Security Proofs for Private-Key Quantum Money: *From Optimal Counterfeiting Bounds to Adaptive Attack Vulnerabilities*

LU Haodong, SUN Leyuan, ZHAO Jiachen & ZHAO Xinpeng

Department of Computer Science & Engineering

The Chinese University of Hong Kong

{hdlu24, lysun0, jczhao23, xpzhao24}@cse.cuhk.edu.hk

ABSTRACT

Wiesner's quantum money, proposed in the late 1960s and published in 1983, represents one of the earliest and most conceptually elegant applications of quantum mechanics to cryptography. Its security hinges on the no-cloning theorem. However, practical unforgeability critically depends on attack scenarios and bank protocols. This review synthesizes findings from two pivotal analyses. First, we discuss the work on optimal non-adaptive counterfeiting bounds (**?**). Given a single genuine banknote, this research establishes that the maximum success probability for a counterfeiter to create two independently verifiable copies is $(3/4)^n$ for Wiesner's original scheme, and $(3/4+\sqrt{2}/8)^n$ for a classical-verification variant, highlighting inherent exponential security against such "simple" attacks. We also note their finding that schemes based on d-dimensional quantum systems face a counterfeiting success probability of at least $2/(d+1)$.

In contrast, we examine research on adaptive attacks within a "strict testing" regime where valid money is returned post-verification (**?**). This work demonstrates that such an operational assumption allows attackers, using strategies analogous to quantum Zeno effect-based measurements (like bomb-testing or protective measurements), to progressively learn the secret quantum state, ultimately compromising the scheme. By juxtaposing these works, this review highlights the critical interplay between quantum-mechanical security foundations, adversarial capabilities, and banking protocol details in determining the true robustness of quantum money.

1 INTRODUCTION

The quest for perfectly unforgeable currency has been a long-standing challenge. Stephen Wiesner's seminal idea of quantum money **?**, conceived around 1970, offered a revolutionary and elegant paradigm by leveraging quantum mechanics (**?**). The core concept involves a bank encoding information in qubits drawn from non-orthogonal bases. Each banknote has a serial number and a secret classical string known to the bank, specifying the preparation basis for each qubit. Verification entails the bank measuring each qubit in its original basis; attempts to measure or perfectly clone an unknown quantum state are fundamentally limited by quantum principles such as the no-cloning theorem (**?**).

While the no-cloning theorem provides a foundational pillar, its direct application does not fully encapsulate the security of a complete quantum money protocol under various operational circumstances and attacker models. This review focuses on two significant contributions that explore these nuances by analyzing Wiesner's quantum money under distinct attack scenarios.

The first major work we consider, by **?**, addresses the fundamental question: given a single authentic Wiesner banknote, what is the maximum probability with which a counterfeiter can produce two banknotes that both independently pass the bank's verification test, assuming no further interaction with the bank during the counterfeiting process? They employ semidefinite programming (SDP) to derive tight bounds for this "simple counterfeiting attack." For Wiesner's original scheme, they determined this optimal success probability to be exactly $(3/4)^n$, where n is the number of qubits. They also analyzed generalizations, including schemes with classical communication for verification, finding, for example, an optimal success probability of $(3/4 + \sqrt{2}/8)^n$ for their classical-verification analogue. Furthermore, **?** showed that for schemes generalized to d-dimensional qudits, any such money scheme is subject to a simple counterfeiting attack with success probability at least $2/(d+1)$, and they described a scheme for which this bound is optimal. These results establish a baseline security for Wiesner's scheme against non-adaptive attackers. As an example of a non-Wiesner scheme, they also analyzed a 6-state variant proposed by Pastawski et al. obtaining a tight bound of $(2/3)^n$.

Contrasting with this non-adaptive model, the second paper, by **?**, investigates the security of Wiesner's scheme under an adaptive attack model within what they term a "strict testing" regime. In this scenario, a crucial operational assumption is made: if a banknote is deemed valid by the bank, it is returned to the user; if invalid, it is destroyed. **?** demonstrate that this policy critically undermines the scheme's security. They propose two main adaptive attack strategies:

1. **The Bomb-Testing (BT) Attack:** Inspired by the Elitzur-Vaidman bomb tester, this attack uses an ancillary probe qubit that is weakly interacted with a qubit of the money state. The money is then sent for bank validation. If the money qubit is in a state that is an eigenstate of the interaction (e.g., $|+\rangle$ for a CNOT controlled by the probe, where the money qubit is the target), the probe rotates, indicating this state. If the money qubit is in a state that would be "flipped" by the interaction (e.g., $|0\rangle$ or $|1\rangle$), the interaction causes a small disturbance. Due to the quantum Zeno effect, if the bank's validation (a projective measurement onto the correct state) is successful, the money state is likely projected back to its original state, and the probe qubit remains largely unaffected, allowing differentiation. This method can identify the quantum money state by repeating tests for each possible state and qubit. For generalized schemes with n qubits chosen from r possible states, and a minimum angular separation $\theta_{\min}$ between states, this attack can succeed with $O(nr^2\theta_{\min}^{-2}f^{-1})$ validations with failure probability f.

2. **The Protective Measurement (PM) Attack:** This alternative tomographic attack also relies on weak coupling between a probe and the system, followed by bank validation. The goal is to estimate the expectation value $\langle A\rangle$ of an observable A on the money state $|\psi\rangle$ without significantly disturbing $|\psi\rangle$. By repeating the weak coupling and validation N times, the probe's state evolves to approximately $e^{-iN\delta\langle A\rangle\sigma_x}|\phi_0\rangle$, allowing $\langle A\rangle$ to be estimated. **?** argue this attack works even when the number of possible states per qubit is infinite (implying $\theta_{\min} \to 0$), a scenario where their BT attack might fail. They provide analysis for the resources required for protective tomography.

Through these adaptive interactions, the attacker can progressively learn the complete state of the banknote with high probability, thus enabling perfect forgery.

This review aims to synthesize the insights from these two papers. By examining their distinct problem settings, methodologies, and conclusions, we can appreciate the multifaceted nature of security in quantum cryptographic protocols.

