Quantum Money from Hidden Subspaces

Scott Aaronson* Paul Christiano†

Abstract

Forty years ago, Wiesner pointed out that quantum mechanics raises the striking possibility of money that cannot be counterfeited according to the laws of physics. We propose the first quantum money scheme that is

(1) *public-key*—meaning that anyone can verify a banknote as genuine, not only the bank that printed it, and

(2) *cryptographically secure*, under a "classical" hardness assumption that has nothing to do with quantum money.

Our scheme is based on *hidden subspaces*, encoded as the zero-sets of random multivariate polynomials. A main technical advance is to show that the "black-box" version of our scheme, where the polynomials are replaced by classical oracles, is *unconditionally* secure. Previously, such a result had only been known relative to a *quantum* oracle (and even there, the proof was never published).

Even in Wiesner's original setting—quantum money that can only be verified by the bank—we are able to use our techniques to patch a major security hole in Wiesner's scheme. We give the first private-key quantum money scheme that allows unlimited verifications and that remains unconditionally secure, even if the counterfeiter can interact adaptively with the bank.

Our money scheme is simpler than previous public-key quantum money schemes, including a knot-based scheme of Farhi et al. The verifier needs to perform only two tests, one in the standard basis and one in the Hadamard basis—matching the original intuition for quantum money, based on the existence of complementary observables.

Our security proofs use a new variant of Ambainis's quantum adversary method, and several other tools that might be of independent interest.

Contents

1 Introduction **2**

1.1 The History of Quantum Money . . . 3

1.2 The Challenge . . . 5

1.3 Our Results . . . 6

1.4 Motivation . . . 8

2 Preliminaries **10**

2.1 Cryptography . . . 10

2.2 Quantum Information . . . 12

2.3 Quantum Search . . . 13

*MIT. Email: aaronson@csail.mit.edu. This material is based upon work supported by the National Science Foundation under Grant No. 0844626. Also supported by a DARPA YFA grant, an NSF STC grant, a TIBCO Chair, and a Sloan Fellowship.

†MIT. Email: paulfchristiano@gmail.com.

3 Formalizing Quantum Money **16**
3.1 Quantum Money Schemes . . . 16
3.2 Mini-Schemes . . . 18
3.3 The Standard Construction . . . 20

4 The Inner-Product Adversary Method **23**
4.1 Idea of Method . . . 24
4.2 The Method . . . 25

5 Quantum Money with a Classical Oracle **27**
5.1 The Hidden Subspace Mini-Scheme . . . 27
5.2 Formal Specification . . . 28
5.3 Analysis . . . 29

6 Explicit Candidate Quantum Money Scheme **33**
6.1 Useful Facts About Polynomials . . . 34
6.2 The Explicit Hidden-Subspace Mini-Scheme . . . 36
6.3 Analysis . . . 36
6.4 Justifying Our Hardness Assumption . . . 38

7 Improving Private-Key Quantum Money **40**

8 Open Problems **43**
8.1 Quantum Copy-Protection and More . . . 43

9 Acknowledgments **44**

10 Appendix: Reducing Completeness Error **47**

11 Appendix: Complexity-Theoretic No-Cloning Theorem **49**

1 Introduction

"Information wants to be free"—this slogan expresses the idea that classical bits, unlike traditional economic goods, can be copied an unlimited number of times. The copyability of classical information is one of the foundations of the digital economy, but it is also a nuisance to governments, publishers, software companies, and others who wish to prevent copying. Today, essentially all electronic commerce involves a trusted third party, such as a credit card company, to mediate transactions. Without such a third party entering at *some* stage, it is impossible to prevent electronic cash from being counterfeited, regardless of what cryptographic assumptions one makes.[1]

Famously, though, quantum bits do *not* "want to be free" in the same sense that classical bits do: in many respects, they behave more like gold, oil, or other traditional economic goods. Indeed, the *No-Cloning Theorem*, which is an immediate consequence of the linearity of quantum mechanics,

[1] The recent Bitcoin system is an interesting illustration of this principle: it gets rid of the centralized third party, but still uses a "third party" distributed over the community of Bitcoin users.

says that there is no physical procedure that takes as input an unknown[2] quantum pure state $|\psi\rangle$, and that produces as output two unentangled copies of $|\psi\rangle$, or even a close approximation thereof. The No-Cloning Theorem is closely related to the *uncertainty principle*, which says that there exist "complementary" properties of a quantum state (for example, its position and momentum) that cannot both be measured to unlimited accuracy.[3]

1.1 The History of Quantum Money

But can one actually *exploit* the No-Cloning Theorem to achieve classically-impossible cryptographic tasks? This question was first asked by Wiesner [43], in a remarkable paper written around 1970 (but only published in 1983) that arguably founded quantum information science. In that paper, Wiesner proposed a scheme for *quantum money* that would be physically impossible to clone. In Wiesner's scheme, each "banknote" would consist of a classical serial number s, together with a quantum state $|\psi_s\rangle$ consisting of n unentangled qubits, each one $|0\rangle$, $|1\rangle$, $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$, or $\frac{|0\rangle-|1\rangle}{\sqrt{2}}$ with equal probability. The issuing bank would maintain a giant database, which stored a classical description of $|\psi_s\rangle$ for each serial number s. Whenever someone wanted to *verify* a banknote, he or she would take it back to the bank—whereupon the bank would use its knowledge of how $|\psi_s\rangle$ was prepared to measure each qubit in the appropriate basis, and check that it got the correct outcomes. On the other hand, it can be proved [35] that someone who did *not* know the appropriate bases could copy the banknote with success probability at most $(3/4)^n$.

Though historically revolutionary, Wiesner's money scheme suffered at least three drawbacks:

(1) **The "Verifiability Problem":** The only entity that can verify a banknote is the bank that printed it.

(2) **The "Online Attack Problem":** A counterfeiter able to submit banknotes for verification, and get them back afterward, can easily break Wiesner's scheme ([32, 3]; see also Section 7).

(3) **The "Giant Database Problem":** The bank needs to maintain a database with an entry for every banknote in circulation.

In followup work in 1982, Bennett, Brassard, Breidbart, and Wiesner [16] (henceforth BBBW) at least showed how to eliminate the giant database problem: namely, by generating the state $|\psi_s\rangle = |\psi_{f_k(s)}\rangle$ using a *pseudorandom function* f_k, with key k known only by the bank. Unlike Wiesner's original scheme, the BBBW scheme is no longer *information-theoretically* secure: a counterfeiter can recover k given exponential computation time. On the other hand, a counterfeiter cannot break the scheme in polynomial time, unless it can *also* distinguish f_k from a random function.

These early ideas about quantum money inspired the field of *quantum cryptography* [15]. But strangely, the subject of quantum money itself lay dormant for more than two decades, even as interest in quantum computing exploded. However, the past few years have witnessed a "quantum

[2]The adjective "unknown" is needed because, if we knew a classical description of a procedure to *prepare* $|\psi\rangle$, then of course we could run that procedure multiple times to prepare multiple copies.

[3]Indeed, if we could copy $|\psi\rangle$, then we could violate the uncertainty principle by measuring one observable (such as position) on some copies, and a complementary observable (such as momentum) on other copies. Conversely, if we could measure all the properties of $|\psi\rangle$ to unlimited accuracy, then we could use the measurement results to create additional copies of $|\psi\rangle$.

money renaissance." Some recent work has offered partial solutions to the verifiability problem: for example, Mosca and Stebila [36] suggested that the bank use a *blind quantum computing* protocol to offload the verification of banknotes to local merchants, while Gavinsky [25] proposed a variant of Wiesner's scheme that requires only *classical* communication between the merchant and bank.

However, most of the focus today is on a more ambitious goal: namely, creating what Aaronson [3] called *public-key quantum money*, or quantum money that *anyone* could authenticate, not just the bank that printed it. As with public-key cryptography in the 1970s, it is far from obvious *a priori* whether public-key quantum money is possible at all. Can a bank publish a description of a quantum circuit that lets people feasibly *recognize* a state $|\psi\rangle$, but does not let them feasibly prepare or even copy $|\psi\rangle$?

Aaronson [3] gave the first formal treatment of public-key quantum money, as well as related notions such as copy-protected quantum software. He proved that there exists a *quantum oracle* relative to which secure public-key quantum money is possible. Unfortunately, that result, though already involved, did not lead in any obvious way to an explicit (or "real-world") quantum money scheme.[4] He raised as an open problem whether secure public-key quantum money is possible relative to a *classical* oracle. In the same paper, Aaronson also proposed an explicit scheme, based on random stabilizer states, but could not offer any evidence for its security. And indeed, the scheme was broken about a year afterward by Lutomirski et al. [34], using an algorithm for finding planted cliques in random graphs due to Alon, Krivelevich, and Sudakov [8].

Recently, Farhi et al. [24] took a completely different approach to public-key quantum money. They proposed a quantum money scheme based on *knot theory*, where each banknote is a superposition over exponentially-many oriented link diagrams. Within a given banknote, all the link diagrams L have the same Alexander polynomial $p(L)$ (a certain knot invariant).[5] This $p(L)$, together with a digital signature of $p(L)$, serves as the banknote's "classical serial number." Besides the unusual mathematics employed, the work of Farhi et al. [24] (building on [34]) also developed an idea that will play a major role in our work. That idea is to construct public-key quantum money schemes by composing two "simpler" ingredients: first, objects that we call *mini-schemes*; and second, classical digital signature schemes.

The main disadvantage of the knot-based scheme, which it shares with every previous scheme, is that no one can say much about its security—other than that it has not yet been broken, and that various known counterfeiting strategies fail. Indeed, even characterizing *which quantum states Farhi et al.'s verification procedure accepts* remains a difficult open problem, on which progress seems likely to require major advances in knot theory! In other words, there might be states that look completely different from "legitimate banknotes," but are still accepted with high probability.

In followup work, Lutomirski [33] proposed an "abstract" version of the knot scheme, which gets rid of the link diagrams and Alexander polynomials, and simply uses a classical oracle to achieve the same purposes. Lutomirski raised the challenge of proving that this *oracle* scheme is secure—in which case, it would have yielded the first public-key quantum money scheme that was proven secure relative to a classical oracle. Unfortunately, proving the security of Lutomirski's

[4] Also, the proof of Aaronson's result never appeared—an inexcusable debt that this paper finally repays, with interest.

[5] Instead of knots, Farhi et al. [24] could also have used, say, superpositions over n-vertex *graphs* having the same eigenvalue spectrum. But in that case, their scheme would have been breakable, the reason being that the *graph isomorphism problem* is easy for random graphs. By contrast, it is not known how to solve *knot isomorphism* efficiently, even with a quantum computer and even for random knots.

scheme remains open, and seems hard.[6]

As alluded to earlier, there is already some research on ways to *break* quantum money schemes. Besides the papers by Lutomirski [32] and Lutomirski et al. [34] mentioned before, let us mention the beautiful work of Farhi et al. on *quantum state restoration* [23]. As we discuss in Section 7, quantum state restoration can be used to break many public-key quantum money schemes: roughly speaking, any scheme where the banknotes contain only limited entanglement, and where verification consists of a rank-1 projective measurement. This fact explains why our scheme, like the knot-based scheme of Farhi et al. [24], will require highly-entangled banknotes.

1.2 The Challenge

Work over the past few years has revealed a surprising richness in the quantum money problem—both in the ideas that have been used to construct public-key quantum money schemes, *and* in the ideas that have been used to break them. Of course, this record also underscores the need for caution! To whatever extent we can, we ought to hold quantum money schemes to modern cryptographic standards, and not be satisfied with "we tried to break it and failed."

It is easy to see that, if public-key quantum money is possible at all, then it must rely on *some* computational assumption, in addition to the No-Cloning Theorem.[7] The best case would be to show that secure, public-key quantum money is possible, *if* (for example) there exist one-way functions resistant to quantum attack. Unfortunately, we seem a long way from showing anything of the kind. The basic problem is that *uncloneability* is a novel cryptographic requirement: something that would not even make sense in a classical context. Indeed, work by Farhi et al. [23] and Aaronson [3] has shown that it is sometimes possible to copy quantum banknotes, via attacks that do not even *measure* the banknotes in an attempt to learn a classical secret! Rather, these attacks simply perform some unitary transformation on a legitimate banknote $|\$\rangle$ together with an ancilla $|0\rangle$, the end result of which is to produce $|\$\rangle^{\otimes 2}$. Given such a strange attack, how can one deduce the failure of any "standard" cryptographic assumption?

Yet despite the novelty of the quantum money problem—or perhaps because of it—it seems reasonable to want *some* non-tautological evidence that a public-key quantum money scheme is secure. A minimal wish-list might include:

(1) Security under *some* plausible assumption, of a sort cryptographers know how to evaluate. Such an assumption should talk only about computing a classical output from a classical input; it should have nothing to do with cloning of quantum states.

(2) A proof that the money scheme is secure against *black-box counterfeiters*: those that do not exploit the structure of some cryptographic function f used in verifying the banknotes.

(3) A "simple" verification process, which accepts all valid banknotes $|\$\rangle$ with probability 1, and rejects all banknotes that are far from $|\$\rangle$.

[6] One way to understand the difficulty is that any security proof for Lutomirski's scheme would need to contain Aaronson's quantum lower bound for the *collision problem* [1] as a (tiny) special case. The lower bound for the collision problem is proved using the polynomial method of Beals et al. [13]. In this work, by contrast, we will only manage to prove the security of *our* oracle scheme using a specially-designed variant of Ambainis's quantum adversary method [9]. Despite great progress in quantum lower bounds over the past decade, it is still not known (except implicitly) how to prove the collision lower bound using an adversary argument.

[7] This is because a counterfeiter with unlimited time could simply search for a state $|\psi\rangle$ that the (publicly-known) verification procedure accepted.

1.3 Our Results

Our main contribution is a new public-key quantum money scheme, which achieves all three items in the wish-list above, and which is the first to achieve (1) or (2). Regardless of whether our particular scheme stands or falls, we introduce at least four techniques that should be useful for the design and analysis of *any* public-key quantum money scheme. These are:

- The "inner-product adversary method," a new variant of Ambainis's quantum adversary method [9] that can be used to rule out black-box counterfeiting strategies.
- A notion of *mini-schemes*, and a proof that (together with standard cryptographic assumptions) these objects imply full-fledged quantum money schemes.
- A method to *amplify* weak counterfeiters into strong ones, so that one only needs to rule out the latter to show security.
- A new connection between the security of quantum money schemes and *direct-product assumptions* in cryptography.

A second contribution is to construct the first *private*-key quantum money schemes that remain *unconditionally* secure, even if the counterfeiter can interact adaptively with the bank. This gives the first solution to the "online attack problem," a major security hole in the Wiesner [43] and BBBW [16] schemes pointed out by Lutomirski [32] and Aaronson [3]. These private-key schemes are direct adaptations of our public-key scheme.

In more detail, our quantum money scheme is based on *hidden subspaces* of the vector space $\mathbb{F}_2^n$. Each of our money states is a uniform superposition of the vectors in a random $n/2$-dimensional subspace $A \leq \mathbb{F}_2^n$. We denote this superposition by $|A\rangle$. Crucially, we can recognize the state $|A\rangle$ using only membership oracles for A and for its dual subspace $A^\perp$. To do so, we apply the membership oracle for A, then a Fourier transform, then the membership oracle for $A^\perp$, and then a second Fourier transform to restore the original state. We prove that this operation computes a rank-1 projection onto $|A\rangle$.

Underlying the security of our money schemes is the assertion that the states $|A\rangle$ are difficult to clone, even given membership oracles for A and $A^\perp$. Or more concretely: *any quantum algorithm that maps $|A\rangle$ to $|A\rangle^{\otimes 2}$ must make $2^{\Omega(n)}$ queries to the $A, A^\perp$ oracles.*

In order to prove this statement, we introduce a new method for proving lower bounds on quantum query complexity, which we call the *inner-product adversary method*. This technique considers a single counterfeiting algorithm being run in parallel to clone two distinct states $|A\rangle$ and $|A'\rangle$, with each having access to the membership oracles for $A, A^\perp$ or $A', A'^\perp$, as appropriate. To measure how much progress the algorithm has made, we consider the inner product between the states produced by the parallel executions: because $\langle A|^{\otimes 2} |A'\rangle^{\otimes 2} < \langle A|A'\rangle$ for many pairs of subspaces A, A', in order to succeed a counterfeiter will have to reduce this inner product substantially. We prove that when averaged over a suitable distribution of pairs A, A', the *expected inner product* between the two states produced by the counterfeiter cannot decrease too much with a single query to the membership oracles. We conclude that in order to produce $|A\rangle^{\otimes 2}$ given $|A\rangle$ and membership oracles for $A, A^\perp$, a counterfeiter must use exponentially many queries.

Having ruled out the possibility of nearly perfect cloning, we introduce a new amplification protocol, which allows us to transform a counterfeiter who succeeds with $\Omega\left(1/\operatorname{poly}(n)\right)$ success

probability into a counterfeiter who succeeds with probability arbitrarily close to 1. This technique is based on combining standard Grover search with a monotonic state amplification protocol of Tulsi, Grover, and Patel [42], to obtain monotonic convergence with the quadratic speedup of Grover search.[8] Combining this amplification with the inner-product adversary method, and applying a random linear transformation to convert the counterfeiter's worst case to its average case, we conclude that no counterfeiting algorithm can succeed with any non-negligible probability on a non-negligible fraction of states $|A\rangle$.

Before constructing a quantum money scheme, we first introduce the notion of a *quantum money mini-scheme*, a formalization of the setting in which the bank issues only a single money state and maintains no secret information. Formally, a mini-scheme is a protocol Bank for outputting pairs (s, ρ_s) and a verification procedure Ver_s for identifying ρ_s. We say a mini-scheme is *complete* if the state ρ_s passes the verification Ver_s with high probability, and we say the scheme is *secure* if furthermore no counterfeiter can take a single state ρ_s, and produce two (possibly-entangled) states ρ_1 and ρ_2 which simultaneously pass the verification procedure with non-negligible probability.

Using our uncloneability result, we produce a secure mini-scheme relative to a classical oracle. The algorithm Bank queries the classical oracle to obtain a serial number s and the description of a subspace A. Using this description it prepares $|A\rangle$, and publishes $(s, |A\rangle)$. The verification procedure uses the serial number s as an index into another classical oracle, which allows it to test membership in A and $A^{\perp}$. We prove that the uncloneability of the states $|A\rangle$ implies that this mini-scheme is secure.

We also present a reduction from quantum money schemes to mini-schemes, based on combining a mini-scheme with a secure signature scheme. The bank maintains a secret key for the signature scheme, and to issue a banknote, it runs Bank to produce a pair (s, ρ_s), then digitally signs the serial number s. Special cases of this reduction appeared in [24, 34], but we provide the first rigorous security proof.

By combining this reduction with our mini-scheme, we are able to obtain a "black-box" public key quantum money scheme relative to a classical oracle, which is unconditionally secure:

Theorem (Security of Hidden Subspace Money)**.** *Relative to some (classical) oracle A, there is a secure public-key quantum money scheme.*

More precisely, there is an algorithm KeyGen^A which outputs pairs $(k_{\text{private}}, k_{\text{public}})$, an algorithm $\mathsf{Bank}^A(k_{\text{private}})$ which generates a "quantum banknote" $|\$\rangle$, and a verification algorithm $\mathsf{Ver}^A(k_{\text{public}}, |\$\rangle)$ which tests the authenticity of a purported banknote. These algorithms have the following properties:

- ***Completeness:*** *If $(k_{\text{private}}, k_{\text{public}})$ is produced by KeyGen^A, then $\mathsf{Ver}^A\left(k_{\text{public}}, \mathsf{Bank}^A(k_{\text{private}})\right)$ accepts with probability 1.*
- ***Soundness:*** *Suppose a would-be counterfeiter with access to A and k_{public} is given q valid banknotes. If this counterfeiter outputs any number of (possibly-entangled) quantum states, there is at most an exponentially-small probability that Ver^A will accept more than q of them.*

By adapting these ideas to the private-key setting, we are also able to provide the first *private*-key quantum money scheme that is unconditionally secure, even if the counterfeiter is able to

[8] Although the "quadratic speedup" part is not strictly necessary for us, it improves our lower bound on the number of queries the counterfeiter needs to make—to the tight one, in fact—and might be of independent interest.

interact adaptively with the bank. This patches a security hole in Wiesner's original scheme which was observed in [32, 3], but which has not previously been addressed in a provably-secure way.

Finally, we provide a candidate cryptographic protocol for obfuscating the indicator functions of subspaces $A \leq \mathbb{F}_2^n$. In order to obfuscate a membership oracle for A, we provide a random system of polynomials $p_1, \ldots, p_m$ that vanish on A. Membership in A can be tested by evaluating the p_i's, but given only the p_i's, we conjecture that it is difficult to recover A. Combining this protocol with the black-box money scheme, we obtain an *explicit* quantum money scheme, the first one whose security is based on a plausible "classical" cryptographic assumption. Here is the assumption:

Conjecture (*). *Suppose A is a uniformly-random $n/2$-dimensional subspace of $\mathbb{F}_2^n$, and that $\{p_1, \ldots, p_{2n}\}$ and $\{q_1, \ldots, q_{2n}\}$ are systems of degree-d polynomials from $\mathbb{F}_2^n$ to $\mathbb{F}_2$, which vanish on A and $A^\perp$ respectively but are otherwise uniformly-random. Then for large enough constant d, there is no polynomial-time quantum algorithm that takes as input descriptions of the p_i's and q_i's, and that outputs a basis for A with success probability $\Omega\left(2^{-n/2}\right)$.*

Note that we can trivially guess a *single* nonzero A element with success probability $2^{-n/2}$, but guessing a whole *basis* for A would succeed with probability only $2^{-\Omega(n^2)}$. Conjecture (*) asserts that it is harder to find many elements of A than to find just one element.

The following theorem says that, *if* a counterfeiter could break our quantum money scheme, then with nontrivial success probability, it could *also* recover a description of A from the p_i's and q_i's alone—even without having access to a bank that provides a valid money state $|A\rangle$.

