, using the key $k = (\mathbf{A}, \mathbf{As} + \mathbf{e})$ and the input bit b the procedure computes

$$q^{-\frac{n}{2}} \sum_{\substack{x \in \mathcal{X} \\ \mathbf{e}_0 \in \mathbb{Z}_q^m}} \sqrt{D_{\mathbb{Z}_q^m, B_P}(\mathbf{e}_0)} |x\rangle |\mathbf{e}_0\rangle |\mathbf{A}x + \mathbf{e}_0 + b \cdot (\mathbf{As} + \mathbf{e})\rangle \ . \tag{24}$$

At this point, observe that $\mathbf{e}_0$ can be computed from x, the last register, b and the key k. The procedure can then uncompute the register containing $\mathbf{e}_0$, yielding

$$\begin{aligned} & q^{-\frac{n}{2}} \sum_{\substack{x \in \mathcal{X} \\ \mathbf{e}_0 \in \mathbb{Z}_q^m}} \sqrt{D_{\mathbb{Z}_q^m, B_P}}(\mathbf{e}_0) |x\rangle |\mathbf{A}x + \mathbf{e}_0 + b \cdot (\mathbf{As} + \mathbf{e})\rangle \\ & \quad = q^{-\frac{n}{2}} \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \sqrt{D_{\mathbb{Z}_q^m, B_P}(y - \mathbf{A}x - b \cdot (\mathbf{As} + \mathbf{e}))} |x\rangle |y\rangle \\ & \quad = q^{-\frac{n}{2}} \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \sqrt{(f'_{k,b}(x))(y)} |x\rangle |y\rangle \ . \end{aligned} \tag{25}$$

4.4 Adaptive Hardcore Bit

This section is devoted to the proof of the adaptive hardcore bit condition. The main statement is provided in Lemma 4.7 in Section 4.4.3. The proof of the lemma proceeds in three steps. First, in Section 4.4.1 we establish some preliminary results on the distribution of the inner product $(\hat{d} \cdot s \bmod 2)$, where $\hat{d} \in \{0,1\}^n$ is a fixed nonzero binary vector and $s \leftarrow_U \{0,1\}^n$ a uniformly random binary vector, conditioned on $\mathbf{Cs} = \mathbf{v}$ for some randomly chosen matrix $\mathbf{C} \in \mathbb{Z}_q^{\ell \times n}$ and arbitrary $\mathbf{v} \in \mathbb{Z}_q^{\ell}$. This condition is combined with the LWE assumption in Section 4.4.2 to argue that $(\hat{d} \cdot s \bmod 2)$ remains computationally indistinguishable from uniform even when the matrix $\mathbf{C}$ is an LWE matrix $\mathbf{A}$, and the adversary is able to choose $\hat{d}$ after being given access to $\mathbf{As} + \mathbf{e}$ for some error vector $\mathbf{e} \in \mathbb{Z}_q^m$. Finally, in Section 4.4.3 the required hardcore bit condition is reduced to the one established in Section 4.4.2 by relating the inner product appearing in the definition of H_k (in condition 4.(b) of Definition 3.1) to an inner product of the form $\hat{d} \cdot s$, where $\hat{d}$ can be efficiently computed from d.

4.4.1 Moderate matrices

The following lemma argues that, provided the matrix $\mathbf{C} \in \mathbb{Z}_q^{\ell \times n}$ is a uniformly random matrix with sufficiently few rows, the distribution $(\mathbf{C}, \mathbf{Cs})$ for arbitrary $s \in \{0,1\}^n$ does not reveal any parity of s.

[5] Specifically, the state can be created using a technique by Grover and Rudolph ([GR02]), who show that in order to create such a state, it suffices to have the ability to efficiently compute the sum $\sum_{x=c}^{d} D_{\mathbb{Z}_q, B_P}(x)$ for any $c, d \in \{-\lfloor\sqrt{B_P}\rfloor, \ldots, \lceil\sqrt{B_P}\rceil\} \subseteq \mathbb{Z}_q$ and to within good precision. This can be done using standard techniques used to sample from the normal distribution.

Lemma 4.2. *Let q be a prime, $\ell, n \geq 1$ integers, and $\mathbf{C} \in \mathbb{Z}_q^{\ell \times n}$ a uniformly random matrix. With probability at least $1 - q^\ell \cdot 2^{-\frac{n}{8}}$ over the choice of $\mathbf{C}$ the following holds. For a fixed $\mathbf{C}$, all $\mathbf{v} \in \mathbb{Z}_q^\ell$ and $\hat{d} \in \{0,1\}^n \setminus \{0^n\}$, the distribution of $(\hat{d} \cdot s \bmod 2)$, where s is uniform in $\{0,1\}^n$ conditioned on $\mathbf{Cs} = \mathbf{v}$, is within statistical distance $O(q^{\frac{3\ell}{2}} \cdot 2^{-\frac{n}{40}})$ of the uniform distribution over $\{0,1\}$.*

To prove the lemma we introduce the notion of a *moderate* matrix.

Definition 4.3. Let $\mathbf{b} \in \mathbb{Z}_q^n$. We say that $\mathbf{b}$ is moderate if it contains at least $\frac{n}{4}$ entries whose unique representative in $(-q/2, q/2]$ has its absolute value in the range $(\frac{q}{8}, \frac{3q}{8}]$. A matrix $\mathbf{C} \in \mathbb{Z}_q^{\ell \times n}$ is moderate if its entire row span (except 0^n) is moderate.

Lemma 4.4. *Let q be prime and ℓ, n be integers. Then*

$$\Pr_{\mathbf{C} \leftarrow_U \mathbb{Z}_q^{\ell \times n}} \big(\mathbf{C} \textit{ is moderate}\big) \geq 1 - q^\ell \cdot 2^{-\frac{n}{8}} .$$

Proof. Consider an arbitrary non zero vector $\mathbf{b}$ in the row-span of a uniform $\mathbf{C}$. Then the marginal distribution of $\mathbf{b}$ is uniform. By Chernoff, $\mathbf{b}$ is moderate with probability at least $1 - e^{-\frac{2n}{16}} \geq 1 - 2^{-\frac{n}{8}}$. Applying the union bound over all at most $q^\ell - 1$ non zero vectors in the row span, the result follows. □

Lemma 4.5. *Let $\mathbf{C} \in \mathbb{Z}_q^{\ell \times n}$ be an arbitrary moderate matrix and let $\hat{d} \in \{0,1\}^n \setminus \{0^n\}$ be an arbitrary non zero binary vector. Let s be uniform over $\{0,1\}^n$ and consider the random variables $\mathbf{v} = \mathbf{Cs} \bmod q$ and $z = \hat{d} \cdot s \bmod 2$. Then $(\mathbf{v}, z)$ is within total variation distance at most $q^{\frac{\ell}{2}} \cdot 2^{-\frac{n}{40}}$ of the uniform distribution over $\mathbb{Z}_q^\ell \times \{0,1\}$.*

Proof. Let f be the probability density function of $(\mathbf{v}, z)$. Interpreting z as an element of $\mathbb{Z}_2$, let $\hat{f}$ be the Fourier transform over $\mathbb{Z}_q^\ell \times \mathbb{Z}_2$. Let U denote the density of the uniform distribution over $\mathbb{Z}_q^\ell \times \mathbb{Z}_2$. Applying the Cauchy-Schwarz inequality,

$$\begin{aligned}
\frac{1}{2}\|f - U\|_1 &\leq \sqrt{\frac{q^\ell}{2}}\|f - U\|_2 \\
&= \frac{1}{2}\|\hat{f} - \hat{U}\|_2 \\
&= \frac{1}{2}\Big(\sum_{(\hat{\mathbf{v}}, \hat{z}) \in \mathbb{Z}_q^\ell \times \mathbb{Z}_2 \setminus \{(\mathbf{0},0)\}} |\hat{f}(\hat{\mathbf{v}}, \hat{z})|^2\Big)^{1/2} ,
\end{aligned} \tag{26}$$

where the second line follows from Parseval's identity, and for the third line we used $\hat{f}(\mathbf{0}, 0) = \hat{U}(0,0) = 1$ and $\hat{U}(\hat{\mathbf{v}}, \hat{z}) = 0$ for all $(\hat{\mathbf{v}}, \hat{z}) \neq (0^\ell, 0)$. To bound (26) we estimate the Fourier coefficients of f. Denoting $\omega_{2q} = e^{-\frac{2\pi i}{2q}}$, for any $(\hat{\mathbf{v}}, \hat{z}) \in \mathbb{Z}_q^\ell \times \mathbb{Z}_2$ we can write

$$\begin{aligned}
\hat{f}(\hat{\mathbf{v}}, \hat{z}) &= \mathrm{E}_{\mathbf{s}}\Big[\omega_{2q}^{(2 \cdot \hat{\mathbf{v}}^T C + q \cdot \hat{z}\hat{\mathbf{d}}^T)\mathbf{s}}\Big] \\
&= \mathrm{E}_{\mathbf{s}}\big[\omega_{2q}^{\mathbf{w}^T \mathbf{s}}\big] \\
&= \prod_i \mathrm{E}_{s_i}\big[\omega_{2q}^{w_i s_i}\big] ,
\end{aligned} \tag{27}$$

where we wrote $\mathbf{w}^T = 2 \cdot \hat{\mathbf{v}}^T \mathbf{C} + q \cdot \hat{z}\hat{\mathbf{d}}^T \in \mathbb{Z}_{2q}^n$. It follows that $\hat{f}(0^\ell, 1) = 0$, since $(d \cdot s \mod 2)$ is uniform for s uniform.

We now observe that for all $i \in \{1,\ldots,n\}$ such that the representative of $(\hat{\mathbf{v}}^T\mathbf{C})_i$ in $(-q/2, q/2]$ has its absolute value in $(\frac{q}{8}, \frac{3q}{8}]$ it holds that $\frac{w_i}{q} \in (\frac{1}{4}, \frac{3}{4}] \bmod 1$, in which case

$$\left|\mathrm{E}_{s_i}[\omega_{2q}^{w_i s_i}]\right| \;=\; \left|\cos\left(\frac{\pi}{2}\cdot\frac{w_i}{q}\right)\right| \;\le\; \cos\left(\frac{\pi}{8}\right) \;\le\; 2^{-\frac{1}{10}}\,. \tag{28}$$

Since $\mathbf{C}$ is moderate, there are at least $\frac{n}{4}$ such entries, so that from (27) it follows that $|\hat{f}(\hat{\mathbf{v}}, \hat{z})| \le 2^{-\frac{n}{40}}$ for all $\hat{\mathbf{v}} \neq \mathbf{0}$. Recalling (26), the lemma is proved. □

We now prove Lemma 4.2 by generalizing Lemma 4.5 to adaptive d (i.e. d can depend on $\mathbf{C}, \mathbf{Cs}$).

Proof of Lemma 4.2. We assume $\mathbf{C}$ is moderate; by Lemma 4.4, $\mathbf{C}$ is moderate with probability at least $1 - q^\ell \cdot 2^{-\frac{n}{8}}$. Let s be uniform over $\{0,1\}^n$, $D_1 = (\mathbf{Cs}, \hat{d}\cdot s \bmod 2)$, and D_2 uniformly distributed over $\mathbb{Z}_q^\ell \times \{0,1\}$. Using that $\mathbf{C}$ is moderate, it follows from Lemma 4.5 that

$$\varepsilon \;=\; \|D_1 - D_2\|_{TV} \;\le\; q^{\frac{\ell}{2}} \cdot 2^{\frac{-n}{40}}\,. \tag{29}$$

Fix $\mathbf{v}_0 \in \mathbb{Z}_q^\ell$ and let

$$\Delta \;=\; \frac{1}{2}\sum_{b\in\{0,1\}} \left| \Pr_{s\leftarrow_U\{0,1\}^n}\left(\hat{d}\cdot s \bmod 2 = b \,\middle|\, \mathbf{Cs} = \mathbf{v}_0\right) - \frac{1}{2}\right|. \tag{30}$$

To prove the lemma it suffices to establish the appropriate upper bound on Δ, for all $\mathbf{v}_0$. By definition,

$$\begin{aligned}
\varepsilon \;=\; \|D_1 - D_2\|_{TV} \;&=\; \frac{1}{2}\sum_{b\in\{0,1\}, v\in\mathbb{Z}_q^\ell} \left|\Pr\left(\mathbf{Cs}=\mathbf{v}\right)\Pr\left(\hat{d}\cdot s \bmod 2 = b \,\middle|\, \mathbf{Cs}=\mathbf{v}\right) - \frac{1}{2q^\ell}\right| \\
&\ge\; \frac{1}{2}\sum_{b\in\{0,1\}} \left|\Pr\left(\mathbf{Cs}=\mathbf{v}_0\right)\Pr\left(\hat{d}\cdot s \bmod 2 = b \,\middle|\, \mathbf{Cs}=\mathbf{v}_0\right) - \frac{1}{2q^\ell}\right| \\
&=\; \frac{1}{2}\sum_{b\in\{0,1\}} \left|\Pr\left(\mathbf{Cs}=\mathbf{v}_0\right)\left(\frac{1}{2} + (-1)^b\Delta\right) - \frac{1}{2q^\ell}\right|,
\end{aligned} \tag{31}$$

where all probabilities are under a uniform choice of $s \leftarrow_U \{0,1\}^n$, and the last line follows from the definition of Δ in (30). Applying the inequality $|a| + |b| \ge \max(|a-b|, |a+b|)$, valid for any real a, b, to (31) it follows that

$$\Pr\left(\mathbf{Cs}=\mathbf{v}_0\right)\cdot\Delta \;\le\; \varepsilon \qquad \text{and} \qquad \Pr\left(\mathbf{Cs}=\mathbf{v}_0\right) \;\ge\; \frac{1}{q^\ell} - 2\varepsilon\,. \tag{32}$$

If $q^{3\ell/2}2^{-\frac{n}{40}} > \frac{1}{3}$ the bound claimed in the lemma is trivial. If $q^{3\ell/2}2^{-\frac{n}{40}} \le \frac{1}{3}$, then $\varepsilon q^\ell \le \frac{1}{3}$, so it follows from (32) that $\Delta \le 3q^\ell\varepsilon$, which together with (29) proves the lemma. □

4.4.2 LWE Hardcore bit

The next step is to use Lemma 4.2 to obtain a form of the hardcore bit statement that is appropriate for our purposes. We use the following notation: we write $s \in \{0,1\}^n$ as $s = (s_0, s_1)$, where $s_0, s_1 \in \{0,1\}^{\frac{n}{2}}$ are the $\frac{n}{2}$-bit prefix and suffix of s respectively (for simplicity, assume n is even; if not, ties can be broken arbitrarily). We will show computational indistinguishability based on the hardness assumption $\mathrm{LWE}_{\ell,q,D_{\mathbb{Z}_q,B_L}}$ specified in Definition 2.5.

