The Black-Box Simulation Barrier Persists in a Fully Quantum World

Nai-Hui Chia[1], Kai-Min Chung[2], Xiao Liang[3], and Jiahui Liu[4]

[1] Rice University, USA
nc67@rice.edu
[2] Academia Sinica, Taiwan
kmchung@iis.sinica.edu.tw
[3] The Chinese University of Hong Kong, Hong Kong
xiaoliang@cuhk.edu.hk
[4] Fujitsu Research of America, USA
jliu@fujitsu.com

Abstract. Zero-Knowledge (ZK) protocols have been a subject of intensive study due to their fundamental importance and versatility in modern cryptography. However, the inherently different nature of quantum information significantly alters the landscape, necessitating a re-examination of ZK designs.

A crucial aspect of ZK protocols is their round complexity, intricately linked to *simulation*, which forms the foundation of their formal definition and security proofs. In the *post-quantum* setting, where honest parties and their communication channels are all classical but the adversaries could be quantum, Chia, Chung, Liu, and Yamakawa [FOCS'21 & QIP'22] demonstrated the non-existence of constant-round *black-box-simulatable* ZK arguments (BBZK) for **NP** unless **NP** $\subseteq$ **BQP**. However, this problem remains widely open in the full-fledged quantum future that will eventually arrive, where all parties (including the honest ones) and their communication are naturally quantum.

Indeed, this problem is of interest to the broader theory of quantum computing. It has been an important theme to investigate how quantum power fundamentally alters traditional computational tasks, such as the *unconditional* security of Quantum Key Distribution and the incorporation of Oblivious Transfers in MiniQCrypt. Moreover, quantum communication has led to round compression for commitments and interactive arguments. Along this line, the above problem is of great significance in understanding whether quantum computing could also change the nature of ZK protocols in some fundamentally manner.

We resolved this problem by proving that only languages in **BQP** admit constant-round *fully-quantum* BBZK. This result holds significant implications. Firstly, it illuminates the nature of quantum zero-knowledge and provides valuable insights for designing future protocols in the quantum realm. Secondly, it relates ZK round complexity with the intriguing problem of **BQP** vs **QMA**, which is out of the reach of previous analogue impossibility results in the classical or post-quantum setting. Lastly, it justifies the need for the *non-black-box* simulation techniques or the relaxed security notions employed in existing constant-round fully-quantum BBZK protocols.

Keywords: Quantum Zero-Knowledge · Black-Box · Impossibility

1 Extended Abstract

Zero-Knowledge (ZK) protocols [GMR85] enable a prover to prove a statement (e.g., one in an **NP** language) without disclosing additional information (e.g., the **NP** witness) beyond the statement's truthfulness. Since their inception, ZK protocols have garnered considerable research interest and emerged as a cornerstone of cryptography. Moreover, they are closely related to computational complexity theory, as evidenced by various studies [IY88, SV97a, SV97b, BGG+90, BMO90, GSY19, GIK+23]. Understanding the nature of ZK protocols is of significant importance.

Round Complexity and Black-Box Simulation. The formal security definition of ZK mandates the presence of a *simulator* $\mathcal{S}$ that interacts with a malicious verifier $\mathcal{V}^*$ and convinces $\mathcal{V}^*$ that it is communicating with the honest prover. Importantly, $\mathcal{S}$ possesses no secret input akin to that of the honest prover. Thus, the success of $\mathcal{S}$ in convincing $\mathcal{V}^*$ encapsulates the intuitive essence of ZK protocols — $\mathcal{V}^*$ gains no extra information (other than the statement's truthfulness) from the interaction with the honest prover. Typically, the simulator $\mathcal{S}$ interacts with $\mathcal{V}^*$ in a *black-box* manner, meaning that $\mathcal{S}$ solely leverages the Input/Output behavior of $\mathcal{V}^*$ and treats it as an oracle. This black-box approach is arguably the most natural choice for simulation, often being simpler, more efficient and more modular compared to non-black-box methods.

Strong impossibility results have been established for ZK protocols employing black-box simulation (dubbed BBZK henceforth). Most relevant to us is the seminal work by Barak and Lindell [BL02], who established the non-existence of constant-round BBZK with *strict polynomial-time* simulation, unless **NP** $\subseteq$ **BPP**. These impossibility results significantly advance our understanding of zero-knowledge and offer valuable guidance for positive results (i.e., constructions).

BBZK in the Quantum Era. As we move to the quantum setting, a fundamental question is posed: *What can we say about the round complexity of ZK protocols in the quantum setting?* To address this question, two models need to be considered.