2 BACKGROUND

The concept of **quantum money** (**?**), first proposed by Stephen Wiesner in the early 1970s and later published in 1983, introduced a radical idea: using quantum states to make physical currency that is provably unforgeable. Wiesner's scheme encodes each bill with a sequence of qubits, each prepared in a randomly chosen basis—either a computational basis $\{|0\rangle, |1\rangle\}$ or a conjugate basis such as $\{|+\rangle, |-\rangle\}$, where $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. These quantum bits are paired with classical metadata, such as a serial number, and their precise states and bases are recorded only at the issuing bank. To verify the authenticity of a bill, the bank performs measurements in the known correct bases and checks that the outcomes match the recorded state. A counterfeiter who attempts to duplicate the bill without knowing the correct measurement bases inevitably disturbs the original quantum states due to the uncertainty principle, rendering their copy invalid.

This intuitive security argument was later formalized by the **no-cloning theorem** (**?**), proven in 1982 by Wootters and Zurek. The theorem shows that no physical process can produce an exact copy of an arbitrary unknown quantum state, due to the linearity of quantum mechanics. Specifically, any attempt to clone a superposition state like $a|0\rangle + b|1\rangle$ results in a final state that deviates from two identical copies, even if cloning succeeds for certain orthogonal states like $|0\rangle$ and $|1\rangle$. This impossibility directly supports the robustness of Wiesner's quantum money scheme, as it ensures that an adversary cannot replicate the embedded qubit states without altering them and being detected during verification. Remarkably, Wiesner anticipated this cryptographic application of quantum mechanics a decade before the no-cloning theorem was rigorously established, making quantum money one of the earliest practical illustrations of how uniquely quantum properties can be harnessed for secure information processing.

3 Optimal Counterfeiting Bounds via Semidefinite Programming

This section delves into the contributions of **?**, who provided some of the first rigorous, quantitative analyses of the security limits of Wiesner's quantum money scheme and its generalizations. Their work moved beyond the qualitative intuition offered by the no-cloning theorem, employing powerful mathematical tools to establish tight bounds on counterfeiting capabilities.

3.1 Motivation and Problem Formulation

Wiesner's quantum money scheme (**?**), while conceptually groundbreaking, existed for decades without a rigorous analysis quantifying its security against optimal adversaries, particularly concerning explicit success probabilities. While the no-cloning principle (**?**) provides the fundamental security intuition, translating this into concrete bounds requires dedicated analysis, much like the security proofs developed for quantum key distribution protocols such as BB84 (**??**). The work by **?** aimed to fill this gap by providing precise mathematical formulations and solutions for counterfeiting attacks against Wiesner's scheme.

Their investigation sought answers to several key questions:

- What is the *optimal* success probability for a counterfeiter attempting a "simple counterfeiting attack" – that is, starting with a single authentic banknote associated with a specific serial number, producing two quantum states that both independently pass the bank's verification procedure for that same serial number?
- Can the security of such schemes be enhanced or modified by generalizing the underlying quantum states? This includes considering different ensembles of single-qubit states beyond Wiesner's original four, or extending the scheme to use higher-dimensional quantum systems (qudits, $d > 2$).
- Is it feasible to design variants of quantum money where the verification process relies solely on *classical* communication between the note holder and the bank, thus avoiding the need to physically send quantum states back? If so, what level of security do these classical-verification schemes offer against counterfeiting attacks?
- Can semidefinite programming (SDP) serve as a unifying mathematical framework to precisely model these counterfeiting scenarios and compute the optimal attack success probabilities?

By addressing these questions, **?** aimed to place the security analysis of Wiesner's scheme on a firm quantitative footing.

3.2 CORE MATHEMATICAL TOOL: SEMIDEFINITE PROGRAMMING (SDP)

The primary mathematical engine driving the analysis in **?** is semidefinite programming (SDP). To understand its application, we first introduce some notation from the paper regarding linear operators on finite-dimensional complex Hilbert spaces. For a Hilbert space $\mathcal{X}$, $L(\mathcal{X})$ denotes the set of linear operators acting on $\mathcal{X}$, $\mathrm{Herm}(\mathcal{X})$ denotes the Hermitian operators, $\mathrm{Pos}(\mathcal{X})$ denotes the positive semidefinite operators, $\mathrm{Pd}(\mathcal{X})$ the positive definite operators, and $\mathrm{D}(\mathcal{X})$ the density operators (states, $\rho \in \mathrm{Pos}(\mathcal{X})$ with $\mathrm{Tr}(\rho) = 1$). The standard inner product between operators $A, B \in L(\mathcal{X})$ is the Hilbert-Schmidt inner product:

$$\langle A, B \rangle = \mathrm{Tr}(A^\dagger B). \tag{1}$$

For any linear map $\Phi : L(\mathcal{X}) \to L(\mathcal{Y})$, there exists a unique adjoint map $\Phi^* : L(\mathcal{Y}) \to L(\mathcal{X})$ satisfying $\langle Y, \Phi(X) \rangle = \langle \Phi^*(Y), X \rangle$ for all $X \in L(\mathcal{X})$ and $Y \in L(\mathcal{Y})$.

SDP is a subfield of convex optimization concerned with optimizing a linear objective function over the cone of positive semidefinite matrices, subject to linear equality constraints. A standard SDP formulation involves a primal and a dual problem. Given a Hermiticity-preserving linear map $\Phi : L(\mathcal{X}) \to L(\mathcal{Y})$ and Hermitian operators $A \in \mathrm{Herm}(\mathcal{X})$, $B \in \mathrm{Herm}(\mathcal{Y})$, the primal problem is typically stated as:

$$\begin{aligned} \sup \quad & \langle A, X \rangle \\ \text{s.t.} \quad & \Phi(X) = B, \\ & X \in \mathrm{Pos}(\mathcal{X}). \end{aligned} \tag{2}$$

Its corresponding dual problem is:

$$\begin{aligned} \inf \quad & \langle B, Y \rangle \\ \text{s.t.} \quad & \Phi^*(Y) \succeq A, \\ & Y \in \mathrm{Herm}(\mathcal{Y}). \end{aligned} \tag{3}$$

Here, $\Phi^*(Y) \succeq A$ means $\Phi^*(Y) - A \in \mathrm{Pos}(\mathcal{X})$.