For reasons that will become clear in the next section, we consider procedures that output a tuple $(b, x, d, c) \in \{0,1\} \times \mathcal{X} \times \{0,1\}^w \times \{0,1\}$.

Lemma 4.6. *Assume a choice of parameters satisfying the conditions* (13)*. Assume the hardness assumption* $\mathrm{LWE}_{\ell,q,D_{\mathbb{Z}_q,B_L}}$ *holds. Let*

$$\mathcal{A}:\ \mathbb{Z}_q^{m\times n}\times\mathbb{Z}_q^m\ \to\ \{0,1\}\times\mathcal{X}\times\{0,1\}^w\times\{0,1\}$$

be a quantum polynomial-time procedure. For $b\in\{0,1\}$ *and* $x\in\mathcal{X}$ *let* $I_{b,x}:\{0,1\}^w\to\{0,1\}^n$ *be an efficiently computable map. For every* $s=(s_0,s_1)\in\{0,1\}^n$ *and* $(b,x)\in\{0,1\}\times\mathcal{X}$*, let* $\hat{G}_{s_{b\oplus 1},b,x}\subseteq\{0,1\}^w$ *be a set depending only on* b,x *and* $s_{b\oplus 1}$ *and such that for all* $d\in\hat{G}_{s_{b\oplus 1},b,x}$ *the first (if* $b=0$*) or last (if* $b=1$*)* $\frac{n}{2}$ *bits of* $I_{b,x}(d)$ *are not all* 0*. Then the distributions*

$$D_0\ =\ \big((\mathbf{A},\mathbf{As}+\mathbf{e})\leftarrow GEN_{\mathcal{F}_{\mathrm{LWE}}}(1^\lambda),\ (b,x,d,c)\leftarrow\mathcal{A}(\mathbf{A},\mathbf{As}+\mathbf{e}),\ I_{b,x}(d)\cdot s\mod 2\big)\tag{33}$$

and

$$D_1\ =\ \big((\mathbf{A},\mathbf{As}+\mathbf{e})\leftarrow GEN_{\mathcal{F}_{\mathrm{LWE}}}(1^\lambda),\ (b,x,d,c)\leftarrow\mathcal{A}(\mathbf{A},\mathbf{As}+\mathbf{e}),\ (\delta_{d\in\hat{G}_{s_{b\oplus 1},b,x}}r)\oplus(I_{b,x}(d)\cdot s\mod 2)\big)\,,\tag{34}$$

where $r\leftarrow_U\{0,1\}$ *and* $\delta_{d\in\hat{G}_{s_{b\oplus 1},b,x}}$ *is* 1 *if* $d\in\hat{G}_{s_{b\oplus 1},b,x}$ *and* 0 *otherwise, are computationally indistinguishable.*

Proof. We prove computational indistinguishability by introducing a sequence of hybrids. For the first step we let

$$D^{(1)}\ =\ \big((\tilde{\mathbf{A}},\tilde{\mathbf{A}}\mathbf{s}+\mathbf{e}),\ (b,x,d,c)\leftarrow\mathcal{A}(\tilde{\mathbf{A}},\tilde{\mathbf{A}}\mathbf{s}+\mathbf{e}),\ I_{b,x}(d)\cdot s\mod 2\big)\,,\tag{35}$$

where $\tilde{\mathbf{A}}=\mathbf{BC}+\mathbf{F}\leftarrow\mathrm{LOSSY}(1^n,1^m,1^\ell,q,D_{\mathbb{Z}_q,B_L})$ is sampled from a lossy sampler (see Definition 2.7). From the definition, $\mathbf{F}\in\mathbb{Z}_q^{m\times n}$ has entries i.i.d. from the distribution $D_{\mathbb{Z}_q,B_L}$ over $\mathbb{Z}_q$. To see that D_0 and $D^{(1)}$ are computationally indistinguishable, first note that the distribution of matrices $\mathbf{A}$ generated by $\mathrm{GEN}_{\mathcal{F}_{\mathrm{LWE}}}$ is negligibly far from the uniform distribution (see Theorem 2.6). Next, by Theorem 2.8, under the $\mathrm{LWE}_{\ell,q,D_{\mathbb{Z}_q,B_L}}$ assumption a uniformly random matrix $\mathbf{A}$ and a lossy matrix $\tilde{\mathbf{A}}$ are computationally indistinguishable. Note that this step, as well as subsequent steps, uses that $\mathcal{A}$ and $I_{b,x}$ are efficiently computable.

For the second step we remove the term $\mathbf{Fs}$ from the lossy LWE sample $\tilde{\mathbf{A}}\mathbf{s}+\mathbf{e}$ to obtain the distribution

$$D^{(2)}\ =\ \big((\mathbf{BC}+\mathbf{F},\mathbf{BCs}+\mathbf{e}),\ (b,x,d,c)\leftarrow\mathcal{A}(\mathbf{BC}+\mathbf{F},\mathbf{BCs}+\mathbf{e}),\ I_{b,x}(d)\cdot s\mod 2\big)\,.\tag{36}$$

Using that $\mathbf{s}$ is binary and the entries of $\mathbf{F}$ are taken from a B_L-bounded distribution, it follows that $\|\mathbf{Fs}\|\le n\sqrt{m}B_L$. Applying Lemma 2.4, the statistical distance between $D^{(1)}$ and $D^{(2)}$ is at most

$$\gamma\ =\ \sqrt{2}\Big(1-e^{\frac{-2\pi mnB_L}{B_V}}\Big)^{1/2}\,,\tag{37}$$

which is negligible, due to the requirement that $\frac{B_V}{B_L}$ is superpolynomial given in (13).

For the third step, observe that the distribution $D^{(2)}$ in (36) only depends on s_b through $\mathbf{Cs}$ and $I_{b,x}(d)\cdot s$, where $\mathbf{C}$ is uniformly random. It follows from Lemma 4.2 that provided $\frac{n}{2}=\Omega(\ell\log q)$, with overwhelming probability over the choice of $\mathbf{C}$, if we fix all variables except for s_b, the distribution of $(I_{b,x}(d)\cdot s \mod 2)$ is statistically indistinguishable from $r\leftarrow_U\{0,1\}$ as long as the $\frac{n}{2}$ bits of $I_{b,x}(d)$ associated with

s_b are not all 0 (i.e. the first $\frac{n}{2}$ bits if $b = 0$ or the last $\frac{n}{2}$ bits if $b = 1$). Using that for $d \in \hat{G}_{s_{b\oplus 1},b,x}$ the $\frac{n}{2}$ bits of $I_{b,x}(d)$ associated with s_b are not all 0, the distribution $D^{(2)}$ in (36) is statistically indistinguishable from

$$D^{(3)} \;=\; \left((\mathbf{BC}+\mathbf{F}, \mathbf{BCs}+\mathbf{e}),\ (b,x,d,c) \leftarrow \mathcal{A}(\mathbf{BC}+\mathbf{F}, \mathbf{BCs}+\mathbf{e}),\ (\delta_{d\in\hat{G}_{s_{b\oplus 1},b,x}} r) \oplus (I_{b,x}(d)\cdot s \mod 2)\right) ,$$

where $r \leftarrow_U \{0,1\}$.

For the fourth step we reinsert the term $F\mathbf{s}$ to obtain

$$D^{(4)} \;=\; \left(\tilde{\mathbf{A}}, \tilde{\mathbf{A}}\mathbf{s}+\mathbf{e},\ (b,x,d,c) \leftarrow \mathcal{A}(\tilde{\mathbf{A}}, \tilde{\mathbf{A}}\mathbf{s}+\mathbf{e}),\ (\delta_{d\in\hat{G}_{s_{b\oplus 1},b,x}} r) \oplus (I_{b,x}(d)\cdot s \mod 2)\right) .$$

Statistical indistinguishability between $D^{(3)}$ and $D^{(4)}$ follows similarly as between $D^{(1)}$ and $D^{(2)}$. Finally, computational indistiguishability between $D^{(4)}$ and D_1 follows similarly to between $D^{(1)}$ and D_0. □

4.4.3 Adaptive hardcore lemma

We now prove that condition 4 of Definition 3.1 holds. Recall that $\mathcal{X} = \mathbb{Z}_q^n$ and $w = n\lceil \log q \rceil$. Let $J : \mathcal{X} \to \{0,1\}^w$ be such that $J(x)$ returns the binary representation of $x \in \mathcal{X}$. For $b \in \{0,1\}$, $x \in \mathcal{X}$, and $d \in \{0,1\}^w$, let $I_{b,x}(d) \in \{0,1\}^n$ be the vector whose each coordinate is obtained by taking the inner product mod 2 of the corresponding block of $\lceil \log q \rceil$ coordinates of d and of $J(x) \oplus J(x-(-1)^b\mathbf{1})$, where $\mathbf{1} \in \mathbb{Z}_q^n$ is the vector with all its coordinates equal to $1 \in \mathbb{Z}_q$. For $k = (\mathbf{A}, \mathbf{As}+\mathbf{e}), b \in \{0,1\}$ and $x \in \mathcal{X}$, we define the set $G_{k,b,x}$ as

$$G_{k,b,x} \;=\; \left\{ d \in \{0,1\}^w \;\middle|\; \exists i \in \left\{ b\frac{n}{2}, \ldots, b\frac{n}{2}+\frac{n}{2} \right\} :\ (I_{b,x}(d))_i \neq 0 \right\} .$$

Observe that for all $b \in \{0,1\}$ and $x \in \mathcal{X}$, if d is sampled uniformly at random, $d \notin G_{k,b,x}$ with negligible probability. This follows simply because for any $b \in \{0,1\}$, $J(x) \oplus J(x-(-1)^b\mathbf{1})$ is non-zero, since J is injective. Observe also that checking membership in $G_{k,b,x}$ is possible given only b, x. This shows condition 4.(a) in the adaptive hardcore bit condition in Definition 3.1.

Given $(x_0, x_1) \in \mathcal{R}_k$ (where $k = (\mathbf{A}, \mathbf{As}+\mathbf{e})$), recall from Section 4.2 that $x_1 = x_0 - \mathbf{s}$. For convenience we also introduce the following set, where $y = f_{k,0}(x_0) = f_{k,1}(x_1)$:

$$\hat{G}_{s_1,0,x_0} \;=\; \hat{G}_{s_0,1,x_1} \;=\; G_{k,0,x_0} \cap G_{k,1,x_1} . \tag{38}$$

The motivation for using two different notation for the same set is to clarify that membership in the set can be decided given $(s_{b\oplus 1}, b, x_b)$, for either $b \in \{0,1\}$. This point is important in the proof of Lemma 4.6.

The following lemma establishes item 4.(b) in Definition 3.1.

Lemma 4.7. *Assume a choice of parameters satisfying the conditions* (13)*. Assume the hardness assumption* $\mathrm{LWE}_{\ell,q,D_{\mathbb{Z}_q,B_L}}$ *holds. Let* $s \in \{0,1\}^n$*. Let* [6]

$$H_s \;=\; \left\{ (b,x,d, d\cdot(J(x)\oplus J(x-(-1)^b\mathbf{s}))) \mid b \in \{0,1\},\ x \in \mathcal{X},\ d \in \hat{G}_{s_{b\oplus 1},b,x} \right\} , \tag{39}$$

$$\overline{H}_s \;=\; \left\{ (b,x,d,c) \mid (b,x,d,c\oplus 1) \in H_s \right\} . \tag{40}$$

Then for any quantum polynomial-time procedure

$$\mathcal{A} :\ \mathbb{Z}_q^{m\times n} \times \mathbb{Z}_q^m \;\to\; \{0,1\} \times \mathcal{X} \times \{0,1\}^w \times \{0,1\}$$

[6] We write the sets as H_s instead of H_k to emphasize the dependence on s.

there exists a negligible function $\mu(\lambda)$ such that

$$\Big|\Pr_{(\mathbf{A},\mathbf{As}+\mathbf{e})\leftarrow GEN_{\mathcal{F}_{\mathrm{LWE}}}(1^\lambda)}\left[\mathcal{A}(\mathbf{A},\mathbf{As}+\mathbf{e})\in H_s\right]-\Pr_{(\mathbf{A},\mathbf{As}+\mathbf{e})\leftarrow GEN_{\mathcal{F}_{\mathrm{LWE}}}(1^\lambda)}\left[\mathcal{A}(\mathbf{A},\mathbf{As}+\mathbf{e})\in \overline{H}_s\right]\Big| \;\leq\; \mu(\lambda)\;. \tag{41}$$

Proof. The proof is by contradiction. Assume that there exists a quantum polynomial-time procedure $\mathcal{A}$ such that the left-hand side of (41) is at least some non-negligible function $\eta(\lambda)$. We derive a contradiction by showing that the two distributions D_0 and D_1 in Lemma 4.6, for $I_{b,x}$ defined at the start of this section and $\hat{G}_{s_{b\oplus 1},b,x}$ defined in (38), are computationally distinguishable, giving a contradiction.

Let $(\mathbf{A},\mathbf{As}+\mathbf{e})\leftarrow \mathrm{GEN}_{\mathcal{F}_{\mathrm{LWE}}}(1^\lambda)$ and $(b,x,d,c)\leftarrow\mathcal{A}(\mathbf{A},\mathbf{As}+\mathbf{e})$. To link $\mathcal{A}$ to the distributions in Lemma 4.6 we relate the inner product condition in (39) to an inner product $\hat{d}\cdot s$ of the form appearing in (33), for $\hat{d}=I_{b,x}(d)$ that can be computed efficiently from b,x and d. This is based on the following claim.

Claim 4.8. *For all $b\in\{0,1\}, x\in\mathcal{X}, d\in\{0,1\}^w$ and $s\in\{0,1\}^n$ the following equality holds:*

$$d\cdot(J(x)\oplus J(x-(-1)^b\mathbf{s}))\;=\;I_{b,x}(d)\cdot s\;. \tag{42}$$

Moreover, the function $I_{b,x}$ is efficiently computable given b,x.