Theorem. *Assuming Conjecture (*), there is a public-key quantum money scheme with perfect completeness and exponentially-small soundness error. That is, the verifier always accepts valid banknotes, and a would-be counterfeiter succeeds only with exponentially-small probability.*[9]

The problem of recovering a subspace A, given a system of equations that vanish on A, is closely related to *algebraic cryptanalysis*, and in particular to the so-called *polynomial isomorphism problem*. In the latter problem, we are given as input two polynomials $p, q : \mathbb{F}^n \to \mathbb{F}$ related by an unknown linear change of basis L; the challenge is to find L. When $\deg(p) = \deg(q) = 3$, the best known algorithms for the polynomial isomorphism problem require exponential time [39, 26, 19]. An attacker might be able to use known techniques to effectively reduce the degree of the polynomials in our scheme by 1, at the expense of an exponentially reduced success probability [19]. Provided the degree is at least 4, however, recovering A seems to be beyond existing techniques.

1.4 Motivation

Unlike the closely-related task of *quantum key distribution* [15] (which is already practical), quantum money currently seems to be a long way off. The basic difficulty is how to maintain the *coherence* of a quantum money state for any appreciable length of time. *All* money eventually loses its value unless it is spent, but money that decohered on a scale of microseconds would be an extreme example!

In principle—and assuming, of course, the validity of quantum mechanics—this problem could be solved using *quantum fault-tolerance* constructions [6, 30], of the same sort that would be used

[9] This theorem remains true even if the statement of Conjecture (*) is weakened by adding random noise to the p_i's and q_i's, so that only a constant fraction of them vanish on A or $A^\perp$. The presence of noise interferes substantially with known techniques for solving systems of equations, though an attacker who was able to recover A from a *single* polynomial would of course not be hindered by such noise.

for universal quantum computing. (In this case, the goal would be to fault-tolerantly quantum-compute the identity function!) Alas, the currently-achievable decoherence rates are well above what known fault-tolerance constructions require.

So one might wonder: why develop rigorous foundations for a cryptographic functionality that seems so far from being practical? One answer is that, just as quantum key distribution uses many of the same ideas as private-key quantum money, but without requiring long-lasting coherence, so it is not hard to imagine protocols that would use many of the same ideas as *public*-key quantum money without requiring long-lasting coherence. Indeed, depending on the problem, rapid decoherence might actually be a *feature* rather than a bug!

As one example, public-key quantum money that decohered quickly could be used to create **non-interactive uncloneable signatures**. These are n-qubit quantum states $|\psi\rangle$ that an agent can efficiently prepare using a private key, then freely hand out to passersby. By feeding $|\psi\rangle$, together with the agent's *public* key, into suitable measuring equipment, anyone can verify on the spot that the agent is who she says she is and not an impostor. Compared with *classical* identification protocols, the novel feature here is that the agent does not need to respond to a *challenge*—for example, digitally signing a random string—but can instead just hand out a fixed $|\psi\rangle$ non-interactively. Furthermore, because

(1) $|\psi\rangle$ decoheres in a matter of seconds, and

(2) recovering a classical *description* of $|\psi\rangle$ from measurements on it is computationally intractable,

someone who is given $|\psi\rangle$ cannot use it later to impersonate the agent.

Of course, if an attacker managed to solve the technological problem of keeping $|\psi\rangle$ coherent for very long times, then he could break this system, by collecting one or more copies of $|\psi\rangle$ that an agent had handed out, and using them to impersonate the agent. But in that case, whatever method the attacker was using to keep the states coherent could *also*—once discovered—be used to create a secure public-key quantum money scheme!

However, we would argue that the "real" reason to study quantum money is basically the same as the "real" reason to study quantum computing as a whole—or for that matter, to study the many interesting aspects of *classical* cryptography that are equally far from application. As theoretical computer scientists, we are in the business of mapping out the inherent capabilities and limits of information processing, according to our best current understanding of the physical world.

In our case, what quantum money provides is a near-ideal playground for understanding the implications of the uncertainty principle and the No-Cloning Theorem. In the early days of quantum mechanics, Bohr [17] and others argued that the uncertainty principle requires us to change our conception of science itself—their basic argument being that, in physics, predictions are only ever as good as our knowledge of a system's initial state $|\psi\rangle$, but the uncertainty principle might mean that the initial state is unknowable even with arbitrarily-precise measurements.

But does this argument have any "teeth"? In other words: among the properties of a quantum state $|\psi\rangle$ that make the state impossible to learn precisely or to duplicate, can any of those properties ever *matter empirically*? To us, quantum money is interesting precisely because it gives one of the clearest examples where the answer to that question is yes.

But why study quantum money *relative to oracles*? Our results themselves supply answers to this question. First, like with any other cryptographic problem, once we have an "abstract" scheme

involving an oracle, we can then try to *instantiate* the oracle with an explicit function. In our case, we actually carry this program out, and give an explicit realization of our oracle scheme, which we prove secure under a plausible assumption about multivariate polynomial cryptography. A second motivation comes from Section 7. There, we take the same techniques that we used to construct provably-secure public-key quantum money with an oracle, and apply them to construct provably-secure *private-key* quantum money *without* an oracle, assuming the counterfeiter can adaptively query the bank.

2 Preliminaries

To begin, we fix some notation. Let $[N] = \{1, \ldots, N\}$. Given a subspace S of a vector space V, let $S^{\perp}$ be the orthogonal complement of S (that is, the set of $y \in V$ such that $x \cdot y = 0$ for all $x \in S$). We call a function $\delta(n)$ *negligible* if $\delta(n) = o(1/p(n))$ for every polynomial p.

By a *classical oracle*, we will mean a unitary transformation of the form $|x\rangle \rightarrow (-1)^{f(x)} |x\rangle$, for some Boolean function $f : \{0,1\}^* \rightarrow \{0,1\}$. Note that, unless specified otherwise, even a classical oracle can be queried in quantum superposition. A *quantum oracle*, by contrast, is an arbitrary n-qubit unitary transformation U (or rather, a collection of such U's, one for each n) that a quantum algorithm can apply in a black-box fashion. Quantum oracles were defined and studied by Aaronson and Kuperberg [5].

2.1 Cryptography

Before we construct quantum money schemes, it will be helpful to have some "conventional" cryptographic primitives in our toolbox. Foremost among these is a *digital signature scheme secure against quantum chosen-message attacks*. We now define digital signature schemes—both for completeness, and to fix the quantum attack model that is relevant for us.

Definition 1 (Digital Signature Schemes)**.** *A (classical, public-key)* ***digital signature scheme*** $\mathcal{D}$ *consists of three probabilistic polynomial-time classical algorithms:*

- KeyGen*, which takes as input a security parameter* 0^n*, and generates a* ***key pair*** $(k_{\text{private}}, k_{\text{public}})$*.*
- Sign*, which takes as input* k_{private} *and a message* x*, and generates a* ***signature*** $\mathsf{Sign}(k_{\text{private}}, x)$*.*[10]
- Ver*, which takes as input* k_{public}*, a message* x*, and a claimed signature* w*, and either accepts or rejects.*

We say $\mathcal{D}$ *has* ***completeness error*** ε *if* $\mathsf{Ver}(k_{\text{public}}, x, \mathsf{Sign}(x, k_{\text{private}}))$ *accepts with probability at least* $1 - \varepsilon$ *for all messages* x *and key pairs* $(k_{\text{private}}, k_{\text{public}})$*. Here the probability is over the behavior of* Ver *and* Sign*.*

Let C *(the* ***counterfeiter****) be a quantum circuit of size* $\text{poly}(n)$ *that takes* k_{public} *as input and does the following:*

(1) Probabilistically generates a classical list of messages $x_1, \ldots, x_m$*, and submits them to a* ***signing oracle*** $\mathcal{O}$*.*

[10] We indulge in slight abuse of notation, since if Sign is randomized then the signature need not be a function of k_{private} and x.

(2) Gets back independently-generated signatures $w_1, \ldots, w_m$, where $w_i := \mathsf{Sign}\left(k_{\text{private}}, x_i\right)$.

(3) Outputs a pair (x, w).

*We say C **succeeds** if $x \notin \{x_1, \ldots, x_m\}$ and $\mathsf{Ver}\left(k_{\text{public}}, x, w\right)$ accepts. We say $\mathcal{D}$ has **soundness error** δ if every counterfeiter C succeeds with probability at most δ. Here the probability is over the key pair $(k_{\text{private}}, k_{\text{public}})$ and the behavior of C, Sign, and Ver.*

*We call $\mathcal{D}$ **secure against (nonadaptive) quantum chosen-message attacks** if it has completeness error $\leq 1/3$ and negligible soundness error.*

Intuitively, we call a signature scheme "secure" if no *quantum* counterfeiter with *nonadaptive, classical* access to a signing oracle $\mathcal{O}$ can forge a signature for any message that it did not submit to $\mathcal{O}$. Depending on the application, one might want to change Definition 1 in various ways: for example, by giving the counterfeiter *adaptive* or *quantum* access to $\mathcal{O}$, or by letting KeyGen, Sign, and Ver be quantum algorithms themselves. For this paper, however, Definition 1 provides all we need.

Do signature schemes secure quantum attack actually exist? Naturally, signature schemes based on RSA or other number-theoretic problems can all be broken by a quantum computer. However, building on earlier work by Naor and Yung [37] (among many others), Rompel [41] showed that a secure public-key signature scheme can be constructed from *any* one-way function—not necessarily a trapdoor function. Furthermore, Rompel's security reduction, from breaking the signature scheme to inverting the one-way function, is *black-box*: in particular, nothing in it depends on the assumption that the adversary is classical rather than quantum. We therefore get the following consequence:

Theorem 2 (Quantum-Secure Signature Schemes [41])**.** *If there exists a (classical) one-way function f secure against quantum attack, then there also exists a digital signature scheme secure against quantum chosen-message attacks.*

Recently, Boneh et al. [18] proved several results similar to Theorem 2, and they needed non-trivial work to do so. However, a crucial difference is that Boneh et al. were (justifiably) concerned with quantum adversaries who can make *quantum* queries to the signing oracle $\mathcal{O}$. By contrast, as mentioned earlier, for our application it suffices to consider adversaries who query $\mathcal{O}$ *classically*—and in that case, the standard security reductions go through essentially without change.

Let us state another consequence of Theorem 2, which will be useful for our oracle construction in Section 5.

Theorem 3 (Relativized Quantum-Secure Signatures)**.** *Relative to a suitable oracle A, there exists a digital signature scheme secure against quantum chosen-message attacks.*

Proof Sketch. It is easy to give an oracle $A : \{0,1\}^* \rightarrow \{0,1\}$ relative to which there exists a one-way function $f_n : \{0,1\}^n \rightarrow \{0,1\}^{p(n)}$ secure against quantum adversaries. Indeed, we can let A be a *random* oracle, and then define

$$f_n\left(x\right) := A\left(x, 1\right) \ldots A\left(x, p\left(n\right)\right)$$

directly in terms of A. Assume $p\left(n\right) \geq n$. Then the lower bound on the quantum query complexity of function inversion, proved by Bennett et al. [14] and Ambainis [9], straightforwardly implies

that any quantum algorithm to invert f_n, with success probability $\varepsilon > 0$, must make $\Omega\left(2^{n/2}\sqrt{\varepsilon}\right)$ quantum queries to A.

Now, the security reduction of Rompel [41] is not only black-box but *relativizing*: that is, it goes through if all legitimate and malicious parties have access to the same oracle A. So by Theorem 2, starting from $\{f_n\}$ one can construct a digital signature scheme relative to the same oracle A, which is secure against quantum chosen-message attacks. ■

2.2 Quantum Information

In this section, we collect various facts about quantum pure and mixed states that are used in the paper. We assume basic familiarity with the formalism of bras, kets, density matrices, unitary transformations, and measurements; see Nielsen and Chuang [38] for a good overview.

Given two mixed states ρ and σ, their *trace distance* is defined as

$$D(\rho,\sigma) := \frac{1}{2}\sum_{i=1}^{N} |\lambda_i|,$$

where $\lambda_1,\ldots,\lambda_N$ are the eigenvalues of $\rho-\sigma$. Trace distance is a metric and satisfies $0 \leq D(\rho,\sigma) \leq 1$. Also, the *fidelity* $0 \leq F(\rho,\sigma) \leq 1$ is defined, in this paper, as the maximum of $|\langle\psi|\varphi\rangle|$ over all purifications $|\psi\rangle$ of ρ and $|\varphi\rangle$ of σ.[11] By extension, given a subspace S, we let $F(\rho,S)$ be the maximum of $|\langle\psi|\varphi\rangle|$ over all purifications $|\psi\rangle$ of ρ and all unit vectors $|\varphi\rangle \in S$. Trace distance and fidelity are related as follows [38]:

Proposition 4. *For all mixed states ρ,σ,*

$$D(\rho,\sigma) \leq \sqrt{1-F(\rho,\sigma)^2},$$

with equality if either ρ or σ is pure.

While fidelity is *not* a metric, it does satisfy the following inequality, which will be helpful in Section 5.

Lemma 5 ("Triangle Inequality" for Fidelity)**.** *Suppose $\langle\psi|\rho|\psi\rangle \geq 1-\varepsilon$ and $\langle\varphi|\sigma|\varphi\rangle \geq 1-\varepsilon$. Then $F(\rho,\sigma) \leq |\langle\psi|\varphi\rangle| + 2\varepsilon^{1/4}$.*

Proof. By Proposition 4,

$$D(\rho,|\psi\rangle) \leq \sqrt{1-\langle\psi|\rho|\psi\rangle} \leq \sqrt{\varepsilon},$$

and likewise $D(\sigma,|\varphi\rangle) \leq \sqrt{\varepsilon}$. Thus, since trace distance satisfies the triangle inequality,

$$\begin{aligned} D(\rho,\sigma) &\geq D(|\psi\rangle,|\varphi\rangle) - D(\rho,|\psi\rangle) - D(\sigma,|\varphi\rangle) \\ &\geq \sqrt{1-|\langle\psi|\varphi\rangle|^2} - 2\sqrt{\varepsilon}. \end{aligned}$$

[11] Some authors instead define "fidelity" as the maximum of $|\langle\psi|\varphi\rangle|^2$.

Then

$$\begin{aligned}
|\langle\Psi|\Phi\rangle| &\leq F(\rho,\sigma) \\
&\leq \sqrt{1-D(\rho,\sigma)^2} \\
&\leq \sqrt{1-\left(\sqrt{1-|\langle\psi|\varphi\rangle|^2}-2\sqrt{\varepsilon}\right)^2} \\
&\leq \sqrt{|\langle\psi|\varphi\rangle|^2+4\sqrt{\varepsilon}} \\
&\leq |\langle\psi|\varphi\rangle|+2\varepsilon^{1/4}.
\end{aligned}$$

■

Finally, the following lemma of Aaronson [2] will imply that, as long as a quantum money scheme has small *completeness error* (i.e., small probability of rejecting a valid banknote), the banknotes can be reused many times.

Lemma 6 ("Almost As Good As New Lemma" [2])**.** *Suppose a measurement on a mixed state ρ yields a particular outcome with probability $1-\varepsilon$. Then after the measurement, one can recover a state $\widetilde{\rho}$ such that $\|\widetilde{\rho}-\rho\|_{\mathrm{tr}} \leq \sqrt{\varepsilon}$.*

2.3 Quantum Search

In our security proof for quantum money, a crucial step will be to *amplify* a counterfeiter who copies a banknote \$ with any non-negligible fidelity to a counterfeiter who copies \$ almost perfectly. Taking the contrapositive, this will imply that to rule out the former sort of counterfeiter, it suffices to rule out the latter.

In this section, we first review two variants of Grover's search algorithm [28] that are useful for amplifying the fidelity of quantum states. We then introduce a new variant that combines the advantages of both.

Throughout, assume we are given a pure initial state $|\mathrm{Init}\rangle$, in some Hilbert space $\mathcal{H}$. Our goal is to map $|\mathrm{Init}\rangle$ to a final state $|\Psi\rangle$ that lies in (or close to) a "good subspace" $G \leq \mathcal{H}$. We have oracle access to two unitary transformations:

- U_{Init}, which maps $|\mathrm{Init}\rangle$ to $-|\mathrm{Init}\rangle$, and acts as the identity on all $|v\rangle$ orthogonal to $|\mathrm{Init}\rangle$.
- U_G, which maps $|v\rangle$ to $-|v\rangle$ for all $|v\rangle \in G$, and acts as the identity on all $|v\rangle$ orthogonal to G.

We are promised that the fidelity of the initial state with G,

$$F(|\mathrm{Init}\rangle, G) = \max_{|\psi\rangle \in G} \langle \mathrm{Init}|\psi\rangle,$$

is at least some $\varepsilon > 0$.

In this scenario, *provided $F(|\mathrm{Init}\rangle, G)$ is known*, the amplitude amplification framework of Brassard, Høyer, Mosca, and Tapp [20] lets us prepare a state close to G using only $\Theta(1/\varepsilon)$ iterations:

Lemma 7 (Amplitude Amplification [20]). *Write* $|\text{Init}\rangle$ *as* $\sin\theta\,|\text{Good}\rangle + \cos\theta\,|\text{Bad}\rangle$*, where* $|\text{Good}\rangle$ *is the unit vector in* G *closest to* $|\text{Init}\rangle$ *and* $|\text{Bad}\rangle$ *is orthogonal to* $|\text{Good}\rangle$*. Then by using* $O(T)$ *oracle calls to* U_{Init} *and* U_G*, we can prepare the state*

$$|\Phi_T\rangle := \sin\left[(2T+1)\,\theta\right]|\text{Good}\rangle + \cos\left[(2T+1)\,\theta\right]|\text{Bad}\rangle$$

Note that Grover's algorithm is simply a special case of Lemma 7, where $|\text{Init}\rangle$ is the uniform superposition over N basis states $|1\rangle, \ldots, |N\rangle$, and G is the subspace spanned by "marked" states.

However, Lemma 7 has an annoying drawback, which it shares with ordinary Grover search. Namely, the algorithm does *not* converge monotonically toward the target subspace G, but could instead "wildly overshoot it," cycling around the 2-dimensional subspace spanned by $|\text{Bad}\rangle$ and $|\text{Good}\rangle$. If we know the fidelity $F(|\text{Init}\rangle, G)$ in advance (rather than just a lower bound on the fidelity), *or* if we can prepare new copies of $|\text{Init}\rangle$ "free of charge" in case of failure, then this overshooting is not a serious problem. Alas, neither of those conditions will hold in our application.

Fortunately, for independent reasons, in 2005 Tulsi, Grover, and Patel [42] introduced a new quantum search algorithm that *does* guarantee monotonic convergence toward G, by alternating unitary transformations with measurements. (Their algorithm was later simplified and improved by Chakraborty, Radhakrishnan, and Raghunathan [21].)

Lemma 8 (Fixed-Point Quantum Search [42, 21]). *By using* T *oracle calls to* U_{Init} *and* U_G*, we can prepare a state* $|\Psi\rangle$ *such that* $F(|\Psi\rangle, G) \geq 1 - \exp\left(-T\varepsilon^2\right)$*.*

Rearranging, Lemma 8 lets us prepare a state $|\Psi\rangle$ such that $F(|\Psi\rangle, G) \geq 1-\delta$ using $T = O\left(\frac{1}{\varepsilon^2}\ln\frac{1}{\delta}\right)$ iterations. On the positive side, the dependence on $1/\delta$ in this bound is logarithmic: we get not only monotonic convergence toward G, but *exponentially-fast* convergence. On the negative side, notice that the dependence on ε has worsened from $1/\varepsilon$ to $1/\varepsilon^2$—negating the quadratic speedup that was the original point of quantum search!

In the rest of this section, we give a "hybrid" quantum search algorithm that *combines* the advantages of Lemmas 7 and 8—i.e., it converges monotonically toward the target subspace G (rather than "overshooting" G), but *also* achieves a quadratic speedup. In the context of our security proof for quantum money, the hybrid algorithm will lead to a quadratically-better lower bound on the number of queries that a counterfeiter needs to make, compared to what we would get from using Lemma 8 by itself. While this quadratic improvement is perhaps only of moderate interest, we include the algorithm in the hope that it will find other applications.

We first give a technical lemma needed to analyze our algorithm.

Lemma 9. *For all* L, β, η, γ*, there are at most* $(L/\beta+1)(2\eta+1)$ *integers* $T \in \{0, \ldots, L\}$ *such that* $|T - (\beta n + \gamma)| < \eta$ *for some integer* n*.*

Proof. The real interval $[0, L]$ can intersect at most $L/\beta+1$ intervals $(\beta n + \gamma - \eta, \beta n + \gamma + \eta)$, and each such interval can contain at most $2\eta+1$ integer points. ■

We now give our hybrid of Lemmas 7 and 8. Note that our hybrid loses the property of *exponentially-fast* convergence toward the target subspace G, but that property will not be important for us anyway. We leave as an open problem whether there exists a hybrid algorithm with exponentially-fast convergence.

Theorem 10 (Faster Fixed-Point Search). *Let* $\delta \geq 2\varepsilon$*. Then by using* $O\left(\frac{\log 1/\delta}{\varepsilon\delta^2}\right)$ *oracle calls to* U_{Init} *and* U_G*, we can prepare a state* ρ *such that* $F(\rho, G) \geq 1-\delta$*.*

Proof. Let $\epsilon := \arcsin \varepsilon$; note that $\varepsilon \leq \epsilon \leq \frac{\pi}{2}\varepsilon$. Also let $L := \lceil 100/\epsilon \rceil$ and $R := \frac{25}{\delta^2}\left(2 + \log \frac{1}{\delta}\right)$. Then the algorithm is as follows:

(1) Choose an integer $T \in \{0, \ldots, L\}$ uniformly at random.

(2) Apply T iterations of amplitude amplification with $|\mathrm{Init}\rangle$ as the initial state and G as the target subspace (as in Lemma 7), to obtain a state $|\Phi_T\rangle$.

(3) Apply R iterations of fixed-point quantum search with $|\Phi_T\rangle$ as the initial state and G as the target subspace (as in Lemma 8), to obtain a state $|\Psi\rangle = |\Psi_T\rangle$.

The final output of the above algorithm is

$$\rho = \mathop{\mathrm{E}}_{T \in \{0,\ldots,L\}} \left[|\Psi_T\rangle \langle \Psi_T| \right].$$

Also, the total number of oracle calls to U_{Init} and U_G is

$$O(TR) = O\left(\frac{\log 1/\delta}{\varepsilon \delta^2}\right).$$

(The reason this number scales like TR rather than $T + R$ is that, in step (3), each time we reflect about the initial state $|\Phi_T\rangle$ we need to rerun step (2). Thus, we need $\Theta(T)$ oracle calls within each of the R iterations.)

