Vulnerabilities of Quantum Lightning

Bhaskar Roberts

UC Berkeley and Princeton University
bhaskarr@berkeley.edu

Abstract

Zhandry recently defined a new cryptographic object called quantum lightning, which has a number of useful applications, including a strong form of quantum money. Further, Zhandry proposed a construction of quantum lightning based on superpositions of low-rank matrices. The scheme is unusual, so it is difficult to analyze whether the scheme is secure and difficult to base the scheme's security on any widespread computational assumptions. Instead, Zhandry proposed a new hardness assumption that, if true, could be used to prove security. While the new hardness assumption is plausible, it has not been rigorously analyzed.

In this work, we analyze the hardness assumption to determine how, if at all, it can be justified. We show that Zhandry's hardness assumption is in fact false, so the proof of security for the quantum lightning scheme does not hold. While the scheme itself has not been proven insecure, our analysis suggests how we might prove it insecure.

Quantum Lightning Scheme

Bolt: The bolt is a superposition that is hard to duplicate. In [Zha19], the bolt is a superposition of pre-images to a hash function that collide.

Hash function: The hash function f_A is used to generate and verify the bolts. f_A maps M, a symmetric, low-rank, m×m matrix, to y, an n-dimensional vector. f_A is defined by a set of matrices, $\{A_1, \dots, A_n\}$. For each i in [n]:

$$y_i = [f_A(M)]_i = \mathrm{Tr}(A_i \cdot M)$$

To prove the scheme secure, we need f_A to be (2k+2)-multi-collision resistant (MCR), for some positive integer k. This means it is infeasible to find 2k+2 inputs to f_A that map to the same output. (We will show later that f_A is not (2k+2)-MCR, so the scheme may not be secure).

Q. Lightning Scheme Cont.

Trapdoor: f_A has an associated matrix R, which is a public trapdoor that is used to verify bolts. R has the following useful property: for all i in [n],

$$\mathrm{Tr}(R \cdot A_i \cdot R^T) = 0$$

Generation: To generate a bolt, we first construct a superposition of all matrices in the domain of f_A. Then, we compute f_A in superposition on the inputs, and finally measure the output of f_A. The value we measure is a random y-value, and the superposition that remains is over all of y's pre-images. This superposition is called a mini-bolt for y. A bolt is a set of k+1 mini-bolts for the same y, and y is the bolt's serial number. It is supposedly hard to duplicate this bolt.

Verification: To verify that a given state is an honestly generated bolt, we make two sets of measurements, one in the computational basis and one in the Fourier basis.

1. First, we check that for each purported mini-bolt, the eigenstates are preimages of y.
2. Next, we apply the quantum Fourier transform. For each eigenstate M in the Fourier domain, we check that $\mathrm{rank}(R \cdot M \cdot R^T)$ is lower than a given threshold.

Hardness Assumption

The scheme presented in [Zha19] is secure if it is hard for an adversary to produce two bolts with the same serial number y. It is unknown whether the scheme is secure, and the main difficulty is that little is known about the hash function f_A.

However, [Zha19] makes the following hardness assumption and proves that their scheme is secure if the assumption holds. Informally, they assume that **for some choice of parameters, f_A is (2k+2)-multi-collision resistant, for some positive integer k, even if R is made public.** This means it is infeasible to find 2k+2 inputs to f_A that map to the same output.

Breaking the Assumption

In fact, the hardness assumption is false. Because R is public, an adversary can use R to construct a large number of colliding inputs. From the properties of R and f_A, we can show that $f_A(R \cdot R^T) = 0$, so $R \cdot R^T$ is in the kernel of f_A.

This opens the door to constructing many other matrices in the kernel of f_A. First, let the rows of R be $\{r_1, \dots, r_e\}$. Then we construct the following set of matrices: $K = \{r_1 \cdot r_1^T, \dots, r_e \cdot r_e^T\}$.

The matrices in K are rank-1 symmetric matrices in the kernel of f_A. Any linear combination of the matrices in K is also in the kernel, and these matrices represent colliding inputs. Since it is easy to construct many colliding inputs to f_A, the hardness assumption is false.

Implications

Since the hardness assumption is broken, the proof of security for [Zha19]'s scheme does not hold. However, this result *does not* prove [Zha19]'s scheme insecure.

As future work, it might be possible to prove the scheme insecure as well. An adversary can construct two bolts with the same serial number if they can find sufficiently many inputs to f_A that collide. Therefore an attack similar to the one used to break the hardness assumption might succeed in breaking the overall scheme.

Works Cited

- [AC13] Scott Aaronson and Paul Christiano. Quantum money from hidden subspaces. Theory of Computing, 9(9):349-401, 2013.
- [FGH+12] Edward Farhi, David Gosset, Avinatan Hassidim, Andrew Lutomirski, and Peter Shor. Quantum money from knots. In Proceedings of the 3rd Innovations in Theoretical Computer Science Conference, ITCS'12, pages 276-289,New York, NY, USA, 2012. ACM.
- [Wie83] Stephen Wiesner. Conjugate coding. SIGACT News, 15(1):78-88, January 1983.
- [Zha19] Mark Zhandry. Quantum lightning never strikes the same state twice. 11478:408-438, 2019.

I thank Mark Zhandry for advising this work.