Post-Quantum ZK. This is the model where honest parties and their communication channels are all classical, but the adversary could be a quantum machine. Exciting progress has been made in this model. The recent breakthrough by Bitansky and Shmueli [BS20] presents a constant-round post-quantum ZK argument for **NP** using *non-black-box simulation*. On the other hand, Chia, Chung, Liu, and Yamakawa [CCLY21] showed that there does not exist a constant-round post-quantum BBZK for **NP** unless **NP** $\subseteq$ **BQP**. It is worth noting that the result by [CCLY21] holds even for *expected* quantum-polynomial-time (QPT) simulation. Thus, it is *not* merely a analogue of [BL02], but rather represents a qualitatively stronger impossibility result due to quantum phenomenon.

Fully-Quantum ZK. This is the full-fledged quantum model where all parties (including the honest ones) and communication channels are allowed to be quantum. This quantum capability enables new possibilities, such as ZK protocols for **QMA**, which are infeasible in the post-quantum setting since classical honest provers cannot make use of quantum **QMA** witnesses. In contrast to the post-quantum model, our understanding of the round complexity of *fully-quantum* BBZK is quite limited. While fully-quantum BBZK protocols do exist, they require at least super-constant rounds. Conversely, it remains unclear whether the aforementioned impossibility results extend to this fully quantum model. Thus, *the central problem in the area remains widely open:*

> **Question:** *Do there exist constant-round fully-quantum BBZK protocols for* **NP** *(or* **QMA***) with strict (or expected) QPT simulation?*

This question is also of interest to the broader theory of quantum computing. It has been an important theme to investigate how quantum power fundamentally alters traditional computational tasks, such as the *unconditional* security of Quantum Key Distribution [BB84] and the incorporation of Oblivious

Transfers in MiniQCrypt [GLSV21, BCKM21]. Additionally, quantum communication has led to round compression for commitments [Yan22] and interactive arguments [KW00, KKMV09, BQSY24].

Along this line, this Question is critical for understanding whether quantum computing could also alter the nature of zero-knowledge protocols in some fundamentally manner. Moreover, answers to this Question are desirable because (1) they would connect zero-knowledge with the intriguing problem of **BQP** vs **QMA**, whereas post-quantum protocols could at most relate to **BQP** vs **NP**, as explained above; (2) they would provide deeper insights into the curious situation where constant rounds have only been achieved either under a relaxed security notion (e.g., ε-simulation) or through non-black-box simulation; (3) lastly, if quantum computing does not enable a constant-round BBZK protocol, it is likely that new techniques will be required to establish the lower bound, which may prove useful beyond this specific problem. We provide further discussion of this point in our technical manuscript.

1.1 Our Results

We address the Question by establishing the following impossibility result.

Theorem 1. *For any language $\mathcal{L}$, if there exists a constant-round fully-quantum BBZK protocol with expected QPT simulation, then it holds that $\mathcal{L} \in$ **BQP**.*

It is worth noticing that Thm. 1 rules out *expected* QPT simulation, and thus it can be viewed as the exact analogue of the [CCLY21] impossibility in the fully-quantum context. However, our techniques to establishing Thm. 1 diverge significantly from [CCLY21] due to the fundamentally different nature inherent in fully-quantum protocols. Further discussion on this matter will be provided in the **Technical Overview** in our technical manuscript.

As a corollary of the above Thm. 1, we can now connect zero-knowledge with the quantum complexity class **QMA**: There does not exist any constant-round fully-quantum BBZK protocol for **QMA** (resp. **NP**) unless **QMA** $\subseteq$ **BQP** (resp. **NP** $\subseteq$ **BQP**).

Also, Thm. 1 justifies the necessity, *in the fully-quantum context*, of (1) the use of non-black-box simulation in [BS20], (2) the relaxation to ε-simulation in [CCY21, CCLY22], and (3) the non-constant round complexity observed in existing BBZK or secure *quantum* computation protocols with black-box simulation, e.g., [DNS12, BJSW16, BG20, DGJ+20, GLSV21, ACL21] etc.

1.2 Technical Summary

Roughly speaking, [CCLY21] follows the paradigm in [BL02] by designing a specific malicious verifier V^* such that if there exists a constant k-round, post-quantum ZK protocol Π with a QPT black-box simulator $\mathcal{S}$, then the interaction of $\mathcal{S}$ and the designed V^* can be turned into a **BQP** algorithm to decide if any given instance x is in language $\mathcal{L}$ — the algorithm simply runs $\mathcal{S}$ (with black-box access to V^*) on input x, and outputs V^*'s final decision (acceptance or rejection) as its result.