By Lemma 7, after step (2) we have a state $|\Phi_T\rangle$ such that

$$F(|\Phi_T\rangle, G) = |\langle \Phi_T | \mathrm{Good}\rangle| = |\sin[(2T+1)\epsilon]|.$$

So for any $\alpha \in (0, 1)$,

$$\begin{aligned}
\Pr_{T \in \{0,\ldots,L\}} \left[F(|\Phi_T\rangle, G) < \alpha \right] &= \Pr_{T \in \{0,\ldots,L\}} \left[|\sin[(2T+1)\epsilon]| < \alpha \right] \\
&= \Pr_{T \in \{0,\ldots,L\}} \left[\exists n \in \mathbb{Z} : |(2T+1)\epsilon - \pi n| < \arcsin \alpha \right] \\
&\leq \frac{\left(\frac{L}{\pi/2\epsilon} + 1\right)\left(\frac{\arcsin \alpha}{\epsilon} + 1\right)}{L+1} \\
&\leq \frac{2}{\pi} \arcsin \alpha + \frac{2\epsilon}{\pi} + \frac{\arcsin \alpha}{100} + \frac{\epsilon}{100} \\
&\leq 1.02(\alpha + \varepsilon),
\end{aligned}$$

where the third line uses Lemma 9.

Now assume $F(|\Phi_T\rangle, G) \geq \alpha$. Then by Lemma 8, after step (3) we have a state $|\Psi_T\rangle$ such that

$$F(|\Phi_T\rangle, G) \geq 1 - \exp\left(-R\alpha^2\right).$$

Let us now make the choice $\alpha := \delta/5$. Then by the union bound, the "average" output $\rho = \mathrm{E}_T \left[|\Psi_T\rangle \langle \Psi_T| \right]$ satisfies

$$\begin{aligned}
1 - F(\rho, G) &\leq 1.02(\alpha + \varepsilon) + \exp\left(-R\alpha^2\right) \\
&\leq 0.204\delta + 0.51\delta + \exp\left(-\frac{R\delta^2}{25}\right) \\
&< \delta.
\end{aligned}$$

■

3 Formalizing Quantum Money

In this section, we first give a formal cryptographic definition of *public-key quantum money schemes*. Our definition is similar to that of Aaronson [3]. However, departing from [3], we next define the new notion of a *quantum money mini-scheme*, which is easier to construct and analyze than a full-blown quantum money scheme. A mini-scheme is basically a quantum money scheme where only *one* banknote ever needs to printed, not many banknotes; and where the procedure for verifying that banknote is treated as given (rather than something that *itself* needs to be authenticated using the bank's public key). We then prove two results:

- **Theorem 15:** Any mini-scheme that is secure against almost-*perfect* copying, and that is *projective* (i.e., the verification procedure consists of a projective measurement), is also secure against copying with any non-negligible fidelity.
- **Theorem 16:** Given a secure mini-scheme, as well as a digital signature scheme secure against quantum chosen-message attacks, one can construct a secure quantum money scheme. The basic idea, which was also used by Lutomirski et al. [34] and Farhi et al. [24], is to append to each quantum banknote $|\$_s\rangle = |s\rangle\,|\psi_s\rangle$ a digital signature of the classical serial number s, thereby proving that $|\$_s\rangle$ came from the bank. We provide the first proof of this construction's security.

As discussed in Section 2.1, the outlook for quantumly-secure signature schemes is extremely bright. For by using known results [37, 41], one can construct such a scheme starting from *any quantumly-secure one-way function* (not even a trapdoor OWF is needed). Certainly, such OWFs exist under the sorts of assumptions that we will already be making anyway in the construction of mini-schemes! Thus, Theorems 15 and 16 tell us that, to solve the quantum money problem, it suffices to construct a projective mini-scheme that is secure against almost-perfect copying. So that is what we proceed to do in the rest of the paper: in the oracle world in Section 5, and in the "real" world (under an assumption about multivariate polynomial cryptography) in Section 6.

3.1 Quantum Money Schemes

Intuitively, a *public-key quantum money scheme* is a scheme by which

(1) a trusted "bank" can feasibly generate an unlimited number of quantum banknotes,

(2) anyone can feasibly verify a valid banknote as having come from the bank, but

(3) no one besides the bank can feasibly map $q = \text{poly}(n)$ banknotes to $r > q$ banknotes with any non-negligible success probability.[12]

We now make the notion more formal.

Definition 11 (Quantum Money Schemes)**.** *A* ***public-key quantum money scheme*** $\mathcal{S}$ *consists of three polynomial-time quantum algorithms:*

[12] Previously, Aaronson [3] required only that no polynomial-time counterfeiter could increase its *expected* number of valid banknotes. However, the stronger condition required here is both achievable, and seemingly more natural from the standpoint of security proofs.

- KeyGen, *which takes as input a security parameter* 0^n, *and probabilistically generates a key pair* $(k_{\text{private}}, k_{\text{public}})$.

- Bank, *which takes as input* k_{private}, *and probabilistically generates a quantum state* $\$$ *called a* ***banknote***. *(Usually* $\$$ *will be an ordered pair* (s, ρ_s), *consisting of a classical* ***serial number*** s *and a* ***quantum money state*** ρ_s, *but this is not strictly necessary.)*

- Ver, *which takes as input* k_{public} *and an alleged banknote* $\not{c}$, *and either accepts or rejects.*

We say $\mathcal{S}$ *has* ***completeness error*** ε *if* $\mathsf{Ver}(k_{\text{public}}, \$)$ *accepts with probability at least* $1-\varepsilon$ *for all public keys* k_{public} *and valid banknotes* $\$$. *If* $\varepsilon = 0$ *then* $\mathcal{S}$ *has* ***perfect completeness***.

Let Count *(the* ***money counter****) take as input* k_{public} *as well as a collection of (possibly-entangled) alleged banknotes* $\not{c}_1, \ldots, \not{c}_r$, *and output the number of indices* $i \in [r]$ *such that* $\mathsf{Ver}(k_{\text{public}}, \not{c}_i)$ *accepts. Then we say* $\mathcal{S}$ *has* ***soundness error*** δ *if, given any quantum circuit* $C(k_{\text{public}}, \$_1, \ldots, \$_q)$ *of size* $\text{poly}(n)$ *(called the* ***counterfeiter****), which maps* $q = \text{poly}(n)$ *valid banknotes* $\$_1, \ldots, \$_q$ *to* $r = \text{poly}(n)$ *(possibly-entangled) alleged banknotes* $\not{c}_1, \ldots, \not{c}_r$,

$$\Pr\left[\mathsf{Count}\left(k_{\text{public}}, C\left(k_{\text{public}}, \$_1, \ldots, \$_q\right)\right) > q\right] \leq \delta.$$

Here the probability is over the key pair $(k_{\text{private}}, k_{\text{public}})$, *valid banknotes* $\$_1, \ldots, \$_q$ *generated independently by* $\mathsf{Bank}(k_{\text{private}})$, *and the behavior of* Count *and* C.

We call $\mathcal{S}$ ***secure*** *if it has completeness error* $\leq 1/3$ *and negligible soundness error.*

In Appendix 10, we will show that the completeness error in any quantum money scheme can be amplified to $1/2^{\text{poly}(n)}$, at the cost of only a small increase in the soundness error. Note that, by Lemma 6 (the "Almost As Good As New Lemma"), once we make the completeness error exponentially small, we can also give our scheme the property that *any banknote* $\$$ *can be verified exponentially many times*, before $\$$ gets "worn out" by repeated measurements. This observation is part of what justifies our use of the term "money."[13] We leave as an open problem whether the soundness error in any quantum money scheme can be similarly amplified.

In this paper, we will often consider **relativized** quantum money schemes, which simply means that the three procedures KeyGen, Bank, Ver—as well as the counterfeiter C—all get access to exactly the same oracle $A : \{0,1\}^* \rightarrow \{0,1\}$. We will also consider relativized digital signature schemes, etc., which are defined analogously.

A **private-key quantum money scheme** is the same as a public-key scheme, except that the counterfeiter C no longer gets access to k_{public}. (Thus, we might as well set $k := k_{\text{public}} = k_{\text{private}}$, since the public and private keys no longer play separate roles.) We call a private-key scheme **query-secure**—a notion "intermediate" between private-key and public-key—if the counterfeiter C is allowed to interact repeatedly with a "naïve and trusting bank." Given any alleged banknote σ, such a bank runs the verification procedure $\mathsf{Ver}(k, \sigma)$, then returns to C both the classical result (i.e., accept or reject) *and* the post-measurement quantum state $\widetilde{\sigma}$.

Let us now use these definitions to place our work in context. For us, just like for Aaronson [3], Lutomirski et al. [34], and Farhi et al. [24], the main goal will be *public-key* quantum money.

In our language, the original work of Wiesner [43] gave a *private-key* quantum money scheme with information-theoretic security, *relative to a random oracle* R. (The oracle R—or rather a

[13] By contrast, BBBW [16] introduced the term "subway tokens" for quantum money states that get destroyed immediately upon verification.

secret part of it—provides the giant database of random qubit polarizations used by the bank.) Also, the later work of BBBW [16] gave a private-key quantum money scheme with no oracle, but assuming quantumly-secure pseudorandom functions.

While a full discussion of private-key quantum money is outside the scope of this paper, Section 7 will take up the crucial issue of *query-security* for private-key schemes. Previously, Lutomirski [32] and Aaronson [3] pointed out that the Wiesner and BBBW schemes are *not* query-secure. However, we will show in Section 7 that the machinery developed in this paper—for giving *public-key* quantum money schemes that are secure relative to a classical oracle—can be readily adapted to solve the query-security problem as well. In particular, we construct:

- A private-key quantum money scheme that is *unconditionally query-secure*, and that uses a random oracle R (i.e., assumes that the bank has a giant database of random numbers, just like Wiesner's scheme [43]).
- A private-key quantum money scheme with *no* oracle, and which is provably query-secure assuming we have a quantumly-secure pseudorandom function family (just like the BBBW scheme [16]).

In both of the above schemes, the quantum money states consist of uniform superpositions $|A\rangle$ over subspaces $A \leq \mathbb{F}_2^n$ with $\dim(A) = n/2$, just like in our public-key scheme, and the security proof uses the same ingredients. We find it striking that we would not know how to close a major security loophole even in *private-key* quantum money, without using tools—including the inner-product adversary method—that were originally developed for the seemingly much more ambitious goal of public-key quantum money.

3.2 Mini-Schemes

While Definition 11 captures our intuitive requirements for a public-key quantum money scheme, experience has shown that it is cumbersome to work with in practice. So in this section, we introduce a simpler primitive called *mini-schemes*, which require only *one* uncopyable banknote. We also prove an amplification theorem for a large class of mini-schemes. Then, in Section 3.3, we will show how mini-schemes can be generically combined with conventional digital signature schemes to create full public-key quantum money schemes.

Definition 12 (Mini-Schemes)**.** *A (public-key)* ***mini-scheme*** $\mathcal{M}$ *consists of two polynomial-time quantum algorithms:*

- Bank*, which takes as input a security parameter* 0^n*, and probabilistically generates a banknote* $\$ = (s, \rho_s)$*, where* s *is a classical* ***serial number****, and* ρ_s *is a quantum money state.*
- Ver*, which takes as input an alleged banknote* $\not{c}$*, and either accepts or rejects.*

We say $\mathcal{M}$ *has* ***completeness error*** ε *if* $\mathsf{Ver}(\$)$ *accepts with probability at least* $1-\varepsilon$ *for all valid banknotes* $\$$*. If* $\varepsilon = 0$ *then* $\mathcal{M}$ *has perfect completeness. If, furthermore,* $\rho_s = |\psi_s\rangle\langle\psi_s|$ *is always a pure state, and* Ver *simply consists of a projective measurement onto the rank-*1 *subspace spanned by* $|\psi_s\rangle$*, then we say* $\mathcal{M}$ *is* ***projective****.*[14]

[14] We similarly call a full quantum money scheme projective, if $\mathsf{Ver}(\$)$ consists of a measurement on one part of $\$$ in the computational basis, followed by a rank-1 projective measurement on the remaining part.

Let Ver_2 *(the* ***double verifier****) take as input a single serial number* s *as well as two (possibly-entangled) states* σ_1 *and* σ_2*, and accept if and only* $\mathsf{Ver}(s, \sigma_1)$ *and* $\mathsf{Ver}(s, \sigma_2)$ *both accept. We say* $\mathcal{M}$ *has* ***soundness error*** δ *if, given any quantum circuit* C *of size* $\mathrm{poly}(n)$ *(the* ***counterfeiter****),* $\mathsf{Ver}_2(s, C(\$))$ *accepts with probability at most* δ*. Here the probability is over the banknote* $\$$ *output by* $\mathsf{Bank}(0^n)$*, as well as the behavior of* Ver_2 *and* C*.*

We call $\mathcal{M}$ ***secure*** *if it has completeness error* $\leq 1/3$ *and negligible soundness error.*

We observe a simple relationship between Definitions 11 and 12:

Proposition 13. *If there exists a secure public-key money scheme* $\mathcal{S} = (\mathsf{KeyGen}_{\mathcal{S}}, \mathsf{Bank}_{\mathcal{S}}, \mathsf{Ver}_{\mathcal{S}})$*, then there also exists a secure mini-scheme* $\mathcal{M} = (\mathsf{Bank}_{\mathcal{M}}, \mathsf{Ver}_{\mathcal{M}})$*.*

Proof. Each banknote output by $\mathsf{Bank}_{\mathcal{M}}(0^n)$ will have the form $(k_{\text{public}}, \mathsf{Bank}_{\mathcal{S}}(k_{\text{private}}))$, where $(k_{\text{private}}, k_{\text{public}})$ is a key pair output by $\mathsf{KeyGen}_{\mathcal{S}}(0^n)$. Then $\mathsf{Ver}_{\mathcal{M}}(s, \rho_s)$ will accept if and only if $\mathsf{Ver}_{\mathcal{S}}(s, \rho_s)$ does. Any counterfeiter $C_{\mathcal{M}}$ against $\mathcal{M}$ can be converted directly into a counterfeiter $C_{\mathcal{S}}$ against $\mathcal{S}$. ■

We can classify mini-schemes according to how the banknotes are generated. In particular, call a mini-scheme $\mathcal{M} = (\mathsf{Bank}, \mathsf{Ver})$ **secret-based** if there exists

(1) a polynomial-time deterministic, classical[15] algorithm $\mathsf{SerialNum}$, which takes as input a uniformly-random string $r \in \{0,1\}^{\mathrm{poly}(n)}$ and outputs a serial number s_r, and

(2) a polynomial-time quantum algorithm $\mathsf{MoneyState}$, which takes r as input and outputs a money state ρ_r,

such that $\$_r := (s_r, \rho_r)$ is indistinguishable from a banknote $\$ = (s, \rho_s)$ output by $\mathsf{Bank}(0^n)$. Otherwise, call $\mathcal{M}$ **non-secret-based**. Intuitively, in a secret-based scheme, the bank can generate many identical banknotes by simply reusing r, while in a non-secret-based scheme, *not even the bank* might be able to generate two identical banknotes. Here is an interesting observation:

Proposition 14. *If there exists a secure, secret-based mini-scheme, then there also exists a one-way function secure against quantum attack.*

Proof. The desired OWF is $\mathsf{SerialNum}(r) := s_r$. If there existed a polynomial-time quantum algorithm to recover r given s_r, then we could use that algorithm to produce an unlimited number of additional banknotes $\$_r$. ■

All of the mini-schemes developed in this paper will be secret-based. By contrast, the earlier schemes of Lutomirski et al. [34] and Farhi et al. [24] are (conjectured to be) non-secret-based, since the serial number s is only obtained as the outcome of a quantum measurement. Understanding the relation between these two types of quantum money, and the relative advantages of each, remains an important problem.

The following result is one of the most useful in the paper. Intuitively, it says that in projective mini-schemes, a counterfeiter that copies a banknote with *any* non-negligible fidelity can be "amplified" to a counterfeiter that copies the banknote almost *perfectly*—or conversely, that to rule out the former sort of counterfeiter, it suffices to rule out the latter. The proof makes essential use of the amplitude amplification results from Section 2.3.

[15] In general, we could also let $\mathsf{SerialNum}$ be a quantum algorithm. However, this freedom is not used in any actual secret-based mini-scheme that we know about, and restricting to classical algorithms has the advantage that it will let us relate secret-based mini-schemes to conventional one-way functions (in Proposition 14).

Theorem 15 (Amplification of Counterfeiters)**.** *Let $\mathcal{M} = (\mathsf{Bank}, \mathsf{Ver})$ be a projective mini-scheme, and let $\$ = (s, \rho)$ be a valid banknote in $\mathcal{M}$. Suppose there exists a counterfeiter C that copies $\$$ with probability $\varepsilon > 0$: that is,*

$$\Pr\left[\mathsf{Ver}_2\left(s, C\left(\$\right)\right) \text{ accepts}\right] \geq \varepsilon.$$

Then for all $\delta > 0$, there is also a modified counterfeiter C' (depending only on ε and δ, not $\$$), which makes $O\left(\frac{\log 1/\delta}{\sqrt{\varepsilon}\left(\sqrt{\varepsilon}+\delta^2\right)}\right)$ queries to C, C^{-1}, and Ver and which satisfies

$$\Pr\left[\mathsf{Ver}_2\left(s, C'\left(\$\right)\right) \text{ accepts}\right] \geq 1 - \delta.$$

Proof. Write $\$$ as a mixture of pure states:

$$\$ = \sum p_i \left|\psi_i\right\rangle \left\langle\psi_i\right|.$$

By linearity, clearly it suffices to show that

$$\Pr\left[\mathsf{Ver}_2\left(s, C'\left(\left|\psi_i\right\rangle\right)\right) \text{ accepts}\right] \geq 1 - \delta$$

for all i such that $p_i > 0$. We focus on $\left|\psi\right\rangle := \left|\psi_1\right\rangle$ without loss of generality.

By assumption, there exists a subspace S such that $\Pr\left[\mathsf{Ver}\left(\rho\right) \text{ accepts}\right] = F\left(\rho, S\right)$ for all ρ. Then $F\left(\$, S\right) = F\left(\left|\psi\right\rangle, S\right) = 1$.

Now, just as Ver is simply a projector onto S, so Ver_2 is a projector onto $S^{\otimes 2}$. Thus

$$F\left(C\left(\left|\psi\right\rangle\right), S^{\otimes 2}\right) \geq \sqrt{\varepsilon}.$$

So consider performing a fixed-point Grover search, with $C\left(\left|\psi\right\rangle\right)$ as the initial state and $S^{\otimes 2}$ as the target subspace. By Lemma 8, this will produce a state ρ such that $F\left(\rho, S^{\otimes 2}\right) \geq 1 - \delta$ using $O\left(\frac{1}{\varepsilon}\log\frac{1}{\delta}\right)$ Grover iterations. Each iteration requires a reflection about $C\left(\left|\psi\right\rangle\right)$ and a reflection about $S^{\otimes 2}$, which can be implemented using $O\left(1\right)$ queries to C, C^{-1} and Ver respectively. Therefore the number of queries to C, C^{-1} and Ver is $O\left(\frac{1}{\varepsilon}\log\frac{1}{\delta}\right)$ as well.

If δ is large compared to ε, then we can instead use Theorem 10, which produces a state ρ such that $F\left(\rho, S^{\otimes 2}\right) \geq 1 - \delta$ using $O\left(\frac{1}{\sqrt{\varepsilon}\delta^2}\log\frac{1}{\delta}\right)$ iterations. Taking the minimum of the two bounds gives us the claimed bound of

$$O\left(\frac{\log 1/\delta}{\sqrt{\varepsilon}\left(\sqrt{\varepsilon}+\delta^2\right)}\right).$$

■

Theorem 15 is unlikely to hold for *arbitrary* (non-projective) mini-schemes, for the simple reason that we can always create a mini-scheme where Ver accepts *any* state with some small nonzero probability ε. We leave it as an open problem to find the largest class of mini-schemes for which Theorem 15 holds.

3.3 The Standard Construction

We are now ready to define the "standard construction" of public-key quantum money schemes from mini-schemes and digital signature schemes, and to prove this construction's security.

Theorem 16 (Standard Construction of Public-Key Quantum Money)**.** *Let $\mathcal{M}$ be any secure mini-scheme, and let $\mathcal{D}$ be any digital signature scheme secure against quantum chosen-message attacks. By combining $\mathcal{M}$ and $\mathcal{D}$, we can create a secure public-key quantum money scheme $\mathcal{S}$.*

Proof. Given $\mathcal{M} = (\mathsf{Bank}_{\mathcal{M}}, \mathsf{Ver}_{\mathcal{M}})$ and $\mathcal{D} = (\mathsf{KeyGen}_{\mathcal{D}}, \mathsf{Sign}_{\mathcal{D}}, \mathsf{Ver}_{\mathcal{D}})$, our quantum money scheme $\mathcal{S} = (\mathsf{KeyGen}_{\mathcal{S}}, \mathsf{Bank}_{\mathcal{S}}, \mathsf{Ver}_{\mathcal{S}})$ is defined as follows:

- $\mathsf{KeyGen}_{\mathcal{S}}$ is simply $\mathsf{KeyGen}_{\mathcal{D}}$ from the digital signature scheme.
- $\mathsf{Bank}_{\mathcal{S}}$ first calls $\mathsf{Bank}_{\mathcal{M}}$ from the mini-scheme to obtain a banknote (s, ρ). It then outputs (s, ρ) together with a digital signature of the serial number s:

$$\mathsf{Bank}_{\mathcal{S}}\left(k_{\text{private}}\right) := \left(s, \mathsf{Sign}_{\mathcal{D}}\left(k_{\text{private}}, s\right), \rho\right).$$

- $\mathsf{Ver}_{\mathcal{S}}$ accepts an alleged banknote (s, w, σ), if and only if $\mathsf{Ver}_{\mathcal{M}}(s, \sigma)$ and $\mathsf{Ver}_{\mathcal{D}}(k_{\text{public}}, s, w)$ both accept.