Let α be the optimal value of the primal problem equation **??** and β be the optimal value of the dual problem equation **??**. Weak duality always holds: $\alpha \leq \beta$. SDPs are particularly powerful because they can often be solved efficiently (**?**), and under mild conditions (such as Slater's condition, which

requires strict feasibility of either the primal or dual problem), strong duality holds ($\alpha = \beta$), and optimal solutions exist and are achievable.

The applicability of SDP to quantum information processing stems fundamentally from the fact that quantum states (density matrices ρ) must be positive semidefinite operators ($\rho \succeq 0$). Furthermore, quantum operations are described by quantum channels, which are linear maps that are completely positive (CP) and trace-preserving (TP). The set of CPTP maps forms a convex set, amenable to optimization techniques like SDP (**?**).

A key mechanism bridging quantum operations and SDP is the **Choi-Jamiolkowski isomorphism** (**??**). This establishes a one-to-one correspondence between linear maps $\Phi : L(\mathcal{H}_X) \to L(\mathcal{H}_Y)$ and linear operators $J(\Phi) \in L(\mathcal{H}_Y \otimes \mathcal{H}_X)$, where $\dim(\mathcal{H}_X) = d$. Fixing an orthonormal basis $\{|1\rangle, \ldots, |d\rangle\}$ for $\mathcal{H}_X$, the Choi operator is defined as:

$$J(\Phi) = \sum_{1 \leq i,j \leq d} \Phi(|i\rangle\langle j|) \otimes |i\rangle\langle j|. \tag{4}$$

This isomorphism translates properties of the map Φ into properties of the operator $J(\Phi)$:

- Φ is completely positive (CP) if and only if $J(\Phi) \in \mathrm{Pos}(\mathcal{H}_Y \otimes \mathcal{H}_X)$, i.e., $J(\Phi) \succeq 0$.

- Φ is trace-preserving (TP) if and only if $\mathrm{Tr}_Y(J(\Phi)) = I_X$, where Tr_Y denotes the partial trace over $\mathcal{H}_Y$ and I_X is the identity operator on $\mathcal{H}_X$.

The action of the channel on a pure state $|\psi\rangle \in \mathcal{H}_X$ can also be related to the Choi operator. For any vector $|\phi\rangle \in \mathcal{H}_Y$, the overlap is given by:

$$\langle\phi|\Phi(|\psi\rangle\langle\psi|)|\phi\rangle = \langle\phi \otimes \psi^T|J(\Phi)|\phi \otimes \psi^T\rangle, \tag{5}$$

where $|\psi^T\rangle$ denotes the vector obtained by taking the complex conjugate of the components of $|\psi\rangle$ in the standard basis $\{|i\rangle\}$.

This isomorphism is crucial for the work of **?**. Optimizing over the set of all physically realizable quantum channels (CPTP maps) Φ that perform a certain task (like counterfeiting) translates directly into optimizing over the set of Choi operators $J(\Phi)$ that satisfy $J(\Phi) \succeq 0$ and $\mathrm{Tr}_Y(J(\Phi)) = I_X$.

These conditions fit precisely into the SDP framework equation **??**-equation **??**, allowing for the computation of optimal strategies and success probabilities for tasks like quantum state cloning and counterfeiting.

3.3 WIESNER'S QUANTUM MONEY SCHEME: QUANTUM VERIFICATION

Wiesner's quantum money scheme [Wie83] stands as a foundational concept in quantum cryptography. It leverages the principles of quantum mechanics to create currency that is, in theory, impossible to counterfeit perfectly.

3.3.1 ORIGINAL SCHEME SETUP AND VERIFICATION

In Wiesner's original proposal, each banknote comprises n quantum bits (qubits). The bank prepares each qubit independently, choosing its state uniformly at random from the set $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$, where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Thus, each of these four states is selected with a probability $p_k = 1/4$. The bank meticulously records this sequence of prepared states (the "key") and associates it with a unique serial number imprinted on the banknote. This record is kept secret by the bank.

Verification of a banknote's authenticity requires it to be returned to the bank. The bank, using the serial number to retrieve the secret key, performs a quantum measurement on each qubit. For a qubit prepared in $|0\rangle$ or $|1\rangle$, the bank measures in the computational basis ($\{|0\rangle, |1\rangle\}$). For a qubit prepared in $|+\rangle$ or $|-\rangle$, the measurement is performed in the Hadamard basis ($\{|+\rangle, |-\rangle\}$). The banknote is deemed authentic if and only if all measurement outcomes correspond to the originally prepared states. The security of this scheme fundamentally relies on the no-cloning theorem [WZ82], which asserts that it is impossible to create an identical copy of an unknown quantum state.

3.3.2 OPTIMAL COUNTERFEITING ATTACKS

A primary concern for any currency scheme is its resilience against counterfeiting. In the context of Wiesner's scheme, a "simple counterfeiting attack" is defined as an attempt by a counterfeiter, who is given a single genuine banknote, to produce two banknotes (associated with the same serial number) that can both independently pass the bank's verification procedure.

Let the original banknote's quantum state be in a register X, and the two alleged copies produced by the counterfeiter be in registers Y and Z. A counterfeiting attempt is described by a quantum channel (a completely positive trace-preserving map) $\Phi : L(X) \to L(Y \otimes Z)$. If the bank originally prepared the state $|\psi_k\rangle$, the probability that both counterfeit notes pass verification is $\langle\psi_k \otimes \psi_k|\Phi(|\psi_k\rangle\langle\psi_k|)|\psi_k \otimes \psi_k\rangle$. Averaging over all possible initial states chosen by the bank (with probabilities p_k), the overall success probability of the counterfeiting attack is:

$$P_{\text{success}} = \sum_{k=1}^{N} p_k \langle \psi_k \otimes \psi_k | \Phi(|\psi_k\rangle\langle\psi_k|) | \psi_k \otimes \psi_k \rangle \quad (1)$$

The paper establishes a crucial result for Wiesner's original scheme:

Theorem 1 (Molina, Vidick, Watrous, 2012)

The optimal simple counterfeiting attack against Wiesner's quantum money scheme has a success probability of exactly $(3/4)^n$, where n is the number of qubits in each banknote.

