Bombs do not have to explode, forgers can get away scot-free (an adaptive attack on Wiesner's quantum money)

Aharon Brodutch[1], Daniel Nagaj[2], Or Sattath[3] and Dominique Unruh[4]

[1]Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, Ontario, Canada
[2]Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, Slovakia
[3]Computer Science Division, University of California, Berkeley, California, USA
[4]University of Tartu, Estonia

November 28, 2014

Abstract

Unlike classical money, which is hard to forge for practical reasons (e.g. producing paper with a certain property), quantum money is attractive because its security may be proved based on the no-cloning theorem. The first quantum money scheme was introduced by Wiesner circa 1970. Although more sophisticated quantum money schemes were proposed, Wiesner's scheme remained appealing because it is both conceptually clean and relatively easy to implement.

We show an efficient adaptive attack on Wiesner's quantum money scheme [15] (and its variant by Bennett et al. [5]), when valid money is accepted and passed on, while invalid money is destroyed. Our approach is inspired by the Elitzur-Vaidman bomb testing problem[7, 11] and the idea of *protective measurements*[3]. It allows us to break Wiesner's scheme with 4 possible states per qubit, and generalizations which use more than 4 states per qubit.

The preprint arXiv:1404.1507 contains the full version of this extended abstract.

Wiesner's Quantum Money. One of the main requirements for any medium of money is that it it will not be easy to copy. For this very reason, it is appealing to construct *quantum money*: its security would follow from the laws of quantum mechanics, or more specifically, the no-cloning theorem [16]. Indeed, quantum money was one of the earliest quantum information protocols, introduced by Stephen Wiesner circa 1970, although it took some time to be published [15].

Wiesner's quantum money scheme uses only single-qubit memories and single-qubit measurements, as follows: A bank creates a note of size n with a public serial number s, and, for each serial number a random (classical) private key $k^{(s)} \in \{0,1,+,-\}^n$. The corresponding banknote contains a *quantum money state* $|\$_s\rangle = |k_1^{(s)}\rangle \otimes |k_2^{(s)}\rangle \otimes \ldots \otimes |k_n^{(s)}\rangle$, where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. The serial number together with the quantum money state, i.e. the pair $(s, |\$_s\rangle)$, form *legitimate* quantum money. In order to validate her money (and pay with it), Alice sends the (potentially forged) banknote $(s, |\psi\rangle)$ to the bank. The bank measures each of the qubits of $|\psi\rangle$ in its respective basis; the i^{th} qubit is measured in the basis $\{|0\rangle, |1\rangle\}$ if $k_i^{(s)} \in \{0,1\}$, and in the $\{|+\rangle, |-\rangle\}$ basis otherwise. The money is declared valid if and only if all[1] the measurement outcomes agree with the measurements on the legal state $|\$_s\rangle$.

One downside of Wiesner's original scheme is that the bank must keep a database containing the secret key for every serial number. In a follow-up paper, Bennett, Brassard, Breidbart and Wiesner used a fixed pseudo-random function for choosing the secret keys for all the serial numbers, which implies that the memory required by the bank does not grow as a function of the number of legitimate serial numbers [5]. All our results apply both to Wiesner's and to the Bennett et al. scheme.

[1]Any realistic implementation must tolerate some incorrect measurements due to noise.

There are two specifications which are important for our work: we have to state whether after a successful validity test, the money state is returned to Alice (or passed to Bob who does business with Alice), or replaced with a new quantum money state and a new serial number. Next, after a failed validity test, is the post-measurement state (a bad bill) returned to Alice for inspection? These distinctions are crucial for the scheme's security. We define *strict testing* to be the variant of Wiesner's scheme in which the valid money state *is returned to the owner* (or passed to another seller/buyer) after a successful test, and the post measurement state of a failed test is confiscated (or the person trying to pass bad money goes to jail). Our attack on Wiesner's scheme works in this setting (we note that this setting may not be robust to noise[14]).

Wiesner proposed his quantum money scheme more than 40 years ago, and it was believed to be secure although a complete security proof was never published. However, Lutomirski [12] and Aaronson [1] independently observed that the scheme is insecure if the bank returns valid notes as well as the post measurement state after detecting an invalid note. Farhi et al. have showed that "single copy tomography" [8] can be performed in a much more general setting. When we own a single copy of an unknown state $|\psi\rangle$ and have access to a projective measurement (validation test) $|\psi\rangle\langle\psi|, \mathbb{I} - |\psi\rangle\langle\psi|$ provided as a black box, we can efficiently estimate the reduced local density matrices of the state $|\psi\rangle$. This approach, based on Jordan's lemma, pointed out the security threat posed by having access to a validation procedure (in particular, the post-measurement state). Lutomirski's conclusion was that so long as the bank does not return the post-measurement state of an invalid note, the scheme is safe.