[CCLY21] needs to show that this algorithm is both *complete* (i.e., working correctly for $x \in \mathcal{L}$) and *sound* (i.e., working correctly for $x \notin \mathcal{L}$). The completeness of the algorithm follows rather easily from the completeness and ZK property of the assumed protocol Π. But the soundness is quite challenging to establish. Firstly, as a ZK simulator, $\mathcal{S}$ has the power to rewind V^*: for example, after getting V^*'s response in the third round of the protocol, $\mathcal{S}$ can rewind to the first round and ask V^*'s response on a different first prover message. The key idea behind [CCLY21, BL02] is to ask V^* to abort and output "rejection" with high probability at each round, so that there are roughly only k $\mathcal{S}$'s messages to the V^* that are eventually not aborted. Such a randomly-aborting V^* essentially nullifies $\mathcal{S}$'s ability of rewinding. If the assumed $\mathcal{S}$ could still fool V^* to accept some $x \notin \mathcal{L}$ in this setting, then $\mathcal{S}$ could be used as a "straight-line" malicious prover to break the soundness of the assumed protocol Π, reaching a contradiction.

Barriers and Solutions for Lifting [CCLY21] to Fully Quantum Scenario. The above framework was initially implemented in the classical setting by [BL02]. Significant additional ideas and efforts are already necessary to elevate it to the post-quantum scenario, as detailed in [CCLY21]. However, as we will elaborate below, the V^* designed in [CCLY21] is essentially *classical* and *deterministic* still. The quantumness lies only in the fact that $\mathcal{S}$ can run V^* *in superpostion*. [CCLY21] resolved related issues via a creative application of the *Measure-and-Reprogram* technique from [DFM20], by measuring k randomly selected quantum queries from $\mathcal{S}$ to execute the above algorithm. The classical nature of the honest protocol in [CCLY21] makes it easier to leverage the [BL02] paradigm, mainly due to the following reasons:

1. V^* uses a (quantum-accessible) hash function, modeled as a quantum random oracle, to evaluate on the classical prover messages at each round, determining whether V^* needs to prematurely abort. Apparently, it is pivotal for the original protocol's messages to be classical to directly adapt this randomly-aborting technique.
2. When not aborting, V^* acts the same as an honest classical verifier. One can thus fix its randomness tape to make it deterministic. Therefore, when $\mathcal{S}$ makes the same query for more than one times, it always gets the same response. This property plays an important role in the soundness proof. However, when the honest protocol is fully quantum, we cannot derandomize the verifier's behavior as above due to the inherent random nature of quantum computation.

The above two reasons, in turn, represent obstacles we need to conquer when further lifting the impossibility result to the fully quantum setting.

To address Item 1, we "dequantize" the communication by transforming any fully quantum honest protocol into a protocol where verifier and prover pre-share EPR pairs and send each other classical messages via quantum teleportation. We work in an intermediate model where $\mathcal{S}$ is required to honestly use these EPRs in its simulation so that any protocol secure in this model will imply a protocol secure in the fully quantum model. It then suffices to show an impossibility result in this intermediate model. In this way, we could recover V^*'s aborting behavior by evaluating the prover messages, which are classical now, with its random oracle. But issues on the inherently random quantum behaviors in the above Item 2 still exist. *Worse, we even complicated the task further by additionally introducing pre-shared entanglement into the honest protocol!*

Fortunately, we managed to develop new analytical tools to characterize the evolution of V^*'s state throughout its interaction with $\mathcal{S}$, *even in the presence of the pre-shared entanglement*. Roughly speaking, we show that the overall state is always in a superposition of a "good" branch and a "bad" branch. The good branch corresponds to the messages and internal verifier state that an honest verifier will execute on to decide whether to accept the input x. The bad branch consists of "garbage" states which result from pre-maturely aborted messages and $\mathcal{S}$'s malicious rewinding behaviors, and we have to handle them carefully.

As a key technical observation, we found that those garbage states maintain certain structures during ZK simulation, and we managed to accurately characterize the that structure in the recursive evolution of the garbage states. This lemma allows us to construct a dummy verifier operator on the bad branch when $\mathcal{S}$ queries V^* or rewinds it. Eventually, these properties guarantee that anything meaningful $\mathcal{S}$ can do appears only in the good branch, which has analogous behaviors to a non-aborting, "straight-line" interaction in the post-quantum setting of [CCLY21].

Due to space constraints, the intuition provided above is a significant oversimplification. Many intricate details must be addressed to complete the proof. We invite readers to refer to the **Technical Overview** in our technical manuscript for a moderately detailed explanation.

