Contents

Chapter 2

Quantum money

In this chapter we put our freshly acquired formalism of qubits and measurements to good use by discussing a first cryptographic application: quantum money! The idea for quantum money was discovered in the first ever paper written on quantum information, by Stephen Wiesner in 1983. Wiesner's key observation was that the possibility to encode information in different bases, such as the standard basis and the Hadamard basis, provides a natural mechanism for copy-protection. In this chapter we explain Wiesner's idea and use it as an opportunity to deepen our understanding of quantum states and measurements.

So what is money? Generally, a bill has two components. First, there is a physical object, such as a piece of paper or metal. Second, there is often some form of identifier associated with the physical object, such as a serial number. The serial number is created on the day that the bill is minted, and it is used as a means to specify all kinds of information about the bill, such as its value, its provenance, the date on which it was minted, etc. This information is kept by the bank as a means to keep track of all valid money in circulation.

The main security guarantee that one wants of money is that it cannot be duplicated. This is what the "paper" part of the bill is meant for: if the bill only consisted of a serial number, this number could be easily copied and the amount of real currency associated to it spent twice. A piece of paper is technically a little harder to duplicate than a mere number... but not impossible!

Remember the *no-cloning principle* from Chapter **??**. Informally, this principle states that there is no quantum operation that can perfectly copy an arbitrary qubit. In other words, qubits cannot be duplicated. You can see where this is going, right? Let's first explore a very simple (but flawed) idea for a quantum money scheme.

2.1 A (too) simple quantum money scheme

Let's give a preliminary definition of a quantum bill as a quantum state $|\psi_\$\rangle$ that is associated with a serial number $\$$ kept by the bank as an identifier for the state. Think of the serial number as a string of bits that is publicly known and states general things about the bill, such as when it was created, how much money it is worth, etc. Generally, we'll take the serial number $\$ \in \{0,1\}^n$ for some integer n that is sufficiently large so that we never run out of serial numbers, such as $n = 1024$.

Here is the simplest quantum money scheme you might think of. To create a quantum bill, first generate a serial number $\$ \in \{0,1\}^n$ uniformly at random. Then, create an n-qubit quantum state such that the i-th qubit is initialized as a standard basis state equal to the i-th bit of $\$$. In other words, create the quantum state $|\psi_\$\rangle = |\$\rangle$. The quantum bill is the pair $(\$, |\psi_\$\rangle)$ of the serial number and the state associated to it. Since qubits (and a fortiori n-qubits) cannot be cloned, the scheme is secure, right?

Of course not! This scheme has no secret information. Given a pair of a bill and serial number, $(|\psi_\$\rangle, \$)$, it is very easy to create an unlimited number of identical copies of it by simply using the serial number to prepare the state $|\psi_\$\rangle$. This does not violate the no-cloning principle, because we are given a classical description of the state: it is the standard basis state associated with the n-bit string $\$$. Given this classical description, and a quantum computer, it is straightforward to create as many copies of $|\psi_\$\rangle$ as desired. In fact, even if we didn't have access to the classical serial number, the scheme would be entirely broken, as an attacker could first measure $|\psi_\$\rangle$ in the standard basis to obtain $\$$, and then re-create as many copies of it as desired.

In case you're not sure why the no-cloning principle does not apply, remember that the impossible task is to design a quantum machine that has the ability to clone *every* state. But there still can be machines that clone specific families of states, such as all standard basis states. An interesting money scheme will necessarily involve states that are more complicated than simple standard basis states!

2.2 Wiesner's quantum money

In Wiesner's scheme, two strings, $x_\$, \theta_\$ \in \{0,1\}^n$, where n is an integer that parametrizes the security of the scheme (for example, think of $n = 1024$), are associated to each serial number $\$$. In the following, we drop the subscript $\$$ and simply write x, θ for $x_\$, \theta_\$$. It is important that the pair (x, θ) is kept secret: it is generated at random by the bank on the day when the quantum bill is minted, but

it is never revealed, even to the honest holder of the bill.

For $x, \theta \in \{0,1\}$ we introduce the notation $|x\rangle_\theta = H^\theta |x\rangle$. Then the money state associated with \$ is an n-qubit state $|\psi_\$\rangle$ such that the i-th qubit is in state $|x_i\rangle_{\theta_i}$:

$$|\psi_\$\rangle = |x_1\rangle_{\theta_1} \otimes \cdots \otimes |x_n\rangle_{\theta_n} .$$

Example 2.2.1. *Suppose that $n = 2$, and consider a serial number \$ such that $x_\$ = 01$ and $\theta_\$ = 10$. Then the associated quantum money state is*

$$|\psi_\$\rangle = (H|0\rangle) \otimes (|1\rangle) = |+\rangle \otimes |1\rangle .$$

■

Can you break this scheme? If you go back to the proof of the no-cloning principle in Section [**sec:nocloning**], you will notice that the theorem alrneady applies in case the only states considered are $|0\rangle, |1\rangle, |+\rangle, |-\rangle$.[1] This seems to rule out a perfect cloning machine. However, notice that if you measure each qubit of $|\psi_\$\rangle$ in either the standard or the Hadamard basis, without knowing which is the correct basis, you expect to get the right answer for approximately half the qubits. So, you "learn" half the state in this way. What if you could learn more? What if you could recover 99% of the qubits? Or all the qubits, 99% of the time? Then, would we still want to consider the scheme to be secure, even though perfect cloning is impossible?