Now, suppose there exists a counterfeiter $C_{\mathcal{S}}$ against $\mathcal{S}$: that is, a polynomial-time quantum algorithm such that

$$\Pr\left[\mathsf{Count}\left(k_{\text{public}}, C_{\mathcal{S}}\left(k_{\text{public}}, \$_1, \ldots, \$_q\right)\right) > q\right] \geq \frac{1}{p(n)}.$$

Here $\$_i := (s_i, w_i, \rho_i)$ is a valid banknote, Count is the money counter from Definition 11, and p is some polynomial. Also, the probability is over the key pair $(k_{\text{private}}, k_{\text{public}})$, the valid banknotes $\$_1, \ldots, \$_q$, and the behavior of Count and $C_{\mathcal{S}}$. Suppose further that $\mathcal{D}$ is secure. Then it suffices to show that, by using $C_{\mathcal{S}}$, we can construct a counterfeiter $C_{\mathcal{M}}$ against the underlying mini-scheme $\mathcal{M}$. Let $\mathsf{New}\left(k_{\text{public}}, \$_1, \ldots, \$_q\right)$ be an algorithm that does the following:

(1) Records the serial numbers $s_1, \ldots, s_q$ of $\$_1, \ldots, \$_q$, and lets $U := \{s_1, \ldots, s_q\}$.

(2) Runs $C_{\mathcal{S}}\left(k_{\text{public}}, \$_1, \ldots, \$_q\right)$, and examines the output states $\phi_1, \ldots, \phi_r$.

(3) Returns the number of $i \in [r]$ such that $\mathsf{Ver}_{\mathcal{S}}(\phi_i)$ accepts, *and* ϕ_i's serial number s'_i does not belong to U.

Then we claim that

$$\Pr\left[\mathsf{New}\left(k_{\text{public}}, \$_1, \ldots, \$_q\right) > 0\right]$$

is negligibly small, where the probability is over the same variables as before. The proof is simply that, if this were not so, then we could easily create a counterfeiter $C_{\mathcal{D}}$ against the digital signature scheme $\mathcal{D}$. With non-negligible probability, $C_{\mathcal{D}}$ would generate a valid signature $\mathsf{Sign}_{\mathcal{D}}\left(k_{\text{private}}, s'_i\right)$, for a message s'_i for which $C_{\mathcal{D}}$ had never before seen a valid signature, by running $C_{\mathcal{S}}\left(k_{\text{public}}, \$_1, \ldots, \$_q\right)$, then measuring $\phi_i = (s'_i, w'_i, \rho'_i)$ for a uniformly random $i \in [r]$. (Note that $C_{\mathcal{D}}$ can generate q money states $\$_1, \ldots, \$_q$, without knowledge of k_{private}, by generating the s_i's and ρ_i's on its own, then calling the signing oracle $\mathcal{O}$ to get the w_i's.)

We now define a counterfeiter $C_{\mathcal{M}}$ against the mini-scheme $\mathcal{M}$, as follows:

(i) Run $\mathsf{KeyGen}_{\mathcal{D}}\left(0^n\right)$, to generate a *new* key pair $\left(k'_{\text{private}}, k'_{\text{public}}\right)$.

(ii) Label the banknote to copied (s_ℓ, ρ_ℓ), for some $\ell \in [q]$ chosen uniformly at random.

(iii) Repeatedly call $\mathsf{Bank}_{\mathcal{M}}(0^n)$ to generate $q-1$ serial numbers and quantum money states, labeled (s_i, ρ_i) for all $i \in [q] \setminus \{\ell\}$. Let $U := \{s_1, \ldots, s_q\}$.

(iv) Generate a digital signature $w_i := \mathsf{Sign}_{\mathcal{D}}\left(k'_{\text{private}}, s_i\right)$ for each $i \in [q]$. Let $\$_i := (s_i, w_i, \rho_i)$.

(v) Run the counterfeiter $C_{\mathcal{S}}(k_{\text{public}}, \$_1, \ldots, \$_q)$, to obtain $r > q$ alleged banknotes $\not{\$}_1, \ldots, \not{\$}_r$ where $\not{\$}_j = \left(s'_j, w'_j, \rho'_j\right)$.

(vi) Choose $j, k \in [r]$ uniformly at random without replacement, and output $\left(\rho'_j, \rho'_k\right)$ as a candidate for two copies of ρ_ℓ.

Suppose that $\mathsf{Count} > q$, as happens with probability at least $\frac{1}{p(n)}$. Also suppose that $\mathsf{New} = 0$, as happens all but a negligible fraction of the time. Then by the pigeonhole principle, there must exist indices $j \neq k$ such that $s'_j = s'_k$. With probability at least $1/\binom{r}{2}$, step (vi) will find such a (j, k) pair. Therefore $C_{\mathcal{D}}$ succeeds with overall probability $\Omega(1/\operatorname{poly}(n))$. ■

Theorem 16 reduced the construction of a public-key quantum money scheme to two "smaller" problems:

(1) constructing a mini-scheme, and

(2) constructing a signature scheme secure against quantum attacks.

In practice, however, the situation is even better than that, since in this paper, all of our constructions of mini-schemes will *also* yield signature schemes "free of charge"! The following proposition explains why:

Proposition 17. *If there exists a secure, secret-based mini-scheme $\mathcal{M}$, then there also exists a secure public-key quantum money scheme $\mathcal{S}$.*

Proof. Starting from $\mathcal{M}$, we can get a one-way function secure against quantum attack from Proposition 14, and hence a digital signature scheme $\mathcal{D}$ secure against quantum chosen-message attack from Theorem 2. Combining $\mathcal{M}$ and $\mathcal{D}$ now yields $\mathcal{S}$ by Theorem 16. ■

Proposition 17 means that in constructing secret-based quantum money schemes, there is no need to postulate signature schemes separately.

Finally, let us make explicit what Theorem 16 means for oracle construction.

Corollary 18. *Suppose there exists a mini-scheme $\mathcal{M}$ that is provably secure relative to some oracle $A_{\mathcal{M}}$ (i.e., any counterfeiter $C_{\mathcal{M}}$ against $\mathcal{M}$ must make superpolynomially many queries to $A_{\mathcal{M}}$). Then there exists a public-key quantum money scheme $\mathcal{S}$ that is provably secure relative to some other oracle $A_{\mathcal{S}}$.*

Proof. By Theorem 3, relative to a suitable oracle $A_{\mathcal{D}}$ (in fact, a *random* oracle suffices), there exists a signature scheme $\mathcal{D}$, such that any quantum chosen-message attack against $\mathcal{D}$ must make superpolynomially many queries to $A_{\mathcal{D}}$.

The oracle $A_{\mathcal{S}}$ will simply be a concatenation of $A_{\mathcal{M}}$ with $A_{\mathcal{D}}$. Relative to $A_{\mathcal{S}}$, we claim that the mini-scheme $\mathcal{M}$ and signature scheme $\mathcal{D}$ *both* remain secure—and therefore, by Theorem 16, we can construct a secure public-key quantum money scheme $\mathcal{S}$.

The only worry is that a counterfeiter $C_{\mathcal{M}}$ against $\mathcal{M}$ might gain some advantage by querying $A_{\mathcal{D}}$; or conversely, a counterfeiter $C_{\mathcal{D}}$ against $\mathcal{D}$ might gain some advantage by querying $A_{\mathcal{M}}$. However, this worry is illusory, for the simple reason that the oracles $A_{\mathcal{D}}$ and $A_{\mathcal{M}}$ are generated independently. Thus, if $C_{\mathcal{M}}$ can break $\mathcal{M}$ by querying $A_{\mathcal{D}}$, then it can *also* break $\mathcal{M}$ by querying a randomly-generated "mock-up" $A'_{\mathcal{D}}$ of $A_{\mathcal{D}}$; and conversely, if $C_{\mathcal{D}}$ can break $\mathcal{D}$ by querying $A_{\mathcal{M}}$, then it can also break $\mathcal{D}$ by querying a randomly-generated mock-up $A'_{\mathcal{M}}$ of $A_{\mathcal{M}}$. Recall that any attack on $\mathcal{M}$ must make superpolynomially many *queries* to $A_{\mathcal{M}}$, while any attack on $\mathcal{D}$ must make superpolynomially many queries to $A_{\mathcal{D}}$. So regardless of the *computational* cost of generating $A'_{\mathcal{D}}$ and $A'_{\mathcal{M}}$, we obtain a break against $\mathcal{D}$ or $\mathcal{M}$ that makes only poly (n) oracle queries, thereby giving the desired contradiction. ■

4 The Inner-Product Adversary Method

At least in the black-box or oracle setting, our goal is to create quantum money (mini-)schemes that we can *prove* are secure—by showing that any counterfeiter would need to make exponentially many queries to some oracle. Proving security results of this kind turns out to require interesting quantum lower bound machinery. In this section, we introduce the *inner-product adversary method*, a new variant of Ambainis's quantum adversary method [9] that is well-adapted to proving the security of quantum money schemes, and that seems likely to find other applications.

Let us first explain the difficulty that our method needs to overcome. In a public-key quantum money scheme, any counterfeiter C has *two* powerful resources available:

(1) One or more copies of a "legitimate" quantum money state $|\psi\rangle$.

(2) Access to a *verification procedure* V, which accepts $|\psi\rangle$ and rejects every state orthogonal to $|\psi\rangle$.

Indeed, for us, the situation will be even better for C (i.e., worse for us!), since C can query not only the verification procedure V itself, but also an underlying *classical* oracle U that the legitimate buyers and sellers use to implement V. But let us ignore that issue for now.

Of course, if we want to show that counterfeiting is hard given resources (1) and (2) *combined*, then as a first step, we should understand how to rule out counterfeiting given (1) or (2) separately. Thus, suppose C had a copy of $|\psi\rangle$, but no oracle access to V. In that case, the impossibility of preparing $|\psi\rangle|\psi\rangle$ would essentially amount to the No-Cloning Theorem. Conversely, suppose C had oracle access to V, but no copy of $|\psi\rangle$. Then *given unlimited time*, C could prepare as many copies of $|\psi\rangle$ as it wanted, by using Grover's algorithm to search for a quantum state that V accepts. The problem is "merely" that, if $|\psi\rangle$ has n qubits, then Grover's algorithm will in general require $\Theta\left(2^{n/2}\right)$ iterations, and the BBBV hybrid argument [14] shows that Grover's algorithm is optimal.

What we need, then, is a theorem showing that any counterfeiter needs exponentially many queries to V to prepare $|\psi\rangle|\psi\rangle$, *even if* the counterfeiter has a legitimate copy of $|\psi\rangle$ to start with. Such a theorem would be a *common generalization* of the No-Cloning Theorem and the BBBV hybrid argument, containing both of those earlier limitations as special cases. Aaronson [3] called the desired generalization the *Complexity-Theoretic No-Cloning Theorem*, and sketched a proof of it using Ambainis's adversary method. Based on that result, Aaronson also argued that there exists

a *quantum oracle* (i.e., a black-box unitary transformation V) relative to which secure public-key quantum money is possible. However, the details were never published.

In this section, we prove a result—Theorem 20—that is much more general than the Complexity-Theoretic No-Cloning Theorem claimed by Aaronson [3]. Then, in Section 5, we apply Theorem 20 to prove the security of public-key quantum money relative to a *classical* oracle. This solves the main open problem of Aaronson [3]. For completeness, in Appendix 11, we also apply Theorem 20 to prove the "original" Complexity-Theoretic No-Cloning Theorem [3], which involves Haar-random n-qubit states $|\psi\rangle$, rather than superpositions $|A\rangle$ over subspaces $A \leq \mathbb{F}_2^n$.[16]

4.1 Idea of Method

So, what *is* the inner-product adversary method? In Ambainis's adversary method [9]—like in the BBBV hybrid argument [14] from which it evolved—the basic idea is to upper-bound how much "progress" a quantum algorithm Q can make at distinguishing pairs of oracles, as the result of a single query. Let $\left|\Psi_t^U\right\rangle$ be Q's state after t queries, assuming that the oracle is U. Then *normally*, before any queries have been made, we can assume that $\left|\Psi_0^U\right\rangle = \left|\Psi_0^V\right\rangle$ for all oracles U and V. By contrast, after the final query T, for all oracle pairs (U, V) that Q is trying to distinguish, we must have (say) $\left|\left\langle\Psi_T^U|\Psi_T^V\right\rangle\right| \leq 1/2$. Thus, *if* we can show that the inner product $\left|\left\langle\Psi_t^U|\Psi_t^V\right\rangle\right|$ can decrease by at most ε as the result of a single query, then it follows that Q must make $\Omega\left(1/\varepsilon\right)$ queries.

Alas, when we try to apply the above framework to quantum money, we run into serious difficulties. Most obviously, it is no longer true that $\left|\Psi_0^U\right\rangle = \left|\Psi_0^V\right\rangle$ for all oracles U, V. Indeed, before Q makes even a *single* query to its oracle V, it already has a great deal of information about V, in the form of a legitimate money state $|\psi\rangle$ that V accepts. The task is "merely" to prepare a second copy of a state that Q already has! Worse yet, once we fix two oracles U and V, we find that Q generally *can* exploit the "head start" provided by its initial state to decrease the inner product $\left|\left\langle\Psi_t^U|\Psi_t^V\right\rangle\right|$ by a constant amount, by making just a single query to U or V respectively.

Our solution is as follows. We first carefully choose a distribution $\mathcal{D}$ over oracle pairs (U, V). We then analyze how much the *expected* inner product

$$\mathop{\mathrm{E}}_{(U,V)\sim\mathcal{D}}\left[\left|\left\langle\Psi_t^U|\Psi_t^V\right\rangle\right|\right]$$

can decrease as the result of a single query to U or V. We will find that, even if Q can substantially increase the angle between $\left|\Psi_t^U\right\rangle$ and $\left|\Psi_t^V\right\rangle$ for certain *specific* (U, V) pairs by making a single query, it cannot do so for *most* pairs. To illustrate, let $|\psi\rangle$ and $|\varphi\rangle$ be two possible quantum money states, which were chosen uniformly at random subject to the constraint that (say) $\langle\psi|\varphi\rangle = 1/2$. Then if a counterfeiting algorithm succeeds perfectly, it must map $|\psi\rangle$ to $|\psi\rangle^{\otimes 2}$, and $|\varphi\rangle$ to $|\varphi\rangle^{\otimes 2}$. Since

$$\langle\psi|^{\otimes 2}|\varphi\rangle^{\otimes 2} = \left(\langle\psi|\varphi\rangle\right)^2 = \frac{1}{4},$$

this means that the counterfeiter must *decrease the corresponding inner product* by at least 1/4. However, while the counterfeiter can indeed "push apart" *some* $|\psi\rangle, |\varphi\rangle$ pairs easily, we will show

[16] For whatever it is worth, we get a lower bound of $\Omega\left(2^{n/2}\right)$ on the number of queries needed to copy a Haar-random state, which is quadratically better than the $\Omega\left(2^{n/4}\right)$ that we get for subspace states.

that the *average* inner product can decrease by at most $1/\exp(n)$ as the result of a single query. From this it will follow that the counterfeiter needs to make $\exp(n)$ queries.

Let us mention that today, there are several "sophisticated" versions of the quantum adversary method, which *can* yield lower bounds for quantum state generation tasks not unlike the ones we consider. See especially Ambainis et al.'s "symmetry-assisted" adversary method [10], and Lee et al.'s adversary method for state conversion [31]. However, a drawback of these methods is that they are extremely hard to apply to concrete problems: doing so typically requires eigenvalue bounds, and often the use of representation theory. For this reason, even if one of the "sophisticated" adversary methods (or a variant thereof) could be applied to the quantum money problem, our approach might still be preferable.

4.2 The Method

We now introduce the inner-product adversary method. Let $\mathcal{O}$ be a set of quantum oracles acting on n qubits each. For each $U \in \mathcal{O}$, assume there exists a subspace $S_U \leq \mathbb{C}^{2^n}$ such that

(i) $U|\psi\rangle = -|\psi\rangle$ for all $|\psi\rangle \in S_U$, and

(ii) $U|\eta\rangle = |\eta\rangle$ for all $|\eta\rangle \in S_U^{\perp}$.

Let $R \subset \mathcal{O} \times \mathcal{O}$ be a symmetric binary relation on $\mathcal{O}$, with the properties that

(i) $(U,U) \notin R$ for all $U \in \mathcal{O}$, and

(ii) for every $U \in \mathcal{O}$ there exists a $V \in \mathcal{O}$ such that $(U,V) \in R$.

Suppose that for all $U \in \mathcal{O}$ and all $|\eta\rangle \in S_U^{\perp}$, we have

$$\mathop{\mathrm{E}}_{V\,:\,(U,V)\in R}\left[F\left(|\eta\rangle, S_V\right)^2\right] \leq \varepsilon,$$

where

$$F\left(|\eta\rangle, S_V\right) = \max_{|\psi\rangle \in S_V} |\langle\eta|\psi\rangle|$$

is the fidelity between $|\eta\rangle$ and S_V. Let Q be a quantum oracle algorithm, and let Q^U denote Q run with the oracle $U \in \mathcal{O}$. Suppose Q^U begins in the state $\left|\Psi_0^U\right\rangle$ (possibly already dependent on U). Let $\left|\Psi_t^U\right\rangle$ denote the state of Q^U immediately after the tth query. Also, define a *progress measure* p_t by

$$p_t := \mathop{\mathrm{E}}_{U,V\,:\,(U,V)\in R}\left[\left|\left\langle\Psi_t^U|\Psi_t^V\right\rangle\right|\right].$$

The following lemma bounds how much p_t can decrease as the result of a single query.

Lemma 19 (Bound on Progress Rate)**.**

$$p_t \geq p_{t-1} - 4\sqrt{\varepsilon}.$$

Proof. Let $\left|\Phi_t^U\right\rangle$ denote the state of Q^U immediately *before* the tth query. Then for all t, it is clear that $\left\langle\Phi_t^U|\Phi_t^V\right\rangle = \left\langle\Psi_{t-1}^U|\Psi_{t-1}^V\right\rangle$: in other words, the unitary transformations that Q performs in between query steps have no effect on the inner products. So to prove the lemma, it suffices to show the following inequality:

$$\mathop{\mathrm{E}}_{U,V\,:\,(U,V)\in R}\left[\left|\left\langle\Phi_t^U|\Phi_t^V\right\rangle\right|\right] - \mathop{\mathrm{E}}_{U,V\,:\,(U,V)\in R}\left[\left|\left\langle\Psi_t^U|\Psi_t^V\right\rangle\right|\right] \leq 4\sqrt{\varepsilon}. \tag{*}$$

Let $\{|i\rangle\}_{i\in[B]}$ be an arbitrary orthonormal basis for Q's workspace register. Then we can write

$$\begin{aligned}\left|\Phi_t^U\right\rangle &= \sum_{i\in[B]} \alpha_{t,i}^U \left|i\right\rangle \left|\Phi_{t,i}^U\right\rangle \\ &= \sum_{i\in[B]} \left|i\right\rangle \left(\beta_{t,i}^U \left|\eta_{t,i}^U\right\rangle + \gamma_{t,i}^U \left|\psi_{t,i}^U\right\rangle\right),\end{aligned}$$

where $\left|\eta_{t,i}^U\right\rangle \in S_U^\perp$ and $\left|\psi_{t,i}^U\right\rangle \in S_U$. (By normalization, $\left|\beta_{t,i}^U\right|^2 + \left|\gamma_{t,i}^U\right|^2 = \left|\alpha_{t,i}^U\right|^2$.) A query transforms the above state to

$$\left|\Psi_t^U\right\rangle = \sum_{i\in[B]} \left|i\right\rangle \left(\beta_{t,i}^U \left|\eta_{t,i}^U\right\rangle - \gamma_{t,i}^U \left|\psi_{t,i}^U\right\rangle\right).$$

So for all $U, V \in \mathcal{O}$,

$$\begin{aligned}\left\langle\Phi_t^U|\Phi_t^V\right\rangle - \left\langle\Psi_t^U|\Psi_t^V\right\rangle =& \sum_{i\in[B]} \left(\overline{\beta}_{t,i}^U \left\langle\eta_{t,i}^U\right| + \overline{\gamma}_{t,i}^U \left\langle\psi_{t,i}^U\right|\right)\left(\beta_{t,i}^V \left|\eta_{t,i}^V\right\rangle + \gamma_{t,i}^V \left|\psi_{t,i}^V\right\rangle\right) \\ &- \sum_{i\in[B]} \left(\overline{\beta}_{t,i}^U \left\langle\eta_{t,i}^U\right| - \overline{\gamma}_{t,i}^U \left\langle\psi_{t,i}^U\right|\right)\left(\beta_{t,i}^V \left|\eta_{t,i}^V\right\rangle - \gamma_{t,i}^V \left|\psi_{t,i}^V\right\rangle\right) \\ =& 2\sum_{i\in[B]} \left(\overline{\beta}_{t,i}^U \gamma_{t,i}^V \left\langle\eta_{t,i}^U|\psi_{t,i}^V\right\rangle + \overline{\gamma}_{t,i}^U \beta_{t,i}^V \left\langle\psi_{t,i}^U|\eta_{t,i}^V\right\rangle\right).\end{aligned}$$

By Cauchy-Schwarz, the above implies that

$$\left|\left\langle\Phi_t^U|\Phi_t^V\right\rangle\right| - \left|\left\langle\Psi_t^U|\Psi_t^V\right\rangle\right| \leq 2\max_{i\in[B]}\left|\left\langle\eta_{t,i}^U|\psi_{t,i}^V\right\rangle\right| + 2\max_{i\in[B]}\left|\left\langle\psi_{t,i}^U|\eta_{t,i}^V\right\rangle\right|.$$

Now fix $U \in \mathcal{O}$ and $i \in [B]$. Then again applying Cauchy-Schwarz,

$$\begin{aligned}\mathop{\mathrm{E}}_{V\,:\,(U,V)\in R}\left[\left|\left\langle\eta_{t,i}^U|\psi_{t,i}^V\right\rangle\right|\right] &\leq \sqrt{\mathop{\mathrm{E}}_{V\,:\,(U,V)\in R}\left[\left|\left\langle\eta_{t,i}^U|\psi_{t,i}^V\right\rangle\right|^2\right]} \\ &\leq \sqrt{\mathop{\mathrm{E}}_{V\,:\,(U,V)\in R}\left[\max_{|\psi\rangle\in S_V}\left|\left\langle\eta_{t,i}^U|\psi\right\rangle\right|^2\right]} \\ &\leq \sqrt{\varepsilon}.\end{aligned}$$

Hence

$$\mathop{\mathrm{E}}_{U,V\,:\,(U,V)\in R}\left[\left|\left\langle\eta_{t,i}^U|\psi_{t,i}^V\right\rangle\right|\right] \leq \sqrt{\varepsilon}$$

as well, and likewise

$$\mathop{\mathrm{E}}_{U,V\,:\,(U,V)\in R}\left[\left|\left\langle\psi_{t,i}^U|\eta_{t,i}^V\right\rangle\right|\right]\leq\sqrt{\varepsilon}$$

by symmetry. Putting everything together,

$$\begin{aligned}p_{t-1}-p_t &= \mathop{\mathrm{E}}_{U,V\,:\,(U,V)\in R}\left[\left|\left\langle\Phi_t^U|\Phi_t^V\right\rangle\right|-\left|\left\langle\Psi_t^U|\Psi_t^V\right\rangle\right|\right]\\ &\leq 2\mathop{\mathrm{E}}_{U,V\,:\,(U,V)\in R}\left[\max_{i\in[B]}\left|\left\langle\eta_{t,i}^U|\psi_{t,i}^V\right\rangle\right|\right]+2\mathop{\mathrm{E}}_{U,V\,:\,(U,V)\in R}\left[\max_{i\in[B]}\left|\left\langle\psi_{t,i}^U|\eta_{t,i}^V\right\rangle\right|\right]\\ &\leq 4\sqrt{\varepsilon}.\end{aligned}$$

This proves inequality (*) and hence the lemma. ■

From Lemma 19 we immediately deduce the following.