For a single-qubit banknote ($n = 1$), the optimal success probability is $3/4$. For an n-qubit banknote, this probability decreases exponentially.

3.3.3 Semidefinite Programming (SDP) Formulation

The problem of determining the optimal success probability for a simple counterfeiting attack can be precisely formulated and solved using semidefinite programming (SDP). This mathematical framework is powerful for optimization problems involving quantum states and operations, yielding tight bounds on security.

The optimal success probability (1) can be found by solving the following primal SDP problem:

- **Primal Problem**:

$$\begin{aligned} \sup \quad & \langle Q, X \rangle \\ \text{s.t.} \quad & \text{Tr}_{Y \otimes Z}(X) = I_X \\ & X \in \text{Pos}(Y \otimes Z \otimes X) \end{aligned}$$

Here, X is the Choi-Jamiołkowski representation $J(\Phi)$ of the counterfeiting channel Φ. The operator $Q \in \text{Pos}(Y \otimes Z \otimes X)$ is defined as:

$$Q = \sum_{k=1}^{N} p_k |\psi_k \otimes \psi_k \otimes \overline{\psi_k}\rangle\langle\psi_k \otimes \psi_k \otimes \overline{\psi_k}| \quad (2)$$

where $\overline{\psi_k}$ denotes the complex conjugate of $|\psi_k\rangle$ (with respect to the standard basis in X). The inner product is $\langle A, B \rangle = \text{Tr}(A^* B)$.

The corresponding dual problem is:

- **Dual Problem**:

$$\inf \quad \text{Tr}(Y)$$

$$\text{s.t.} \quad I_{Y\otimes Z} \otimes Y \geq Q$$

$$Y \in \text{Herm}(X)$$

Strong duality holds for this problem, meaning the optimal values of the primal and dual problems are equal.

For Wiesner's original single-qubit scheme, $N = 4$, $p_k = 1/4$, and the states are $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$. The operator Q becomes:

$$Q = \frac{1}{4}(|000\rangle\langle 000| + |111\rangle\langle 111| + |+++\rangle\langle +++| + |---\rangle\langle ---|).$$

The optimal value of this SDP is $3/4$. This can be shown by finding explicit feasible solutions:

- A primal feasible solution $X = J(\Phi)$ achieving $3/4$ is given by the channel below:

  $$\Phi(\rho) = A_0 \rho A_0^* + A_1 \rho A_1^*,$$

  with specific Kraus operators below (mapping to a 4×1 vector space for $Y \otimes Z$):

  $$A_0 = \frac{1}{\sqrt{12}} \begin{bmatrix} 3 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \text{and} \quad A_1 = \frac{1}{\sqrt{12}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 3 \end{bmatrix},$$

- A dual feasible solution achieving $3/4$ is $Y = \frac{3}{8} I_X$.

The analysis further reveals that a counterfeiter gains no advantage by attempting to correlate attacks across multiple qubits of an n-qubit note. The optimal strategy involves attacking each qubit independently. Consequently, if the optimal success probability for a single repetition (one qubit) is α, for n repetitions (an n-qubit note), it is α^n. This confirms the $(3/4)^n$ result for Wiesner's scheme. This SDP formulation provides a rigorous confirmation of Wiesner's original intuition and quantifies the security precisely.

3.4 Generalizations and Optimizations

In this section, we extend our analysis beyond Wiesner's original quantum money scheme, exploring various generalizations and optimized schemes. Specifically, we discuss improvements via optimized single-qubit states, the impact of parallel repetitions, threshold results, and higher-dimensional (qudit-based) quantum money schemes.

3.4.1 SINGLE-QUBIT OPTIMAL SCHEMES

While Wiesner's original scheme offers a certain level of security, researchers have explored whether different ensembles of single-qubit states could provide better protection. Pastawski et al. [PYJ+11] investigated a scheme using a 6-state ensemble. In this scheme, each qubit is prepared in one of the six eigenstates of the Pauli operators ($\sigma_x, \sigma_y, \sigma_z$): $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle, |y+\rangle, |y-\rangle\}$, where $|y\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm i|1\rangle)$, each chosen with probability $1/6$. For this 6-state scheme, the optimal simple counterfeiting success probability per qubit is reduced to $2/3$.

Molina, Vidick, and Watrous (the authors of the reference paper) showed that this $2/3$ bound can also be achieved with a 4-state ensemble, specifically if the four states form a Symmetric Informationally Complete Positive Operator-Valued Measure (SIC-POVM) [RBKSC04]. For such schemes, the operator Q in the SDP (Equation 2) has a specific structure related to the transposition map, and the optimal value is $2/3$. An explicit primal solution (channel) achieving $2/3$ is $\Phi(\rho) = A_0\rho A_0^* + A_1\rho A_1^*$, with $A_0 = \frac{1}{\sqrt{6}}\begin{bmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $A_1 = \frac{1}{\sqrt{6}}\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \end{bmatrix}$. A dual solution is $Y = \frac{1}{3}I_X$. This success probability of $2/3$ is significant because it matches the optimal success probability for the universal 1-to-2 qubit cloner (Bužek-Hillery cloner [BH96]). This cloner succeeds with probability $2/3$ for *any* input qubit state, implying that $2/3$ is a fundamental lower bound on the counterfeiting probability for any single-qubit money scheme of this type. Thus, these 6-state and 4-state SIC-POVM schemes are optimal among single-qubit schemes.

Thus, optimal single-qubit schemes outperform Wiesner's original four-state scheme, providing stronger security per qubit.

3.4.2 PARALLEL REPETITIONS AND THRESHOLD SCHEMES

Quantum money schemes typically use n-qubit banknotes, relying on the idea that security increases exponentially with n. As discussed for Wiesner's scheme, if the optimal counterfeiting probability for a single instance of a scheme (e.g., one qubit) is α, then for an n-fold parallel repetition (an n-qubit note where each qubit is an independent instance of the scheme), the optimal success probability for the counterfeiter to succeed on all n instances is α^n. This multiplicative behavior holds generally for these types of schemes, meaning correlated attacks across qubits offer no advantage to the counterfeiter.