Proof. We do the proof in case $n=1$ and $w=\lceil\log q\rceil$, as the case of general n follows by linearity. In this case s is a single bit. If $s=0$ then both sides of (42) evaluate to zero, so the equality holds trivially. It then suffices to define $I_{b,x_b}(d)$ so that the equation holds when $s=1$. A choice of either of

$$I_{0,x_0}(d)\;=\;d\cdot(J(x_0)\oplus J(x_0-\mathbf{1}))\;,\quad I_{1,x_1}(d)\;=\;d\cdot(J(x_1)\oplus J(x_1+\mathbf{1}))$$

satisfies all requirements. It is clear from the definition of $I_{b,x}$ that it can be computed efficiently given b,x. □

The procedure $\mathcal{A}$, the function $I_{b,x}$ defined at the start of this section and the sets $\hat{G}_{s_{b\oplus1},b,x}$ in (38) fully specify D_0 and D_1. To conclude we construct a distinguisher $\mathcal{A}'$ between D_0 and D_1. Consider two possible distinguishers, $\mathcal{A}'_u$ for $u\in\{0,1\}$. Given a sample $w=((\mathbf{A},\mathbf{As}+\mathbf{e}),(b,x,d,c),t)$, $\mathcal{A}'_u$ returns 0 if $c=t\oplus u$, and 1 otherwise. First note that

$$\begin{aligned}&\sum_{u\in\{0,1\}}\Big|\Pr_{w\leftarrow D_0}\left[\mathcal{A}'_u(w)=0\right]-\Pr_{w\leftarrow D_1}\left[\mathcal{A}'_u(w)=0\right]\Big|\\ &\qquad=\sum_{u\in\{0,1\}}\Big|\Pr_{w\leftarrow D_0}\left[\mathcal{A}'_u(w)=0\wedge d\in\hat{G}_{s_{b\oplus1},b,x}\right]-\Pr_{w\leftarrow D_1}\left[\mathcal{A}'_u(w)=0\wedge d\in\hat{G}_{s_{b\oplus1},b,x}\right]\Big|\end{aligned} \tag{43}$$

since if $d\notin\hat{G}_{s_{b\oplus1},b,x}$ the distributions D_0 and D_1 are identical by definition. Next, if the sample held by $\mathcal{A}'_u$ is from the distribution D_0 and if $(b,x,d,c)\in H_s$, then by the definition of H_s and (42) it follows that $c=d\cdot(J(x)\oplus J(x-(-1)^b\mathbf{s}))=I_{b,x}(d)\cdot s=t$. If instead $(b,x,d,c)\in\overline{H}_s$ then $c\oplus1=d\cdot(J(x)\oplus$

$J(x-(-1)^b\mathbf{s}) = I_{b,x}(d)\cdot s = t$. The expression in (43) is thus equal to:

$$\begin{aligned}(43) &= \Big|\Pr_{(\mathbf{A},\mathbf{As}+\mathbf{e})\leftarrow \mathrm{GEN}_{\mathcal{F}_{\mathrm{LWE}}}(1^\lambda)}\big[\mathcal{A}(\mathbf{A},\mathbf{As}+\mathbf{e})\in H_s\big] - \frac{1}{2}\Pr_{w\leftarrow D_1}\big[d\in \hat{G}_{s_{b\oplus 1},b,x}\big]\Big| \\ &\qquad + \Big|\Pr_{(\mathbf{A},\mathbf{As}+\mathbf{e})\leftarrow \mathrm{GEN}_{\mathcal{F}_{\mathrm{LWE}}}(1^\lambda)}\big[\mathcal{A}(\mathbf{A},\mathbf{As}+\mathbf{e})\in \overline{H}_s\big] - \frac{1}{2}\Pr_{w\leftarrow D_1}\big[d\in \hat{G}_{s_{b\oplus 1},b,x}\big]\Big| \\ &\geq \Big|\Pr_{(\mathbf{A},\mathbf{As}+\mathbf{e})\leftarrow \mathrm{GEN}_{\mathcal{F}_{\mathrm{LWE}}}(1^\lambda)}\big[\mathcal{A}(\mathbf{A},\mathbf{As}+\mathbf{e})\in H_s\big] - \Pr_{(\mathbf{A},\mathbf{As}+\mathbf{e})\leftarrow \mathrm{GEN}_{\mathcal{F}_{\mathrm{LWE}}}(1^\lambda)}\big[\mathcal{A}(\mathbf{A},\mathbf{As}+\mathbf{e})\in \overline{H}_s\big]\Big| \\ &\geq \eta\ .\end{aligned}$$

Therefore, at least one of $\mathcal{A}'_0$ or $\mathcal{A}'_1$ must successfully distinguish between D_0 and D_1 with advantage at least $\frac{\eta}{2}$, a contradiction with the statement of Lemma 4.6. □

5 Protocol description

We introduce two protocols. The first we call the *(general) randomness expansion protocol*, or Protocol 1. This is our main randomness expansion protocol. It is introduced in Section 5.1, and summarized in Figure 1. The protocol describes the interaction between a *verifier* and *prover*. Ultimately, we aim to obtain the guarantee that any computationally bounded prover that is accepted with non-negligible probability by the verifier in the protocol must generate transcripts that contain information-theoretic randomness.

The second protocol is called the *simplified protocol*, or Protocol 2. It is introduced in Section 5.2, and summarized in Figure 2. This protocol abstracts some of the main features Protocol 1, and will be used as a tool in the analysis (it is not meant to be executed literally).

5.1 The randomness expansion protocol

Our randomness expansion protocol, Protocol 1, is described in Figure 1. The protocol is parametrized by a security parameter λ and a number of rounds N. The other parameters, the error tolerance parameter $\gamma \geq 0$ and the testing parameter $q \in (0,1]$, are assumed to be specified as a function of λ and N. For intuition, γ can be thought of as a small constant and q as a parameter that scales as $\mathrm{poly}(\lambda)/N$.

At the start of the protocol, the verifier executes $(k,t_k) \leftarrow \textsc{Gen}_{\mathcal{F}}(1^\lambda)$ to obtain the public key k and trapdoor t_k for a pair of functions $\{f_{k,b}:\mathcal{X}\to\mathcal{D}_\mathcal{Y}\}_{b\in\{0,1\}}$ from the NTCF family (see Definition 3.1). The verifier sends the public key k to the prover and keeps the associated trapdoor private.

In each of the N rounds of the protocol, the prover is first required to provide a value $y\in\mathcal{Y}$. For each $b\in\{0,1\}$, the verifier uses the trapdoor to compute $\hat{x}_b \leftarrow \textsc{Inv}_{\mathcal{F}}(t_k,b,y)$. (If the inversion procedure fails, the verifier requests another sample from the prover.) For convenience, introduce a set

$$\hat{G}_y \;=\; G_{k,0,x_0}\cap G_{k,1,x_1}\ , \tag{44}$$

where for $b\in\{0,1\}$ the set G_{k,b,x_b} is defined in 4.(a) of Definition 3.1. The verifier then chooses a round type $G\in\{0,1\}$ according to a biased distribution: either a *test round*, $G=0$, chosen with probability $\Pr(G=0)=q$, or a *generation round*, $G=1$, chosen with the remaining probability $\Pr(G=1)=1-q$. The former type of round is less frequent, as the parameter q will eventually be set to a very small value, that goes to 0 with the number of rounds of the protocol. The prover is not told the round type.

Depending on the round type, the verifier chooses a challenge $C\in\{0,1\}$ that she sends to the prover. In the case of a test round the challenge is chosen uniformly at random; in the case of a generation round

the challenge is always $C = 1$. In case $C = 0$ the prover is asked to return a pair $(u, d) \in \{0,1\} \times \{0,1\}^w$. The pair is called valid if $u = d \cdot (J(\hat{x}_0) \oplus J(\hat{x}_1))$ and $d \in \hat{G}_y$, where the function J is as in 4.(b) of Definition 3.1. If $d \in \hat{G}_y$, the verifier sets a decision bit $W = 1$ if the answer is valid, and $W = 0$ if not. If $d \notin \hat{G}_y$, the verifier sets the decision bit $W \in \{0,1\}$ uniformly at random.[7] In case $C = 1$, the prover should return a pair $(b, x) \in \{0,1\} \times \mathcal{X}$. The pair is called valid if $\mathrm{CHK}_{\mathcal{F}}(k, b, x, y) = 1$. The verifier sets a decision bit $W = 1$ in case the pair is valid, and $W = 0$ otherwise. The set of valid pairs on challenge $C = c \in \{0,1\}$ is denoted $V_{y,c}$.

After each test round the verifier samples a fresh $(k, t_k) \leftarrow \mathrm{GEN}_{\mathcal{F}}(1^\lambda)$ and communicates the new public key k to the prover.

At the end of the protocol, the verifier computes the fraction of test rounds in which the decision bit has been set to 1. If this fraction is smaller than $(1 - \gamma)$, the verifier aborts. Otherwise, the verifier returns the concatenation of the bits b obtained from the prover in generation rounds. (These bits are recorded in the verifier's output string $O_1 \cdots O_N$, such that $O_i = 0$ whenever the round is a test round.)

5.2 The simplified protocol

For purposes of analysis only we introduce a simplified variant of Protocol 1, which is specified in Figure 2. We call it the *simplified protocol*, or Protocol 2. The protocol is very similar to the randomness expansion protocol described in Figure 1, except that the prover's answers and the verifier's checks are simplified, and in test rounds there is an additional challenge bit $T \in \{0,1\}$. This new challenge asks the prover to perform a projective measurement on its private space that indicates whether the state lies in a "good subspace" (indicated by an outcome $K = 0$) or in the complementary "bad subspace" (outcome $K = 1$). The "good" and "bad" subspaces represent portions of space where the device's other two measurements, M and Π are anti-aligned and aligned respectively; see the definition of a simplified device in Section 6.2 for details.

For the case of a challenge $C = 0$, in Protocol 1 the prover returns an equation (u, d). In the simplified protocol the prover returns a single bit $e \in \{0,1\}$ that is meant to directly indicate the verifier's decision (i.e. the bit W). If moreover $T = 1$ the prover is required to reply with an additional bit $k \in \{0,1\}$. In this case, the verifier makes the decision to accept, i.e. sets $W = 1$, if and only if $e = 1$ and $k = 0$. For the case of a challenge $C = 1$, in Protocol 1 the prover returns a pair (b, x). In the simplified protocol the prover returns a value $v \in \{0,1,2\}$ that is such that $v = b$ in case (b, x) is valid, i.e. $(b, x) \in V_{y,1}$, and $v = 2$ otherwise.

Note that this "honest" behavior for the prover is not necessarily efficient. Moreover, it is easy for a "malicious" prover to succeed in Protocol 2, e.g. by always returning $u = 1$ (valid equation), $k = 0$ (good subspace) and $v \in \{0,1\}$ (valid pre-image). Our analysis will not consider arbitrary provers in Protocol 2, but instead provers whose measurements satisfy certain constraints that arise from the analysis of Protocol 1. For such provers, it will be impossible to succeed in the simplified protocol without generating randomness. Further details are given in Section 7.

5.3 Completeness

We describe the intended behavior for the prover in Protocol 1. Fix an NTCF family $\mathcal{F}$ and a key $k \in \mathcal{K}_{\mathcal{F}}$. In each round, the "honest" prover performs the following actions.

[7] This choice if made for technical reasons that have to do with the definition of the adaptive hardcore bit property; see Section 7 and the proof of Proposition 7.4 for details.

Let λ be a security parameter, $N \geq 1$ a number of rounds, and $\gamma, q > 0$ functions of λ and N. Let $\mathcal{F}$ be an NTCF family.
At the start of the protocol, the verifier communicates N to the prover. In addition, the verifier samples an initial key $(k, t_k) \leftarrow \text{GEN}_{\mathcal{F}}(1^\lambda)$, sends k to the prover and keeps the trapdoor information t_k private.

1. For $i = 1, \ldots, N$:
   (a) The prover returns a $y \in \mathcal{Y}$ to the verifier. For $b \in \{0,1\}$ the verifier uses the trapdoor to compute $\hat{x}_b \leftarrow \text{INV}_{\mathcal{F}}(t_k, b, y)$.
   (b) The verifier selects a round type $G_i \in \{0,1\}$ according to a Bernoulli distribution with parameter q: $\Pr(G_i = 0) = q$ and $\Pr(G_i = 1) = 1 - q$. In case $G_i = 0$ (*test round*), she chooses a challenge $C_i \in \{0,1\}$ uniformly at random. In case $G_i = 1$ (*generation round*), she sets $C_i = 1$. The verifier keeps G_i private, and sends C_i to the prover.
      i. In case $C_i = 0$ the prover returns $(u, d) \in \{0,1\} \times \{0,1\}^w$. If $d \notin \hat{G}_y$, the set defined in (44), the verifier sets W to a uniformly random bit. Otherwise, the verifier sets $W = 1$ if $d \cdot (J(\hat{x}_0) \oplus J(\hat{x}_1)) = u$ and $W = 0$ if not.
      ii. In case $C_i = 1$ the prover returns $(b, x) \in \{0,1\} \times \mathcal{X}$. The verifier sets W as the value returned by $\text{CHK}_{\mathcal{F}}(k, b, x, y)$.
   (c) In case $G_i = 1$, the verifier sets $O_i = b$. In case $G_i = 0$, she sets $W_i = W$.
   (d) In case $G_i = 0$, the verifier samples a new key $(k, t_k) \leftarrow \text{GEN}_{\mathcal{F}}(1^\lambda)$. She sends k to the prover and keeps the trapdoor information t_k private. This key will be used until the next test round, included.
2. If $\sum_{i:G_i=0} W_i < (1-\gamma)qN$, the verifier aborts. Otherwise, she returns the string O obtained by concatenating the bits O_i for all $i \in \{1, \ldots, N\}$ such that $G_i = 1$.

Figure 1: The randomness expansion protocol, Protocol 1. See Definition 3.1 for notation associated with the NTCF family $\mathcal{F}$.

Let λ be a security parameter, $N \geq 1$ a number of rounds, and $\gamma, \eta, \kappa, q > 0$ functions of λ and N.

1. For $i = 1, \ldots, N$:

   (a) The verifier selects a round type $G_i \in \{0,1\}$ according to a Bernoulli distribution with parameter q: $\Pr(G_i = 0) = q$ and $\Pr(G_i = 1) = 1 - q$. In case $G_i = 0$ (*test round*), she chooses $C_i \in \{0,1\}$ uniformly at random and $T_i \in \{0,1\}$ such that $\Pr(T_i = 0) = 1 - \kappa$ and $\Pr(T_i = 1) = \kappa$. In case $G_i = 1$ (*generation round*), she sets $C_i = 1$ and $T_i = 0$. The verifier keeps G_i private, and sends (C_i, T_i) to the prover.

      i. In case $C_i = 0$ the prover returns $e \in \{0,1\}$. If $T_i = 1$ the prover in addition reports $k \in \{0,1\}$.[8] If $T_i = 0$ the verifier sets $W_i = e$. If $T_i = 1$ the verifier sets $W_i = e(1-k)$.

      ii. In case $C_i = 1$ the prover returns $v \in \{0,1,2\}$. The verifier sets $O_i = v$ and $W_i = 1_{v \in \{0,1\}}$.

2. If $\sum_{i:G_i=0 \wedge T_i=1} W_i < (1 - \frac{\gamma}{\kappa} - \eta)\kappa q N$, the verifier rejects the interaction. Otherwise, she returns the string O obtained by concatenating the bits O_i for all $i \in \{1, \ldots, N\}$ such that $G_i = 1$.

Figure 2: The simplified protocol, Protocol 2.