References

ACL21. Prabhanjan Ananth, Kai-Min Chung, and Rolando L. La Placa. On the concurrent composition of quantum zero-knowledge. In Tal Malkin and Chris Peikert, editors, *Advances in Cryptology – CRYPTO 2021, Part I*, volume 12825 of *Lecture Notes in Computer Science*, pages 346–374, Virtual Event, August 16–20, 2021. Springer, Heidelberg, Germany. 2

BB84. Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, 1984. 1

BCKM21. James Bartusek, Andrea Coladangelo, Dakshita Khurana, and Fermi Ma. One-way functions imply secure computation in a quantum world. In Tal Malkin and Chris Peikert, editors, *Advances in Cryptology – CRYPTO 2021, Part I*, volume 12825 of *Lecture Notes in Computer Science*, pages 467–496, Virtual Event, August 16–20, 2021. Springer, Heidelberg, Germany. 2

BG20. Anne Broadbent and Alex B. Grilo. QMA-hardness of consistency of local density matrices with applications to quantum zero-knowledge. In *61st Annual Symposium on Foundations of Computer Science*, pages 196–205, Durham, NC, USA, November 16–19, 2020. IEEE Computer Society Press. 2

BGG$^+$90. Michael Ben-Or, Oded Goldreich, Shafi Goldwasser, Johan Håstad, Joe Kilian, Silvio Micali, and Phillip Rogaway. Everything provable is provable in zero-knowledge. In Shafi Goldwasser, editor, *Advances in Cryptology – CRYPTO'88*, volume 403 of *Lecture Notes in Computer Science*, pages 37–56, Santa Barbara, CA, USA, August 21–25, 1990. Springer, Heidelberg, Germany. 1

BJSW16. Anne Broadbent, Zhengfeng Ji, Fang Song, and John Watrous. Zero-knowledge proof systems for QMA. In Irit Dinur, editor, *57th Annual Symposium on Foundations of Computer Science*, pages 31–40, New Brunswick, NJ, USA, October 9–11, 2016. IEEE Computer Society Press. 2

BL02. Boaz Barak and Yehuda Lindell. Strict polynomial-time in simulation and extraction. In *34th Annual ACM Symposium on Theory of Computing*, pages 484–493, Montréal, Québec, Canada, May 19–21, 2002. ACM Press. 1, 2, 3

BMO90. Mihir Bellare, Silvio Micali, and Rafail Ostrovsky. The (true) complexity of statistical zero knowledge. In *22nd Annual ACM Symposium on Theory of Computing*, pages 494–502, Baltimore, MD, USA, May 14–16, 1990. ACM Press. 1

BQSY24. John Bostanci, Luowen Qian, Nicholas Spooner, and Henry Yuen. An efficient quantum parallel repetition theorem and applications, 2024. Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024. 2

BS20. Nir Bitansky and Omri Shmueli. Post-quantum zero knowledge in constant rounds. In Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy, editors, *52nd Annual ACM Symposium on Theory of Computing*, pages 269–279, Chicago, IL, USA, June 22–26, 2020. ACM Press. 1, 2

CCLY21. Nai-Hui Chia, Kai-Min Chung, Qipeng Liu, and Takashi Yamakawa. On the impossibility of post-quantum black-box zero-knowledge in constant round. In *IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS 2021). Also appeared in the 25th Annual Conference on Quantum Information Processing (QIP 2022)*, 2021. 1, 2, 3

CCLY22. Nai-Hui Chia, Kai-Min Chung, Xiao Liang, and Takashi Yamakawa. Post-quantum simulatable extraction with minimal assumptions: Black-box and constant-round. In Yevgeniy Dodis and Thomas Shrimpton, editors, *Advances in Cryptology – CRYPTO 2022, Part III*, volume 13509 of *Lecture Notes in Computer Science*, pages 533–563, Santa Barbara, CA, USA, August 15–18, 2022. Springer, Heidelberg, Germany. 2

CCY21. Nai-Hui Chia, Kai-Min Chung, and Takashi Yamakawa. A black-box approach to post-quantum zero-knowledge in constant rounds. In Tal Malkin and Chris Peikert, editors, *Advances in Cryptology – CRYPTO 2021, Part I*, volume 12825 of *Lecture Notes in Computer Science*, pages 315–345, Virtual Event, August 16–20, 2021. Springer, Heidelberg, Germany. 2