To answer this question we have to go through one of the most important exercises in cryptography: introducing a security definition! Until now we have been arguing about security at a very intuitive level; to make progress we need to establish firm foundations to support our investigation.

2.2.1 Definition of quantum money

n

To specify a quantum money scheme, we need to provide annswers to the following questions: How (and by whom) is a quantum bill generated? And what is the procedure for determining the validity of a purported bill? Defining a quantum money scheme requires us to specify the following two procedures, each meant to answer one of these two questions:

- A *state generation procedure* $\mathrm{GEN}(n)$: This is the procedure applied by the bank to mint money. It takes as input an integer n called the "security parameter" (intuitively, the larger n is, the more secure the scheme). The procedure

[1] These states are often referred to as "BB84" states — we will see why in Chapter **??**.

returns a triple $(\$, |\psi_\$\rangle, k_\$)$ of a quantum state $|\psi_\$\rangle$, a classical serial number $\$$, and a classical "private key" $k_\$$ that specifies secret information about the bill that is to be kept by the bank.

- A *bill verification procedure* $\mathrm{VER}(\$, |\psi\rangle, k)$: This is the procedure executed by the bank to verify a quantum state. It takes as input a pair $(|\psi\rangle, \$)$ of a quantum state and a serial number, as well as a key k, and returns either "accept" or "reject".

Note that the state generation procedure GEN does not explicitly specify a denomination for the quantum bills. The simplest implementation of the scheme will associate an identical value to each money state, such as 1€. It is also possible to associate different values to the bills: in this case, we can imagine that a bill's value is specified as a (classical) integer accompanying the state, and is also kept together with the serial number in the bank's records (so that a user cannot arbitrarily change the value of their money state). Here, we stay with the simpler definition of assuming that all money states have the same value.

Note also that our specification of the verification procedure implicitly destroys the money state: VER takes as input $|\psi\rangle$ and $\$$ and only returns "accept" or "reject". In general it may seem desirable that valid money states be returned to the user, so that it is possible to verify a state without being forced to destroy it. This, however, creates security risks that we explore in Section 2.4 below. For the time being, we stick with the definition and consider that verification always entirely destroys the money state (if need be, the bank can always generate a fresh bill to compensate the user).

Let's see how this abstract formalism looks like for the case of Wiesner's scheme:

- The state generation procedure $\mathrm{GEN}_W(1^n)$ first selects a serial number in an arbitrary way (for example, the serial numbers can be chosen sequentially, or they can be sampled at random and contain a time stamp, etc.). Then, it selects two strings $x, \theta \in \{0,1\}^n$ uniformly at random. The bank records the information $k_\$ = (x, \theta)$, and it creates the state

$$|\psi_\$\rangle = \bigotimes_{i=1}^{n} \left(H^{\theta_i} |x_i\rangle \right) .$$

  Finally, GEN_W returns the triple $(\$, |\psi_\$\rangle, k_\$)$.

- The state generation procedure $\mathrm{VER}_W(|\psi\rangle, \$, k)$ proceeds as follows. First, it interprets the key k as a pair of n-bit strings (x, θ). Then, it applies a

Hadamard gate to each of the n qubits of $|\psi\rangle$, for $i = 1, \ldots, n$, such that $\theta_i = 1$. Finally, it measures all qubits in the standard basis to obtain a string $y \in \{0,1\}^n$. It returns "accept" if $y = x$, and "reject" otherwise.

Since all the operations that we described can be implemented as quantum maps, this is a well-defined quantum money scheme. But this doesn't make it an interesting quantum money scheme! Indeed, the basic scheme from the previous section is also "well-defined". To make it interesting, we need to introduce *correctness* and *security* requirements for the scheme.

2.2.2 Security of quantum money

The first property that a quantum money scheme should satisfy is called the *correctness* property: valid money states should always be accepted by the verification procedure. Formally,

$$\forall n \geq 1 \ , \quad \mathrm{VER}\big(\mathrm{GEN}(1^n)\big) = \text{“accept”} \ . \tag{2.1}$$

This seems like an absolute minimum requirement, as otherwise the bank wouldn't accept its own correctly minted bills. Note, however, that it doesn't prevent the verification procedure from accepting *all* states! This would still be a correct money scheme according to our definition. However, intuitively it would be far from secure, since any user could create bills out of anything and still pass verification. To prevent this and other more subtle attacks, we need to introduce a *security condition* for quantum money.

How should we define security? Informally, we would like it to be impossible to "duplicate" a quantum money state: given a valid quantum bill, a user should be able to spend it once (this is guaranteed by (2.1)), but not twice. In cryptography it is often convenient to formalize security through a "game" that expresses precisely the kind of situation that the scheme should prevent from happening. Here, the forbidden situation is that an adversary to the scheme manages to copy a quantum bill. Let's formulate this requirement as the following game, played between an "adversary" and a (trusted) "challenger":

- The challenger executes the procedure $(\$, |\psi_\$\rangle\,, k_\$) \leftarrow \mathrm{GEN}(1^n)$. It keeps $k_\$$ to itself and provides $(\$, |\psi_\$\rangle)$ to the adversary.
- The adversary returns *two* quantum states σ and σ', each of the same number of qubits as $|\psi_\$\rangle$. (It is up to the adversary how these states are obtained.)
- The challenger executes $\mathrm{VER}(\$, \sigma, k_\$)$ and $\mathrm{VER}(\$, \sigma', k_\$)$. It accepts if and only if both verification attempts accept.