1. The prover executes the efficient procedure $\mathrm{SAMP}_{\mathcal{F}}$ in superposition to obtain the state

$$|\psi^{(1)}\rangle \;=\; \frac{1}{\sqrt{|\mathcal{X}|}} \sum_{x \in \mathcal{X}, y \in \mathcal{Y}, b \in \{0,1\}} \sqrt{(f'_{k,b}(x))(y)}\, |b, x\rangle |y\rangle \;.$$

2. The prover measures the last register to obtain an $y \in \mathcal{Y}$. Using item 2. from the definition of an NTCF, the prover's re-normalized post-measurement state is

$$|\psi^{(2)}\rangle \;=\; \frac{1}{\sqrt{2}}\big(|0, x_0\rangle + |1, x_1\rangle\big)|y\rangle \;,$$

   where for $b \in \{0,1\}$, $x_b = \mathrm{INV}_{\mathcal{F}}(t_k, b, y)$.

   (a) In case $C_i = 0$, the prover evaluates the function J on the second register, containing x_b, and then applies a Hadamard transform to all $w+1$ qubits in the first two registers. Tracing out the register that contains y, this yields the state

$$\begin{aligned} |\psi^{(3)}\rangle &= 2^{-\frac{w+2}{2}} \sum_{d,b,u} (-1)^{d \cdot J(x_b) \oplus ub} |u\rangle |d\rangle \\ &= (-1)^{J(x_0)} 2^{-\frac{w}{2}} \sum_{d \in \{0,1\}^w} |d \cdot (J(x_0) \oplus J(x_1))\rangle |d\rangle \;. \end{aligned}$$

   The prover measures both registers to obtain an (u, d) that it sends back to the verifier.

   (b) In case $C_i = 1$, the prover measures the first two registers of $|\psi^{(2)}\rangle$ in the computational basis, and returns the outcome (b, x_b) to the verifier.

[8] The bit k should not be confused with the public key k for the NTCF that is used in Protocol 1. In Protocol 2, there is no NTCF, and no key.

Lemma 5.1. *For any λ and $k \leftarrow \mathrm{GEN}_{\mathcal{F}}(1^\lambda)$, the strategy for the honest prover (on input k) in one round of the protocol can be implemented in time polynomial in λ and is accepted with probability negligibly close to* 1.

Proof. Both efficiency and correctness of the prover follow from the definition of an NTCF (Definition 3.1). The prover fails only if it obtains an outcome $d \notin \hat{G}_y$, which by item 4(a) in the definition happens with negligible probability. □

6 Devices

We model an arbitrary prover in the randomness expansion protocol (Protocol 1 in Figure 1) as a *device* that implements the actions of the prover: the device first returns an $y \in \mathcal{Y}$; then, depending on the challenge $C \in \{0,1\}$, it either returns an equation (u,d) (case $C=0$), or a candidate pre-image (b,x) (case $C=1$). For simplicity we assume that the device makes the same set of measurements in each round of the protocol. This is without loss of generality, as we allow the state of the device to change from one round to the next; in particular the device is allowed to use a quantum memory as a control register for the measurements.

In Section 6.1 we introduce our notation for modeling provers in Protocol 1 as devices. In Section 6.2 we consider a simplified form of device, that is appropriate for modeling a prover in the simplified protocol, Protocol 2. In Section 7 we give a reduction showing how to associate a specific simplified device to any computationally efficient general device, such that the randomness generation properties of the two devices can be related to each other (this is done in Section 8).

For the remainder of this section we fix an NTCF family $\mathcal{F}$ satisfying the conditions of Definition 3.1, and use notation introduced in the definition.

6.1 General devices

The following notion of device models the behavior of an arbitrary prover in the randomness expansion protocol, Protocol 1 (Figure 1).

Definition 6.1. Given $k \in \mathcal{K}_{\mathcal{F}}$, a device $D = (\phi, \Pi, M)$ (implicitly, compatible with k) is specified by the following:

1. A (not necessarily normalized) positive semidefinite $\phi \in \mathrm{Pos}(\mathcal{H}_{\mathsf{D}} \otimes \mathcal{H}_{\mathsf{Y}})$. Here $\mathcal{H}_{\mathsf{D}}$ is an arbitrary space private to the device, and $\mathcal{H}_{\mathsf{Y}}$ is a space of the same dimension as the cardinality of the set $\mathcal{Y}$, also private to the device. For every $y \in \mathcal{Y}$, define

$$\phi_y \;=\; (\mathrm{Id}_{\mathsf{D}} \otimes \langle y|_{\mathsf{Y}})\, \phi\, (\mathrm{Id}_{\mathsf{D}} \otimes |y\rangle_{\mathsf{Y}}) \;\in\; \mathrm{Pos}(\mathcal{H}_{\mathsf{D}})\,.$$

   Note that ϕ_y is not normalized, and $\sum_{y\in\mathcal{Y}} \mathrm{Tr}(\phi_y) = \mathrm{Tr}(\phi)$.

2. For every $y \in \mathcal{Y}$, a projective measurement $\{M_y^{(u,d)}\}$ on $\mathcal{H}_{\mathsf{D}}$, with outcomes $(u,d) \in \{0,1\} \times \{0,1\}^w$.

3. For every $y \in \mathcal{Y}$, a projective measurement $\{\Pi_y^{(b,x)}\}$ on $\mathcal{H}_{\mathsf{D}}$, with outcomes $(b,x) \in \{0,1\} \times \mathcal{X}$. For each y, this measurement has two designated outcomes $(0,x_0)$ and $(1,x_1)$, which are the answers that are accepted on challenge $C=1$ in the protocol; recall that we use the notation $V_{y,1}$ for this set. For $b \in \{0,1\}$ we use the shorthand $\Pi_y^b = \Pi_y^{(b,x_b)}$, $\Pi_y = \Pi_y^0 + \Pi_y^1$, and $\Pi_y^2 = \mathrm{Id} - \Pi_y^0 - \Pi_y^1$.

By Naimark's theorem, up to increasing the dimension of $\mathcal{H}_{\mathsf{D}}$ the assumption that $\{\Pi_y^{(b,x)}\}$ and $\{M_y^{(u,d)}\}$ are projective is without loss of generality.

We explain the connection between the notion of device in Definition 6.1 and a prover in Protocol 1. Given a device $D = (\phi, \Pi, M)$, we can define actions for the prover in Protocol 1 as follows. The prover is initialized in state ϕ. When a round of the protocol is initiated, the prover measures register Y in the computational basis and returns the outcome $y \in \mathcal{Y}$. We always assume that the prover directly measures the register, as any pre-processing unitary can be incorporated in the definition of the state ϕ. When sent challenge $C = 0$ (resp. $C = 1$), the prover measures register D using the device's projective measurement $\{M_y^{(u,d)}\}$ (resp. $\{\Pi_y^{(b,x)}\}$), and returns the outcome to the verifier.

Definition 6.2. We say that a device $D = (\phi, \Pi, M)$ is *efficient* if

1. There is a polynomial-size circuit to prepare ϕ, given the NTCF key k;
2. For every $y \in \mathcal{Y}$, the measurements $\{M_y^{(u,d)}\}$ and $\{\Pi_y^{(b,x)}\}$ can be implemented by polynomial-size circuits.

Using the definition of an NTCF family (Definition 3.1), it is straightforward to verify that the device associated with the "honest" prover described in Section 5.3 is efficient.

We introduce notation related to the post-measurement states generated by a device in Protocol 1. An execution of Protocol 1 involves a choice of round types $g \in \{0,1\}^N$ and challenges $c \in \{0,1\}^N$ by the verifier, and a sequence of outputs $o \in \{0,1,2\}^N$ computed by the verifier as a function of the answers provided by the device. Here, in case $g = 0$ (test round) we use $o \in \{0,1\}$ to denote the outcome of the test (called W in the protocol description), and in case $g = 1$ (generation round) we use $o \in \{0,1,2\}$ such that $o = 2$ in case $W = 0$, and $o = O$ as recorded by the verifier in case $W = 1$. We call the tuple (g, c, o) the transcript of the protocol; it contains all the information relevant to the verifier's final acceptance decision and to the extraction of randomness. Additional information such as the choice of NTCF key and the prover's complete answers (including the value y) is discarded for ease of presentation. We let ACC denote the set of transcripts (g, c, o) that are accepted by the verifier in the last step of the protocol, i.e. such that $\sum_{i:g_i=0} o_i \geq (1-\gamma)qN$.

Definition 6.3. Let $D = (\phi, \Pi, M)$ be a device. For any transcript (g, c, o) for an execution of Protocol 1 with D, let ϕ_{D}^{co} be the post-measurement state of the device, conditioned on having received challenges c and returned outcomes o. The joint state of the transcript and the device at the end of the N rounds (but before the verifier's decision to abort) is

$$\phi_{\mathsf{COD}}^{(N)} = \sum_{g,c,o} q(g,c)\, |c\rangle\langle c|_{\mathsf{C}} \otimes |o\rangle\langle o|_{\mathsf{O}} \otimes \phi_{\mathsf{D}}^{co}\,, \tag{45}$$

where $q(g,c)$ is the probability that the sequence of round types and challenges (g,c) is chosen by the verifier in the protocol.

We write $|\phi\rangle_{\mathsf{DE}}$ for a purification of the initial state ϕ_{D} of the device, with E the purifying register, and ρ_{E}^{co} for the post-measurement state on register E conditioned on the transcript being (c, o).

6.2 Simplified devices

Next we introduce a simplified notion of device, that can be used to model the actions of a prover in the simplified protocol, Protocol 2 (Figure 2).

Definition 6.4. A *simplified device* is a tuple (ϕ, Π, M, K) such that:

1. $\phi = \{\phi_y\}_{y\in\mathcal{Y}} \subseteq \mathrm{Pos}(\mathcal{H}_\mathsf{D})$ is a family of positive semidefinite operators on an arbitrary space $\mathcal{H}_\mathsf{D}$;
2. For each $y \in \mathcal{Y}$, $\{M_y^0, M_y^1 = \mathrm{Id} - M_y^0\}$, $\{\Pi_y^0, \Pi_y^1, \Pi_y^2 = \mathrm{Id} - \Pi_y^0 - \Pi_y^1\}$, and $\{K_y^0, K_y^1 = \mathrm{Id} - K_y^0\}$ are projective measurements on $\mathcal{H}_\mathsf{D}$;
3. For each $y \in \mathcal{Y}$, the measurement operators K_y commute with the M_y and with the Π_y. (M_y and Π_y do not necessarily commute with each other.)

We introduce a quantity called *overlap* that measures how "incompatible" a simplified device's measurements are. This measure is analogous to the measure of overlap used to quantify incompatibility in the derivation of entropic uncertainty relations (see e.g. [MU88]).

Definition 6.5. Given a simplified device $D = (\phi, \Pi, M, K)$, the *overlap* of D is

$$\Delta(D) \;=\; \max_{y\in\mathcal{Y}} \left\| K_y^0 \left(\Pi_y^0 M_y^1 \Pi_y^0 + \Pi_y^1 M_y^1 \Pi_y^1\right) \right\| .$$

Note that the overlap only quantifies the measurement incompatibility in the "good subspace" K_y^0.
To any simplified device $D = (\phi, \Pi, M, K)$ we associate the post-measurement states

$$\begin{aligned}
\forall e \in \{0,1\}, \quad \phi_{00}^e &= \sum_{y\in\mathcal{Y}} |y\rangle\langle y| \otimes M_y^e \phi_y M_y^e \,, \\
\forall e,k \in \{0,1\}, \quad \phi_{01}^{ek} &= \sum_{y\in\mathcal{Y}} |y\rangle\langle y| \otimes K_y^k M_y^e \phi_y M_y^e K_y^k \,, \\
\forall v \in \{0,1,2\}, \quad \phi_1^v &= \sum_{y\in\mathcal{Y}} |y\rangle\langle y| \otimes \Pi_y^v \phi_y \Pi_y^v \,.
\end{aligned} \tag{46}$$

A simplified device can be used in the simplified protocol in a straightforward way: upon receipt of a challenge $C = 0$ (resp. $C = 1$), the device first samples an $y \in \mathcal{Y}$ according to the distribution with weights $\mathrm{Tr}(\phi_y)$. It then performs the projective measurement $\{M_y^0, M_y^1\}$ followed by, if $T = 1$, $\{K_y^0, K_y^1\}$ (resp. $\{\Pi_y^0, \Pi_y^1, \Pi_y^2\}$) on ϕ_y, and returns the outcomes $e, k \in \{0,1\}$ (resp. $v \in \{0,1,2\}$) to the verifier.

Definition 6.6. Let $D = (\phi, \Pi, M, K)$ be a simplified device. For any transcript (g, c, t, o, k) for an execution of Protocol 1 with D, let ϕ_D^{ctok} be the post-measurement state of the device, conditioned on having received challenges (c, t) and returned outcomes (o, k). The joint state of the transcript and the device at the end of the N rounds (but before the verifier's decision to abort) of the protocol is

$$\phi_{\mathsf{CTOKD}}^{(N)} \;=\; \sum_{g,c,t,o,k} q(g,c,t)\, |c,t\rangle\langle c,t|_\mathsf{CT} \otimes |o,k\rangle\langle o,k|_\mathsf{OK} \otimes \phi_\mathsf{D}^{ctok} \,, \tag{47}$$

where $q(g,c,t) = q(g,c)\kappa(t)$ with $\kappa(t) = \prod_i \kappa^{t_i}(1-\kappa)^{1-t_i}$ is the probability that the sequence of round types and challenges (g,c,t) is chosen by the verifier in the protocol.

7 Single-round analysis

In this section we consider the behavior of an arbitrary device D in a single round of the randomness expansion protocol, Protocol 1 in Figure 1. Our goal is to introduce a simplified device D' such that analyzing

the randomness generation properties of D' is easier than it is for D, and such that bounds on the amount of randomness generated by D' in the simplified protocol, Protocol 2 in Figure 2, imply bounds on the amount of randomness generated by D in Protocol 1. Throughout the section we fix an NTCF family $\mathcal{F}$ (Definition 3.1) and a key $k \in \mathcal{K}_{\mathcal{F}}$ sampled according to $\text{GEN}(1^\lambda)$, for a parameter λ that plays the role of security parameter.

7.1 A constraint on the measurements of any efficient device

We start with a lemma showing that for any efficient device $D = (\phi, \Pi, M)$, the measurements Π and M must be strongly incompatible, in the sense that if the device first measures Π, and then measures M, it is unable to determine if the pair (u, d) returned by M corresponds to a valid pair, i.e. $(u, d) \in V_{y,0}$. Indeed, if this were the case the device could be used to violate the hardcore bit property (12). Recall the definition of the set $\hat{G}_y \subseteq \{0,1\}^w$ in (44).