Some schemes might employ a "threshold" verification: the bank declares a banknote valid if at least t out of n qubit measurements are correct (where $t \leq n$). For such schemes, if each repetition

is attacked independently and optimally with success probability α, the probability of passing the threshold test would be $\sum_{j=t}^{n} \binom{n}{j} \alpha^j (1-\alpha)^{n-j}$. The paper shows that this binomial probability is indeed the optimal counterfeiting probability for the threshold scheme, provided two conditions hold:

1. The ensemble average state is maximally mixed: $\sum_{k=1}^{N} p_k |\psi_k\rangle\langle\psi_k| = \frac{1}{d} I_X$, where d is the dimension of the Hilbert space for a single repetition (e.g., $d = 2$ for qubits). (Equation 5 in the paper)

2. $Y = \frac{\alpha}{d} I_X$ is an optimal dual solution to the single-repetition SDP.

These conditions are met by Wiesner's original scheme and the other specific schemes discussed. This implies that even for threshold verification, independent attacks on each component are optimal.

3.4.3 Higher-Dimensional (Qudit) Schemes

An alternative way to potentially enhance security is to use quantum systems of dimension $d > 2$, known as qudits. For a d-dimensional quantum system, Werner [Wer98] showed that a universal quantum cloner can produce two copies from one original with a success probability of $2/(d+1)$ for any input state. This sets a fundamental limit: any quantum money scheme based on d-dimensional states is vulnerable to a counterfeiting attack with success probability at least $2/(d+1)$.

The paper demonstrates the existence of qudit-based quantum money schemes that achieve this optimal security bound.

Proposition 4 (Molina, Vidick, Watrous, 2012)

Let $\mathcal{E} = \{p_k, |\psi_k\rangle\}$ be an ensemble of d-dimensional states. If the operator Q (defined in Equation 2) is given by:

$$Q = \frac{1}{\text{rank}(\Pi)} (I_{L(C^d)} \otimes I_{L(C^d)} \otimes T)(\Pi)$$

where T is the transposition mapping (with respect to the standard basis of C^d) and Π is the orthogonal projector onto the symmetric subspace of $C^d \otimes C^d \otimes C^d$, then no simple counterfeiting strategy can succeed against the money scheme derived from $\mathcal{E}$ with probability more than $2/(d+1)$.

Ensembles $\mathcal{E}$ derived from quantum 3-designs satisfy the condition on Q in Equation (3) and thus lead to optimal d-dimensional money schemes. For $d = 2$ (qubits), this optimal probability is $2/(2 + 1) = 2/3$, consistent with the optimal single-qubit schemes. For $d = 3$ (qutrits), the probability drops to $2/(3+1) = 1/2$. As d increases, the security $2/(d+1)$ improves significantly, highlighting the potential of higher-dimensional systems.

3.5 Quantum Money with Classical Verification

A practical drawback of Wiesner's original scheme is the requirement for quantum communication with the bank for verification. This involves physically sending the quantum banknote to the bank, which can be cumbersome and technologically demanding.

3.5.1 Motivation and Scheme Description

To address this, variants of quantum money have been proposed that only require classical communication for verification. In such a "quantum ticket" scheme, the physical quantum state (the ticket) remains with the user during verification. The process is typically as follows:

1. The bank prepares a quantum ticket consisting of n qudits. Each qudit i is prepared in a state $|\psi_{k_i}\rangle$ based on a secret key k_i known only to the bank. The ticket has a unique serial number.

2. To verify the ticket, the user (holder) provides the serial number to the bank.

3. The bank sends the user a randomly chosen classical "challenge" $c \in \mathcal{C}$ for each qudit (or a single challenge c for the entire ticket). $\mathcal{C}$ is a fixed finite set of possible challenges.

4. The user performs a measurement $\Pi_c = \{\Pi_a^c\}_{a \in A}$ on their ticket, where the choice of measurement basis (or POVM) depends on the challenge c. The user then reports the classical outcome a to the bank.

5. The bank, knowing the original secret key k associated with the serial number and the challenge c it sent, checks if the reported outcome a is valid. The ticket is accepted if the triple (a, c, k) falls into a predefined, publicly known set S of valid triples. An honest user, knowing the state preparation, would always pass if their ticket is genuine.

A key aspect is that verification may alter or destroy the quantum state of the ticket.

3.5.2 SDP Formulation for Classical Verification

A counterfeiter in this scenario, given one genuine ticket, attempts to successfully answer two independent challenges from the bank for that same ticket. The counterfeiter's strategy can be modeled by a collection of POVMs $\{A_{a_1 a_2}^{c_1 c_2}\}_{a_1 a_2}$ for each pair of challenges (c_1, c_2). The success probability, averaged over the bank's choice of initial states $|\psi_k\rangle$ (with probability p_k) and its choice of two independent challenges c_1, c_2 (chosen uniformly from $\mathcal{C}$), is:

$$P_{\text{success}} = \sum_{k=1}^{N} p_k \frac{1}{|\mathcal{C}|^2} \sum_{c_1, c_2 \in \mathcal{C}} \sum_{\substack{(a_1,a_2):\\ (a_1,c_1,k)\in S\\ (a_2,c_2,k)\in S}} \langle \psi_k | A_{a_1 a_2}^{c_1 c_2} | \psi_k \rangle \quad (4)$$

Maximizing this probability can also be cast as an SDP. The corresponding operator Q_{cv} for this SDP is:

$$Q_{cv} = \sum_{k=1}^{N} p_k \frac{1}{|\mathcal{C}|^2} \sum_{c_1, c_2} \sum_{\substack{(a_1,a_2):\\ (a_1,c_1,k)\in S\\ (a_2,c_2,k)\in S}} |a_1\rangle|a_2\rangle|c_1, c_2, \psi_k\rangle\langle a_1|\langle a_2|\langle c_1, c_2, \psi_k|$$

This operator Q_{cv} acts on a space that includes classical registers for the outcomes a_1, a_2 and challenges c_1, c_2, in addition to the quantum state $|\psi_k\rangle$.