We simply call this game CLONE. Note that in the game we do not assume that the adversary returns two pure single-qubit states. The reason is that it is physically impossible to tell if a state is pure or mixed: indeed, a mixed state is nothing but a distribution over pure states, so allowing mixed states is similar to allowing the adversary to apply a randomized strategy to implement its attack. In particular, it is allowed that σ and σ' are each a qubit taken out of a bigger state ρ. The procedure $\mathrm{VER}(\sigma, \$, k_\$)$ is still well-defined: it applies whatever measurement VER would apply on a pure state $|\psi\rangle$ to the mixed state σ.

Definition 2.2.1. *For any $\varepsilon \geq 0$ we say that a quantum money scheme is* ε-secure *if the maximum probability with which any adversary can succeed in the game* CLONE *is at most ε, where the probability is taken over the randomness in* GEN*,* VER*, and any randomness used by the adversary.*

It is important to realize that the above is a *definition*. The definition captures a very specific type of attack, modeled in the security game. If a scheme is shown secure (for a small enough ε, the smaller the better), it exactly means that such attacks are impossible; it means no more and no less. Indeed, we have to be careful with words, and the definition does not capture *all* attacks! For example, if the adversary is allowed to break into the bank and steal the database of secret keys $k_\$$, then all guarantees are off. The way this is captured in the security game is when we write that "[The challenger] keeps $k_\$$ to itself" in the first step. If the challenger (the bank) does *not* "keep $k_\$$ to itself" (e.g. the adversary manages to steal it), then this scenario falls outside of the rules of the game, and the security definition no longer provides any guarantees.

In the remainder of this chapter we study the security of Wiesner's quantum money scheme $(\mathrm{GEN}_W, \mathrm{VER}_W)$. Before trying to prove security, we explore some simple "attacks" on the scheme.

2.3 Cloning attacks

It is always instructive to try to *break* a cryptographic scheme, by designing the best possible attack on it. By trying hard enough, you can have surprises! Many cryptographic proposals since the birth of cryptography could have been broken by their inventor, had they tried hard enough, and this would have avoided some painful realizations!

In general an "attack" on a cryptographic system is a procedure that breaks the security definition. In the case of quantum money, our goal is to define actions for the "adversary" in the game CLONE such that the adversary's success probability is as high as possible.

For simplicity let's first consider the case of the single-qubit version of Wiesner's money. This corresponds to choosing $n = 1$ in CLONE. In this case the adversary is given by the challenger a single-qubit state $|\psi_\$\rangle \in \{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$, together with the serial number. Since, however, the serial number is chosen by GEN independently of anything else, we may as well ignore it. The adversary does not know which of the four options is the case, yet it has to return two single-qubit density matrices σ and σ' (or equivalently, a two-qubit density matrix ρ) such that the probability of both qubits of ρ passing verification is maximized.

Since we know exactly what the verification procedure in Wiesner's scheme does, we can write out explicitly the adversary's maximum success probability as a function of the state ρ that it returns. Let ρ_0, ρ_1, ρ_+, ρ_- be the two-qubit density matrix returned by the adversary on challenge $|0\rangle$, $|1\rangle$, $|+\rangle$, $|-\rangle$ respectively. Then the probability that ρ_0 is accepted is the probability that a measurement of its two qubits in the computational basis yields the outcome $|00\rangle$, which by the Born rule $\langle 00| \rho |00\rangle$. Similarly, the probability that ρ_1 is accepted is $\langle 11| \rho |11\rangle$. For ρ_+ and ρ_-, it is $\langle ++| \rho_+ |++\rangle$ and $\langle --| \rho_- |--\rangle$ respectively. Since in CLONE the quantum bill is chosen uniformly at random among the four possibilities, the adversary's success probability evaluates to the average of these four quantities, i.e.

$$p_{succ} = \frac{1}{4}\big(\langle 0| \langle 0| \rho_0 |0\rangle |0\rangle + \langle 1| \langle 1| \rho_1 |1\rangle |1\rangle + \langle +| \langle +| \rho_+ |+\rangle |+\rangle + \langle -| \langle -| \rho_- |-\rangle |-\rangle \big) . \tag{2.2}$$

Our goal in designing an attack on the scheme is to define a quantum operation that sends single-qubit states $|\psi\rangle \in \{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$ to two-qubit density matrices $\rho_0, \rho_1, \rho_+, \rho_-$ in a way that (2.2) is maximized. Note that it is important here that these density matrices are obtained from $|\psi_\$\rangle$ by a valid quantum operation; otherwise we could just define ρ to be the correct 2-qubit state in all four cases and be done. What prevents us from doing so is the *no-cloning principle*, that we already discussed in Section [**sec:nocloning**]: if you remember, a special case of the principle states precisely that the operation $|\psi\rangle \mapsto |\psi\rangle |\psi\rangle$ for all $\psi \in \{0, 1, +, -\}$ is *not* a valid quantum map (it does not respect linearity).