DFM20. Jelle Don, Serge Fehr, and Christian Majenz. The measure-and-reprogram technique 2.0: Multi-round fiat-shamir and more. In Daniele Micciancio and Thomas Ristenpart, editors, *Advances in Cryptology – CRYPTO 2020, Part III*, volume 12172 of *Lecture Notes in Computer Science*, pages 602–631, Santa Barbara, CA, USA, August 17–21, 2020. Springer, Heidelberg, Germany. 3

DGJ$^+$20. Yfke Dulek, Alex B. Grilo, Stacey Jeffery, Christian Majenz, and Christian Schaffner. Secure multi-party quantum computation with a dishonest majority. In Anne Canteaut and Yuval Ishai, editors, *Advances in Cryptology – EUROCRYPT 2020, Part III*, volume 12107 of *Lecture Notes in Computer Science*, pages 729–758, Zagreb, Croatia, May 10–14, 2020. Springer, Heidelberg, Germany. 2

DNS12. Frédéric Dupuis, Jesper Buus Nielsen, and Louis Salvail. Actively secure two-party evaluation of any quantum operation. In Reihaneh Safavi-Naini and Ran Canetti, editors, *Advances in Cryptology – CRYPTO 2012*, volume 7417 of *Lecture Notes in Computer Science*, pages 794–811, Santa Barbara, CA, USA, August 19–23, 2012. Springer, Heidelberg, Germany. 2

GIK+23. Riddhi Ghosal, Yuval Ishai, Alexis Korb, Eyal Kushilevitz, Paul Lou, and Amit Sahai. Hard languages in NP ∩ conp and NIZK proofs from unstructured hardness. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 1243–1256. ACM, 2023. 1

GLSV21. Alex B. Grilo, Huijia Lin, Fang Song, and Vinod Vaikuntanathan. Oblivious transfer is in MiniQCrypt. In Anne Canteaut and François-Xavier Standaert, editors, *Advances in Cryptology – EUROCRYPT 2021, Part II*, volume 12697 of *Lecture Notes in Computer Science*, pages 531–561, Zagreb, Croatia, October 17–21, 2021. Springer, Heidelberg, Germany. 2

GMR85. Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof-systems (extended abstract). In *17th Annual ACM Symposium on Theory of Computing*, pages 291–304, Providence, RI, USA, May 6–8, 1985. ACM Press. 1

GSY19. Alex Bredariol Grilo, William Slofstra, and Henry Yuen. Perfect zero knowledge for quantum multiprover interactive proofs. In David Zuckerman, editor, *60th Annual Symposium on Foundations of Computer Science*, pages 611–635, Baltimore, MD, USA, November 9–12, 2019. IEEE Computer Society Press. 1

IY88. Russell Impagliazzo and Moti Yung. Direct minimum-knowledge computations. In Carl Pomerance, editor, *Advances in Cryptology – CRYPTO'87*, volume 293 of *Lecture Notes in Computer Science*, pages 40–51, Santa Barbara, CA, USA, August 16–20, 1988. Springer, Heidelberg, Germany. 1

KKMV09. Julia Kempe, Hirotada Kobayashi, Keiji Matsumoto, and Thomas Vidick. Using entanglement in quantum multi-prover interactive proofs. *Comput. Complex.*, 18(2):273–307, 2009. 2

KW00. Alexei Kitaev and John Watrous. Parallelization, amplification, and exponential time simulation of quantum interactive proof systems. In *32nd Annual ACM Symposium on Theory of Computing*, pages 608–617, Portland, OR, USA, May 21–23, 2000. ACM Press. 2

SV97a. Amit Sahai and Salil P. Vadhan. A complete promise problem for statistical zero-knowledge. In *38th Annual Symposium on Foundations of Computer Science*, pages 448–457, Miami Beach, Florida, October 19–22, 1997. IEEE Computer Society Press. 1

SV97b. Amit Sahai and Salil P. Vadhan. Manipulating statistical difference. In Panos M. Pardalos, Sanguthevar Rajasekaran, and José Rolim, editors, *Randomization Methods in Algorithm Design, Proceedings of a DIMACS Workshop, Princeton, New Jersey, USA, December 12-14, 1997*, volume 43 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 251–270. DIMACS/AMS, 1997. 1

Yan22. Jun Yan. General properties of quantum bit commitments (extended abstract). In Shweta Agrawal and Dongdai Lin, editors, *Advances in Cryptology - ASIACRYPT 2022 - 28th International Conference on the Theory and Application of Cryptology and Information Security, Taipei, Taiwan, December 5-9, 2022, Proceedings, Part IV*, volume 13794 of *Lecture Notes in Computer Science*, pages 628–657. Springer, 2022. 2