On the other hand, Molina, Watrous and Vidick proved that Wiesner's scheme is secure against *simple counterfeiting attacks* [13]. In this setting, an attacker is given a single copy of an authenticated state, and attempts to create two banknotes with the same serial number which, independently, pass the bank's validity test. During the counterfeiting process, the attacker does not have access to the validity test. They showed that the success probability of the optimal attack on Wiesner's quantum money scheme is $\left(\frac{3}{4}\right)^n$. The question whether the scheme is secure under general attacks (i.e. non-simple attack) was left open. Similarly, in the scheme of Pastawski et al. [14] which is secure also against a certain level of noise, a good or bad token (money) is destroyed after every use and never returned. However, in a general attack (as in our case), the attacker could make some use of the validation procedure, even if the bank strictly discourages failed tests and does not return bad banknotes.

Main results (arXiv:1404.1507). We show that in Wiesner's *strict testing* scheme (that is, if only valid money is returned to the owner), given a single valid quantum money state $(s, |\$_s\rangle)$, a counterfeiter can efficiently create as many copies of $|\$_s\rangle$ as he wishes (hence, the scheme is insecure). We rely on the quantum Zeno effect for protection – if we disturb the quantum money state only slightly, we are likely to be projected back to the original state after a test. Interestingly, this allows us to distinguish between the four different qubit states without ever being caught. The basic trick is to apply the operation X in a controlled way – this operation leaves the state $|+\rangle$ intact, and does something nontrivial to the states $|0\rangle, |1\rangle, |-\rangle$.

The simplest attack is based on the Zeno assisted Elitzur-Vaidman bomb test [11]. Our goal is to answer the question whether the first qubit of our quantum money is in the state $|+\rangle$ or not. By repeating this process for each of the money qubits, and for each of the four possible states, we can identify the quantum money state. We use an ancillary *probe* qubit, initialized to the $|0\rangle$ state. We repeat the following steps $N = \frac{\pi}{2\delta}$ times, as depicted in Fig. 1: (i) apply a rotation of a small angle δ to the probe. (ii) Apply a C-NOT from the probe qubit to the money qubit. (iii) Send the quantum money to the bank for verification.

If the money qubit is in the $|+\rangle$ state, it stays invariant under the NOT operation, and therefore also by the C-NOT operation. Hence, at the end of the procedure the probe qubit will be in the state $|1\rangle$. The bank's verification will always pass the bill since it is unchanged. If the quantum money state is in either the $|0\rangle$ or $|1\rangle$ state, the probe will be in the $|0\rangle$ state, using the same analysis of the Elitzur-Vaidman bomb tester. In these two cases the maximal rotation induced on the money state at any one time is at most δ and the bank's probability of detecting a counterfeiter is at most δ^2; hence, the overall probability of detection by the bank is $O(\delta)$. The last case is somewhat different: after the first iteration, the probe has angle $-\delta$. At the end of the second iteration, the probe return to state $|0\rangle$, etc. Therefore, at the end of the procedure, the probe is in the $|0\rangle$ (as long as N is even). Here again the money state is left invariant.

One might hope that a simple generalization of Wiesner's strict testing scheme, using r money states instead

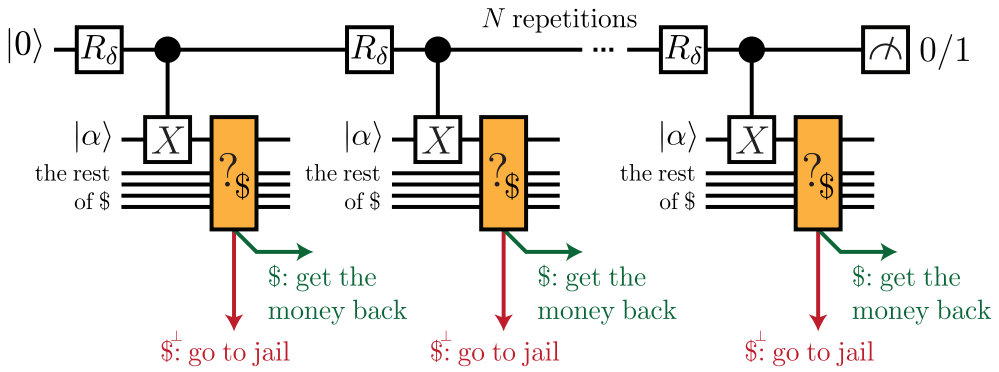

Figure 1: An adaptive attack on Wiesner's quantum money with a strict testing procedure. We can identify whether the qubit $|\alpha\rangle$ is in the state $|+\rangle$ without going to jail (being detected).

of 4 states will be able to hold off our attack. However, this is not the case. A natural generalization works as long as the set of states used is finite, and is known in advance. An alternative, tomographic, attack based on protective measurements [3] is used when the set of states is unknown, or if the number of possible money states is infinite.