Of course this impossibility result does not say that we can't *try* to clone, and see how well we manage to do! So let's give it a try.

2.3.1 Measure-and-prepare attacks

A first kind of attack for the adversary is to choose a basis in which to measure $|\psi_\$\rangle$, and then attempt to prepare two copies of it based on the information obtained from the measurement outcome.

As a warm-up, consider the case where the adversary measures in the standard basis. Let's say that if they obtain outcome 0, they return $\rho = |0\rangle\langle 0| \otimes |0\rangle\langle 0|$, and if they get outcome 1, they return $\rho = |1\rangle\langle 1| \otimes |1\rangle\langle 1|$. What is the success probability of this attack? If $\psi_\$ \in \{0, 1\}$ then the state is perfectly cloned. But if $\psi_\$ \in \{+, -\}$ then the outcome from the measurement is a uniformly random bit, so $\rho_+ = \rho_- = \frac{1}{4}\mathbb{I} \otimes \mathbb{I}$. Overall,

$$p_{succ} = \frac{1}{4}\big(1 + 1 + \frac{1}{4} + \frac{1}{4}\big) = \frac{5}{8} \,.$$

Exercise 2.3.1. *Evaluate the success probability of an attack where the attacker measures $|\psi\rangle$ in the Hadamard basis to obtain an outcome $x \in \{+, -\}$, and returns the Hadamard basis state $\rho = |x\rangle\langle x| \otimes |x\rangle\langle x|$.* ∎

What about measurements in other bases? The following exercise guides you through the determination of the optimal attacks among all attacks of a certain form.

Exercise 2.3.2. *Consider all attacks that take the following form. The adversary decides on an arbitrary orthonormal basis $(|u_0\rangle, |u_1\rangle)$ for the single-qubit space $\mathbb{C}^2$. It then measures the challenger's state $|\psi_\$\rangle$ in the basis $(|u_0\rangle, |u_1\rangle)$ to obtain an outcome $b \in \{0, 1\}$. Finally, the adversary returns the density matrix $\rho = |u_b\rangle\langle u_b| \otimes |u_b\rangle\langle u_b|$. Express the success probability of this attack as a function of the coefficients α, β of $|u_0\rangle = \alpha|0\rangle + \beta|1\rangle$. (Since $|u_1\rangle$ is orthogonal to $|u_0\rangle$, without loss of generality $|u_1\rangle = \beta|0\rangle - \alpha|1\rangle$.) Find the choice of α, β that maximizes the success probability (don't forget about complex numbers!). Did you find an attack that is better than the ones consider above?* ∎

We could consider even more general attacks. For example, there is no reason to limit the adversary to a basis measurement: it can also apply a more general POVM, with more than two outcomes. How can we classify such a broad class of attacks? This points you to the difficulty of proving security of a cryptographic scheme in general: there is no limit to the ingenuity of adversaries! In the next section we consider another class of attacks.

2.3.2 Cloning attacks

In general there is no reason to require the adversary to measure their money state. Instead, they could attempt to "clone" the state right away. Since quantum maps are unitary, and the adversary has to transform a single-qubit state into a two-qubit state, this may seem impossible. There is a simple solution however: the adversary may of course use any number of "ancilla" qubits, initialized to an arbitrary state

independent of the challenge $|\psi_\$\rangle$, in their own private workspace. In the simplest case of a single ancilla qubit we are looking for a unitary map U such that the adversary returns the two-qubit state $|\varphi\rangle = U(|\psi_\$\rangle\,|0\rangle)$ as its answer to the verifier's challenge. Starting from (2.2) and using that here $\sigma_x = U\,|x\rangle\,|x\rangle\,\langle x|\,\langle x|\,U^*$ for $x \in \{0,1,+,-\}$ we can rewrite the success probability of such a "simple cloning attack" as

$$p_s = \frac{1}{4}\Big(\big|\,\langle 0|\,\langle 0|\,U\,|0\rangle\,|0\rangle\,\big|^2 + \big|\,\langle 1|\,\langle 1|\,U\,|1\rangle\,|0\rangle\,\big|^2 + \big|\,\langle +|\,\langle +|\,U\,|+\rangle\,|0\rangle\,\big|^2 + \big|\,\langle -|\,\langle -|\,U\,|-\rangle\,|0\rangle\,\big|^2\Big)\,. \tag{2.3}$$

But there are even more general attacks! Why would the adversary limit themselves to preparing a *pure* two-qubit state: in general, they may as well prepare a mixed state. This could arise because the adversary finds it helpful to flip some random bits, and depending on the outcomes, prepare a different state. Or because they apply a unitary map that sends the single-qubit money state to a three-qubit pure state, and then trace out one of the qubits before returning the remaining two as their attempted clone. To see what such an attack could look like, let's consider a specific example.