Lemma 7.1. *Let $D = (\phi, \Pi, M)$ be an efficient device. Define a sub-normalized density*

$$\begin{aligned}\tilde{\phi}_{\mathsf{YBXD}} &= \sum_{y \in \mathcal{Y}} |y\rangle\langle y|_{\mathsf{Y}} \otimes \sum_{b \in \{0,1\}} |b, x_b\rangle\langle b, x_b|_{\mathsf{BX}} \otimes \Pi_y^{(b,x_b)} \, \phi_y \, \Pi_y^{(b,x_b)} \\ &= \sum_{b \in \{0,1\}} |b, x_b\rangle\langle b, x_b|_{\mathsf{BX}} \otimes \tilde{\phi}_{\mathsf{YD}}^{(b)} \,. \end{aligned} \tag{48}$$

Let

$$\begin{aligned}\sigma_0 &= \sum_{b \in \{0,1\}} |b, x_b\rangle\langle b, x_b|_{\mathsf{BX}} \otimes \sum_{(u,d) \in V_{y,0}} |u, d\rangle\langle u, d|_{\mathsf{U}} \otimes (\mathrm{Id}_{\mathsf{Y}} \otimes M_y^{(u,d)}) \tilde{\phi}_{\mathsf{YD}}^{(b)} (\mathrm{Id}_{\mathsf{Y}} \otimes M_y^{(u,d)}) \,, \\ \sigma_1 &= \sum_{b \in \{0,1\}} |b, x_b\rangle\langle b, x_b|_{\mathsf{BX}} \otimes \sum_{(u,d) \notin V_{y,0}} 1_{d \in \hat{G}_y} |u, d\rangle\langle u, d|_{\mathsf{U}} \otimes (\mathrm{Id}_{\mathsf{Y}} \otimes M_y^{(u,d)}) \tilde{\phi}_{\mathsf{YD}}^{(b)} (\mathrm{Id}_{\mathsf{Y}} \otimes M_y^{(u,d)}) \,. \end{aligned} \tag{49}$$

Then σ_0 and σ_1 are computationally indistinguishable.

Proof. Suppose for contradiction that there exists an efficient observable O such that

$$\mathrm{Tr}(O(\sigma_0 - \sigma_1)) \geq \mu \,, \tag{50}$$

for some non-negligible function $\mu(\lambda)$. Consider the following efficient procedure. The procedure first prepares the state $\tilde{\phi}_{\mathsf{YBXD}}$ in (48). This can be done efficiently by first preparing ϕ_{YD}, then measuring a $y \in \mathcal{Y}$, then applying the measurement $\{\Pi_y^{(b,x)}\}$ to ϕ_y, and returning a special abort symbol if the outcome is invalid, i.e. $\text{CHK}_{\mathcal{F}}(k, b, x, y) = 0$.

The procedure then applies the measurement $\{M_y^{(u,d)}\}$ to $\tilde{\phi}_{\mathsf{YBXD}}$, obtaining an outcome (u, d). At this point, conditioned on the event that $d \in \hat{G}_y$, depending on whether $(u, d) \in V_{y,0}$ or $(u, d) \notin V_{y,0}$ the procedure has either prepared σ_0 or σ_1. Finally, the procedure measures O to obtain a bit v, and returns $(b, x, d, v \oplus u)$. This defines an efficient procedure. Moreover, using (50) it follows that the procedure violates the hardcore bit property (12). (The cases where $d \notin \hat{G}_y$ are not taken into account by the hardcore bit property, so it is sufficient to have a good distinguishing ability conditioned on $d \in \hat{G}_y$.) □

7.2 Angles between incompatible measurements

We show a general lemma that argues about the principal angles between two binary-outcome measurements that have a certain form of incompatibility.

Lemma 7.2. *Let Π, M be two orthogonal projections on $\mathcal{H}$ and ϕ a state on $\mathcal{H}$. Let $\gamma = 1 - \mathrm{Tr}(M\phi)$ and*

$$\mu \;=\; \Big|\frac{1}{2} - \mathrm{Tr}\big(M\Pi\phi\Pi\big) - \mathrm{Tr}\big(M(\mathrm{Id}-\Pi)\phi(\mathrm{Id}-\Pi)\big)\Big| \,.$$

Let $\frac{1}{2} < \omega \leq 1$. Let K be the orthogonal projection on the direct sum of eigenspaces of $\Pi M \Pi + (\mathrm{Id}-\Pi)M(\mathrm{Id}-\Pi)$ with associated eigenvalue in $[1-\omega, \omega]$. Then

$$\mathrm{Tr}\big((\mathrm{Id}-K)\phi\big) \;\leq\; \frac{2\mu + 10\sqrt{\gamma}}{1 - 4\omega(1-\omega)} \,.$$

Proof. Using Jordan's lemma we find a basis of $\mathcal{H}$ in which

$$M = \oplus_j \begin{pmatrix} c_j^2 & c_j s_j \\ c_j s_j & s_j^2 \end{pmatrix} \quad \text{and} \quad \Pi = \oplus_j \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \,, \tag{51}$$

where $c_j = \cos\theta_j$, $s_j = \sin\theta_j$, for some angles θ_j. There may be 1-dimensional blocks in the Jordan decomposition, but up to adding a few dimensions these can be identified with two-dimensional blocks such that $c_j^2 \in \{0,1\}$. Let K be the orthogonal projection on those 2-dimensional blocks such that $\min(c_j^2, s_j^2) \geq 1-\omega$. Note that K commutes with both M and Π, but not necessarily with ϕ. It is easy to verify that this definition of K coincides with the definition given in the lemma.

Suppose first that $\gamma = 0$. Then ϕ is supported on the range of M. For any block j, let P_j be the projection on the block and $\alpha_j = \mathrm{Tr}(P_j\phi)$. It follows from the decomposition in (51) and the definition of μ that

$$\Big|\frac{1}{2} - \sum_j \alpha_j \left(c_j^4 + s_j^4\right)\Big| \;\leq\; \mu \,. \tag{52}$$

Using that for j such that $\min(c_j^2, s_j^2) \geq 1-\omega$ we have

$$c_j^4 + s_j^4 \;=\; 1 - 2\max(c_j^2, s_j^2)\big(1 - \max(c_j^2, s_j^2)\big) \;\geq\; \frac{1}{2} + \Big(\frac{1}{2} - 2\omega(1-\omega)\Big) \,,$$

and $c_j^4 + s_j^4 \geq \frac{1}{2}$ always, it follows from (52) that for any $\omega > \frac{1}{2}$,

$$\mathrm{Tr}\big((\mathrm{Id}-K)\phi\big) \;\leq\; \frac{2\mu}{1 - 4\omega(1-\omega)} \,. \tag{53}$$

Next consider the case where $\gamma > 0$. Assume $\mathrm{Tr}(M\phi) > 0$, as otherwise the lemma is trivial. Let $\phi' = M\phi M / \mathrm{Tr}(M\phi)$. By the gentle measurement lemma (see e.g. [Wil13, Lemma 9.4.1]),

$$\big\|\phi' - \phi\big\|_1 \;\leq\; 2\sqrt{\gamma} \,. \tag{54}$$

Using the definition of μ, it follows that

$$\Big|\frac{1}{2} - \mathrm{Tr}\big(M\Pi\phi'\Pi\big) - \mathrm{Tr}\big(M(\mathrm{Id}-\Pi)\phi'(\mathrm{Id}-\Pi)\big)\Big| \;\leq\; \mu + 4\sqrt{\gamma} \,.$$

Applying the same reasoning as for the case $\gamma = 0$ yields an analogue of (53), with ϕ' instead of ϕ on the left-hand side and $\mu + 4\sqrt{\gamma}$ instead of μ on the right-hand side. Finally, using again (54) the same bound transfers to ϕ up to an additional loss of $2\sqrt{\gamma}$. □

7.3 Simulating an efficient device using a simplified device

Recall the definitions of a simplified device (Definition 6.4) and of the overlap of a simplified device (Definition 6.5). Recall also the definition of post-measurement states $\{\phi^{co}\}$ associated with a device $D = (\phi, \Pi, M)$ given in Definition 6.3, and of post-measurement states $\{(\phi')^{ctok}\}$ associated with a simplified device $D' = (\phi', \Pi', M', K)$ given in Definition 6.6. These ensembles of states provide a means to meaningfully compare a device D and a simplified device D'. We record this in the following definition.

Definition 7.3. Let $D = (\phi, \Pi, M)$ be a device and $D' = (\phi', \Pi', M', K)$ a simplified device. We say that D' *simulates* D if for every $(c, o) \in \{0,1\}^N \times \{0,1,2\}^N$ and $t = 0^N$ the states ϕ^{co} and $(\phi')^{cto}$ are identical.

The following proposition shows that any efficient device can be simulated by a simplified device whose measurements generally make an angle that is bounded away from 1. As in Lemma 7.1, the only assumption required on the efficient device is that it does not break the hardcore bit property (12).

Proposition 7.4. *Let $D = (\phi, \Pi, M)$ be an efficient device and $\frac{1}{2} < \omega \leq 1$. Then there is a (not necessarily efficient) simplified device $\tilde{D} = (\phi, \tilde{\Pi}, \tilde{M}, K)$ such that the following hold:*

1. *$\tilde{D}$ has overlap $\Delta(\tilde{D}) \leq \omega$;*
2. *The simplified device $\tilde{D}$ simulates the device D;*
3. *For any advice states $\phi' = \{\phi'_y\}$ that are independent from the key $k \in \mathcal{K}_{\mathcal{F}}$ (see Definition 2.5) it holds that*
$$\sum_y \mathrm{Tr}(K^1_y \phi'_y) \;\leq\; C \sqrt{\sum_y \mathrm{Tr}(\tilde{M}^1_y \phi'_y)} + \mathrm{negl}(\lambda)\,, \tag{55}$$
where $C > 0$ is a constant depending only on ω.

Proof. For each $y \in \mathcal{Y}$ let

$$\hat{M}_y = \sum_{(u,d)\colon d \notin \hat{G}_y} M_y^{(u,d)}\,, \qquad M_y = \sum_{(u,d)\in V_{y,0}} M_y^{(u,d)} + \frac{1}{2}\hat{M}_y\,,$$

and for $b \in \{0,1\}$, $\Pi^b_y = \Pi^{(b,x_b)}_y$. By introducing an isometry $U_y : \mathcal{H}_{\mathsf{D}} \to \mathcal{H}_{\mathsf{D}'}$ into a larger space, we can embed M_y into a projection $\overline{M}_y$ such that $M_y = U_y^\dagger \overline{M}_y U_y$. For $b \in \{0,1\}$ let $\overline{\Pi}^b_y$ be such that $U_y^\dagger \overline{\Pi}^b_y U_y = \Pi^b_y$.

The device $\tilde{D}$ is defined as follows. The device first measures an $y \in \mathcal{Y}$ exactly as D would. It then applies the isometry U_y. This defines the states $\{\phi'_y\}$.

- The measurement $\{\tilde{M}^0_y, \tilde{M}^1_y\}$ is defined as follows. The device coherently performs the measurement $\{\overline{M}^{(u,d)}_y\}$. If $d \notin \hat{G}_y$ the device returns a random outcome. Otherwise, if $(u,d) \in V_{y,0}$ it returns a 0, and 1 if not.
- The measurement $\{\tilde{\Pi}^0_y, \tilde{\Pi}^1_y, \tilde{\Pi}^2_y\}$ is defined as follows. The device first coherently performs the measurement $\{\overline{\Pi}^{(b,x)}_y\}$. If an outcome $(b,x) \in V_{y,1}$ is obtained the device returns $v = b$. Otherwise the device returns $v = 2$.

- Let K_y be the projection obtained by applying Lemma 7.2 to the projections $\Pi = \overline{\Pi}_y^0$ and $M = \overline{M}_y$ and the state

$$\phi \;=\; \frac{(\overline{\Pi}_y^0 + \overline{\Pi}_y^1)\phi'_y(\overline{\Pi}_y^0 + \overline{\Pi}_y^1)}{\mathrm{Tr}\big((\overline{\Pi}_y^0 + \overline{\Pi}_y^1)\phi'_y\big)} \,.$$

  The measurement $\{K_{y'}^0, K_y^1\}$ is defined by setting

$$K_y^0 = (\overline{\Pi}_y^0 + \overline{\Pi}_y^1)K + (\mathrm{Id} - \overline{\Pi}_y^0 - \overline{\Pi}_y^1) \quad \text{and} \quad K_y^1 = (\overline{\Pi}_y^0 + \overline{\Pi}_y^1)(\mathrm{Id} - K) \,.$$

The first two conditions on D' claimed in the lemma follow by definition. The overlap property holds by definition of K_y^0. For the simulation property, note that it is possible for D' to further measure the post-measurement states to locally obtain an equation, or a pre-image, as D would have; this guarantees that the post-measurement states of the two devices are identical in each round.

It remains to show the third item. It follows from computational indistinguishability of σ_0 and σ_1 shown in Lemma 7.1 that both operators have a trace that is within negligible of each other. Using the notation introduced here, and in particular the definition of $\overline{M}_y$, this implies that the difference

$$\Big| \sum_{b\in\{0,1\}} \mathrm{Tr}\big(\overline{M}_y \overline{\Pi}_y^b \phi'_y \overline{\Pi}_y^b\big) - \sum_{b\in\{0,1\}} \mathrm{Tr}\big((\mathrm{Id} - \overline{M}_y)\overline{\Pi}_y^b \phi'_y \overline{\Pi}_y^b\big) \Big|$$

is negligible. Since the two expressions sum to $\mathrm{Tr}\big((\mathrm{Id} - \overline{\Pi}_y^2)\phi'_y\big)$, it follows that, letting

$$\tilde{\phi}_y \;=\; \frac{(\mathrm{Id} - \overline{\Pi}_y^2)\phi'_y(\mathrm{Id} - \overline{\Pi}_y^2)}{\mathrm{Tr}\big((\mathrm{Id} - \overline{\Pi}_y^2)\phi'_y\big)} \,, \tag{56}$$

we get that

$$\mathrm{Tr}(\overline{M}_y \overline{\Pi}_y^0 \tilde{\phi}_y \overline{\Pi}_y^0) + \mathrm{Tr}(\overline{M}_y \overline{\Pi}_y^1 \tilde{\phi}_y \overline{\Pi}_y^1)$$

is within negligible of $\frac{1}{2}$. To conclude we apply Lemma 7.2 to the operators $\Pi = \overline{\Pi}^0$ and $M = \overline{M}_y$. The conclusion of the lemma gives that

$$\mathrm{Tr}\big((\mathrm{Id} - K_y)\tilde{\phi}_y\big) \;\leq\; C\sqrt{\mathrm{Tr}((\mathrm{Id} - \tilde{M}_y^0)\tilde{\phi}_y)} + \mathrm{negl}(\lambda) \,, \tag{57}$$

for some universal constant C (depending on ω). Using the definition (56) of $\tilde{\phi}_y$ and $\mathrm{Tr}\big((\mathrm{Id} - \overline{\Pi}_y^2)\phi_y\big) \leq 1$, (57) implies

$$\mathrm{Tr}\big(K_y^1 \phi'_y\big) \;\leq\; C\sqrt{\mathrm{Tr}((\mathrm{Id} - \tilde{M}_y^0)\phi'_y)} + \mathrm{negl}(\lambda) \,.$$

Summing this bound over all y and using concavity of the square root concludes the proof. □

8 Accumulating randomness across multiple rounds

To analyze the randomness generated by a device in the randomness expansion protocol we proceed in two steps. First, we show that the randomness generated by the device can be related to the randomness generated by the simplified device $\tilde{D}$ that is associated to it by Proposition 7.4, when it is used as a device in the simplified protocol, Protocol 2. This is done in Section 8.1. Then, in Section 8.2 we analyze the randomness generated in a single round of the simplified protocol, and in Section 8.3 we analyze multiple rounds of the protocol.