3.5.3 Optimality and Performance Analysis

The paper provides tight bounds for such classical verification schemes. For an n-qubit scheme analogous to Wiesner's (where qubits are prepared in one of two bases, e.g., computational or Hadamard, and the challenge selects which basis to measure), the optimal success probability for a counterfeiter to pass two independent verifications is given by: **Theorem 2 (Molina, Vidick, Watrous, 2012):** For the classical-verification analogue of Wiesner's quantum money scheme, the optimal simple counterfeiting attack has success probability exactly $(3/4 + \sqrt{2}/8)^n$.

This value is $((6 + \sqrt{2})/8)^n \approx (0.9267)^n$, which is higher than the $(3/4)^n = (0.75)^n$ for the quantum verification scheme, indicating that classical verification is inherently less secure for the same underlying state ensemble, though still offering exponential security.

The analysis can be extended to d-dimensional qudits. Consider a scheme where each qudit is prepared using one of two bases, $B_0 = \{|e_s^0\rangle\}$ and $B_1 = \{|e_t^1\rangle\}$. The challenge $c \in \{0, 1\}$ dictates which basis is measured.

Lemma 5 (Molina, Vidick, Watrous, 2012)

For such an n-qudit classical-verification scheme, the success probability of any simple counterfeiting attack is at most $(3/4 + \sqrt{c_{overlap}}/4)^n$, where $c_{overlap} = \max_{s,t} |\langle e_s^0 | e_t^1 \rangle|^2$ is the maximum squared overlap between basis vectors from the two different bases. If $d = 2$ (qubits), this bound is achievable. For qubits using the computational and Hadamard bases, $c_{overlap} = (1/\sqrt{2})^2 = 1/2$, leading to $(3/4 + \sqrt{1/2}/4)^n = (3/4 + 1/(4\sqrt{2}))^n = (3/4 + \sqrt{2}/8)^n$, matching Theorem 2.

The paper further shows a matching lower bound for a specific d-dimensional scheme. This scheme uses the computational basis $\{|t\rangle\}$ and its Quantum Fourier Transform (QFT) basis $\{F|t\rangle\}$. For these bases, the overlap $c_{overlap} = 1/d$ for all pairs of basis vectors.

Lemma 6 (Molina, Vidick, Watrous, 2012)

There is a cloner for the n-qudit ticket scheme (using computational and QFT bases) which successfully answers both challenges with probability $(3/4 + 1/(4\sqrt{d}))^n$.

This result demonstrates that for these specific bases, the success probability is $(3/4 + 1/(4\sqrt{d}))^n$. As d increases, $1/(4\sqrt{d})$ decreases, and the security improves. For large d, the success probability approaches $(3/4)^n$.

These findings underscore that quantum money with classical verification can offer robust, exponentially decaying counterfeiting probabilities, making them a more practical alternative to schemes requiring quantum communication, despite a slight reduction in security compared to their quantum-verified counterparts.

4 AN ADAPTIVE ATTACK ON WIESNER'S QUANTUM MONEY

4.1 MOTIVATION

Quantum money, first introduced by Wiesner in the 1970s (**?**), stands as a foundational proposal for leveraging quantum mechanics to achieve unforgeable currency. At the heart of its security lies the *no-cloning theorem* (**?**), which prohibits perfect duplication of arbitrary quantum states. In Wiesner's scheme, each quantum banknote is associated with a *serial number* and a corresponding *quantum state*—typically a tensor product of single-qubit states, each randomly chosen from the set $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$. The pair $(s, |\$s\rangle)$, comprising the serial number and quantum state, constitutes valid money (**?**). Verification of a banknote is performed by measuring each qubit in its original preparation basis, information known only to the issuing bank.

A critical yet often overlooked aspect of the protocol is the *post-verification behavior*—what happens to a banknote after it passes or fails the bank's validity test. One primary model arises where successful validation leads to the *return of the same banknote*, while failed attempts result in the state being *destroyed*. This regime, termed *strict testing* (**?**), aligns with intuitive expectations: valid money remains in circulation; forgeries are confiscated. However, an important security question emerges: *Does this model inadvertently enable an attacker to extract information without being caught?*

The paper (**?**) investigates the vulnerabilities of Wiesner's scheme under *strict testing*, revealing that even without access to the post-measurement state of failed tests, a counterfeiter can still compromise the scheme. The authors demonstrate that under strict testing, the quantum Zeno effect can be exploited to perform *non-destructive, weak interactions* with the quantum money state. Such interactions enable the attacker to infer the identity of each qubit state while minimizing the risk of triggering a failed test, thus learning the full money state with high confidence and negligible disturbance. Once learned, the attacker can efficiently create counterfeit copies of the original state $|\$s\rangle$, undermining the presumed security of the scheme.

4.2 PROBLEM FORMULATION

In the paper (**?**), there are two different post-validation policies for analyzing the *security of Wiesner's quantum money*.

- **Strict Testing Regime:** The bank returns the same valid quantum banknote after a successful test, and destroys it after a failed one.
- **Token Replacement Regime:** The bank issues a *fresh new quantum banknote with a new serial number* after each successful validation, akin to a *single-use token*. The bank destroys the invalid quantum banknote after a failed test.

The adaptive attack presented in the paper (**?**) is effective *only under the strict testing regime*. The authors introduce two main techniques:

- **Zeno-assisted Elitzur-Vaidman Bomb Test**, which allows the identification of individual qubit states with low risk of detection.
- **Protective Measurement (PM) Attack**, which enables a form of tomography on the unknown state using weak measurements and repeated validation. This attack remains effective even when the number of possible states per qubit is infinite (implying $\theta_{\min} \to 0$), a setting in which the bomb-testing attack fails.

4.3 Attack Procedure

In this section, we describe the attack procedure of both proposed techniques in (**?**).