Example 2.3.1. *Consider the following map T from single-qubit quantum money states to two-qubit density matrices, where we write $|\phi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$:*

$$\begin{aligned}
|0\rangle\langle 0| \mapsto \rho_0 &= \frac{2}{3}\,|00\rangle\,\langle 00| + \frac{1}{3}|\phi^-\rangle\langle\phi^-|\,,\\
|1\rangle\langle 1| \mapsto \rho_1 &= \frac{2}{3}|11\rangle\langle 11| + \frac{1}{3}|\phi^-\rangle\langle\phi^-|\,,\\
|+\rangle\langle +| \mapsto \rho_+ &= \frac{1}{12}\big(2\,|00\rangle + \sqrt{2}\,|\phi^-\rangle\big)\big(2\,\langle 00| + \sqrt{2}\,\langle\phi^-|\big)\\
&\quad + \frac{1}{12}\big(2\,|11\rangle + \sqrt{2}\,|\phi^-\rangle\big)\big(2\,\langle 11| + \sqrt{2}\,\langle\phi^-|\big)\,,\\
|-\rangle\langle -| \mapsto \rho_- &= \frac{1}{12}\big(2\,|00\rangle - \sqrt{2}\,|\phi^-\rangle\big)\big(2\,\langle 00| - \sqrt{2}\,\langle\phi^-|\big)\\
&\quad + \frac{1}{12}\big(2\,|11\rangle - \sqrt{2}\,|\phi^-\rangle\big)\big(2\,\langle 11| - \sqrt{2}\,\langle\phi^-|\big)\,.
\end{aligned}$$

■

As a first step we should verify that the map defined in Example 2.3.1 is a valid quantum map. The next exercise guides you through this.

Exercise 2.3.3. *Verify that the map T defined in Example 2.3.1 is a valid quantum map. For this, you can start with some simple "sanity checks", such as verifying*

that each of the four matrices ρ_0, ρ_1, ρ_+ and ρ_- is a valid density matrix. However, this is not enough: for example, these conditions are satisfied by the "optimal cloning map" $|\psi_\$\rangle\langle\psi_\$| \mapsto |\psi_\$\rangle\langle\psi_\$| \otimes |\psi_\$\rangle\langle\psi_\$|$, but we know that there exists no such map! To see that T is a well-defined map, we verify that it can be implemented by the combination of a unitary transformation, followed by a tracing-out operation. Consider the following map V:

$$|0\rangle_A |00\rangle_{BC} \mapsto \frac{2}{\sqrt{6}} |00\rangle_{AB} |0\rangle_C + \frac{1}{\sqrt{3}} |\phi^-\rangle_{AB} |1\rangle_C \ ,$$
$$|1\rangle_A |00\rangle_{BC} \mapsto \frac{1}{\sqrt{3}} |\phi^-\rangle_{AB} |0\rangle_C + \frac{2}{\sqrt{6}} |11\rangle_{AB} |1\rangle_C \ .$$

*Using Remark **??**, as long as the two states on the right-hand side, call them $|v_0\rangle$ and $|v_1\rangle$, are orthonormal (check this!), it is possible to extend V to a valid unitary operation on the entire 3-qubit space $\mathbb{C}^8$.*

Now, show that the map T is identical to the composition of V, followed by the tracing out operation applied to the third qubit. That is, for all states $|\psi\rangle$,

$$T(|\psi\rangle\langle\psi|) = \mathrm{Tr}_C \left(V \big(|\psi\rangle\langle\psi| \otimes |00\rangle\langle 00| \big) V^\dagger \right) .$$

This justifies that T is a valid quantum map, because it can be written as a sequence of three valid operations.[2] ∎

Now that we've verified that Example 2.3.1 specifies a valid quantum map T that could, at least in principle, be implemented by a malicious adversary to the quantum cloning scheme, let's see how well this adversary does. For this, we just need to evaluate its success probability using (2.2). For the case of the density matrices from Example 2.3.1, this is straightforward. Working through the calculation, we find

$$p_s = \frac{1}{4}\Big(\frac{2}{3} + \frac{2}{3} + \frac{2}{3} + \frac{2}{3}\Big) = \frac{2}{3} \, .$$

(Make sure that you are able to verify that each of the four $\frac{2}{3} \approx 0.667$ is correct!) As you can see, this attack is only very marginally better than the simple prepare-and-measure attack we considered earlier, that achieves a success probability of $\frac{5}{8} = 0.625$. After so many calculations, this may come as a disappointment. However, we studied the map T for a good reason. Indeed, it is possible to verify by

[2] We won't show it here, but a general theorem in quantum information shows that *any* valid quantum map can be written as the composition of these three simple operations: adding an ancilla, applying a unitary, and tracing out some qubits.

direct calculation that T has the property that for *any* pure single-qubit state $|\psi\rangle$ it holds that

$$\langle\psi|\,\langle\psi|\,T(|\psi\rangle\langle\psi|)\,|\psi\rangle\,|\psi\rangle \;=\; \frac{2}{3}\;.$$

In words, the quantum map T has the ability to "clone" any single-qubit state with success probability $\frac{2}{3}$ (and not only the four BB84 states that appear in Wiesner's quantum money scheme).