Theorem 20 (Inner-Product Adversary Method)**.** *Suppose that initially* $\left|\left\langle\Psi_0^U|\Psi_0^V\right\rangle\right|\geq c$ *for all* $(U,V)\in R$*, whereas by the end we need* $\left|\left\langle\Psi_T^U|\Psi_T^V\right\rangle\right|\leq d$ *for all* $(U,V)\in R$*. Then* Q *must make* $T=\Omega\left(\frac{c-d}{\sqrt{\varepsilon}}\right)$ *oracle queries.*

5 Quantum Money with a Classical Oracle

In this section, we construct a mini-scheme, called the *Hidden Subspace Mini-Scheme*, that requires only a classical oracle. We then use the inner-product adversary method from Section 4 to show that our mini-scheme is secure—indeed, that any counterfeiter must make $\Omega\left(2^{n/4}\right)$ oracle queries to copy a banknote. By the results of Sections 3.3 and 2.1, our mini-scheme will lead automatically to a full-blown public-key quantum money scheme, which requires only a classical oracle and is unconditionally secure.

Later, in Section 6, we will propose a way to "instantiate" the oracles in our mini-scheme, using multivariate polynomials. For now, though, we focus on the black-box setting.

5.1 The Hidden Subspace Mini-Scheme

We identify n-bit strings $x\in\{0,1\}^n$ with elements of the vector space $\mathbb{F}_2^n$ in the standard way. Then in our mini-scheme, each n-qubit money state will have the form

$$|A\rangle:=\frac{1}{\sqrt{|A|}}\sum_{x\in A}|x\rangle\,,$$

where A is some randomly-chosen subspace of $\mathbb{F}_2^n$ (i.e., a set of codewords of a linear code), with $\dim A=n/2$. Let $A^\perp$ be the orthogonal complement of A, so that $\dim A^\perp=n/2$ as well. Notice that we can transform $|A\rangle$ to $\left|A^\perp\right\rangle$ and vice versa by simply applying $H_2^{\otimes n}$: a Hadamard gate on each of the n qubits, or equivalently a quantum Fourier transform over $\mathbb{F}_2^n$.

The basic idea of the mini-scheme is as follows: the bank can easily prepare the quantum money state $|A\rangle$, starting from a classical description $\langle A\rangle$ of A (e.g., a list of $n/2$ generators). The bank distributes the state $|A\rangle$, but keeps the classical description $\langle A\rangle$ secret. Along with $|A\rangle$ itself, the bank also publishes details of how to *verify* $|A\rangle$ by querying two classical oracles, U_A and $U_{A^\perp}$. The first oracle, U_A, decides membership in A: for all n-qubit basis states $|x\rangle$,

$$U_A\,|x\rangle=\begin{cases}-\,|x\rangle & \text{if } x\in A\\ |x\rangle & \text{otherwise}\end{cases}$$

The second oracle, $U_{A^\perp}$, decides membership in $A^\perp$ in the same way.

Using U_A, it is easy to implement a projector $\mathbb{P}_A$ onto the set of basis states in A. To do so, simply initialize a control qubit to $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$, then apply U_A conditioned on the control qubit being in state $|1\rangle$, then measure the control qubit in the $\{|+\rangle, |-\rangle\}$ basis, and postselect on getting the outcome $|-\rangle$. Likewise, using $U_{A^\perp}$, it is easy to implement a projector $\mathbb{P}_{A^\perp}$ onto the set of basis states in $A^\perp$. Then V_A, the public verification algorithm for the money state $|A\rangle$, will simply consist of a projection onto A, followed by a Fourier transform, a projection onto $A^\perp$, and finally a second Fourier transform to return the legitimate money state back to $|A\rangle$:

$$V_A := H_2^{\otimes n}\mathbb{P}_{A^\perp}H_2^{\otimes n}\mathbb{P}_A.$$

We will show in Lemma 21 that $V_A = |A\rangle\langle A|$ is just a projector onto $|A\rangle$ itself. This means, in particular, that $V_A|A\rangle = |A\rangle$, and that V_A accepts an arbitrary state $|\psi\rangle$ with probability equal to $|\langle\psi|A\rangle|^2$. Thus, our mini-scheme is *projective* and has *perfect completeness*.

But what about security? Intuitively, a counterfeiter could query U_A or $U_{A^\perp}$ to find an explicit generating set for A or $A^\perp$—but that would require an exponentially-long Grover search, since $|A| = \left|A^\perp\right| = 2^{n/2} \ll 2^n$. Alternatively, the counterfeiter could measure $|A\rangle$ in the standard or Hadamard bases—but that would reveal just *one* random element of A or $A^\perp$. Neither ability seems useful for *copying* $|A\rangle$, let alone recovering a full classical description of A, in polynomial time.[17]

And indeed, using the inner-product adversary method plus some other tools, we will prove the following tight lower bound (Theorem 25):

> *Even if given a single copy of $|A\rangle$, as well as oracle access to U_A and $U_{A^\perp}$, a counterfeiter still needs $\Omega\left(\epsilon 2^{n/4}\right)$ queries to prepare a state that has fidelity ϵ with $|A\rangle^{\otimes 2}$.*

This will imply that our mini-scheme has *exponentially-small soundness error*.

5.2 Formal Specification

We are not quite done, since we never explained how the bank provides access to U_A and $U_{A^\perp}$. Thus, in our "final" mini-scheme $\mathcal{M} = (\mathsf{Bank}_{\mathcal{M}}, \mathsf{Ver}_{\mathcal{M}})$, the bank, the verifier, and the counterfeiter will all have access to a *single* classical oracle U, which consists of four components:

- A *banknote generator* $\mathcal{G}(r)$, which takes as input a random string $r \in \{0,1\}^n$, and outputs a set of linearly independent generators $\langle A_r\rangle = \left\{x_1, \ldots, x_{n/2}\right\}$ for a subspace $A_r \leq \mathbb{F}_2^n$, as well as a unique $3n$-bit *serial number* $s_r \in \{0,1\}^{3n}$. The function $\mathcal{G}$ is chosen uniformly at random, subject to the constraint that the serial numbers are all distinct.[18]

- A *serial number checker* $\mathcal{H}(s)$, which outputs 1 if $s = s_r$ is a valid serial number for some $\langle A_r\rangle$, and 0 otherwise.[19]

[17] Obviously, if the counterfeiter had $\Omega(n)$ copies of $|A\rangle$, then it *could* recover a generating set for A, by simply measuring each copy independently in the standard basis. That is why, in our full quantum money scheme, the counterfeiter will *not* have $\Omega(n)$ copies of $|A\rangle$. Instead, each banknote will involve a completely different subspace $A_s \leq \mathbb{F}_2^n$ (parameterized by its unique serial number s), so that measuring one banknote reveals nothing about the others.

[18] Note that one can implement $\mathcal{G}$ using an ordinary random oracle. In that case, the requirement that the serial numbers are distinct will be satisfied with probability $1 - O\left(2^{-n}\right)$.

[19] Strictly speaking, this component can be omitted, since one can simulate $\mathcal{H}(s)$ using either $\mathcal{T}_{\text{primal}}(s,0)$ or $\mathcal{T}_{\text{dual}}(s,0)$. We include it for ease of presentation.

- A *primal subspace tester* $\mathcal{T}_{\text{primal}}$, which takes an input of the form $|s\rangle|x\rangle$, applies U_{A_r} to $|x\rangle$ if $s = s_r$ is a valid serial number for some $\langle A_r\rangle$, and does nothing otherwise.
- A *dual subspace tester* $\mathcal{T}_{\text{dual}}$, which takes an input of the form $|s\rangle|x\rangle$, applies $U_{A_r^\perp}$ to $|x\rangle$ if $s = s_r$ is a valid serial number for some $\langle A_r\rangle$, and does nothing otherwise.

Then the mini-scheme $\mathcal{M} = (\mathsf{Bank}_{\mathcal{M}}, \mathsf{Ver}_{\mathcal{M}})$ is defined as follows:

- $\mathsf{Bank}_{\mathcal{M}}(0^n)$ chooses $r \in \{0,1\}^n$ uniformly at random. It then looks up $\mathcal{G}(r) = (s_r, \langle A_r\rangle)$, and outputs the banknote $|\$_r\rangle = |s_r\rangle|A_r\rangle$.
- $\mathsf{Ver}_{\mathcal{M}}(\not{c})$ first uses $\mathcal{H}$ to check that $\not{c}$ has the form (s,ρ), where $s = s_r$ is a valid serial number. If so, then it uses $\mathcal{T}_{\text{primal}}$ and $\mathcal{T}_{\text{dual}}$ to apply $V_{A_r} = H_2^{\otimes n}\mathbb{P}_{A_r^\perp}H_2^{\otimes n}\mathbb{P}_{A_r}$, and accepts if and only if $V_{A_r}(\rho)$ accepts.

5.3 Analysis

We now analyze the mini-scheme defined in Sections 5.1 and 5.2. For convenience, we assume for most of the proof that the subspace $A \leq \mathbb{F}_2^n$ is *fixed*, and that the counterfeiter (who does not know A) only has access to the oracles U_A and $U_{A^\perp}$. Then, at the end, we will explain how to generalize the conclusions to the "final" mini-scheme $\mathcal{M}$.

It will be convenient to consider the subset $A^* \subset \{0,1\}^{n+1}$, defined by

$$A^* := (0, A) \cup (1, A^\perp).$$

Let S_{A^*} be the subspace of $\mathbb{C}^{2^{n+1}}$ that is spanned by basis states $|x\rangle$ such that $x \in A^*$. Then we can think of the pair of oracles $(U_A, U_{A^\perp})$ as being a *single* oracle U_{A^*}, which satisfies $U_{A^*}|\psi\rangle = -|\psi\rangle$ for all $|\psi\rangle \in S_{A^*}$, and $U_{A^*}|\eta\rangle = |\eta\rangle$ for all $|\eta\rangle \in S_{A^*}^\perp$.

Recall the definition of the verifier V_A:

$$V_A := H_2^{\otimes n}\mathbb{P}_{A^\perp}H_2^{\otimes n}\mathbb{P}_A,$$

where $\mathbb{P}_A$ and $\mathbb{P}_{A^\perp}$ denote projective measurements that accept a basis state $|x\rangle$ if and only if x belongs to A or $A^\perp$ respectively. The following lemma shows that V_A "works," in the sense that it accepts a state $|\psi\rangle$ with certainty if $|\psi\rangle = |A\rangle$, and rejects with high probability if $|\psi\rangle$ is far from $|A\rangle$.

Lemma 21. *$V_A = |A\rangle\langle A|$ is simply a projector onto $|A\rangle$. So in particular,*

$$\Pr[V_A(|\psi\rangle) \text{ accepts}] = |\langle\psi|A\rangle|^2.$$

Proof. It suffices to show that $V_A|A\rangle = |A\rangle$ and that $V_A|\psi\rangle = 0$ for all $|\psi\rangle$ orthogonal to $|A\rangle$. First,

$$\begin{aligned}
V_A|A\rangle &= H_2^{\otimes n}\mathbb{P}_{A^\perp}H_2^{\otimes n}\mathbb{P}_A|A\rangle \\
&= H_2^{\otimes n}\mathbb{P}_{A^\perp}H_2^{\otimes n}|A\rangle \\
&= H_2^{\otimes n}\mathbb{P}_{A^\perp}\left|A^\perp\right\rangle \\
&= H_2^{\otimes n}\left|A^\perp\right\rangle \\
&= |A\rangle.
\end{aligned}$$

Second, if $\langle\psi|A\rangle = 0$ then we can write

$$|\psi\rangle = \sum_{x\in 2^n} c_x |x\rangle$$

where $\sum_{x\in A} c_x = 0$. Then

$$\begin{aligned}
V_A |\psi\rangle &= H_2^{\otimes n}\mathbb{P}_{A^\perp} H_2^{\otimes n}\mathbb{P}_A \sum_{x\in 2^n} c_x |x\rangle \\
&= H_2^{\otimes n}\mathbb{P}_{A^\perp} H_2^{\otimes n} \sum_{x\in A} c_x |x\rangle \\
&= \frac{1}{\sqrt{2^n}} H_2^{\otimes n}\mathbb{P}_{A^\perp} \sum_{x\in A} c_x \sum_{y\perp x} |y\rangle \\
&= \frac{1}{\sqrt{2^n}} H_2^{\otimes n} \sum_{y\in A^\perp} |y\rangle \sum_{x\in A} c_x \\
&= 0.
\end{aligned}$$

■

An immediate consequence of Lemma 21, which will be useful later, is that our mini-scheme is *projective*.

We now show that perfect counterfeiting requires exponentially many queries to U_{A^*}.

Theorem 22 (Lower Bound for Perfect Counterfeiting)**.** *Given one copy of $|A\rangle$, as well as oracle access to U_{A^*}, a counterfeiter needs $\Omega\left(2^{n/4}\right)$ queries to prepare $|A\rangle^{\otimes 2}$ with certainty (for a worst-case $|A\rangle$).*

Proof. We will apply Theorem 20. Let the set $\mathcal{O}$ contain U_{A^*} for every possible subspace $A \leq \mathbb{F}_2^n$ with $\dim A = n/2$. Also, put $(U_{A^*}, U_{B^*}) \in R$ if and only if $\dim(A\cap B) = n/2 - 1$. Then given $U_{A^*} \in \mathcal{O}$ and $|\eta\rangle \in S_{A^*}^\perp$, let

$$|\eta\rangle = \sum_{x\in\{0,1\}^{n+1}\setminus A^*} \alpha_x |x\rangle .$$

We have

$$\begin{aligned}
\mathop{\mathrm{E}}_{U_{B^*}\,:\,(U_{A^*},U_{B^*})\in R}\left[F\left(|\eta\rangle, S_{B^*}\right)^2\right] &= \mathop{\mathrm{E}}_{B\,:\,\dim(B)=n/2,\dim(A\cap B)=n/2-1}\left[\sum_{x\in B^*\setminus A^*} |\alpha_x|^2\right] \\
&\leq \max_{x\in\{0,1\}^{n+1}\setminus A^*}\left(\Pr_{B\,:\,\dim(B)=n/2,\dim(A\cap B)=n/2-1}\left[x\in B^*\right]\right) \\
&= \max_{x\in\{0,1\}^{n}\setminus A}\left(\Pr_{B\,:\,\dim(B)=n/2,\dim(A\cap B)=n/2-1}\left[x\in B\right]\right) \\
&= \frac{|B\setminus A|}{|\{0,1\}^n\setminus A|} \quad (\text{for } \dim(B) = n/2,\ \dim(A\cap B) = n/2-1) \\
&= \frac{2^{n/2-1}}{2^n - 2^{n/2}} \\
&\leq \frac{1}{2^{n/2}}.
\end{aligned}$$

Here the first line uses the definition of fidelity, the second line uses the easy direction of the minimax theorem, the third line uses the symmetry between A and $A^{\perp}$, and the fourth line uses the symmetry among all $2^n - 2^{n/2}$ strings $x \in \{0,1\}^n \setminus A$. The conclusion is that we can set $\varepsilon := 2^{-n/2}$.

Fix $(U_{A^*}, U_{B^*}) \in R$. Then $|\langle A|B\rangle| = 1/2$. On the other hand, if the counterfeiter succeeds, it must map $|A\rangle$ to some state $|f_A\rangle := |A\rangle\,|A\rangle\,|\text{garbage}_A\rangle$, and $|B\rangle$ to some state $|f_B\rangle := |B\rangle\,|B\rangle\,|\text{garbage}_B\rangle$. Therefore $|\langle f_A|f_B\rangle| \leq 1/4$. So setting $c = 1/2$ and $d = 1/4$, Theorem 20 tells us that the counterfeiter must make

$$\Omega\left(\frac{c-d}{\sqrt{\varepsilon}}\right) = \Omega\left(2^{n/4}\right)$$

queries to U_{A^*}. ■

A simple modification to the proof of Theorem 22 shows that even to counterfeit money *almost* perfectly, one still needs exponentially many queries to U_{A^*}.

Corollary 23 (Lower Bound for Small-Error Counterfeiting)**.** *Given one copy of $|A\rangle$, as well as oracle access to U_{A^*}, a counterfeiter needs $\Omega\left(2^{n/4}\right)$ queries to prepare a state ρ such that $\langle A|^{\otimes 2}\,\rho\,|A\rangle^{\otimes 2} \geq 0.9999$ (for a worst-case $|A\rangle$).*

Proof. Let $|\langle A|B\rangle| = c$, and let $\epsilon = 0.00001$. If the counterfeiter succeeds, it must map $|A\rangle$ to some state ρ_A, and $|B\rangle$ to some state ρ_B, such that $\langle A|^{\otimes 2}\,\rho_A\,|A\rangle^{\otimes 2}$ and $\langle B|^{\otimes 2}\,\rho_B\,|B\rangle^{\otimes 2}$ are both at least $1-\epsilon$. So letting $|f_A\rangle$ and $|f_B\rangle$ be purifications of ρ_A and ρ_B respectively, we have

$$\begin{aligned}
|\langle f_A|f_B\rangle| &\leq F\left(\rho_A, \rho_B\right) \\
&\leq \left|\langle A|^{\otimes 2}\,|B\rangle^{\otimes 2}\right| + 2\epsilon^{1/4} \\
&= c^2 + 2\epsilon^{1/4}
\end{aligned}$$

where the second line follows from Lemma 5. So setting $d := c^2 + 2\epsilon^{1/4}$, Theorem 20 tells us that the counterfeiter must make

$$\Omega\left(\frac{c - c^2 - 2\epsilon^{1/4}}{\sqrt{2^{-n/2}}}\right)$$

queries to U_{A^*}. Fixing $c := 1/2$, the above is $\Omega\left(2^{n/4}\right)$. ■

Since the verifier V_A is projective, we can now combine Corollary 23 with Theorem 15 to obtain the following "amplified" lower bound.

Corollary 24 (Lower Bound for High-Error Counterfeiting)**.** *Let $1/\varepsilon = o\left(2^{n/2}\right)$. Given one copy of $|A\rangle$, as well as oracle access to U_{A^*}, a counterfeiter needs $\Omega\left(\sqrt{\varepsilon}2^{n/4}\right)$ queries to prepare a state ρ such that $\langle A|^{\otimes 2}\,\rho\,|A\rangle^{\otimes 2} \geq \varepsilon$ (for a worst-case $|A\rangle$).*

Proof. Suppose we have a counterfeiter C that makes $o\left(\sqrt{\varepsilon}2^{n/4}\right)$ queries to U_{A^*}, and prepares a state σ such that $\langle A|^{\otimes 2}\,\sigma\,|A\rangle^{\otimes 2} \geq \varepsilon$. Let $\delta := 0.00001$. Then by Theorem 15, there exists an amplified counterfeiter C' that makes

$$O\left(\frac{\log 1/\delta}{\sqrt{\varepsilon}\left(\sqrt{\varepsilon} + \delta^2\right)}\right) = O\left(\frac{1}{\sqrt{\varepsilon}}\right)$$

calls to C and V_A, and that prepares a state ρ such that $\langle A|^{\otimes 2} \rho |A\rangle^{\otimes 2} \geq 1-\delta$. Now, counting the $o\left(\sqrt{\varepsilon}2^{n/4}\right)$ queries from each C invocation and $O(1)$ queries from each V_A invocation, the total number of queries that C' makes to U_{A^*} is

$$\left[o\left(\sqrt{\varepsilon}2^{n/4}\right)+O(1)\right]\cdot O\left(\frac{1}{\sqrt{\varepsilon}}\right)=o\left(2^{n/4}\right).$$

But this contradicts Corollary 23. ■

So far, we have only made statements about the *worst* case for a would-be counterfeiter. But such guarantees are clearly not enough: it could be that *most* money states $|A\rangle$ are easy to duplicate, without contradicting any of the results we have seen so far.

We will show that the problem faced by a counterfeiter is *random self-reducible*: if a counterfeiter could duplicate a uniformly-random money state $|A\rangle$, then it could duplicate *any* $|A\rangle$. Thus the bank can ensure security by creating uniformly-random money states.

In what follows, let $\mathcal{S}$ be the set of all subspaces $A \leq \mathbb{F}_2^n$ such that $\dim A = n/2$. Also, let $V_A^{\otimes 2} = (|A\rangle\langle A|)^{\otimes 2}$ be the projector onto $|A\rangle^{\otimes 2}$: in other words, the result of applying the verification procedure V_A to each of two n-qubit registers, and accepting if and only if *both* are found to be in the state $|A\rangle$.

Theorem 25 (Lower Bound for Average-Case Counterfeiting)**.** *Let $A \leq \mathbb{F}_2^n$ be a uniformly-random element of $\mathcal{S}$. Then given one copy of $|A\rangle$, as well as oracle access to U_{A^*}, a counterfeiter C needs $\Omega\left(\sqrt{\varepsilon}2^{n/4}\right)$ queries to prepare a $2n$-qubit state ρ that $V_A^{\otimes 2}$ accepts with probability at least ε, for all $1/\varepsilon = o\left(2^{n/2}\right)$. Here the probability is taken over the choice of $A \in \mathcal{S}$, as well as the behavior of C and $V_A^{\otimes 2}$.*

Proof. Suppose we had a counterfeiter C that violated the above. Using C as a black box, we will show how to construct a new counterfeiter C' that violates Corollary 24.