4.3.1 The BT(Bomb-testing) Attack

The bomb-testing attack, based on the work of (**?**), serves as a straightforward analogy for understanding quantum money. In this scenario, quantum money is likened to the bomb, while the bank's verification process parallels the bomb-checking procedure. If the money is determined to be counterfeit, it is reported, akin to the bomb's explosion leading to dire consequences. Our goal is to ascertain the state of the money without altering it. However, unlike the bomb-testing scenario, quantum money operates with four possible states per qubit, hence modification of the algorithm is required. The modified algorithm is outlined below.

To understand the BT attack for quantum money, we first introduce the Elitzur-Vaidman's bomb quality tester. Here we pick a large N, then we define:

$$\delta = \frac{\pi}{2N}, \qquad R_\delta = \begin{bmatrix} \cos\delta & -\sin\delta \\ \sin\delta & \cos\delta \end{bmatrix} \tag{6}$$

We also define a controlled interaction between probe and system qubits:

$$C_P = 00 \otimes +11 \otimes P. \tag{7}$$

4.3.2 The PM(Protective Measurement) Attack

4.4 Mathematical Analysis

In this section, we provide a thorough analysis of how these two attack techniques can effectively counterfeit quantum money.

4.4.1 The BT(Bomb-testing) Attack

4.4.2 The PM(Protective Measurement) Attack

4.5 Security Discussion

While both of these attack techniques can counterfeit quantum money, they are contingent upon the strict testing variant of Wiesner's scheme. A common mitigation strategy involves issuing a new quantum bill with a unique serial number to the owner after a valid test. In that case, the quantum

money scheme remains secure. Consequently, these techniques serve as a cautionary reminder about the risks associated with reusing quantum money.

5 CONCLUSION

Wiesner's quantum money scheme, built upon the no-cloning principle, marked a pioneering step towards unconditionally secure cryptographic primitives. The initial promise lay in the inherent difficulty of duplicating unknown quantum states. However, as the analyses by **?** and **?** demonstrate, the journey from a fundamental quantum principle to a robust, practical security protocol is complex.

The work of **?** provided a crucial quantitative understanding of the scheme's baseline security against non-adaptive "simple counterfeiting" attacks. By establishing optimal success probabilities, such as $(3/4)^n$ for Wiesner's original proposal and $2/(d+1)$ as a lower bound for d-dimensional systems, they confirmed the exponential difficulty an attacker faces when attempting to forge notes without intermediate bank interaction. This underscores the inherent strength from the quantum nature of banknotes under a restricted attacker model.

However, **?** offered a starkly different perspective by altering assumptions about the bank's operational procedures. Their investigation into a "strict testing" regime, where valid money is returned, revealed critical vulnerabilities. The adaptive attacks they proposed, leveraging concepts like bomb-testing and protective measurements, showed that an attacker could interactively query the bank's verification system to progressively learn the secret quantum state, ultimately leading to a complete break of unforgeability. This powerfully illustrates that even protocols secured by fundamental quantum laws can be compromised if the broader system allows for exploitable feedback.

In synthesizing these two contributions, it becomes clear that the security of Wiesner's quantum money cannot be assessed in isolation from its operational context and adversarial capabilities. While the no-cloning theorem provides a strong start, practical unforgeability necessitates meticulous protocol design that anticipates a wide range of attack vectors. Specifically for Wiesner's scheme, the implication is that secure implementation must preclude the return of the exact same quantum state after validation; validated notes should arguably be replaced.

The insights from **?** and **?** highlight the ongoing evolution in understanding quantum security. The contrast between the $(3/4)^n$ security against simple attacks and the complete vulnerability under adaptive scenarios underscores the paramount importance of the attacker model and protocol specification in cryptographic security analysis. Future research must continue to explore sophisticated

attacker models and the nuances of practical implementation to realize the full potential of quantum cryptography.

6 SUBMISSION OF CONFERENCE PAPERS TO ICLR 2025

ICLR requires electronic submissions, processed by `https://openreview.net/`. See ICLR's website for more instructions.

If your paper is ultimately accepted, the statement `\iclrfinalcopy` should be inserted to adjust the format to the camera ready requirements.

The format for the submissions is a variant of the NeurIPS format. Please read carefully the instructions below, and follow them faithfully.

6.1 STYLE

6.2 RETRIEVAL OF STYLE FILES

The style files for ICLR and other conference information are available online at:

`http://www.iclr.cc/`

The file `iclr2025_conference.pdf` contains these instructions and illustrates the various formatting requirements your ICLR paper must satisfy. Submissions must be made using LaTeX and the style files `iclr2025_conference.sty` and `iclr2025_conference.bst` (to be used with LaTeX2e). The file `iclr2025_conference.tex` may be used as a "shell" for writing your paper. All you have to do is replace the author, title, abstract, and text of the paper with your own.

The formatting instructions contained in these style files are summarized in sections **??**, **??**, and **??** below.

6.3 FOOTNOTES

Indicate footnotes with a number[1] in the text. Place the footnotes at the bottom of the page on which they appear. Precede the footnote with a horizontal rule of 2 inches (12 picas).[2]

[1] Sample of the first footnote
[2] Sample of the second footnote

6.4 FIGURES

All artwork must be neat, clean, and legible. Lines should be dark enough for purposes of reproduction; art work should not be hand-drawn. The figure number and caption always appear after the figure. Place one line space before the figure caption, and one line space after the figure. The figure caption is lower case (except for first word and proper nouns); figures are numbered consecutively.

Make sure the figure caption does not get separated from the figure. Leave sufficient space to avoid splitting the figure and figure caption.

You may use color figures. However, it is best for the figure captions and the paper body to make sense if the paper is printed either in black/white or in color.

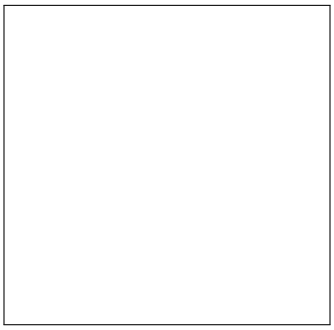

Figure 1: Sample figure caption.

6.5 TABLES

All tables must be centered, neat, clean and legible. Do not use hand-drawn tables. The table number and title always appear before the table. See Table **??**.