8.1 Reduction to the simplified protocol

Let $D = (\phi, \Pi, M, K)$ be a simplified device. The main difference between the behavior of the simplified device and the original device it is derived from is that the simplified device (sometimes) performs an additional projective measurement $\{K^0, K^1\}$, in addition to the "equation" measurement $\{M^0, M^1\}$. (Recall that in protocol 2, the device performs the measurement whenever the verifier sends a challenge bit $T = 1$, which happens with probability $\Pr(T = 1) = \kappa$ in the test rounds.)

In order to analyze the randomness generated by the original device in Protocol 1, it will be convenient to obtain the guarantee that, in most test rounds of Protocol 1, the state of the device lies largely within the "good subspace" $K = K^0$. Recall the definition of the states $\{\phi^{ctok}\}$ associated with the simplified device in Definition 6.6. Let

$$|\phi^{co}\rangle \;=\; \sum_k |\phi^{ctok}\rangle \;=\; \sum_k P_{ct}^{ok}\,|\phi\rangle\;, \tag{58}$$

where P_{ct}^{ko} is notation for the operator that corresponds to applying the device's (projective) measurement operators M, Π and K indicated by c and t respectively, and obtaining the sequence of outcomes o and k respectively. The fact that $|\phi^{co}\rangle$ does not depend on t is justified by the fact that $\{K, \mathrm{Id} - K\}$ is a projective measurement.

Our goal is to bound the contribution to (58) of terms $P_{ct}^{ok}|\phi\rangle$ that correspond to a large fraction of $(\mathrm{Id} - K)$ ("bad subspace") outcomes, i.e. such that the Hamming weight $|k|$ of the string k is large. Establishing the right bound is made delicate by the possibility of interference between the branches. We first state and prove a general lemma, and then show how the lemma can be applied in our context.

Lemma 8.1. *Let n be an integer, $0 < \kappa < 1$, and $T = (T_1, \ldots, T_n)$ a sequence of independent Bernoulli random variables such that for any $t \in \{0,1\}^n$, $\Pr(T = t) = \kappa(t) = \prod_i \kappa^{t_i}(1-\kappa)^{1-t_i}$. Let $M = (M_1, \ldots, M_n)$ and $K = (K_1, \ldots, K_{|T|})$ be arbitrary sequences of random variables over $\{0,1\}$, that may be correlated between themselves and with T. For an integer $i \in \{1, \ldots, n\}$ write $(T, M, K)_{<i}$ for the triple formed by the length-$(i-1)$ prefixes of T and M, and the length-$|T_{<i}|$ prefix of K.*[9]

Assume that there is a monotone concave function $g : [0,1] \to [0,1]$ such that $g(0) = 0$, $g(x) \geq x$ for all $x \in [0,1]$, and for any $i \in \{1, \ldots, n\}$ and any sequences $t, m \in \{0,1\}^i$ and $k \in \{0,1\}^{|t|}$ it holds that

$$\begin{aligned}\Pr\left(K_{|t|+1} = 1 \,\middle|\, (T,M,K)_{<i} = (t,m,k),\ T_{i+1} = 1\right)& \\ \leq g\Big(\Pr\left(M_{i+1} = 0 \,\middle|\, (T,M,K)_{<i} = (t,m,k),\ T_{i+1} = 1\right)\Big)\;.&\end{aligned} \tag{59}$$

Then for any $0 < \gamma, \kappa, \eta < 1$ such that $g(c_1(\sqrt{\kappa} + \gamma/\kappa)) \leq c_2\eta$ for large enough constants $c_1, c_2 > 0$,

$$\sum_{t \in \{0,1\}^n} \kappa(t) \sum_{m:\,|m| \geq (1-\gamma)n} \Big(\sum_{k:\,|k| > \eta\kappa n} \sqrt{\Pr\left((T,M,K) = (t,m,k)\right)}\Big)^2 \;\leq\; C_0\, 2^{-\kappa n}\;, \tag{60}$$

where $C_0 > 0$ is a constant depending on γ, κ, η.

Intuitively, the lemma holds because the condition $|m| \geq (1-\gamma)n$ ensures that the outcome $M_i = 0$ is fairly unlikely, in which case (59) implies that whenever $T_i = 1$ the outcome $K_j = 1$, where j is the number of nonzero entries of T in indices less or equal to i, should also be unlikely. The proof is made a little difficult by the square roots, whose presence is motivated by the application to norms of quantum states

[9] Recall that we write $|T|$ for the Hamming weight of the string T. Here, we think of each K_j as a random variable that is correlated with the random variable M_i, where i is the index of the j-th non-zero entry of T.

detailed later. Nevertheless, to understand the statement of the lemma it may be useful to consider the case when all M_i (resp. K_j) are independent and identically distributed, and the square root and the square are not present. In this case, the lemma reduces to showing that if

$$\mathcal{G} \;=\; \Big\{ m \in \{0,1\}^n \,:\, |m| \geq (1-\gamma)n \Big\}\,, \quad \mathcal{B} = \Big\{ (t,m) \in \{0,1\}^{2n},\, k \in \{0,1\}^{|t|} \,:\, |k| \geq \eta\kappa n \Big\}\,, \tag{61}$$

then $\Pr(\mathcal{G} \wedge \mathcal{B}) \leq C_0 2^{-\kappa n}$. As a first step, note that we may safely assume that $\Pr(M_i = 0) \leq 2\gamma$, as otherwise by a Chernoff bound $\Pr(\mathcal{G}) \leq e^{-\Omega(\eta^2 n)} \leq C_0 2^{-\kappa n}$ provided $\eta^2 \gg \kappa$. Using (59) it follows that $\Pr(K_j = 1) \leq g(2\gamma)$, so that applying Bennett's inequality,

$$\Pr(\mathcal{B}) \;\leq\; e^{-\Omega(h(\eta/g(2\gamma))\kappa n)} \;\leq\; C_0\, 2^{-\kappa n}\,,$$

where $h(x) = (1+u)\log(1+u) - u$ and the second inequality holds provided $\eta \gg g(2\gamma)$. This completes the argument. To extend it to the general case, we use two tail bounds for martingales that replace the use of the Chernoff bound and Bennett's inequality respectively. The first is Azuma's inequality.

Theorem 8.2 (Azuma's inequality)**.** *Let $(\xi_i, \mathcal{F}_i)_{0\leq i\leq n}$ be a martingale difference sequence such that $\xi_0 = 0$ and $\xi_i \leq 1$ for each $i \in \{1,\ldots,n\}$. Then for any $t \geq 0$,*

$$\Pr\Big(\Big| \sum_{i=1}^n \xi_i \Big| \;\geq\; tn \Big) \;\leq\; 2e^{-\frac{t^2}{2}n}\,.$$

The second is a version of Bennett's inequality for martingales.

Theorem 8.3 (Corollary 2.2 in [FGL12])**.** *Let $(\xi_i, \mathcal{F}_i)_{0\leq i\leq n}$ be a supermartingale difference sequence such that $\xi_0 = 0$ and $\xi_i \leq 1$ for each $i \in \{1,\ldots,n\}$. Let*

$$X_n \;=\; \sum_{i=1}^n \xi_i \quad \textit{and} \quad \langle X\rangle_n \;=\; \sum_{i=1}^n \mathrm{E}\big[\xi_i^2 \big|\, \mathcal{F}_{i-1}\big]\,.$$

Then for any $t \geq 0$ and $v > 0$,

$$\Pr\Big(\big|X_n\big| \geq tn \;\textit{and}\; \langle X\rangle_n \leq v^2 n \Big) \;\leq\; e^{-\frac{t}{2}\operatorname{arcsinh}\left(\frac{t}{2v^2}\right)n}\,.$$

We give the proof of Lemma 8.1.

Proof of Lemma 8.1. We reduce the proof of (60) to a sequence of martingale tail bounds. Define a filtration $(\mathcal{F}_1,\ldots,\mathcal{F}_i,\ldots,\mathcal{F}_n)$ where $\mathcal{F}_i$ is the σ-algebra generated by $(M,T,K)_i$. Let $\mathcal{F}_{<i} = \cap_{j<i}\mathcal{F}_j$. Recall the definition of the events $\mathcal{G}$ and $\mathcal{B}$ in (61). The proof proceeds in 3 steps.

First step: conditional expectations of M. [Uses the assumption: $\delta_1^2 \gg \kappa$.]
For $i \in \{1,\ldots,n\}$ let $Z_i = M_i - \mathrm{E}[M_i|\mathcal{F}_{<i}]$ and $W_i = Z_1 + \cdots + Z_i$. Then the sequence $(W_1,\ldots,W_n)$ is a martingale such that $|W_i - W_{i-1}| \leq 1$. Applying Azuma's inequality, it follows that for any $\delta_1 > 0$,

$$\Pr\Big(\Big| \sum_{i=1}^n Z_i \Big| \geq \delta_1 n \Big) \;\leq\; 2\, e^{-\frac{\delta_1^2}{2}n}\,. \tag{62}$$

Let $\delta_1 > 0$ be large enough such that the right-hand side of (62) is less than $2^{-(C+1)\kappa n}$, for some constant C to be determined below. Let $\delta_1' = \delta_1 + \gamma$ and

$$\mathcal{B}' \;=\; \Big\{(t,m,k) : \sum_{i=1}^{n} \mathrm{E}\big[M_i \big| \, (T,M,K)_{<i} = (t,m,k)_{<i}\big] \;\le\; (1-\delta_1')n\Big\} \,.$$

Then,

$$\begin{aligned}\sum_{m\in\mathcal{G}} \sum_{t} \kappa(t) \Big(\sum_{k:(t,m,k)\in\mathcal{B}'} \sqrt{\Pr\big((M,K)=(m,k)|T=t\big)}\Big)^2 \\ &\le 2^{C\kappa n}\Big(\sum_{m\in\mathcal{G}} \sum_{t,k:\,(t,m,k)\in\mathcal{B}'} \Pr\big((T,M,K)=(t,m,k)\big)\Big) + 2^{-\kappa n} \\ &\le 2\cdot 2^{-\kappa n}\,, \end{aligned} \tag{63}$$

where the first inequality uses the Cauchy-Schwarz inequality and the fact that by the Chernoff bound, for C large enough, $\sum_{|t|\ge C\kappa n} \kappa(t) 2^{|t|} \le 2^{-\kappa n}$, and the second inequality follows from (62) since the event that $m \in \mathcal{G}$ and $(t,m,k) \in \mathcal{B}'$ implies $|\sum Z_i| \ge \delta_1 n$.

Second step: conditional expectations of $T(1-M)$. [Uses the assumption: $\delta_1' \ll \delta_2$.]
For $i \in \{1,\dots,n\}$ let $Z_i' = T_i(1-M_i) - \mathrm{E}[T_i(1-M_i)|\mathcal{F}_{<i}]$ and $W_i' = Z_1' + \cdots + Z_i'$. Then the sequence $(W_1',\dots,W_n')$ is a martingale such that $|W_i' - W_{i-1}'| \le 1$. Let $v_{Z'}^2 = \sum_i \mathrm{E}[|Z_i'|^2|\mathcal{F}_{<i}]$. For $(t,m,k) \notin \mathcal{B}'$, using that T_i is independent from M_i and $\mathrm{E}[T_i] = \kappa$ it holds that $v_{Z'}^2 \le \delta_1'\kappa n$. Let $v^2 = \delta_1'\kappa n$. Applying Theorem 8.3, for any $\delta_2 > 0$,

$$\Pr\Big(\Big|\sum Z_i'\Big| \ge \delta_2\kappa n \;\wedge\; v_{Z'}^2 \le v^2 n\Big) \;\le\; e^{-\frac{1}{2}\delta_2\kappa \operatorname{arcsinh}\left(\frac{\delta_2}{2\delta_1'}\right)n} \,. \tag{64}$$

Assume δ_1' small enough, as a function of δ_2, such that the right-hand side in (64) is less than $2^{-(C+1)\kappa n}$. Let $\delta_2' = \delta_2 + \gamma/\kappa$ and

$$\mathcal{B}'' = \Big\{(t,m,k) \notin \mathcal{B}' : \sum_i \mathrm{E}\big[T_i(1-M_i)\big| \, (T,M,K)_{<i} = (t,m,k)_{<i}\big] \;\ge\; \delta_2'\kappa n\Big\} \,.$$

Then similarly to (63) we get

$$\sum_{m\in\mathcal{G}} \sum_{t} \kappa(t) \Big(\sum_{k:(t,m,k)\in\mathcal{B}''} \sqrt{\Pr\big((M,K)=(m,k)|T=t\big)}\Big)^2 \le 2^{-\kappa n} \,. \tag{65}$$

Third step: conditional expectations of $T(1-M)K$. [Uses the assumption: $g(\delta_2') \ll \delta_3$.]
Using assumption (59) and concavity of g, for any $(t,m,k) \notin (\mathcal{B}'' \cup \mathcal{B}')$ it holds that

$$\sum_i \mathrm{E}\big[T_i(1-M_i)K_i\big| \, (T,M,K)_{<i} = (t,m,k)_{<i}\big] \;\le\; g(\delta_2')\kappa n \,. \tag{66}$$

For $i \in \{1,\dots,n\}$ let $Z_i'' = T_i(1-M_i)K_i - \mathrm{E}[T_i(1-M_i)K_i|\mathcal{F}_{<i}]$ and $W_i'' = Z_1'' + \cdots + Z_i''$. Then the sequence $(W_1'',\dots,W_n'')$ is a martingale such that $|W_i'' - W_{i-1}''| \le 1$ and by (66), $v_{Z''}^2 = \sum_i \mathrm{E}[|Z_i''|^2|\mathcal{F}_{<i}] \le g(\delta_2')\kappa n$. Applying Theorem 8.3 once more, for any $\delta_3 > 0$,

$$\Pr\Big(\Big|\sum Z_i''\Big| \ge \delta_3\kappa n \;\wedge\; \overline{\mathcal{B}'' \cup \mathcal{B}'}\Big) \;\le\; e^{-\frac{1}{2}\delta_3\kappa \operatorname{arcsinh}\left(\frac{\delta_3}{2g(\delta_2')}\right)n} \,.$$

Assume δ_2' small enough, as a function of δ_3, such that the right-hand side is less than $2^{-(C+1)\kappa n}$. Assume further that $\delta_3 + \gamma/\kappa \leq \eta$. Let $\mathcal{B}''' = \overline{\mathcal{B}'' \cup \mathcal{B}'} \cap \mathcal{B}$. Then it follows as in (63), (65) that

$$\sum_{m\in\mathcal{G}} \sum_{t} \Pr(T=t) \Big(\sum_{k:(t,m,k)\in\mathcal{B}'''} \sqrt{\Pr\left((M,K)=(m,k)|T=t\right)} \Big)^2 \leq 2^{-\kappa n} \,. \tag{67}$$

Combining (63), (65) and (67) with the triangle inequality proves the lemma. □

Recall the definition of the states $|\phi^{ctok}\rangle$ in (58). For a parameter $\eta > 0$ and any $t \in \{0,1\}^N$ let

$$|\tilde{\phi}^{cto}\rangle \;=\; \sum_{k:|k|\leq \eta\kappa q N} |\phi^{ctok}\rangle \,, \tag{68}$$

and $\tilde{\phi}^{cto}$ the sub-normalized density obtained by taking the partial trace of $|\tilde{\phi}^{cto}\rangle$ over register E.