This is relevant for the following reasons. First, although we won't prove it here, the success probability of $\frac{3}{4}$ is optimal for Wiesner's scheme. However, this immediately raises the question of whether it is possible to do better. And indeed, it is possible! By considering a 6-state scheme, where the two additional states are

$$\frac{1}{\sqrt{2}}\big(\,|0\rangle + i\,|1\rangle\,\big)\;, \qquad \frac{1}{\sqrt{2}}\big(\,|0\rangle - i\,|1\rangle\,\big)\;,$$

it is possible to get a better scheme, such that the optimal cloning attack only has success probability $\frac{2}{3}$. Our calculation using the map T shows that this is optimal: adding more states than this will not help, because however many states the scheme uses, there is always a cloning attack, given by the map T, that succeeds with probability $\frac{2}{3}$. (The only way to do better is to move away from single-qubit states: in the next section we explore improvements based on considering bills made of multiple qubits.)

Finally, let us give a little intuition for the definition of T. It turns out that, from a mathematical point of view, the map can be expressed as follows:

$$T(\rho) = \frac{1}{2}\Pi_s(\rho\otimes\mathbb{I})\Pi_s\;,$$

where Π_s is the orthogonal projection onto the symmetric subspace of the two-qubit space $\mathbb{C}^4$, i.e. the 3-dimensional subspace spanned by the vectors $|00\rangle$, $|11\rangle$ and $|\phi^-\rangle$. At first it is not obvious that this is a valid quantum map, but it is, and you can verify that it is identical to the map T defined earlier. Intuitively, what T does is that it "maximally symmetrizes" its input state by adding an ancilla qubit initialized in the totally mixed state and then projecting both qubits, the quantum money qubit and the ancilla qubit, into the symmetric subspace. This has the effect of "smearing out" the quantum information present in $|\psi_\$\rangle\langle\psi_\$|$ across both qubits and results in the optimal way to approximately clone an arbitrary qubit.

2.3.3 n-qubit cloning attacks

So far we have considered cloning attacks on the single-qubit version of Wiesner's scheme, and observed that the single-qubit scheme is $\frac{3}{4}$-secure according to Definition 2.2.1. In general, a security of $\varepsilon = \frac{3}{4}$ is not very satisfactory, as it means

that an adversary still has a rather large chance of succeeding in the security game. We would like this probability to be as small as possible, without increasing the complexity of the scheme by too much.

A possibility to achieve this is to consider higher-dimensional states, as described at the end of the previous section. The simplest way to create a high-dimensional state is to put many qubits together. If it is "$3/4$-hard" to clone a quantum bill made of a single qubit, how hard is it to clone a bill made of n such qubits, where n is a large integer, say $n = 256$?

Let's consider the associated security game. In the game, the adversary receives a single copy of an n-qubit quantum money state $|\psi_\$\rangle$, such that $|\psi_\$\rangle$ is a tensor product of single-qubit states $|x_1\rangle_{\theta_1}, \ldots, |x_n\rangle_{\theta_n}$ and x, θ are uniformly random n-bit strings chosen by the challenger. Suppose further that the adversary attempts to clone the qubits one by one, applying a map T_1 on the first qubit, T_2 on the second, etc., such that each map sends one qubit to two qubits. This produces the state

$$T(|\$\rangle\langle\$|) = T_1 \otimes \cdots \otimes T_n(|\$\rangle\langle\$|) = T_1(|x_1\rangle\langle x_1|_{\theta_1}) \otimes \cdots \otimes T_n(|x_n\rangle\langle x_n|_{\theta_n}) \ . \quad (2.4)$$

Averaging over all possible random choices of the challenger, the success probability of this map is

$$\begin{aligned} &\frac{1}{2^{2n}} \sum_{\$=(x,\theta)} \langle\$| \langle\$| T(|\$\rangle\langle\$|) |\$\rangle |\$\rangle \\ &\qquad = \Big(\frac{1}{4} \sum_{x_1,\theta_1} \langle x_1|_{\theta_1} \langle x_1|_{\theta_1} T(|x_1\rangle\langle x_1|_{\theta_1}) |x_1\rangle_{\theta_1} |x_1\rangle_{\theta_1} \Big) \\ &\qquad\quad \cdots \Big(\frac{1}{4} \sum_{x_n,\theta_n} \langle x_n|_{\theta_n} \langle x_n|_{\theta_n} T(|x_n\rangle\langle x_n|_{\theta_n}) |x_n\rangle_{\theta_n} |x_n\rangle_{\theta_n} \Big) \ , \end{aligned}$$

where to get the second expression we used that $|\psi_\$\rangle$ is a product state, (2.4), the distributive identity $(A \otimes B) \cdot (C \otimes D) = (AC \otimes BD)$ for any (not necessarily square) matrices, and we re-ordered the $2n$ qubits $(1 \cdots n)(1' \cdots n')$ as $(11') \cdots (nn')$. Since we know that each term on the right-hand side is at most $\frac{3}{4}$, we immediately get that the success probability of this type of attack is at most $(\frac{3}{4})^n$. This number goes to 0 very fast (*exponentially fast*) as $n \to \infty$, so it is a security level we are happy with: for example, choosing $n = 256$ already drives the success probability to a number that is so small that even if the adversary was able to try an attack every nanosecond, it would take them more than the age of the universe (which is of order 10^{25} nanoseconds) to successfully break the scheme.