Given a (deterministically-chosen) money state $|A\rangle$ and oracle access to U_{A^*}, first choose an invertible linear map $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$ uniformly at random. Then $f(A)$, the image of A under f, is a uniformly-random element of $\mathcal{S}$. Furthermore, the state $|A\rangle$ can be transformed into $|f(A)\rangle$ straightforwardly, and the oracle $U_{f(A)^*}$ can be simulated by composing f with U_{A^*}. So by using the counterfeiter C for uniformly-random states, we can produce a state ρ_f that $V_{f(A)}^{\otimes 2}$ accepts with probability at least ε. By applying f^{-1} to both registers of ρ_f, we can then obtain a state ρ that $V_A^{\otimes 2}$ accepts with probability at least ε, thereby contradicting Corollary 24. ■

We are now ready to prove security for the "final" mini-scheme $\mathcal{M}$ defined in Section 5.2.

Theorem 26 (Security of Mini-Scheme)**.** *The mini-scheme $\mathcal{M} = (\mathsf{Bank}_{\mathcal{M}}, \mathsf{Ver}_{\mathcal{M}})$, which is defined relative to the classical oracle U, has perfect completeness and exponentially-small soundness error.*

Proof. That $\mathcal{M}$ has perfect completeness follows from its definition and from Lemma 21. That $\mathcal{M}$ has exponentially-small soundness error *essentially* follows from Theorem 25. We only need to show that, given a banknote of the form $|\$_r\rangle = |s_r\rangle|A_r\rangle$, a polynomial-time counterfeiter C can gain no additional advantage by querying the "full" oracles $\mathcal{G}, \mathcal{H}, \mathcal{T}_{\text{primal}}, \mathcal{T}_{\text{dual}}$, beyond what it gains from querying $U_{A_r^*} = \left(U_{A_r}, U_{A_r^\perp}\right)$.

Let $r \in \{0,1\}^n$ be the random string chosen by the bank, so that $\mathcal{G}(r) = (s_r, \langle A_r\rangle)$. Then observe that, even conditioned on s_r and A_r, as well as complete descriptions of $\mathcal{T}_{\text{primal}}, \mathcal{T}_{\text{dual}}$, and $\mathcal{H}$, the string r remains uniformly random. Nor can querying $\mathcal{G}(r')$ for $r' \neq r$ reveal any

information about r, since the values of $\mathcal{G}$ are generated independently. So suppose we *modify* $\mathcal{G}$ by setting $\mathcal{G}(r) := (s', \langle A' \rangle)$, for some new $3n$-bit serial number s' and list of generators $\langle A' \rangle$ chosen uniformly at random. Then the BBBV hybrid argument [14] tells us that, in expectation over r, this can alter the final state output by the counterfeiter $C(|\$_r\rangle)$ by at most $\text{poly}(n)/2^{n/2}$ in trace distance. So in particular, if C succeeded with non-negligible probability before, then C must *still* succeed with non-negligible probability after we set $\mathcal{G}(r) := (s', \langle A' \rangle)$.

However, once we make this modification, an adversary trying to counterfeit $|A\rangle$ given U_A and $U_{A^\perp}$ can easily "mock up" a serial number s, as well as the oracles $\mathcal{G}, \mathcal{H}, \mathcal{T}_{\text{primal}}$ and $\mathcal{T}_{\text{dual}}$, for itself. For s, $\mathcal{G}$, and $\mathcal{H}$ are now drawn from a distribution completely independent of A. The oracles $\mathcal{T}_{\text{primal}}$ and $\mathcal{T}_{\text{dual}}$ are likewise independent of A, *except* that $\mathcal{T}_{\text{primal}}|s\rangle|v\rangle = |s\rangle U_A|v\rangle$ and $\mathcal{T}_{\text{dual}}|s\rangle|v\rangle = |s\rangle U_{A^\perp}|v\rangle$—behaviors that an adversary can easily simulate using U_A and $U_{A^\perp}$, together with its knowledge of s. Just like in Corollary 18, since our security guarantees are query complexity bounds, we do not care about the *computational* complexity of creating the mock-ups.

By using the mock-ups, one can convert any successful attack on $\mathcal{M}$ into successful counterfeiting of $|A\rangle$, given oracle access to U_A and $U_{A^\perp}$ only. But the latter contradicts Theorem 25. ■

Finally, using Theorem 26 together with Corollary 18, we can obtain a secure public-key quantum money scheme, relative to a classical oracle.

Theorem 27 (Security of Hidden Subspace Money)**.** *By combining the mini-scheme $\mathcal{M}$ with a digital signature scheme, it is possible to construct a public-key quantum money scheme $\mathcal{S} = (\mathsf{KeyGen}_{\mathcal{S}}, \mathsf{Bank}_{\mathcal{S}}, \mathsf{Ver}_{\mathcal{S}})$, defined relative to some classical oracle U', which has perfect completeness and exponentially-small soundness error.*

6 Explicit Candidate Quantum Money Scheme

We have shown how to construct a provably-secure public-key quantum money scheme, when an appropriate classical oracle is available. In this section, we propose a way to obtain the same functionality without an oracle. The key challenge is this:

> *Given a subspace $A \leq \mathbb{F}_2^n$, how can a bank distribute an "obfuscated program" P_A, which legitimate buyers and sellers can use to decide membership in both A and $A^\perp$, but which does not reveal anything else about A that might facilitate counterfeiting?*

Note that, aside from the detail that we need security against quantum adversaries, the above challenge is purely "classical"; it and its variants seem interesting even apart from our quantum money application.

We will suggest a candidate protocol to achieve the challenge, based on *multivariate polynomial cryptography*. Given a collection $p_1, \ldots, p_m : \mathbb{F}_2^n \to \mathbb{F}_2$ of multivariate polynomials over $\mathbb{F}_2$, it is generally hard to find a point $v \in \mathbb{F}_2^n$ on which all of the p_i's vanish. On the other hand, it is easy to check whether a *particular* point v has that property. To "hide" a subspace A, we will provide uniformly-random low-degree polynomials $p_1, \ldots, p_m$ that vanish on each point of A. This information is sufficient to decide membership in A. On the other hand, there is no known efficient algorithm to *find* A given the polynomials, and current techniques seem unlikely to yield even a quantum algorithm.

We can also introduce a constant fraction of *noise* into our scheme without interfering with its completeness. In other words, if only $(1-\epsilon)m$ of the polynomials $p_1, \ldots, p_m$ are chosen to vanish

on A, and the remaining ϵm are random, then counting the number of p_i's that vanish at a point v still suffices to determine whether $v \in A$. Although we know of no attack even against our noise-free scheme, adding noise in this way might improve security.

Crucially, we will state a "classical" conjecture about the security of multivariate polynomial cryptography, and show that the conjecture *implies* the security of our explicit money scheme. For the benefit of cryptographers, let us now state an "abstract" version of our conjecture, which implies what we need, and which might hold even if our concrete conjecture about multivariate polynomials fails.

Conjecture 28 (Subspace-Hiding Conjecture, Sufficient for Quantum Money)**.** *There exists a polynomial-time algorithm that takes as input a description of a uniformly-random subspace $A \leq \mathbb{F}_2^n$ with $\dim(A) = n/2$, and that outputs circuits C_A and $C_{A^\perp}$, such that the following holds.*

(i) $C_A(v)$ decides whether $v \in A$, and $C_{A^\perp}(v)$ decides whether $v \in A^\perp$, for all $v \in \mathbb{F}_2^n$.

(ii) Given descriptions of C_A and $C_{A^\perp}$, no polynomial-time quantum algorithm can find a complete list of generators for A with success probability $\Omega\left(2^{-n/2}\right)$.

Later, Conjecture 34 will specialize Conjecture 28 to the setting of multivariate polynomials.

6.1 Useful Facts About Polynomials

By viewing elements of $\mathbb{F}_2^n$ as n-tuples $(x_1, \ldots, x_n)$, we can evaluate a polynomial $p(x_1, \ldots, x_n)$ on points of $\mathbb{F}_2^n$.

Given a subspace $A \leq \mathbb{F}_2^n$ and a positive integer d, let $\mathcal{I}_{d,A}$ be the set of degree-d polynomials (not necessarily homogeneous) that vanish on A. Since we are working over $\mathbb{F}_2$, note that $x_i^2 = x_i$, so it suffices to consider *multilinear* polynomials (in which no x_i is ever raised to a higher power than 1).

Before presenting our scheme, we need to establish some basic properties of polynomials over $\mathbb{F}_2^n$. First, we observe that the set of polynomials does not depend on the choice of basis.

Proposition 29. *Let L be any invertible linear transformation on $\mathbb{F}_2^n$. Then the map $p(v) \mapsto p(Lv)$ defines a permutation on the set of degree-d polynomials, which maps $\mathcal{I}_{d,A}$ to $\mathcal{I}_{d,L^{-1}A}$.*

Implementing our scheme will require sampling uniformly from $\mathcal{I}_{d,A}$, which the next lemma shows is possible.

Lemma 30. *It is possible to sample a uniformly-random element of $\mathcal{I}_{d,A}$ in time $O(n^d)$.*

Proof. By Proposition 29, we can instead sample from the space of polynomials which vanish on $\operatorname{span}\left(x_1, \ldots, x_{n/2}\right)$, and then apply an appropriate change of basis to obtain a sample from $\mathcal{I}_{d,A}$. So assume without loss of generality that $A = \operatorname{span}\left(x_1, \ldots, x_{n/2}\right)$.

We claim that a polynomial is zero on A if and only if each monomial is divisible by one of $x_{n/2+1}, \ldots, x_n$. This will immediately give an $O\left(n^d\right)$-time sampling algorithm, because we can consider each of the $O\left(n^d\right)$ degree-d monomials in turn, and include each one independently with probability 1/2 if it is divisible by one of $x_{n/2+1}, \ldots, x_n$.

To prove the claim, let p be a polynomial. If each monomial is divisible by one of $x_{n/2+1}, \ldots, x_n$ then clearly p is identically zero on A. Otherwise, let m be a minimal monomial which is not

divisible by $x_{n/2+1}, \ldots, x_n$. Consider the vector v with i^{th} coordinate equal to 1 if and only if x_i divides m. Since m is not divisible by $x_{n/2+1}, \ldots, x_n$, clearly $v \in A$. Also, since m is minimal, every other monomial must evaluate to 0 on v. Thus $p(v) = m(v) = 1$, so p is not identically zero on A. ■

In addition to sampling polynomials that vanish on A, we would like to guarantee that a sufficiently large system of such polynomials uniquely determines the space A, so that such a system can be effectively used as a membership oracle.

Lemma 31. *Fix $A \leq \mathbb{F}_2^n$ and $\beta > 1$, and choose βn independent uniformly-random samples from $\mathcal{I}_{d,A}$. With probability $1 - 2^{-\Omega(n)}$, the set of points on which they are all zero is exactly A.*

Proof. Fix some $v \notin A$. Then there must be some $w \in A^{\perp}$ such that $w \cdot v = 1$. Then the map $p(v) \mapsto p(v) + w \cdot v$ is an involution of $\mathcal{I}_{d,A}$. But within each orbit, exactly one element is zero at v. This means that exactly half of the polynomials in $\mathcal{I}_{d,A}$ vanish at v. So the probability that βn independent samples from $\mathcal{I}_{d,A}$ *all* vanish at v is $2^{-\beta n}$. By the union bound, the probability that there are any shared zeroes outside of A is therefore at most $2^n 2^{-\beta n} = 2^{-\Theta(n)}$. ■

In our scheme we will use noisy samples from $\mathcal{I}_{d,A}$, defined as follows: let $\mathcal{Z}_{d,A,\epsilon}$ be the distribution which sets A' to be A with probability $1-\epsilon$, or a random subspace of dimension $\dim(A)$ with probability ϵ, and then returns a random element of $\mathcal{I}_{d,A'}$. We would like to show that enough samples from $\mathcal{Z}_{d,A,\epsilon}$ also uniquely define the subspace A.

Lemma 32. *Fix $A \leq \mathbb{F}_2^n$ and $\epsilon < 1$, and let $\beta = \frac{25}{(1-\epsilon)^2}$ and $\delta = \frac{1-\epsilon}{3}$. Choose βn independent samples $p_1, \ldots, p_{\beta n}$ from $\mathcal{Z}_{d,A,\epsilon}$. Then with probability $1 - 2^{-\Omega(n)}$, the set of points on which at least $\left(\frac{1}{2} + \delta\right)\beta n$ of the p_i's vanish is exactly A.*

Proof. First observe that with high probability, each point $v \notin A$ is a zero of little more than half of the p_i's. If p_i was drawn from $\mathcal{I}_{d,A}$, then $\mathrm{E}[p_i(v)] = 1/2$, by the argument in Lemma 31. If p_i was drawn from a uniformly-random $\mathcal{I}_{d,A'}$, then

$$\begin{aligned}\mathrm{E}[p_i(v)] &\geq \frac{1}{2} - \Pr\left[v \in A'\right] \\ &= \frac{1}{2} - \frac{1}{2^{n/2}}.\end{aligned}$$

Hence, by a Chernoff bound,

$$\begin{aligned}\Pr\left[\sum_{i=1}^{\beta n} p_i(v) < \left(\frac{1}{2} - \delta\right)\beta n\right] &\leq \exp\left(-\frac{1}{2}\frac{\beta n}{2}(2\delta)^2\right) \\ &= \exp\left(-\frac{25}{(1-\epsilon)^2}\frac{(1-\epsilon)^2}{9}n\right) \\ &< 2^{-4n}.\end{aligned}$$

Next observe that each point $v \in A$ is a zero of *every* p_i drawn from $\mathcal{I}_{d,A}$, as well as each p_i drawn from a random $\mathcal{I}_{d,A'}$ with independent probability at least 1/2. Hence

$$\mathrm{E}[p_i(v)] \leq (1-\epsilon) \cdot 0 + \frac{\epsilon}{2}.$$

So by another Chernoff bound,

$$\begin{aligned}\Pr\left[\sum_{i=1}^{\beta n} p_i(v) \geq \left(\frac{1}{2}-\delta\right)\beta n\right] &\leq \exp\left(-\frac{1}{4}\frac{\beta n\epsilon}{2}\left(\frac{2}{\epsilon}\left(\frac{1}{2}-\delta\right)-1\right)^2\right)\\ &= \exp\left(-\frac{25n\epsilon}{8(1-\epsilon)^2}\left(\frac{1-\epsilon}{3\epsilon}\right)^2\right)\\ &= \exp\left(-\frac{25n}{72\epsilon}\right)\\ &< 2^{-n/2}.\end{aligned}$$

So by the union bound, with $1-2^{-\Omega(n)}$ probability we obtain $\sum_{i=1}^{\beta n} p_i(v) \geq \left(\frac{1}{2}-\delta\right)\beta n$ for all $v \notin A$ and $\sum_{i=1}^{\beta n} p_i(v) < \left(\frac{1}{2}-\delta\right)\beta n$ for all $v \in A$, as desired. ■

6.2 The Explicit Hidden-Subspace Mini-Scheme

In our explicit mini-scheme, the bank chooses a subspace A randomly and publishes βn samples from $\mathcal{Z}_{d,A,\epsilon}$ and βn samples from $\mathcal{Z}_{d,A^\perp,\epsilon}$, along with the quantum money state $|A\rangle$. By Lemma 32 a user can use this information to test membership in A and $A^\perp$, and can therefore implement the oracle mini-scheme in Section 5.1.

Formally, the mini-scheme $\mathcal{E}$ is defined as follows:

- Parameters $\epsilon < 1$ and $d \geq 4$ are fixed. The security and complexity of the scheme may increase for larger d and ϵ. Define $\beta := \frac{25}{(1-\epsilon)^2}$ and $\delta = \frac{1-\epsilon}{3}$.
- $\mathsf{Bank}(0^n)$ selects an $n/2$-dimensional subspace $A \leq \mathbb{F}_2^n$ uniformly at random, say by selecting $n/2$ random linearly-independent generators. It then sets $s := (s_A, s_{A^\perp})$, where s_A is a list of βn independent samples from $\mathcal{Z}_{d,A,\epsilon}$, and $s_{A^\perp}$ is a list of βn independent samples from $\mathcal{Z}_{d,A^\perp,\epsilon}$. It prepares the money state $|A\rangle$ and outputs the banknote $|\$_s\rangle := |s\rangle\,|A\rangle$.
- $\mathsf{Ver}(\not{c})$ first checks that $\not{c}$ has the form $(s_A, s_{A^\perp}, \rho)$ where s_A and $s_{A^\perp}$ are both lists of βn polynomials over $\mathbb{F}_2^n$. If not, it rejects. If so, it can evaluate the polynomials in s_A and check that a $\frac{1}{2}+\delta$ fraction of them are zero to test membership in A, and can similarly use $s_{A^\perp}$ to test membership in $A^\perp$. Using these subroutines, $\mathsf{Ver}(\not{c})$ applies $V_A = H_2^{\otimes n}\mathbb{P}_{A^\perp}H_2^{\otimes n}\mathbb{P}_A$ and accepts $\not{c}$ if and only if $V_A(\rho)$ accepts.

6.3 Analysis

We first show that the mini-scheme $\mathcal{E}$ has exponentially-small completeness error.

Theorem 33. *$\mathcal{E}$ has completeness error $2^{-\Omega(n)}$.*

Proof. By Lemma 32, with probability $1-2^{-\Omega(n)}$ we have that for all points v:

- $v \in A$ if and only if a $\frac{1}{2}+\delta$ fraction of the polynomials in s_A are zero at v.
- $v \in A^\perp$ if and only if a $\frac{1}{2}+\delta$ fraction of the polynomials in $s_{A^\perp}$ are zero at v.

In this case, the verifier can evaluate $\mathbb{P}_A$ and $\mathbb{P}_{A^\perp}$, so verification always works because the original hidden subspace mini-scheme had perfect completeness. ■

We now wish to argue about the soundness of $\mathcal{E}$. Naturally, we can only hope to prove soundness assuming some computational hardness conjecture. What is nice, though, is that we can base $\mathcal{E}$'s soundness on a conjecture that talks only about the hardness of a "classical" cryptographic problem (i.e., a problem with classical inputs and outputs). Let us now state that conjecture, which is simply the abstract Conjecture 28 specialized to the setting of multivariate polynomials.

Conjecture 34 (Direct Product for Finding Subspace Elements)**.** *Let $\beta := \frac{25}{(1-\epsilon)^2}$. Given βn samples from $\mathcal{Z}_{d,A,\epsilon}$ and $\mathcal{Z}_{d,A^\perp,\epsilon}$, no polynomial-time quantum algorithm can find a complete list of generators for A with success probability $\Omega\left(2^{-n/2}\right)$.*

Note that it is easy to find *one* nonzero element of A with success probability $2^{-n/2}$, by choosing $x \in \mathbb{F}_2^n$ randomly. Conjecture 34 asserts both that it is impossible to do too much better using $\mathcal{Z}_{d,A,\epsilon}$ and $\mathcal{Z}_{d,A^\perp,\epsilon}$, and that finding multiple elements of A is significantly harder than finding one element.

The security of mini-scheme $\mathcal{E}$ follows easily from Conjecture 34, *despite* the fact that a would-be counterfeiter has access to a valid quantum banknote, whereas Conjecture 34 involves no such assumption.

Theorem 35 (Security Reduction for Explicit Mini-Scheme)**.** *Unless Conjecture 34 is violated, the mini-scheme $\mathcal{E}$ is secure.*

Proof. Let $C_{\mathcal{E}}$ be a counterfeiter against $\mathcal{E}$. Then we need to show that, using $C_{\mathcal{E}}$, we can find a complete list of generators for A with $\Omega\left(2^{-n/2}\right)$ success probability.

Given $A \leq \mathbb{F}_2^n$ with $\dim(A) = n/2$, let s_A be a set of βn samples from $\mathcal{Z}_{d,A,\epsilon}$, let $s_{A^\perp}$ be a set of βn samples from $\mathcal{Z}_{d,A^\perp,\epsilon}$, and let $s := (s_A, s_{A^\perp})$. Using s, we can decide membership in A, and can therefore apply the projective measurement $\mathbb{P}_A$. So let us prepare the uniform superposition over all 2^n elements of $\mathbb{F}_2^n$, and then apply $\mathbb{P}_A$ to it. With probability $2^{-n/2}$, this produces the state $|A\rangle$.

Once we have s and $|A\rangle$, we can then form the banknote $|\$\rangle := |s\rangle\,|A\rangle$, and provide this banknote to the counterfeiter $C_{\mathcal{E}}$. By hypothesis, $C_{\mathcal{E}}$ outputs a (possibly-entangled) state ρ on two registers, such that

$$\langle A|^{\otimes 2}\,\rho\,|A\rangle^{\otimes 2} \geq \Delta$$

for some $\Delta = \Omega\left(1/\operatorname{poly}(n)\right)$. But now, because the mini-scheme $\mathcal{E}$ is projective, Theorem 15 applies, and we can *amplify* ρ to increase its fidelity with $|A\rangle^{\otimes 2}$. After $O\left(\frac{1}{\Delta^2}\log n\right)$ calls to $C_{\mathcal{E}}$, this gives us a state σ such that

$$\langle A|^{\otimes 2}\,\sigma\,|A\rangle^{\otimes 2} \geq 1 - \frac{1}{n^2}.$$

More generally, by alternating counterfeiting steps and amplification steps, we can produce as many registers as we like that each have large overlap with $|A\rangle$. In particular, we can produce a state ξ such that

$$\langle A|^{\otimes n}\,\xi\,|A\rangle^{\otimes n} \geq 1 - o(1).$$

If we now run Ver on each of the registers of ξ, the probability that every invocation accepts is $1 - o(1)$. Furthermore, supposing that happens, the state we are left with is simply $|A\rangle^{\otimes n}$. Finally, we measure each register of $|A\rangle^{\otimes n}$ in the standard basis. This gives us n elements $x_1, \ldots, x_n \in A$,

which are independent and uniformly random. So by standard estimates, the probability that $x_1, \ldots, x_n$ do *not* contain a complete generating set for A is $1/\exp(n)$.

Overall, the procedure above succeeded with probability $2^{-n/2}(1 - o(1))$, thereby giving us the desired contradiction with Conjecture 34. ■

Using the standard construction of quantum money schemes, we can now produce a complete explicit money scheme, whose security follows from Conjecture 34.