Corollary 8.4. *Let $D = (\phi, \Pi, M, K)$ be a simplified device such that condition* (55) *from Proposition 7.4 holds. Then for any $0 < \gamma, \kappa, \eta < 1$ such that $\gamma \ll \kappa^{3/2}$ and $\kappa \ll \eta^2$,*

$$\sum_{g,c\in\{0,1\}^N} q(g,c) \sum_{t\in\{0,1\}^{N-|g|}} \kappa(t) \sum_{o:\,(g,c,o)\in \textsc{Acc}} \left\| \phi^{co} - \tilde{\phi}^{cto} \right\|_1 \;=\; O\big(2^{-\kappa q N}\big) \,. \tag{69}$$

Proof. We apply Lemma 8.1. Fix $g, c \in \{0,1\}^N$, let $n = |\{i : c_i = 0\}|$ and let M and K be distributed as the measurement outcomes associated with the measurements $\{\mathrm{Id} - M^0, \mathrm{Id} - M^1\}$ and $\{K^0, K^1\}$ made by the device in those rounds $i \in \{0, \ldots, N\}$ such that $c_i = 0$. Using (55) from Proposition 7.4 it follows that these random variables satisfy the assumptions of Lemma 8.1 for a choice of the function $g(x) = C\sqrt{x}$, for a large enough constant C. The conclusion (60) of the lemma gives (69). □

We conclude with a lemma that relates the randomness in the states $\tilde{\phi}^{cto}$ to randomness in the states $\tilde{\phi}^{ctok}$, for k such that $|k| \leq \eta\kappa q N$, as these are the post-measurement states associated with the simplified device in Protocol 2. The lemma relies on the following variant of the Cauchy-Schwarz inequality.

Lemma 8.5. *Let $\ell \geq 1$ be an integer and $|v_1\rangle, \ldots, |v_\ell\rangle$ arbitrary vectors in $\mathbb{C}^d$. Then*

$$\Big(\sum_{i=1}^{\ell} |v_i\rangle \Big) \Big(\sum_{i=1}^{\ell} |v_i\rangle \Big)^\dagger \;\leq\; \ell \sum_{i=1}^{\ell} |v_i\rangle\langle v_i| \,.$$

Using the lemma, we show the following.

Lemma 8.6. *Let $D = (\phi, \Pi, M, K)$ be a simplified device, and $\tilde{\phi}^{cto}$ the ensemble of states associated with D as described in* (68)*. Then*

$$\begin{aligned} \sum_{g,c\in\{0,1\}^N} q(g,c) \sum_{\substack{t\in\{0,1\}^{N-|g|} \\ |t|\leq 2\kappa q N}} \kappa(t) \sum_{o:\,(g,c,o)\in \textsc{Acc}} \big\langle \tilde{\phi}^{cto} \big\rangle_{1+\varepsilon} \\ \leq 2^{O(H(\eta))\kappa q N} \sum_{g,c\in\{0,1\}^N} q(g,c) \sum_{t\in\{0,1\}^{N-|g|}} \kappa(t) \sum_{o,k:\,(g,c,t,o,k)\in \textsc{Acc}_2} \big\langle \phi^{ctok} \big\rangle_{1+\varepsilon} \,, \end{aligned}$$

where $\textsc{Acc}_2$ denotes the set of transcripts that are accepted by the verifier in Protocol 2.

Proof. The proposition follows from the definition of $\tilde{\phi}^{cto}$, Lemma 8.5, and the fact that for any t such that $|t| \leq 2\kappa q N$ there are at most $2^{O(H(\eta))\kappa q N}$ sequences $k \in \{0,1\}^{|t|}$ such that $|k| \leq \eta\kappa q N$. Note that the conditions that $(g,c,o) \in \textsc{Acc}$ and $|k| \leq \eta\kappa q N$ imply $(g,c,t,o,k) \in \textsc{Acc}_2$. □

8.2 Randomness accumulation in the simplified protocol

In this section we consider the behavior of a simplified device $D = (\phi, \Pi, M, K)$ in a single round of Protocol 2. The following lemma shows that, provided the device has overlap $\Delta(D)$ bounded away from 1, then if the state ϕ of the device has high overlap with the projection operator M^1, performing a measurement of $\{\Pi^0, \Pi^1, \Pi^2\}$ on ϕ necessarily perturbs the state (hence generates randomness). The proof is based on a "measurement-disturbance trade-off" from [MS14], itself a consequence of uniform convexity for certain matrix p-norms.

Lemma 8.7. *Let $D = (\phi, \Pi, M, K)$ be a simplified device with overlap $\Delta(D) \leq \omega$, for some $\omega < 1$. Let $0 \leq \varepsilon \leq \frac{1}{2}$ and*

$$t = \frac{\langle \phi_G \rangle_{1+\varepsilon}}{\langle \phi \rangle_{1+\varepsilon}}, \qquad \text{where} \quad G = \frac{1}{2}(\Pi^0 + \Pi^1) + \frac{1}{2} M^1 K^0 \quad \text{and} \quad \phi_G = \sqrt{G}\phi\sqrt{G}. \tag{70}$$

Then

$$\frac{\langle \phi_1^0 \rangle_{1+\varepsilon} + \langle \phi_1^1 \rangle_{1+\varepsilon} + \langle \phi_1^2 \rangle_{1+\varepsilon}}{\langle \phi \rangle_{1+\varepsilon}} \leq 2^{-\varepsilon \lambda_\omega(t)} + O(\varepsilon^2),$$

where the post-measurement states ϕ_1^v, $v \in \{0,1,2\}$, are introduced in (46)*, and*

$$\lambda_\omega(t) = 2\log(e)\Big(t - \frac{1}{2} - \frac{\omega}{2}\Big)^2 \tag{71}$$

if $t \geq \frac{1}{2} + \frac{\omega}{2}$, and 0 otherwise.

Proof. The proof uses ideas from [MS14]. Let ϕ be as in the lemma and $\phi' = \sum_v \Pi^v \phi \Pi^v$. Then

$$\begin{aligned}
\Big\langle \sum_v \sqrt{G}\Pi^v \phi \Pi^v \sqrt{G} \Big\rangle_{1+\varepsilon} &\leq \sum_v \langle \phi^{1/2} \Pi^v G \Pi^v \phi^{1/2} \rangle_{1+\varepsilon} + O(\varepsilon) \\
&\leq \Big(\frac{1}{2} + \frac{\omega}{2}\Big) \langle \phi^{1/2} (\Pi^0 + \Pi^1) \phi^{1/2} \rangle_{1+\varepsilon} + \frac{1}{2} \langle \phi^{1/2} \Pi^2 \phi^{1/2} \rangle_{1+\varepsilon} + O(\varepsilon) \\
&\leq \Big(\frac{1}{2} + \frac{\omega}{2}\Big) \langle \phi' \rangle_{1+\varepsilon} + O(\varepsilon),
\end{aligned}$$

where the first and last lines use the approximate linearity relations (8), and the second line uses the definition of K and $G \leq \mathrm{Id}$. This allows us to proceed as in the proof of [MS14, Theorem 6.3] to obtain

$$\langle \phi - \phi' \rangle_{1+\varepsilon} \geq 2\Big(t - \frac{1}{2} - \frac{\omega}{2}\Big) \langle \phi \rangle_{1+\varepsilon} - O(\varepsilon),$$

and conclude by applying [MS16, Proposition 5.3]. □

Using Lemma 8.7 we proceed to quantify the accumulation of randomness across multiple rounds of the simplified protocol, when it is executed with a simplified device that has overlap bounded away from 1. The following proposition provides a measure of the randomness present in the transcript, conditioned on the verifier not aborting the protocol at the end, i.e. on $(g, c, t, o, k) \in \mathrm{ACC}_2$. (To see the connection with entropy, recall the definition of the $(1+\varepsilon)$ conditional Rényi entropy in Definition 2.9. The connection will be made precise in Section 8.3.)

Proposition 8.8. *Let $D = (\phi, \Pi, M, K)$ be a simplified device such that $\Delta(D) \leq \omega$ for some $\omega < 1$. Let $0 < \varepsilon \leq \frac{1}{2}$. Let $\gamma, \eta, \kappa, q > 0$ and N an integer be parameters for an execution of Protocol 2 (Figure 2) with D. Then*

$$-\frac{1}{\varepsilon N}\log\Big(\frac{\sum_{(g,c,t,o,k)\in \mathrm{ACC}_2} q(g,c)\kappa(t)\,\langle \phi^{ctok}\rangle_{1+\varepsilon}}{\langle\phi\rangle_{1+\varepsilon}}\Big) \;\geq\; \lambda_\omega\Big(1-\frac{\gamma}{\kappa}-\eta\Big) - O\Big(q+\frac{\varepsilon}{\kappa q}\Big)\,, \tag{72}$$

where the states ϕ^{ctok} are introduced in Definition 6.6, λ_ω is the function defined in (71)*, and $q(g,c)$ and $\kappa(t)$ are the distributions on N-bit strings (g,c) and t as selected by the verifier in Protocol 2.*

Proof. Let $t = \frac{\langle \phi_G\rangle_{1+\varepsilon}}{\langle\phi\rangle_{1+\varepsilon}}$ be as defined in Lemma 8.7. Recall the notation for the post-measurement states introduced in (46). After one round of Protocol 2 is executed, the post-measurement state of the device can be decomposed into three components. First, in case $G_i = 1$, which happens with probability $(1-q)$, the round is a generation round. The randomness generated in such a round is captured by the bound from Lemma 8.7,

$$(1-q)\big(\langle\phi_1^0\rangle_{1+\varepsilon} + \langle\phi_1^1\rangle_{1+\varepsilon} + \langle\phi_1^2\rangle_{1+\varepsilon}\big) \;\leq\; (1-q)\big(1-\ln(2)\varepsilon\lambda_\omega(t) + O(\varepsilon^2)\big)\langle\phi\rangle_{1+\varepsilon}\,. \tag{73}$$

The second case corresponds to $G_i = 0$, which happens with probability q. In this case, for reasons that will become clear later in this proof we weigh the "success" and "failure" components of the post-measurement state differently. For the "failure" part we simply write

$$\frac{q}{2}\big((1-\kappa)\langle\phi_{00}^0\rangle_{1+\varepsilon} + \kappa\langle\phi_{01}^{00}\rangle_{1+\varepsilon} + \kappa\langle\phi_{01}^{01}\rangle_{1+\varepsilon} + \kappa\langle\phi_{01}^{11}\rangle_{1+\varepsilon} + \langle\phi_1^2\rangle_{1+\varepsilon}\big)\,. \tag{74}$$

For the "success" part we add a weight of $2^{\frac{\varepsilon s}{\kappa q}}$, where $s = O(1)$ is a real parameter to be determined later, to the cases where $T_i = 1$:

$$\begin{aligned}\frac{(1-\kappa)q}{2}\big(\langle\phi_{00}^1\rangle_{1+\varepsilon} + \langle\phi_1^0\rangle_{1+\varepsilon} + \langle\phi_1^1\rangle_{1+\varepsilon}\big) + \frac{\kappa q}{2}2^{\frac{\varepsilon s}{\kappa q}}\big(\langle\phi_{01}^{10}\rangle_{1+\varepsilon} + \langle\phi_1^0\rangle_{1+\varepsilon} + \langle\phi_1^1\rangle_{1+\varepsilon}\big)\\ \leq \frac{(1-\kappa)q}{2}\big(\langle\phi_{00}^1\rangle_{1+\varepsilon} + \langle\phi_1^0\rangle_{1+\varepsilon} + \langle\phi_1^1\rangle_{1+\varepsilon}\big) + \kappa q\Big(1+\ln(2)\frac{\varepsilon s}{\kappa q} + O\Big(\frac{\varepsilon^2}{\kappa^2 q^2}\Big)\Big)\,t\,\langle\phi\rangle_{1+\varepsilon}\,,\end{aligned} \tag{75}$$

where the inequality follows from the definition of t. Using the first inequality in (8) and regrouping terms, the sum of the left-hand sides of (73), (74) and (80) is at most

$$(73)\;+\;(74)\;+\;(80)\;\leq\;\Big(1-\varepsilon\ln(2)\Big(\lambda_\omega(t) - st + O\Big(q+\frac{\varepsilon}{\kappa q}\Big)\Big)\Big)\,\langle\phi\rangle_{1+\varepsilon}\,. \tag{76}$$

A convenient choice of s is to take the derivative $s = \lambda'_\omega(r)$ for some $r \in [0,1]$ to be determined. With this choice, using that λ_ω is convex it follows that $\min_{t\in[0,1]}\lambda_\omega(t) - st = \lambda_\omega(r) - \lambda'_\omega(r)r$. By chaining the inequality (76) N times, where at each step the density ϕ is updated with the one obtained from the previous round, and using that ACC_2 contains those sequences (g,c,t,o,k) such that the number of occurrences of $(c,t,o,k) \in \{(0,1,1,0),(1,*,0,*),(1,*,1,*)\}$ is at least $(1-\gamma/\kappa-\eta)\kappa qN$ we obtain

$$\begin{aligned}-\frac{1}{\varepsilon N}\log\Big(\frac{\sum_{(g,c,t,o,k)\in \mathrm{ACC}_2} q(g,c)\kappa(t)\,\langle \phi^{ctok}\rangle_{1+\varepsilon}}{\langle\phi\rangle_{1+\varepsilon}}\Big) \;\geq\; (\lambda_\omega(r) - \lambda'_\omega(r)r) + \big(1-\frac{\gamma}{\kappa}-\eta\big)\lambda'_\omega(r)\\ - O\Big(q+\frac{\varepsilon}{\kappa q}\Big)\,,\end{aligned}$$

with the term $(1-\frac{\gamma}{\kappa}-\eta)\lambda'_\omega(r)$ on the right-hand side correcting for the weights $2^{\frac{\varepsilon s}{\kappa q}}$ that would appear on the left-hand side with an exponent derived from the acceptance criterion. Choosing $r = (1-\frac{\gamma}{\kappa}-\eta)$ completes the proof. □

8.3 Randomness accumulation in the general protocol

In this section we combine the results obtained in the previous two sections to analyze the randomness generated in Protocol 1. The main step is given in the following proposition.