Unfortunately, this analysis only considers a very specific type of attack: attacks that attempt to clone the qubits one by one, independently of each other.

Could there be a better attack, one that simultaneously takes all the qubits into account? In general this would be modeled by an arbitrary quantum transformation T from n to $2n$ qubits. As you can imagine, analyzing such a general attack can take a lot of work. Oftentimes the idea is to give an argument that show that no such attack can succeed with substantially higher probability than the independent attacks considered above. Although we will not show it here, in the case of Wiesner's scheme this can be done, and it is known that no attack on the n-qubit scheme can have higher success probability than the one that consists in applying the basic "measure-and-prepare" attack from Section **??** independently on each qubit.

2.4 The Elitzur-Vaidman bomb tester

In the previous section we considered "cloning attacks" on Wiesner's quantum money scheme, and argued that for the n-qubit scheme, the best such attack has success probability at most $(3/4)^n$ in the security game. This should make us confident in the security of the scheme. However, we have to be careful. A security definition in cryptography only covers the kind of attacks that are captured by the underlying security game...but real life can be more complex!

In particular, note that the definition of quantum money requires to specify a verification procedure, but it does not say what should happen with a bill once it has been verified. So what should the bank do with the user's bill, once it has executed the verification procedure on it?

It is natural to consider that, if the bill is accepted, the bank returns it to the user, while if it is rejected, then the bank destroys the bill (and even fines the malicious user, or puts her in prison). We'll soon see that this is *not* a reasonable assumption: as soon as the bank returns valid bills to their user, even if it only does so when the bill has successfully passed verification, a malicious user can take advantage of this fact to break the scheme and clone an arbitrary valid bill!

But wait, didn't we prove security? We did, but in our model the adversary does not have access to what is sometimes called a "verification oracle". In our security game, the adversary receives one bill and she has to produce two; in order to do so her only resource is whatever quantum operation she can apply in her laboratory — there is no interaction with the "bank" or the challenger in-between the two phases. Let's now see how the adversary has a cloning attack if in addition it is allowed to submit "candidate bills" to the bank for verification and the bank returns bills that were declared valid.

2.4.1 A cloning attack against permissive banks

Let's first consider a simpler scenario where we assume that verification *always* returns a bill to the user, even if the bill failed verification. At first it may seem like this is not really a problem: if the bill is invalid, it will remain invalid, so what is Alice going to do with it anyways? We'll call a bank that returns invalid bills to the user a *permissive bank*.

Let's play the security game CLONE in this context. Suppose that the adversary Eve got her hands on a valid quantum bill $(|\psi_\$\rangle, \$)$. Recall that in Wiesner's scheme the bill is made of n qubits, in state $|x_1\rangle_{\theta_1}, \ldots, |x_n\rangle_{\theta_n}$. If Eve sends this state for verification to the bank, the bank accepts it. Now suppose Eve sets the first qubit $|x_1\rangle_{\theta_1}$ aside, replaces it with a $|0\rangle$, and sends the modified bill for verification. If the bill is rejected, she tries with $|1\rangle$, $|+\rangle$, $|-\rangle$. At least one of these must be accepted. As soon as her attempt is accepted, she keeps the qubit, and proceeds in the same way with the second qubit, etc. Once she is done with all n qubits, she has two copies of the valid bill: the one made of the qubits set aside at each step, and the one made of the "guessed" qubits. Moreover, this attack only required going through $4n$ verifications!

Observe that this attack relies on the fact that, once a qubit is accepted, it will be accepted in all future attempts. For example, if the correct first qubit is $|+\rangle$, but Eve submits a $|0\rangle$, she has probability $1/2$ of bring accepted, since $|\langle 0|+\rangle|^2 = \frac{1}{2}$. But if she is accepted, then her qubit is projected to the post-measurement state $|+\rangle$, so she will succeed with probability 1 in any subsequent verification attempt.

This suggests a slightly more streamlined attack. Instead of going through the four possibilities $|0\rangle, |1\rangle, |+\rangle, |-\rangle$ in sequence, Alice can guess a uniformly random state for the qubit, effectively replacing it by the density matrix $\rho = \mathbb{I}/2$. Then, verification will succeed with probability $\frac{1}{2}$ in all cases, since

$$\langle 0| \frac{\mathbb{I}}{2} |0\rangle = \langle 1| \frac{\mathbb{I}}{2} |1\rangle = \langle +| \frac{\mathbb{I}}{2} |+\rangle = \langle -| \frac{\mathbb{I}}{2} |-\rangle = \frac{1}{2}.$$

Moreover, whenever verification succeeds, the state is automatically projected onto the correct one. On expectation, this attack requires only $2n$ verifications to succeed. However, note that since Eve will fail n times on expectation, for the attack to be effective it is required that the bank return invalid bills without putting Eve in prison.

2.4.2 Non-permissive banks

In retrospect you may think that allowing the bank to return invalid bills is an obvious mistake: clearly, if Eve submits an invalid bill, she should be sent to prison

right away! In this situation we can think of verification as a "bomb", that explodes any time it is given the wrong state. Observe however that if Eve submits a state that is "slightly invalid", such a state obtained by applying a small unitary rotation to a valid quantum bill, then her changes of succeeding in verification remain relatively high, depending on the angle of the rotation: this observation will be crucial to her attack!