Theorem 36 (Security Reduction for Explicit Scheme)**.** *Assuming Conjecture 34, there exists a secure public-key quantum money scheme with completeness error* $2^{-\Omega(n)}$*.*

Proof. We apply the standard construction of Theorem 16 with the mini-scheme $\mathcal{E}$, whose completeness and soundness follow from Theorems 33 and 35 respectively, assuming Conjecture 34. ■

6.4 Justifying Our Hardness Assumption

Though our hardness assumption is new, it is closely related to standard assumptions in *multivariate polynomial cryptography*. Given a system of multivariate quadratics over $\mathbb{F}_2$, finding a common zero is known to be NP-hard; moreover, it is strongly believed that the problem remains hard even for *random* systems of multivariate polynomials, and cryptosystems based on this hardness assumption are considered promising candidates for post-quantum cryptography [22]. Therefore, if Conjecture 34 fails, it will almost certainly be because some additional structure in this problem facilitates a new attack.

There are several ways in which Conjecture 34 is stronger than the assumption that solving random systems of multivariate polynomials is hard:

(1) Our systems have large, well-structured solution spaces A and $A^{\perp}$. Systems with many solutions are not normally considered in the literature, and while there appear to be no known attacks that exploit this structure, the possibility is not ruled out.

(2) We provide two related systems, one with zeroes in A and one with zeroes in $A^{\perp}$. Again, this is a very specific structural property which has not been considered, and there might be some unexpected attacks exploiting it.

(3) Conjecture 34 asserts that no adversary can succeed with probability $2^{-n/2}$, which seems significantly easier than succeeding with non-negligible probability.

On the other hand, Conjecture 34 is *weaker* than typical assumptions in multivariate polynomial cryptography in at least one respect: a would-be counterfeiter needs to solve a system of polynomial equations with a *constant fraction of noise*.

Solving noisy systems of *linear* equations over $\mathbb{F}_2$ is called the *learning parity with noise* problem, and is generally believed to be difficult even for quantum computers [40]. If true, this suggests that Gaussian elimination is fundamentally hard to adapt to the presence of noise. But computing a Gröbner basis is a strict generalization of Gaussian elimination to higher degree, and involves a nearly identical process of elimination. Existing superficially-distinct approaches to algebraic cryptanalysis, such as the *XL algorithm*, are also fundamentally based on a similar reduction process [11]. It therefore seems unlikely that these approaches can be efficiently adapted to the setting

with noise. The problem of solving polynomials with noise has been studied recently, and the best-known approaches involve performing an exponential time search to determine which equations are noisy [7].

But if solving linear systems with noise is already hard, why do we even use higher-degree polynomials in our scheme? The reason is that, alas, the structure of our money scheme facilitates a straightforward attack in the case $d = 1$.

Claim 37. *For every $\epsilon < 1$, one can recover A efficiently given a linear number of samples from $\mathcal{Z}_{1,A,\epsilon}$ and $\mathcal{Z}_{1,A^{\perp},\epsilon}$.*

Proof. Let $p_1, \ldots, p_m$ and $q_1, \ldots, q_m$ be homogeneous linear polynomials, of which a $1 - \epsilon$ fraction vanish on A and $A^{\perp}$ respectively. Then the key observation is that each p_i vanishes on A if and only if it has the form $p_i(v) = u_i \cdot v$ for some $u_i \in A^{\perp}$, while each q_i vanishes on $A^{\perp}$ if and only if it has the form $q_i(v) = w_i \cdot v$ for some $w_i \in A$. But for each $i \in [m]$, we can efficiently *decide* whether $u_i \in A^{\perp}$ by counting the number of j's for which $q_j(u_i) = 0$, and we can likewise decide whether $w_i \in A$ by counting the number of j's for which $p_j(w_i) = 0$. Thus we can learn $\Theta(n)$ random elements of A or $A^{\perp}$, and thereby recover bases for them. ■

Notice that this attack was possible because of the very simple relationship between A and $A^{\perp}$ which was reflected in the structure of the linear equations. There *might* be a more sophisticated attack for higher degrees, but this is suggested only weakly by the existence of such an algorithm in the linear case. Indeed, the relation between complementary linear subspaces is precisely the type of structure which should be preserved by linear maps but *not* by higher-degree polynomials.

For degree-2 polynomials, it is possible to obtain a similar attack which nearly recovers A from only a *single sample*. This attack relies on the observation that quadratics have an easily-computed canonical form [19], from which the description of A can be extracted. The essential problem is that quadratic polynomials are very closely related to bilinear forms, and that powerful methods from linear algebra can therefore be applied to them.

Fortunately, this linear structure seems to be computationally obscured when $d \geq 3$. This phenomenon is related to the sharp discontinuity in the difficulty of tensor problems with order 3 and higher. More concretely, the coefficients of a degree-d polynomial can be viewed as the entries of an order-d tensor, and the existence of an attack in the degree $d = 2$ case corresponds to the possibility of efficient operations on order-2 tensors. Basic operations on order-3 tensors are NP-hard [29], however, and this suggests that analogous attacks may not exist against degree-3 polynomials.

This state of affairs is reflected in existing attacks on a standard cryptographic assumption called *polynomial isomorphism with one secret*. In this problem, we are given two polynomials which are related to each other by an unknown linear change of coordinates, and the task is to find such a change of coordinates. For degree-2 polynomials, this problem can be easily solved in polynomial time [19], but for degree-3 polynomials the best known attacks already take exponential time [39, 26, 19]. However, if an attacker is given n bits of partial information about the linear transformation, then even in the $d = 3$ case, it becomes possible to find the linear transformation that relates the polynomials [19]. This does not facilitate an attack on our assumption directly, but it suggests that a similar attack might be possible when $d = 3$, because an attacker is only required to succeed with probability $2^{-n/2}$.

Fortunately, this attack seems to rely on the particular structure of degree 2 and 3 polynomials. Of course it is entirely possible that similar algorithms may be discovered for higher-degree polynomials, but this would represent an advance in algebraic cryptanalysis.

7 Improving Private-Key Quantum Money

Recall that a *private-key* quantum money scheme is one where only the bank itself is able to verify banknotes, using an n-bit key $k = k_{\text{private}} = k_{\text{public}}$ that it keeps a closely-guarded secret. Compensating for this disadvantage, it is known how to construct private-key schemes with much stronger security guarantees than seem possible for public-key schemes.

The "original" private-key scheme was due to Wiesner [43]. There, each banknote has the form $|\$_s\rangle := |s\rangle \left|\psi_{R(k,s)}\right\rangle$, where s is an n-bit classical serial number, R is a classical random oracle, and $\left|\psi_{R(k,s)}\right\rangle$ is a quantum money state consisting of n unentangled qubits, each prepared in one of the four states $|0\rangle, |1\rangle, |+\rangle, |-\rangle$. Given a serial number s, as well as the secret key k, the bank can look up a $2n$-bit *classical description* $R(k,s)$ of $\left|\psi_{R(k,s)}\right\rangle$. Having done so, the bank can then *verify* the banknote $|\$_s\rangle$, by measuring each qubit of $\left|\psi_{R(k,s)}\right\rangle$ in the appropriate basis ($\{|0\rangle, |1\rangle\}$ or $\{|+\rangle, |-\rangle\}$), and checking that it gets the right outcomes. On the other hand, it seems "intuitively obvious" that given $|s\rangle \left|\psi_{R(k,s)}\right\rangle$, a counterfeiter who does not know $R(k,s)$ can prepare two states, both of which pass the bank's verification test, with success probability at most c^n for some constant $c < 1$. If so, then Wiesner's scheme is information-theoretically secure.

Strangely, we are unaware of any rigorous proof of the above statement until quite recently. However, answering a question by one of us,[20] Molina, Vidick and Watrous [35] have now supplied the key ingredient for a proof of security for Wiesner's scheme. In particular, they show that, if a counterfeiter C tries to copy a state of the form $\left|\psi_{R(k,s)}\right\rangle$ (knowing nothing about k or s), then C's two output states will both pass the verification test with probability at most $(3/4)^n$. Furthermore, this is tight: there *exists* a (non-obvious) counterfeiting strategy with $(3/4)^n$ success probability.

To prove the security of Wiesner's scheme rigorously, one also needs to show that, even given *many* states of the form

$$|\$_{s_1}\rangle = |s_1\rangle \left|\psi_{R(k,s_1)}\right\rangle, \ldots, \left|\$_{s_q}\right\rangle = |s_q\rangle \left|\psi_{R(k,s_q)}\right\rangle,$$

a counterfeiter cannot prepare any additional states of the form $|s\rangle \left|\psi_{R(k,s)}\right\rangle$, even with new values of s. A full proof of this fact will be given in a forthcoming paper [4]. However, we can already sketch the idea, since it closely follows the methods in this paper. One first defines a notion of *private-key mini-schemes*, which require only a single banknote to be protected, and which treat the serial number s as "hardwired" into the verification procedure. One then observes that the result of Molina et al. [35] implies the security of the "Wiesner mini-scheme." Finally, one gives a general reduction—directly analogous to Theorem 16—showing that a full-blown private-key quantum money scheme can be constructed from the following two primitives:

(1) any private-key mini-scheme, and

(2) any random or pseudorandom function family $R_k(s) = R(k,s)$.

The construction is exactly the one suggested by Wiesner: that is, each banknote has the form $|\$_s\rangle = |s\rangle \left|\psi_{R_k(s)}\right\rangle$, for a uniformly-random serial number s. The theorem one has to prove is that any attack against the resulting money scheme can be converted *either* into an attack against the mini-scheme, or else into a way to distinguish the functions $\{R_k\}_k$ from random. The proof is broadly similar to the proof of Theorem 16, though the details turn out to be more complicated.

[20] See http://theoreticalphysics.stackexchange.com/questions/370/rigorous-security-proof-for-wiesners-quantum-money

A variant of Wiesner's scheme, proposed by BBBW [16], is exactly the same, except that we replace the random oracle $R(k,s)$ by a *pseudorandom function* $f_k(s)$—something much more realistically available. Unlike Wiesner's scheme, the BBBW scheme is not unconditionally secure.[21] On the other hand, any break of the BBBW scheme would imply a polynomial-time quantum algorithm to distinguish f_k from a truly random function. The proof is simple: if f_k were random, then we would be back in Wiesner's scenario, where counterfeiting is information-theoretically impossible. Therefore, *if* we can copy $|\$_s\rangle = |s\rangle \left|\psi_{f_k(s)}\right\rangle$ with non-negligible fidelity, for s and k chosen uniformly at random, then we have also distinguished f_k from a random function.

Unfortunately, as pointed out by Lutomirski [32] and Aaronson [3], the Wiesner and BBBW schemes both have a serious security hole. Namely, suppose that a counterfeiter C can repeatedly submit alleged banknotes to a "naïve and trusting bank" for verification. Given a quantum state σ, such a bank not only tells C whether the verification procedure $\mathsf{Ver}(k,\sigma)$ accepted or rejected, but also, in either case, *gives the post-measurement state* $\widetilde{\sigma}$ *back to* C. Then starting from a single legitimate banknote $|\$_s\rangle = |s\rangle \left|\psi_{f(s)}\right\rangle$, we claim that C can recover a complete classical description of $\left|\psi_{f(s)}\right\rangle$, using only $O(n \log n)$ queries to the bank—after which C can of course prepare as many copies of $\left|\psi_{f(s)}\right\rangle$ as it likes.

The strategy is simple: let $\left|\psi_{f(s)}\right\rangle = |\theta_1\rangle \cdots |\theta_n\rangle$. Then for each $i \in [n]$, the counterfeiter tries "swapping out" the ith qubit $|\theta_i\rangle$ and replacing it with $|b\rangle$, for each of the four possibilities $|b\rangle \in \{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$. It then uses $O(\log n)$ queries to the bank to estimate

$$\Pr\left[\mathsf{Ver}\left(k, |s\rangle |\theta_1\rangle \cdots |\theta_{i-1}\rangle |b\rangle |\theta_{i+1}\rangle \cdots |\theta_n\rangle\right) \text{ accepts}\right],$$

so that it learns the correct classical description of $|\theta_i\rangle$ with success probability $1 - o(1/n)$. The crucial point is that none of these queries damage the qubits *not* being investigated ($|\theta_j\rangle$ for $j \neq i$), since the bank measures those qubits in the correct bases. Therefore C can reuse the same banknote $|\$_s\rangle$ for each query.

More generally, recall from Section 3.1 that we call a private-key quantum money scheme *query-secure*, if it remains secure even assuming the counterfeiter C can make adaptive queries to the verification procedure $\mathsf{Ver}(k,\cdot)$. Given a state σ, such queries return

(i) the classical information of whether $\mathsf{Ver}(k,\sigma)$ accepted or rejected, and

(ii) the post-measurement quantum state $\widetilde{\sigma}$.

Any *public-key* quantum money scheme yields a query-secure scheme as well, since a counterfeiter who knows the actual *code* of Ver can certainly simulate oracle access to Ver. On the other hand, we saw above that the Wiesner and BBBW schemes are *not* query-secure. Recently, Farhi et al. [23] proved a much more general "no-go" theorem:

Theorem 38 (Adaptive Attack on Wiesner-Like Schemes [23])**.** *No quantum money scheme can be query-secure, if*

(i) the banknotes have the form $|\$_s\rangle = |s\rangle |\psi_s\rangle$*,*

(ii) verification of (s,ρ) *consists of projecting* ρ *onto* $|\psi_s\rangle\langle\psi_s|$*, and*

[21] Indeed, we will show in forthcoming work [4] that *no* quantum money scheme is unconditionally secure, unless it involves an oracle or a giant random number table.

(iii) $|\psi_s\rangle$ *can be reconstructed uniquely from the statistics of* $T = \text{poly}(n)$ *efficiently-implementable measurements* $M_1, \ldots, M_T$*, each of which has at most* $\text{poly}(n)$ *possible outcomes.*

Intuitively, Theorem 38 tells us that, if we want a query-secure quantum money scheme, then the banknotes will have to hide information in the "global correlations" between large numbers of qubits. This raises an extremely interesting question: *can* we construct a query-secure quantum money scheme—one whose security is either *unconditional* (as with Wiesner's scheme [43]), or else based on a pseudorandom function (as with the BBBW scheme [16])?

Below, we sketch a proof that the answer is "yes"—and moreover, that this answer follows in an appealing way from work we have already done in this paper, on constructing *public-key* quantum money schemes that are provably secure relative to a classical oracle.

Theorem 39 (Query-Secure Variant of Wiesner's Scheme)**.** *There exists an unconditionally query-secure quantum money scheme, with perfect completeness and exponentially-small soundness error, relative to a random oracle* R *(i.e., assuming the bank has access to a giant table of random numbers).*

Proof Sketch. Our scheme $\mathcal{S}$ will be modeled closely after the hidden subspace scheme from Section 5. Given a key k and a serial number s, let $A_{k,s}$ be a subspace of $\mathbb{F}_2^n$ chosen uniformly at random, subject to $\dim(A_{k,s}) = n/2$. For each k, s, we will think of the random oracle R as encoding a *classical description* $R(k,s)$ of $A_{k,s}$: for example, a list of $n/2$ generators. Also, let $|A_{k,s}\rangle$ be a uniform superposition over $A_{k,s}$.

We now define our money scheme $\mathcal{S} = (\mathsf{KeyGen}, \mathsf{Bank}, \mathsf{Ver})$.

- $\mathsf{KeyGen}(0^n)$ simply outputs an n-bit secret key $k \in \{0,1\}^n$ uniformly at random.
- $\mathsf{Bank}(k)$ outputs a banknote of the form $|\$_s\rangle := |s\rangle |A_{k,s}\rangle$, for a serial number $s \in \{0,1\}^n$ chosen uniformly at random.
- $\mathsf{Ver}(k,(s,\rho))$ applies the measurement $V_{A_{k,s}}$ to ρ, where $V_{A_{k,s}} := |A_{k,s}\rangle\langle A_{k,s}|$ is a projector onto $|A_{k,s}\rangle$, and accepts if and only if $V_{A_{k,s}}(\rho)$ does.

Clearly the bank can both prepare the money states $|A_{k,s}\rangle$ and verify them in polynomial time, using its knowledge of k together with its oracle access to R. On the other hand, consider a counterfeiter C who has oracle access to Ver, but no knowledge of k, and (for the time being, at least) no access to R either. Given a legitimate banknote $|s\rangle |A_{k,s}\rangle$, suppose C output two (possibly-entangled) registers ϕ_1 and ϕ_2, such that the double-verifier $\mathsf{Ver}_2(k, \phi_1, \phi_2)$ accepted with non-negligible probability. Then we claim that C could *also* be used to break the public-key mini-scheme $\mathcal{M}$ defined in Section 5! Intuitively, this is because any query to the bank, of the form $\mathsf{Ver}(k,(s,\rho))$, can easily be simulated using queries to $U_{A_{k,s}}$ and $U_{A_{k,s}^\perp}$, the membership oracles for $A_{k,s}$ and $A_{k,s}^\perp$ respectively that are available to a counterfeiter C' against $\mathcal{M}$.

To make the reduction rigorous, one needs to confirm three further facts. First, given any serial number $s' \neq s$, a polynomial-time counterfeiter C cannot find *any* state σ such that $\mathsf{Ver}(k,(s',\sigma))$ accepts with non-negligible probability. Second, access to the random oracle R is useless to C. And third, being given *multiple* banknotes $|\$_{s_1}\rangle, \ldots, |\$_{s_q}\rangle$ does not help C. All three facts are "obvious," given C's lack of knowledge of the secret key k together with the BBBV lower bound [14]. A formal proof will be given in the forthcoming paper [4]. ■

Theorem 40 (Query-Secure Variant of BBBW Scheme)**.** *If there exists a pseudorandom function family $\mathcal{F} = \{f_k\}_k$ secure against $2^{n^{\Omega(1)}}$-time quantum adversaries, then there exists a query-secure quantum money scheme with perfect completeness and exponentially small soundness error.*

Proof Sketch. In the construction of Theorem 39, we simply replace the random oracle $R(k, s)$ by the pseudorandom function $f_k(s)$, to obtain a new scheme $\mathcal{S}'$ that does not require an oracle. Now, suppose there existed a counterfeiter C against $\mathcal{S}'$. Then we claim that we could use C to distinguish f_k from a truly random function, with non-negligible bias. The reason is simply that, if f_k *were* a truly random function, then $\mathcal{S}'$ would be information-theoretically secure by Theorem 39! And therefore, any successful attack against $\mathcal{S}'$ would yield a distinguisher for f_k. Once again, details will be given in the forthcoming paper [4]. ■

8 Open Problems

The "obvious" problem is to better understand the security of our explicit scheme based on polynomials. Are there nontrivial attacks, for example using Gröbner-basis algorithms? Can we base the security of our scheme—or a related scheme—on some cryptographic assumption that does *not* involve exponentially-small success probabilities? What happens as we change the finite field size, the polynomial degree, or the amount of noise added? Does "hiding" a subspace $A \leq \mathbb{F}_2^n$ in the way we suggest, as the set of common zeroes of multivariate polynomials $p_1, \ldots, p_m : \mathbb{F}_2^n \to \mathbb{F}_2$, have other cryptographic applications, for example in *program obfuscation* [12]?

Of course, public-key quantum money is still a young subject, and there is tremendous scope for inventing new schemes, which might be based on different assumptions, and have different strengths and weaknesses.

Let us now move on some general questions about public-key quantum money. First, is there an unconditionally-secure public-key quantum money scheme relative to a *random* oracle R? (Recall that Wiesner's original scheme [43] was unconditionally-secure and used only a random oracle, but was private-key. Meanwhile, our scheme from Section 5 is unconditionally-secure and public-key, but requires a non-random oracle.) Second, is there a public-key quantum money scheme where the banknotes consist of *single, unentangled qubits*, as in Wiesner's scheme? Note that the results of Farhi et al. [23] imply that, if such a scheme exists, then it cannot be projective. Third, is there a general way to amplify soundness error in quantum money schemes?[22] (We show how to amplify *completeness* error in Appendix 10.)

The tools we develop, in Section 2, raise some interesting questions in their own right. For example, is there an amplitude amplification algorithm that achieves a quadratic (Grover) speedup, and also converges monotonically and *exponentially-quickly* toward the target subspace G? Also, can we give general tools for securing classical cryptographic constructions—such as the GGM construction of pseudorandom functions from pseudorandom generators [27]—against quantum attack?

8.1 Quantum Copy-Protection and More

Quantum money is just *one* novel cryptographic use for the No-Cloning Theorem, or for the No-Cloning Theorem combined with computational assumptions. Given essentially *any* object of

[22] Theorem 15 gives *some* soundness amplification for projective schemes: namely, from constant to $1/\operatorname{poly}(n)$. Here we are asking whether one can do anything better.

cryptographic interest, one can ask whether quantum mechanics lets us make the object uncloneable. Section 1.4 already discussed one example—*uncloneable signatures*—but there are many others, such as commitments and zero-knowledge proofs.[23]

Along those lines, Aaronson [3] recently proposed a task that, if achievable, would arguably be an even more dramatic application of the No-Cloning Theorem than quantum money: namely, *quantum software copy-protection*. He gave explicit schemes—which have not yet been broken!—for copy-protecting a restricted class of functions, namely the *point functions*. In these schemes, given a "password" $s \in \{0,1\}^n$, a software vendor can prepare a quantum state $|\psi_s\rangle$, which allows its holder to *recognize* s: in other words, to decide whether $x = s$ given $x \in \{0,1\}^n$ as input. On the other hand, given $|\psi_s\rangle$, it seems intractable not only to *find* s for oneself, but even to prepare a *second* quantum state with which s can be recognized.

Admittedly, recognizing passwords is an extremely restricted functionality. However, relative to a quantum oracle, Aaronson [3] also described a scheme to quantumly copy-protect *arbitrary* programs, exactly as well as if the software vendor were able to hand out uncloneable black boxes.[24] In the spirit of this paper, we can now ask: is there likewise a way to quantumly copy-protect arbitrary programs relative to a *classical* oracle? We conjecture that the answer is yes, and in fact we have plausible candidate constructions, which are directly related to the hidden-subspace money scheme of Section 5. However, the security of those constructions seems to hinge on the following conjecture.