Proposition 8.9. *Let $D = (\phi, \Pi, M)$ be an efficient device. Let $|\phi\rangle_{\mathsf{DE}}$ denote an arbitrary purification of ϕ_{D}, and $\overline{\rho}_{\mathsf{COE}}$ the joint state of the verifier's choice of challenges, the outputs computed by the verifier, and the adversary's system E, restricted to transcripts that are accepted by the verifier in the protocol.*[10] *Then there is a $\delta' = 2^{-\Omega(\gamma^{2/3} qN)}$ and a constant $C > 0$ such that for any $\delta > 0$,*

$$\frac{1}{N} H_\infty^{\delta+\delta'}(O|CE)_{\overline{\rho}} \geq \lambda_\omega\big(1 - C\gamma^{1/3}\big) - O\Big(q + \gamma^{1/6} + \frac{1 + \log(2/\delta)}{\gamma^{5/6} qN}\Big) . \tag{77}$$

Proof. Let $\tilde{D} = (\phi, \tilde{\Pi}, \tilde{M}, K)$ be the elementary device obtained by applying Proposition 7.4 to the device D, for a choice of $\omega = \frac{3}{4}$. Let $\tilde{\phi} = \phi^{\frac{1}{1+\varepsilon}}$, where $\varepsilon > 0$ is a small parameter to be specified later. We apply Proposition 8.8 to $\tilde{D}$, with ϕ replaced by $\tilde{\phi}$. Then (72) gives

$$-\frac{1}{\varepsilon N} \log\Big(\frac{\sum_{(g,c,t,o,k)\in \mathrm{ACC}_2} q(g,c)\kappa(t)\, \langle \tilde{\phi}^{ctok}\rangle_{1+\varepsilon}}{\langle \tilde{\phi}\rangle_{1+\varepsilon}}\Big) \geq \lambda_\omega\Big(1 - \frac{\gamma}{\kappa} - \eta\Big) - O\Big(q + \frac{\varepsilon}{\kappa q}\Big) . \tag{78}$$

Next we apply Lemma 8.6 to obtain

$$-\frac{1}{\varepsilon N} \log\Big(\frac{\sum_{(g,c,t,o):(g,c,o)\in \mathrm{ACC}} q(g,c)\kappa(t)\, \langle \tilde{\phi}^{cto}\rangle_{1+\varepsilon}}{\langle \tilde{\phi}\rangle_{1+\varepsilon}}\Big) \geq \lambda_\omega\Big(1 - \frac{\gamma}{\kappa} - \eta\Big) - O\Big(H(\eta)\kappa\frac{q}{\varepsilon} + q + \frac{\varepsilon}{\kappa q}\Big) , \tag{79}$$

where the correction $H(\eta)\kappa\frac{q}{\varepsilon}$ comes from the exponential prefactor in the bound from Lemma 8.6. The left-hand side of the bound in Lemma 8.6 only considers those sequences such that $|t| \leq 2\kappa qN$, but adding those sequences back only incurs a negligible error $2^{-\Omega(\kappa qN)}$ (inside the logarithm), due to the Chernoff bound.

We make one ultimate re-writing step. For any fixed t, the post-measurement state $\tilde{\phi}^{cto}$ can be expressed as

$$P_N \cdots P_1 \tilde{\phi} P_1 \cdots P_N ,$$

where P_i is the measurement operator associated with challenge c_i and outcome o_i. Using $\langle XX^*\rangle_{1+\varepsilon} = \langle X^*X\rangle_{1+\varepsilon}$ for any X, and recalling the definition of $\tilde{\phi} = \phi^{\frac{1}{1+\varepsilon}}$,

$$\langle P_N \cdots P_1 \tilde{\phi} P_1 \cdots P_N\rangle_{1+\varepsilon} = \langle \phi^{\frac{-\varepsilon}{2(1+\varepsilon)}} \phi^{\frac{1}{2}} P_1 \cdots P_N^2 \cdots P_1 \phi^{\frac{1}{2}} \phi^{\frac{-\varepsilon}{2(1+\varepsilon)}}\rangle_{1+\varepsilon} .$$

Introduce a sub-normalized density

$$\rho_{\mathsf{E}}^{cto} = \phi^{\frac{1}{2}} P_1 \cdots P_N^2 \cdots P_1 \phi^{\frac{1}{2}} ,$$

that corresponds to the post-measurement state of register E (recall we assumed a purification $|\phi\rangle_{\mathsf{DE}}$ of ϕ) at the end of Protocol 1, for a given transcript (c, o) for the interaction.

We are in a position to apply Theorem 2.12, with

$$\rho_{\mathsf{CTOE}}^{o} = \sum_{(g,c,t):\,(g,c,o)\in \mathrm{ACC}} q(g,c)\kappa(t)\, |c,t\rangle\langle c,t|_{\mathsf{CT}} \otimes |o\rangle\langle o|_{\mathsf{O}} \otimes \rho_{\mathsf{E}}^{cto} ,$$

[10] The state $\overline{\rho}$ is sub-normalized.

and $\sigma_{\mathsf{CTE}} = \sum_{(g,c,t)} q(g,c)\kappa(t)|c,t\rangle\langle c,t| \otimes \phi$. Applying the theorem and using (79) and $\langle\tilde{\phi}\rangle_{1+\varepsilon} = 1$ by definition, we get that for any $\delta > 0$,

$$\frac{1}{N}H^{\delta}_{\infty}(O|CTE)_{\rho} \geq \lambda_{\omega}\big(1 - \frac{\gamma}{\kappa} - \eta\big) - O\Big(H(\eta)\kappa\frac{q}{\varepsilon} + q + \frac{\varepsilon}{\kappa q}\Big) - \frac{1 + 2\log(1/\delta)}{\varepsilon N}\,. \tag{80}$$

Using that the bound in (77) only considers registers C and O (the transcript) and E, by Corollary 8.4 for any choice of parameters κ, η such that $\kappa \ll \eta^2$ and $\kappa^{3/2} \gg \gamma$, the bound (80) extends to a lower bound on the entropy $H^{\delta+\delta'}_{\infty}(O|CE)_{\overline{\rho}}$ at the cost of an additional $\delta' = O(2^{-\kappa q N})$ in the smoothing parameter.

Choose κ, η to be sufficiently large constant multiples of $\gamma^{2/3}$ and $\gamma^{1/3}$ respectively, so that the constraints $\kappa \ll \eta^2$ and $\kappa^{3/2} \gg \gamma$ are satisfied. Let ε to be a sufficiently small constant multiple of $\gamma^{5/6}q$. With this choice of parameters, the term in the $O(\cdot)$ on the right-hand side of (80) is $O(q + \gamma^{1/6})$. □

Making an appropriate choice of parameters q, γ for an execution of Protocol 1, Proposition 8.9 gives our main result.

Theorem 8.10. *Let $\mathcal{F}$ be an NTCF family and λ a security parameter. Let N be a polynomially bounded function of λ such that $N = \Omega(\lambda^2)$. Set $q = \lambda/N$. Then there is a $\delta = 2^{-\Omega(\gamma q N)}$ such that for any small enough $\gamma > 0$, any efficient prover, and side information E correlated with the prover's initial state,*

$$H^{N\delta}_{\infty}(O|CE)_{\overline{\rho}} \geq (\xi - O(\gamma^{1/6}))N\,,$$

where $\overline{\rho}$ is the final state of the output, challenge, and adversary registers, restricted to transcripts that are accepted by the verifier in the protocol and ξ is a positive constant.[11]

Assume that an execution of $\textsc{Gen}(1^{\lambda})$ requires $O(\lambda^r)$ bits of randomness, for some constant r. (For example, for the case of our construction of a NTCF family based on LWE, we have $r = 2$.) Then an execution of the protocol using the parameters in Theorem 8.10 requires only $\mathrm{poly}(\lambda, \log N)$ bits of randomness for the verifier to generate the key k and select the challenges. Taking N to be slightly sub-exponential in λ, e.g. $N = 2^{\sqrt{\lambda}}$, yields sub-exponential randomness expansion.

Proof of Theorem 8.10. Let D be a device that is accepted with non-negligible probability in Protocol 1, where the parameters are a stated in the theorem. Applying Proposition 8.9 to D and choosing δ to be a negligible function of N such that δ^{-1} is sub-exponential gives the result. □

References

[AFDF+18] Rotem Arnon-Friedman, Frédéric Dupuis, Omar Fawzi, Renato Renner, and Thomas Vidick. Practical device-independent quantum cryptography via entropy accumulation. *Nature communications*, 9(1):459, 2018.

[Ajt99] Miklós Ajtai. Generating hard instances of the short basis problem. In *International Colloquium on Automata, Languages, and Programming*, pages 1–9. Springer, 1999.

[AKPW13] Joël Alwen, Stephan Krenn, Krzysztof Pietrzak, and Daniel Wichs. Learning with rounding, revisited. In *Advances in Cryptology–CRYPTO 2013*, pages 57–74. Springer, 2013.

[11] The constant ξ is at least some positive universal constant of order $1/10$, for all small enough γ.

[BKG+18] Peter Bierhorst, Emanuel Knill, Scott Glancy, Yanbao Zhang, Alan Mink, Stephen Jordan, Andrea Rommal, Yi-Kai Liu, Bradley Christensen, Sae Woo Nam, et al. Experimentally generated randomness certified by the impossibility of superluminal signals. *arXiv preprint arXiv:1803.06219*, 2018.

[BLP+13] Zvika Brakerski, Adeline Langlois, Chris Peikert, Oded Regev, and Damien Stehlé. Classical hardness of learning with errors. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, June 1-4, 2013*, pages 575–584. ACM, 2013.

[CCKW18] Alexandru Cojocaru, Léo Colisson, Elham Kashefi, and Petros Wallden. Delegated pseudo-secret random qubit generator. *arXiv preprint arXiv:1802.08759*, 2018.

[Col06] R. Colbeck. *Quantum And Relativistic Protocols For Secure Multi-Party Computation*. PhD thesis, Trinity College, University of Cambridge, November 2006.

[FGL12] Xiequan Fan, Ion Grama, and Quansheng Liu. Hoeffdings inequality for supermartingales. *Stochastic Processes and their Applications*, 122(10):3545–3559, 2012.

[GKPV10] Shafi Goldwasser, Yael Tauman Kalai, Chris Peikert, and Vinod Vaikuntanathan. Robustness of the learning with errors assumption. In Andrew Chi-Chih Yao, editor, *ICS*, pages 230–240. Tsinghua University Press, 2010.

[GMR84] Shafi Goldwasser, Silvio Micali, and Ronald L. Rivest. A "paradoxical'"solution to the signature problem (abstract). In G. R. Blakley and David Chaum, editors, *Advances in Cryptology, Proceedings of CRYPTO '84, Santa Barbara, California, USA, August 19-22, 1984, Proceedings*, volume 196 of *Lecture Notes in Computer Science*, page 467. Springer, 1984.

[GR02] Lov Grover and Terry Rudolph. Creating superpositions that correspond to efficiently integrable probability distributions. *arXiv preprint quant-ph/0208112*, 2002.

[GSW13] Craig Gentry, Amit Sahai, and Brent Waters. Homomorphic encryption from learning with errors: Conceptually-simpler, asymptotically-faster, attribute-based. In *Advances in Cryptology–CRYPTO 2013*, pages 75–92. Springer, 2013.

[HM17] Aram W Harrow and Ashley Montanaro. Quantum computational supremacy. *Nature*, 549(7671):203, 2017.

[LP11] Richard Lindner and Chris Peikert. Better key sizes (and attacks) for lwe-based encryption. In *Cryptographers Track at the RSA Conference*, pages 319–339. Springer, 2011.

[Mah17a] Urmila Mahadev. Classical homomorphic encryption for quantum circuits. *Arxiv preprint arXiv:1708.02130v1*, 2017.

[Mah17b] Urmila Mahadev. Classical homomorphic encryption for quantum circuits. *arXiv preprint arXiv:1708.02130*, 2017.

[Mah18] Urmila Mahadev. Classical verification of quantum computations. *arXiv preprint arXiv:1804.01082*, 2018.

[MP12] Daniele Micciancio and Chris Peikert. Trapdoors for lattices: Simpler, tighter, faster, smaller. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 700–718. Springer, 2012.

[MS14] Carl A Miller and Yaoyun Shi. Universal security for randomness expansion from the spot-checking protocol. *arXiv preprint arXiv:1411.6608*, 2014.

[MS16] Carl A Miller and Yaoyun Shi. Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices. *Journal of the ACM (JACM)*, 63(4):33, 2016.

[MU88] Hans Maassen and Jos BM Uffink. Generalized entropic uncertainty relations. *Physical Review Letters*, 60(12):1103, 1988.

[PAM+10] S. Pironio, A. Acin, S. Massar, A. Boyer De La Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and et al. Random numbers certified by Bell's theorem. *Nature*, 464(7291), 2010.

[Pei09] Chris Peikert. Public-key cryptosystems from the worst-case shortest vector problem. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 333–342. ACM, 2009.

[PRS17] Chris Peikert, Oded Regev, and Noah Stephens-Davidowitz. Pseudorandomness of ring-LWE for any ring and modulus. In Hamed Hatami, Pierre McKenzie, and Valerie King, editors, *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 461–473. ACM, 2017.

[Reg05] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In Harold N. Gabow and Ronald Fagin, editors, *STOC*, pages 84–93. ACM, 2005.

[TCR09] Marco Tomamichel, Roger Colbeck, and Renato Renner. A fully quantum asymptotic equipartition property. *IEEE Transactions on Information Theory*, 55(12):5840–5847, 2009.

[Tom15] Marco Tomamichel. *Quantum Information Processing with Finite Resources: Mathematical Foundations*, volume 5. Springer, 2015.

[VV11] Umesh Vazirani and Thomas Vidick. Certifiable quantum dice: or, true random number generation secure against quantum adversaries. In *Proceedings of the 44th symposium on Theory of Computing*, STOC '12, pages 61–76. ACM, 2011. Also available as arXiv:1111.6054.

[Wil13] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013.