Eve's goal is, given one copy of the "magic state" that is the only input on which the "bomb" does *not* explode, to find two such inputs... while staying alive (i.e. out of prison). Let's see how it is possible to do this. Informally, the idea is to make use of a quantum variant of the "Zeno effect". Starting from her valid quantum bill, Eve is going to make very small modifications of it, and "test" these modifications against the verification procedure. Since the modifications are small, they are very likely to be accepted. Yet, Eve will make them in a way that she still learns information about her quantum bill.

Let's see how to do this one qubit at a time. Eve does the following.

1. Eve initializes an ancilla qubit to state $|0\rangle$. Including the quantum money qubit, her entire state is $|x\rangle_\theta |0\rangle$, for unknown $x, \theta \in \{0, 1\}$.

2. Eve chooses a small rotation parameter $\delta \in (0, \pi)$ and applies the unitary rotation $R_\delta = \begin{pmatrix} \cos\delta & \sin\delta \\ -\sin\delta & \cos\delta \end{pmatrix}$ to her second qubit (the ancilla qubit).

3. Eve applies a control-X operation from the second qubit to the first. That is, if the ancilla qubit is in state $|1\rangle$ then she applies an X-flip on the money qubit, and otherwise she does nothing.

4. Finally, Eve submits the money qubit for verification to the bank.

Note that, intuitively, for small δ this procedure modifies the money state only a little bit (since $R_\delta |0\rangle = \cos\delta |0\rangle + \sin\delta |1\rangle$ only has a small amplitude $\sin\delta \approx \delta^2$ on $|1\rangle$), so that Eve's chances of succeeding in verification should be high (in particular, if $\delta = 0$ then the money state is unchanged and she succeeds with probability 1).

Let's examine what happens to the money state in all possible cases. The following exercise asks you to consider the case when $\theta = 1$.

Exercise 2.4.1. *Suppose that $\theta = 1$, i.e., $|\psi_\$\rangle = |+\rangle$ or $|\psi_\$\rangle = |-\rangle$. Show that after one step of the procedure described above,*

- *If $|\psi\rangle = |+\rangle$ then the state is $|+\rangle \otimes (R_\delta |0\rangle)$;*
- *If $|\psi\rangle = |-\rangle$ then the state is $|-\rangle \otimes (R_{-\delta} |0\rangle)$.*

Suppose that Eve repeats the procedure an even number $2N$ of times, with an angle $\delta = \frac{\pi}{4N}$. What is the final state of the control qubit, in case $|\psi_\$\rangle = |+\rangle$ or in case $|\psi_\$\rangle = |-\rangle$? Show that the two cases can be perfectly distinguished. ∎

Now suppose that $\theta = 0$, so $|\psi\rangle = |x\rangle$, for some unknown $x \in \{0, 1\}$. In this case, right before the verification attempt, Eve's state is

$$(\cos\delta)\,|x\rangle \otimes |0\rangle + (\sin\delta)(X\,|x\rangle) \otimes |1\rangle \ .$$

Verification measures in the standard basis and accepts if and only if the outcome is $|x\rangle$. Hence in our setting it accepts with probability $\cos^2\delta$, in which case the state gets projected to $|x\rangle\,|0\rangle$, i.e. the same state as originally. With probability $\sin^2\delta \approx \delta^2$ (for small δ), verification fails and Eve goes to jail.

Now, suppose δ is very small, say $\delta = \frac{\pi}{4N}$ for some large integer N, as in Exercise 2.4.1. Suppose Eve executes the procedure described above $(1/\delta)$ times. If $|\psi\rangle$ is $|0\rangle$ or $|1\rangle$, then (unless Eve has been caught) the control qubit is $|0\rangle$. Moreover, the chance that she is caught is at most $(1/\delta)\sin^2\delta \leq \pi/(4N)$, which can be small arbitrarily small by choosing the number of iterations large enough.

Next, if the state is $|\psi\rangle = |-\rangle$, since the number of repetitions is even, the control qubit is in state $|0\rangle$. Finally, if it is $|\psi\rangle = |+\rangle$, the final state of the control qubit is $R_{\pi/2\delta\cdot\delta}\,|0\rangle = R_{\pi/2}\,|0\rangle = |1\rangle$.

Overall, Eve observes a $|1\rangle$ when measuring the control qubit if and only if the state is $|\psi\rangle = |+\rangle$: in this case, she has perfectly identified it! Replacing the controlled operation in step 2 by a control $(-X)$, Z, or $(-Z)$ she can similarly perfectly identify the cases where $|\psi\rangle$ is $|-\rangle$, $|0\rangle$ or $|1\rangle$ respectively. Moreover, since in each case the chance that she is caught is less than $1/(400n)$, she manages to succeed in identifying the first qubit, perfectly, while getting caught only with probability $(1/100n)$. Repeating this procedure independently for all n qubits yields a perfect classical description of the money state, with only a 1% chance of going to jail. Moreover, clearly this 1% can be made much lower if Eve is willing to be just a little more careful: by generalizing our analysis you can easily show that, if she wishes to be caught with probability at most p, then $\sim n/p$ total verification attempts will be enough.