Conjecture 41 (Direct Product for Finding Black-Box Subspace Elements)**.** *Let A be a uniformly-random subspace of $\mathbb{F}_2^n$ satisfying $\dim(A) = n/2$. Then given membership oracles for both A and $A^\perp$, any quantum algorithm needs $2^{\Omega(n)}$ queries to find two distinct nonzero elements $x, y \in A$, with success probability $\Omega\left(2^{-n/2}\right)$.*

Besides its applications for copy-protection, a proof of Conjecture 41 would be important piece of formal evidence for Conjecture 34, on which we based the security of our explicit money scheme. The conjecture is also of interest in its own right, as a natural generalization of our quantum lower bounds in Section 5.3.

9 Acknowledgments

We thank Andris Ambainis, Boaz Barak, Dmitry Gavinsky, Daniel Gottesman, Aram Harrow, Yuval Ishai, Shaunak Kishore, Greg Kuperberg, Andy Lutomirski, Abel Molina, Rafi Ostrovsky, Amit Sahai, Peter Shor, and John Watrous for helpful discussions and correspondence; and the anonymous reviewers for their comments.

References

[1] S. Aaronson. Quantum lower bound for the collision problem. In *Proc. ACM STOC*, pages 635–642, 2002. quant-ph/0111102.

[23] Even within complexity theory, it would be interesting to study the class QMA (Quantum Merlin-Arthur) subject to the constraint that witnesses must be hard to clone—or alternatively, that witnesses must be *easy* to clone!

[24] As usual, full details of the security proof have not yet appeared.

[2] S. Aaronson. Limitations of quantum advice and one-way communication. *Theory of Computing*, 1:1–28, 2005. quant-ph/0402095. Conference version in Proceedings of CCC'2004.

[3] S. Aaronson. Quantum copy-protection and quantum money. In *Proc. IEEE Conference on Computational Complexity*, pages 229–242, 2009.

[4] S. Aaronson. On the security of private-key quantum money, 2011. In preparation.

[5] S. Aaronson and G. Kuperberg. Quantum versus classical proofs and advice. *Theory of Computing*, 3(7):129–157, 2007. Previous version in Proceedings of CCC 2007. quant-ph/0604056.

[6] D. Aharonov and M. Ben-Or. Fault-tolerant quantum computation with constant error. In *Proc. ACM STOC*, pages 176–188, 1997. quant-ph/9906129.

[7] M. Albrecht and C. Cid. Cold boot key recovery by solving polynomial systems with noise. In Javier Lopez and Gene Tsudik, editors, *Applied Cryptography and Network Security - ACNS 2011*, volume 6715 of *Lecture Notes in Computer Science*, pages 57–72, 2011.

[8] N. Alon, M. Krivelevich, and B. Sudakov. Finding a large hidden clique in a random graph. In *Proc. ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 594–598, 1998.

[9] A. Ambainis. Quantum lower bounds by quantum arguments. *J. Comput. Sys. Sci.*, 64:750–767, 2002. Earlier version in ACM STOC 2000. quant-ph/0002066.

[10] A. Ambainis, L. Magnin, M. Roetteler, and J. Roland. Symmetry-assisted adversaries for quantum state generation. In *Proc. IEEE Conference on Computational Complexity*, pages 167–177, 2011. arXiv:1012.2112.

[11] G. Ars, J.-C. Faugère, H. Imai, M. Kawazoe, and M. Sugita. Comparison between XL and Gröbner basis algorithms. In *Proceedings of ASIACRYPT*, pages 338–353. LNCS, Springer-Verlag, 2004.

[12] B. Barak, O. Goldreich, R. Impagliazzo, S. Rudich, A. Sahai, S. Vadhan, and K. Yang. On the (im)possibility of obfuscating programs. In *Proceedings of CRYPTO*, pages 1–18, 2001. ECCC TR01-057.

[13] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, 2001. Earlier version in IEEE FOCS 1998, pp. 352-361. quant-ph/9802049.

[14] C. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. Strengths and weaknesses of quantum computing. *SIAM J. Comput.*, 26(5):1510–1523, 1997. quant-ph/9701001.

[15] C. H. Bennett and G. Brassard. Quantum cryptography: public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers Systems and Signal Processing*, pages 175–179, 1984.

[16] C. H. Bennett, G. Brassard, S. Breidbart, and S. Wiesner. Quantum cryptography, or unforgeable subway tokens. In *Proceedings of CRYPTO*, pages 267–275. Plenum Press, 1982.

[17] N. Bohr. *Atomic Physics and Human Knowledge*. Dover, 2010. First published 1961.

[18] D. Boneh, Ö. Dagdelen, M. Fischlin, A. Lehmann, C. Schaffner, and M. Zhandry. Random oracles in a quantum world. In *Proceedings of ASIACRYPT*, pages 41–69, 2011. arXiv:1008.0931.

[19] C. Bouillaguet, J.-C. Faugère, P.-A. Fouque, and L. Perret. Practical cryptanalysis of the identification scheme based on the isomorphism of polynomial with one secret problem. In *Public Key Cryptography*, pages 473–493, 2011.

[20] G. Brassard, P. Høyer, M. Mosca, and A. Tapp. Quantum amplitude amplification and estimation. In S. J. Lomonaco and H. E. Brandt, editors, *Quantum Computation and Information*, Contemporary Mathematics Series. AMS, 2002. quant-ph/0005055.

[21] S. Chakraborty, J. Radhakrishnan, and N. Raghunathan. Bounds for error reduction with few quantum queries. In *APPROX-RANDOM*, pages 245–256, 2005.

[22] J. Ding and B.-Y. Yang. Multivariate public key cryptography. In D. J. Bernstein, J. Buchmann, and E. Dahmén, editors, *Post-Quantum Cryptography*, pages 198–242. Springer-Verlag, 2009.

[23] E. Farhi, D. Gosset, A. Hassidim, A. Lutomirski, D. Nagaj, and P. Shor. Quantum state restoration and single-copy tomography. *Phys. Rev. Lett.*, 105(190503), 2010. arXiv:0912.3823.

[24] E. Farhi, D. Gosset, A. Hassidim, A. Lutomirski, and P. Shor. Quantum money from knots. In *Proc. Innovations in Theoretical Computer Science (ITCS)*, pages 276–289, 2012. arXiv:1004.5127.

[25] D. Gavinsky. Quantum money with classical verification. arXiv:1109.0372, 2011.

[26] W. Geiselmann, W. Meier, and R. Steinwandt. An attack on the isomorphisms of polynomials problem with one secret. *Int. J. Inf. Sec*, 2(1):59–64, 2003.

[27] O. Goldreich, S. Goldwasser, and S. Micali. How to construct random functions. *J. ACM*, 33(4):792–807, 1984.

[28] L. K. Grover. A fast quantum mechanical algorithm for database search. In *Proc. ACM STOC*, pages 212–219, 1996. quant-ph/9605043.

[29] C. Hillar and L.-H. Lim. Most tensor problems are NP hard, 2009. arXiv:0911.1393.

[30] E. Knill, R. Laflamme, and W. Zurek. Resilient quantum computation. *Science*, 279:342–345, 1998. quant-ph/9702058.

[31] T. Lee, R. Mittal, B. W. Reichardt, R. Špalek, and M. Szegedy. Quantum query complexity of state conversion. In *Proc. IEEE FOCS*, pages 344–353, 2011. arXiv:1011.3020.

[32] A. Lutomirski. An online attack against Wiesner's quantum money. arXiv:1010.0256, 2010.

[33] A. Lutomirski. Component mixers and a hardness result for counterfeiting quantum money. arXiv:1107.0321, 2011.

[34] A. Lutomirski, S. Aaronson, E. Farhi, D. Gosset, A. Hassidim, J. Kelner, and P. Shor. Breaking and making quantum money: toward a new quantum cryptographic protocol. In *Proc. Innovations in Theoretical Computer Science (ITCS)*, pages 20–31, 2010. arXiv:0912.3825.

[35] A. Molina, T. Vidick, and J. Watrous. Optimal counterfeiting attacks and generalizations for Wiesner's quantum money. arXiv:1202.4010, 2012.

[36] M. Mosca and D. Stebila. Quantum coins. In *Error-Correcting Codes, Finite Geometries and Cryptography*, volume 523, pages 35–47. American Mathematical Society, 2010. arXiv:0911.1295.

[37] M. Naor and M. Yung. Universal one-way hash functions and their cryptographic applications. In *Proc. ACM STOC*, pages 33–43, 1989.

[38] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[39] J. Patarin, L. Goubin, and N. Courtois. Improved algorithms for isomorphisms of polynomials. In *Proceedings of EUROCRYPT*, pages 184–200, 1998.

[40] O. Regev. On lattices, learning with errors, random linear codes, and cryptography. In *Proc. ACM STOC*, pages 84–93, 2005.

[41] J. Rompel. One-way functions are necessary and sufficient for secure signatures. In *Proc. ACM STOC*, pages 387–394, 1990.

[42] T. Tulsi, L. Grover, and A. Patel. A new algorithm for fixed point quantum search. *Quantum Information and Computation*, 6(6):483–494, 2006. quant-ph/0505007.

[43] S. Wiesner. Conjugate coding. *SIGACT News*, 15(1):78–88, 1983. Original manuscript written circa 1970.

10 Appendix: Reducing Completeness Error

When we defined quantum money schemes and mini-schemes in Section 3, we allowed the verifier to reject a *legitimate* money state with probability up to 1/3. But of course, a money scheme with completeness error $\varepsilon = 1/3$ is not very useful in practice! So in this appendix, we prove that the completeness error ε can be made exponentially small, at the cost of only a modest increase in the soundness error δ (i.e., the probability of successful counterfeiting).

Theorem 42 (Completeness Amplification for Mini-Schemes)**.** *Let $\mathcal{M} = (\mathsf{Bank}, \mathsf{Ver})$ be a quantum money mini-scheme with completeness error $\varepsilon < 1/2$ and soundness error $\delta < 1 - 2\varepsilon$. Then for all polynomials p and all $\delta' > \frac{\delta}{1-2\varepsilon}$, we can construct an amplified mini-scheme $\mathcal{M}' = \left(\mathsf{Bank}', \mathsf{Ver}'\right)$ with completeness error $1/2^{p(n)}$ and soundness error δ'.*

Proof. Let $k = \text{poly}(n)$ and $\eta > 0$ be parameters to be determined later. Our construction of $\mathcal{M}'$ is the "obvious" one based on repetition:

- $\mathsf{Bank}'(0^n)$ outputs a composite banknote $\$' := (s_1 \ldots s_k, \rho_{s_1} \ldots \rho_{s_k})$, where $(s_1, \rho_{s_1}), \ldots, (s_k, \rho_{s_k})$ are banknotes output independently by $\mathsf{Bank}(0^n)$.
- $\mathsf{Ver}'(¢)$ runs $\mathsf{Ver}(¢_1), \ldots, \mathsf{Ver}(¢_k)$, where $¢_1, \ldots, ¢_k$ are the (s, ρ_s) pairs in the alleged composite banknote $¢$, and accepts if and only if at least $(1 - \varepsilon - \eta)\, k$ invocations accept.

Note that Ver_2', the amplified double verifier, then takes as input a state of the form

$$(s_1 \ldots s_k, \sigma_1 \ldots \sigma_k, \xi_1 \ldots \xi_k),$$

and accepts if and only if $\mathsf{Ver}'(s_1 \ldots s_k, \sigma_1 \ldots \sigma_k)$ and $\mathsf{Ver}'(s_1 \ldots s_k, \xi_1 \ldots \xi_k)$ both accept. By choosing k sufficiently large and applying a Chernoff bound, it is clear that we can make the completeness error $1/2^{p(n)}$ for any polynomial p.

Meanwhile, suppose $\mathcal{M}'$ has soundness error δ': in other words, there exists a counterfeiter C' such that $\mathsf{Ver}_2'(s_1 \ldots s_k, C'(\$'))$ accepts with probability δ', given a valid composite banknote $\$'$. Then to prove the theorem, it suffices to construct a counterfeiter C for the original mini-scheme $\mathcal{M}$, such that $\mathsf{Ver}_2(s, C(\$))$ accepts with probability $\delta \geq (1 - 2\varepsilon - \eta)\delta'$, given a valid banknote $\$ = (s, \rho_s)$. This C works as follows:

(1) By calling $\mathsf{Bank}'(0^n)$, generate a new composite banknote $\$' = (\$_1, \ldots, \$_k)$.

(2) Let $\$'_{\text{new}}$ be the result of starting with $\$'$, then swapping out $\$_i$ for the banknote $\$$ to be copied, for some $i \in [k]$ chosen uniformly at random.

(3) Let $(s_1 \ldots s_k, \sigma_1 \ldots \sigma_k, \xi_1 \ldots \xi_k) := C'(\$'_{\text{new}})$.

(4) Output (s_i, σ_i, ξ_i).

By assumption,

$$\Pr\left[\mathsf{Ver}_2'(s_1 \ldots s_k, \sigma_1 \ldots \sigma_k, \xi_1 \ldots \xi_k) \text{ accepts}\right] \geq \delta'.$$

Now, suppose Ver_2' does accept. Then by the definition of Ver_2', at least $(1 - \varepsilon - \eta)k$ of

$$\mathsf{Ver}(s_1, \sigma_1), \ldots, \mathsf{Ver}(s_k, \sigma_k)$$

must have accepted, along with at least $(1 - \varepsilon - \eta)k$ of

$$\mathsf{Ver}(s_1, \xi_1), \ldots, \mathsf{Ver}(s_k, \xi_k).$$

So there must be at least $(1 - 2\varepsilon - 2\eta)k$ indices $j \in [k]$ such that $\mathsf{Ver}(s_j, \sigma_j)$ and $\mathsf{Ver}(s_j, \xi_j)$ *both* accepted. Therefore

$$\begin{aligned}\Pr\left[\mathsf{Ver}_2(s_i, \sigma_i, \xi_i) \text{ accepts}\right] &= \Pr\left[\mathsf{Ver}(s_i, \sigma_i) \text{ and } \mathsf{Ver}(s_i, \xi_i) \text{ accept}\right] \\ &\geq (1 - 2\varepsilon - 2\eta)\delta'.\end{aligned}$$

Taking $\eta > 0$ sufficiently small now yields the theorem. ■

A direct counterpart of Theorem 42, with exactly the same parameters, can be proved for public-key quantum money schemes. Once again, the main idea is to consider "composite banknotes" $\$' = (\$_1, \ldots, \$_k)$—and this time, to associate with each $\$_i$ a *different*, independently-chosen public/private key pair. Another counterpart of Theorem 42 can be proved for digital signature schemes, indeed with slightly better parameters ($\delta' > \frac{\delta}{1-\varepsilon}$ instead of $\delta' > \frac{\delta}{1-2\varepsilon}$). We omit the details.

11 Appendix: Complexity-Theoretic No-Cloning Theorem

In Section 5, we applied the inner-product adversary method to show that a uniform superposition $|A\rangle$ over a random subspace $A \leq \mathbb{F}_2^n$ requires $\Omega\left(2^{n/4}\right)$ quantum queries to duplicate, even if we are given access to an oracle that decides membership in both A and $A^\perp$. For completeness, in this appendix we present a simpler application of the inner-product adversary method: namely, we show that a Haar-random n-qubit state $|\psi\rangle$ requires $\Omega\left(2^{n/2}\right)$ queries to duplicate, if we are given access to an oracle U_ψ that accepts $|\psi\rangle$ and that rejects every state orthogonal to $|\psi\rangle$. The latter is the original result that Aaronson [3] called the "Complexity-Theoretic No-Cloning Theorem," though a proof has not appeared until now.

In Section 5, we used the lower bound for copying subspace states to construct a quantum money mini-scheme that was provably secure relative to a classical oracle. In the same way, one can use the Complexity-Theoretic No-Cloning Theorem to construct a mini-scheme that is provably secure relative to a *quantum* oracle. We omit the details of that construction, not only because it is superseded by the classical oracle construction in Section 5, but because the two constructions are essentially the same. The one real difference is that the quantum oracle construction benefits from a quadratically better lower bound on the number of queries needed to counterfeit: $\Omega\left(2^{n/2}\right)$ rather than $\Omega\left(2^{n/4}\right)$.

Choose an n-qubit pure state $|\psi\rangle$ uniformly from the Haar measure, and fix $|\psi\rangle$ in what follows. Let U_ψ be a unitary transformation such that $U_\psi|\psi\rangle = -|\psi\rangle$ and $U_\psi|\eta\rangle = |\eta\rangle$ for all $|\eta\rangle$ orthogonal to $|\psi\rangle$. The following is the direct analogue of Theorem 22.

Theorem 43 (Complexity-Theoretic No-Cloning)**.** *Given one copy of $|\psi\rangle$, as well as oracle access to U_ψ, a counterfeiter needs $\Omega\left(2^{n/2}\right)$ queries to prepare $|\psi\rangle^{\otimes 2}$ with certainty (for a worst-case $|\psi\rangle$).*

Proof. We will apply Theorem 20. Let the set $\mathcal{O}$ contain U_ψ for every possible n-qubit state $|\psi\rangle$. Then S_{U_ψ} is just the 1-dimensional subspace corresponding to $|\psi\rangle$. Also, put $(U_\psi, U_\varphi) \in R$ if and only if $|\langle\psi|\varphi\rangle| = c$, for some $0 < c < 1$ to be specified later. Then for all $U_\psi \in \mathcal{O}$ and $|\eta\rangle \in S_{U_\psi}^\perp$, we have

$$\begin{aligned}
\mathop{\mathrm{E}}_{U_\varphi \,:\, (U_\psi, U_\varphi) \in R}\left[|\langle\eta|\varphi\rangle|^2\right] &= \mathop{\mathrm{E}}_{|\varphi\rangle \,:\, |\langle\psi|\varphi\rangle| = c}\left[|\langle\eta|\varphi\rangle|^2\right] \\
&= \mathop{\mathrm{E}}_{|v\rangle \in S_{U_\psi}^\perp}\left[\left|\langle\eta|\left(c\,|\psi\rangle + \sqrt{1-c^2}\,|v\rangle\right)\right|^2\right] \\
&= \left(1-c^2\right) \mathop{\mathrm{E}}_{|v\rangle \in S_{U_\psi}^\perp}\left[|\langle\eta|v\rangle|^2\right] \\
&= \frac{1-c^2}{2^n-1}.
\end{aligned}$$

So set $\varepsilon := \frac{1-c^2}{2^n-1}$. If the counterfeiter succeeds, it must map $|\psi\rangle$ to some state $|f_\psi\rangle := |\psi\rangle|\psi\rangle\left|\mathrm{garbage}_\psi\right\rangle$, and $|\varphi\rangle$ to $|f_\varphi\rangle := |\varphi\rangle|\varphi\rangle\left|\mathrm{garbage}_\varphi\right\rangle$. Note that $|\langle f_\psi|f_\varphi\rangle| \leq c^2$. So setting $d := c^2$, Theorem 20 tells us that the counterfeiter must make

$$\Omega\left(\left(c-c^2\right)\sqrt{\frac{2^n-1}{1-c^2}}\right)$$

queries to U_ψ. Fixing (say) $c = 1/2$, this is $\Omega\left(2^{n/2}\right)$. ■

Like Theorem 22, Theorem 43 is easily seen to be tight, since one can use the amplitude amplification algorithm (Lemma 7) to find $|\psi\rangle$, and thereby prepare $|\psi\rangle^{\otimes 2}$, using $O\left(2^{n/2}\right)$ queries to U_ψ.

For completeness, we observe the following generalization of Theorem 43.

Theorem 44. *Given k copies of $|\psi\rangle$, as well as oracle access to U_ψ, a counterfeiter needs $\Omega\left(2^{n/2}/\sqrt{k}\right)$ queries to prepare $|\psi\rangle^{\otimes k+1}$ with certainty (for a worst-case $|\psi\rangle$).*

Proof. If the counterfeiter succeeds, it must map $|\psi\rangle^{\otimes k}$ to some state $|f_\psi\rangle := |\psi\rangle^{\otimes k+1}\left|\text{garbage}_\psi\right\rangle$, and $|\varphi\rangle^{\otimes k}$ to $|f_\varphi\rangle := |\varphi\rangle^{\otimes k+1}\left|\text{garbage}_\varphi\right\rangle$. Note that $|\langle f_\psi|f_\varphi\rangle| \leq c^{k+1}$. So setting $d := c^{k+1}$, Theorem 20 tells us that the counterfeiter must make

$$\Omega\left(\left(c^k - c^{k+1}\right)\sqrt{\frac{2^n-1}{1-c^2}}\right)$$

queries to U_ψ. Fixing $c := 1 - \frac{1}{k}$, the above is

$$\Omega\left(\left(\frac{1}{e} - \left(1-\frac{1}{k}\right)\frac{1}{e}\right)\sqrt{\frac{2^n}{1-(1-1/k)^2}}\right) = \Omega\left(\frac{2^{n/2}}{\sqrt{k}}\right).$$

■

We end this appendix by stating, without proof, three stronger lower bounds that are the direct analogues of Corollary 23, Corollary 24, and Theorem 25 respectively.

Corollary 45. *Given one copy of $|\psi\rangle$, as well as oracle access to U_ψ, a counterfeiter needs $\Omega\left(2^{n/2}\right)$ queries to prepare a state ρ such that $\langle\psi|^{\otimes 2}\rho|\psi\rangle^{\otimes 2} \geq 0.9999$ (for a worst-case $|\psi\rangle$).*

Corollary 46. *Let $1/\varepsilon = o\left(2^n\right)$. Given one copy of $|\psi\rangle$, as well as oracle access to U_ψ, a counterfeiter needs $\Omega\left(\sqrt{\varepsilon}2^{n/2}\right)$ queries to prepare a state ρ such that $\langle\psi|^{\otimes 2}\rho|\psi\rangle^{\otimes 2} \geq \varepsilon$ (for a worst-case $|\psi\rangle$).*

Theorem 47. *Let $|\psi\rangle$ be an n-qubit pure state chosen uniformly from the Haar measure. Given one copy of $|\psi\rangle$, as well as oracle access to U_ψ, a counterfeiter C needs $\Omega\left(\sqrt{\varepsilon}2^{n/2}\right)$ queries to prepare a $2n$-qubit state ρ that a projector $V_\psi^{\otimes 2}$ onto $|\psi\rangle^{\otimes 2}$ accepts with probability at least ε, for all $1/\varepsilon = o\left(2^n\right)$. Here the probability is taken over the choice of $|\psi\rangle$, as well as the behavior of C and $V_\psi^{\otimes 2}$.*