Protective measurements allow us to estimate the expectation value of an operator A of a state $|\psi\rangle$, without disturbing the state, by preparing a probe in the initial state $|0\rangle$ and repeating the following procedure N times. i) Weakly couple the probe and the system ii) send the state to the bank for verification:

$$|0\rangle|\psi\rangle \xrightarrow{e^{i\delta\sigma_x\otimes A}} \approx |0\rangle|\psi\rangle + i\delta|1\rangle A|\psi\rangle \xrightarrow{\text{Bank Measures } \{|\psi\rangle\langle\psi|, I-|\psi\rangle\langle\psi|\}}$$
$$\approx [\cos(\langle A\rangle\delta)|0\rangle + i\sin(\langle A\rangle\delta)|1\rangle] \otimes |\psi\rangle \xrightarrow{\text{Repeat N times}} \approx \left[\cos(\frac{\pi}{2}\langle A\rangle)|0\rangle + i\sin(\frac{\pi}{2}\langle A\rangle)|1\rangle\right] \otimes |\psi\rangle.$$

The two approaches for the attack require very different analysis, and have complementary properties: The "bomb" attack fails when a guessed state is too close to the real state. Otherwise, it unambiguously identifies the correct state. The "protective" method works in general but only produces a (classical) estimate for the quantum money state.

Discussion and applications. Our attack applies in the strict-testing regime, where good banknotes are returned or passed on, while failed tests result in confiscation of the banknote (or us being sent to jail). However, the attack does not work if after a valid test, a new quantum money and a new serial number are returned to the owner. Does this affect the advantages of Wiesner's scheme? In order to answer this question, we first need to understand its two main advantages (see a more detailed analysis in Ref. [10]):

- The data needed for validation is static. Therefore, after the money has been issued, many bank branches can validate the money without any need for communication.
- The hardware requirements (single-qubit memory and single-qubit measurements) are less demanding than the modern schemes [9, 2].

These advantages remain when the quantum money is replaced by a new one, after each successful validity test.

Our results raise an an important warning about quantum money (and other variants, such as quantum copy protection [1]) constructions – we need to be cautious about reusing valid bills (even though false bills are destroyed).

We believe the greatest potential of this work is in the context of weak measurements. The framework of weak measurement has been proven an important concept in numerous cases [6, 4] including protective measurements and precision metrology. Our "bomb" attack, shares a lot of the properties of weak measurements, but not all. We believe that further investigation of these different approaches in a broad perspective will give practical applications for weak measurement methods.

Acknowledgments. We would like to thank Scott Aaronson, Guy Kindler, Carl Miller, David Gosset and the Simons institute Quantum Hamiltonian Complexity program.

References

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

[2] S. Aaronson and P. Christiano. Quantum money from hidden subspaces. In *Proceedings of the 44th Symposium on Theory of Computing*, pages 41–60. ACM, 2012.

[3] Y. Aharonov, J. Anandan, and L. Vaidman. Meaning of the wave function. *Physical Review A*, 47(6):4616, 1993.

[4] Y. Aharonov and L. Vaidman. The two-state vector formalism: an updated review. In *Time in Quantum Mechanics*, pages 399–447. Springer, 2007.

[5] C. H. Bennett, G. Brassard, S. Breidbart, and S. Wiesner. Quantum cryptography, or unforgeable subway tokens. In *Advances in Cryptology*, pages 267–275. Springer, 1983.

[6] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd. Colloquium: Understanding quantum weak values: Basics and applications. *Rev. Mod. Phys.*, 86(1):307–316, Mar 2014.

[7] A. C. Elitzur and L. Vaidman. Quantum mechanical interaction-free measurements. *Foundations of Physics*, 23(7):987–997, 1993.

[8] E. Farhi, D. Gosset, A. Hassidim, A. Lutomirski, D. Nagaj, and P. Shor. Quantum state restoration and single-copy tomography for ground states of hamiltonians. *Physical review letters*, 105(19):190503, 2010.

[9] E. Farhi, D. Gosset, A. Hassidim, A. Lutomirski, and P. Shor. Quantum money from knots. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, pages 276–289. ACM, 2012.

[10] D. Gavinsky. Quantum money with classical verification. In *IEEE 27th Annual Conference on Computational Complexity*, pages 42–52. IEEE, 2012.

[11] P. Kwiat, H. Weinfurter, T. Herzog, A. Zeilinger, and M. A. Kasevich. Interaction-free measurement. *Physical Review Letters*, 74(24):4763, 1995.

[12] A. Lutomirski. An online attack against wiesner's quantum money. *arXiv preprint arXiv:1010.0256*, 2010.

[13] A. Molina, T. Vidick, and J. Watrous. Optimal counterfeiting attacks and generalizations for wiesners quantum money. In *Theory of Quantum Computation, Communication, and Cryptography*, pages 45–64. Springer, 2013.

[14] F. Pastawski, N. Y. Yao, L. Jiang, M. D. Lukin, and J. I. Cirac. Unforgeable noise-tolerant quantum tokens. *Proceedings of the National Academy of Sciences*, 109(40):16079–16082, 2012.

[15] S. Wiesner. Conjugate coding. *ACM Sigact News*, 15(1):78–88, 1983.

[16] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, 1982.