Place one line space before the table title, one line space after the table title, and one line space after the table. The table title must be lower case (except for first word and proper nouns); tables are numbered consecutively.

7 DEFAULT NOTATION

In an attempt to encourage standardized notation, we have included the notation file from the textbook, *Deep Learning* **?** available at `https://github.com/goodfeli/dlbook_notation/`. Use of this style is not required and can be disabled by commenting out `math_commands.tex`.

Table 1: Sample table title

PART	DESCRIPTION
Dendrite	Input terminal
Axon	Output terminal
Soma	Cell body (contains cell nucleus)

Numbers and Arrays

a	A scalar (integer or real)
$\boldsymbol{a}$	A vector
$\boldsymbol{A}$	A matrix
$\mathsf{\mathbf{A}}$	A tensor
$\boldsymbol{I}_n$	Identity matrix with n rows and n columns
$\boldsymbol{I}$	Identity matrix with dimensionality implied by context
$\boldsymbol{e}^{(i)}$	Standard basis vector $[0, \dots, 0, 1, 0, \dots, 0]$ with a 1 at position i
$\text{diag}(\boldsymbol{a})$	A square, diagonal matrix with diagonal entries given by $\boldsymbol{a}$
a	A scalar random variable
$\mathbf{a}$	A vector-valued random variable
$\mathbf{A}$	A matrix-valued random variable

Sets and Graphs

$\mathbb{A}$	A set
$\mathbb{R}$	The set of real numbers
$\{0, 1\}$	The set containing 0 and 1
$\{0, 1, \dots, n\}$	The set of all integers between 0 and n
$[a, b]$	The real interval including a and b
$(a, b]$	The real interval excluding a but including b
$\mathbb{A} \backslash \mathbb{B}$	Set subtraction, i.e., the set containing the elements of $\mathbb{A}$ that are not in $\mathbb{B}$
$\mathcal{G}$	A graph
$Pa_{\mathcal{G}}(\mathrm{x}_i)$	The parents of x_i in $\mathcal{G}$

Indexing

a_i	Element i of vector $\boldsymbol{a}$, with indexing starting at 1
a_{-i}	All elements of vector $\boldsymbol{a}$ except for element i
$A_{i,j}$	Element i, j of matrix $\boldsymbol{A}$
$\boldsymbol{A}_{i,:}$	Row i of matrix $\boldsymbol{A}$
$\boldsymbol{A}_{:,i}$	Column i of matrix $\boldsymbol{A}$
$\mathsf{A}_{i,j,k}$	Element (i, j, k) of a 3-D tensor A
$\mathsf{A}_{:,:,i}$	2-D slice of a 3-D tensor
a_i	Element i of the random vector $\mathbf{a}$

Calculus

$\frac{dy}{dx}$	Derivative of y with respect to x
$\frac{\partial y}{\partial x}$	Partial derivative of y with respect to x
$\nabla_{\boldsymbol{x}} y$	Gradient of y with respect to $\boldsymbol{x}$
$\nabla_{\boldsymbol{X}} y$	Matrix derivatives of y with respect to $\boldsymbol{X}$
$\nabla_{\mathbf{X}} y$	Tensor containing derivatives of y with respect to $\mathbf{X}$
$\frac{\partial f}{\partial \boldsymbol{x}}$	Jacobian matrix $\boldsymbol{J} \in \mathbb{R}^{m \times n}$ of $f : \mathbb{R}^n \to \mathbb{R}^m$
$\nabla_{\boldsymbol{x}}^2 f(\boldsymbol{x})$ or $\boldsymbol{H}(f)(\boldsymbol{x})$	The Hessian matrix of f at input point $\boldsymbol{x}$
$\int f(\boldsymbol{x}) d\boldsymbol{x}$	Definite integral over the entire domain of $\boldsymbol{x}$
$\int_{\mathbb{S}} f(\boldsymbol{x}) d\boldsymbol{x}$	Definite integral with respect to $\boldsymbol{x}$ over the set $\mathbb{S}$

Probability and Information Theory

$P(\mathrm{a})$	A probability distribution over a discrete variable
$p(\mathrm{a})$	A probability distribution over a continuous variable, or over a variable whose type has not been specified
$\mathrm{a} \sim P$	Random variable a has distribution P
$\mathbb{E}_{\mathrm{x} \sim P}[f(x)]$ or $\mathbb{E} f(x)$	Expectation of $f(x)$ with respect to $P(\mathrm{x})$
$\mathrm{Var}(f(x))$	Variance of $f(x)$ under $P(\mathrm{x})$
$\mathrm{Cov}(f(x), g(x))$	Covariance of $f(x)$ and $g(x)$ under $P(\mathrm{x})$
$H(\mathrm{x})$	Shannon entropy of the random variable x
$D_{\mathrm{KL}}(P \Vert Q)$	Kullback-Leibler divergence of P and Q
$\mathcal{N}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$	Gaussian distribution over $\boldsymbol{x}$ with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$

Functions

$f : \mathbb{A} \to \mathbb{B}$	The function f with domain $\mathbb{A}$ and range $\mathbb{B}$
$f \circ g$	Composition of the functions f and g
$f(\boldsymbol{x}; \boldsymbol{\theta})$	A function of $\boldsymbol{x}$ parametrized by $\boldsymbol{\theta}$. (Sometimes we write $f(\boldsymbol{x})$ and omit the argument $\boldsymbol{\theta}$ to lighten notation)
$\log x$	Natural logarithm of x
$\sigma(x)$	Logistic sigmoid, $\dfrac{1}{1+\exp(-x)}$
$\zeta(x)$	Softplus, $\log(1+\exp(x))$
$\|\|\boldsymbol{x}\|\|_p$	L^p norm of $\boldsymbol{x}$
$\|\|\boldsymbol{x}\|\|$	L^2 norm of $\boldsymbol{x}$
x^+	Positive part of x, i.e., $\max(0, x)$
$\mathbf{1}_{\text{condition}}$	is 1 if the condition is true, 0 otherwise

A APPENDIX

You may include other additional sections here.