Toward Quantum Proof-of-Knowledge from Anticoncentration in Deep Quantum Circuits

Darrell L. Young
Darwin Deason Institute for Cyber Security
Southern Methodist University
Dallas, TX, USA
darrell.young@smu.edu

Mitchell A. Thornton
Darwin Deason Institute for Cyber Security
Southern Methodist University
Dallas, TX, USA
mitch@smu.edu

***Abstract*—Quantum circuits that exhibit strong *anticoncentration*—where output probabilities are broadly distributed and difficult to reproduce classically under standard hardness assumptions—provide a promising source of computational complexity for quantum-native cryptographic design. We investigate whether this property can be harnessed for authentication and proof-of-knowledge style protocols based on parameterized quantum circuits. Numerical experiments using four-qubit, ten-layer circuits demonstrate stable anticoncentration-inspired behavior together with reproducible entropy-based verification metrics. The resulting prototype achieves 95% authentication success in simulation with low observed false-positive rates under the tested conditions. We outline a path toward quantum-secure proof and authentication protocols suitable for fuller journal treatment.**

***Index Terms*—Anticoncentration, quantum circuits, proof-of-knowledge, quantum cryptography, unitary designs.**

I. Introduction

Random quantum circuits have become an important tool in quantum information science, including benchmarking, scrambling studies, and proposals for quantum computational advantage. A key phenomenon in this setting is *anticoncentration*, where output probabilities become broadly distributed in a way associated with chaotic circuit behavior. That regime can also motivate useful cryptographic protocol design.

More broadly, hash-like primitives underpin a wide range of modern cryptographic and systems applications, including commitments, authentication, digital signatures, integrity checking, authenticated data structures, and blockchain-style ledgers. For that reason, even a modest proof-of-concept in this direction is potentially interesting beyond the immediate proof-of-knowledge setting. The present conference paper does not claim a fully formal quantum hash construction; rather, it asks whether secret-parameterized scrambling circuits and their resulting measurement statistics might serve as a useful design ingredient for future commitment-like and proof-oriented quantum primitives in that broader landscape.

The same properties that make quantum circuits difficult to simulate classically suggest their potential utility for cryptography: (1) output probabilities spread across exponentially many bitstrings, (2) reproducing the correct distribution may be computationally difficult under standard hardness assumptions, and (3) recovering secret-dependent circuit behavior from limited observations appears nontrivial. These features motivate the use of quantum circuit outputs as candidate commitment-like objects in authentication and proof-of-knowledge style protocols.

979-8-3315-5193-3/26/$31.00 © 2026 IEEE

Our Contributions. This paper positions anticoncentration in deep quantum circuits as a promising design ingredient for quantum authentication and proof-of-knowledge style protocols. We implement a proof-of-concept protocol achieving 95% authentication success in ∼15 ms on 4-qubit simulators under the tested simulation conditions. We view these results as an outline for fuller journal formalization rather than a complete simulator-based zero-knowledge treatment.

II. Background

A. Anticoncentration in Quantum Circuits

Definition 1 (Anticoncentration [1])**.** *Let U be a unitary drawn from a distribution μ on $U(2^n)$, and let $|0\rangle = |0\rangle^{\otimes n}$. The output probabilities $p_x = |\langle x|U|0\rangle|^2$ for $x \in \{0,1\}^n$ anticoncentrate if there exist universal constants $\alpha, \beta > 0$ such that:*

$$\Pr_{U\sim\mu}\left[p_x \geq \frac{\alpha}{2^n}\right] > \beta. \tag{1}$$

For Haar-random unitaries, output probabilities follow the Porter–Thomas distribution with exponential tail, ensuring anticoncentration. Brandão et al. [2] showed that random circuits with nearest-neighbor gates form ϵ-approximate 2-designs in depth $O(n\log(1/\epsilon))$, which implies anticoncentration via the following cited result from Hangleiter et al. [1] which connects approximate unitary 2-designs to anticoncentration:

Theorem 2 (2-Designs Anticoncentrate [1])**.** *Let μ be an ϵ-approximate unitary 2-design on $U(2^n)$. Then for $0 \leq \alpha \leq 1$:*

$$\Pr_{U\sim\mu}\left[|\langle x|U|0\rangle|^2 > \frac{\alpha(1-\epsilon)}{2^n}\right] \geq \frac{(1-\alpha)^2(1-\epsilon)^2}{2(1+\epsilon)}. \tag{2}$$

In the present conference version, this theorem is used primarily to motivate the circuit architecture and scrambling regime rather than to claim a complete hardness characterization for the specific small-scale implementation studied experimentally.

B. Zero-Knowledge Proofs

A zero-knowledge proof [3] allows a prover $\mathcal{P}$ to convince a verifier $\mathcal{V}$ that a statement is true without revealing additional information. In the standard formal setting, such protocols are defined through completeness, soundness, and a simulator-based zero-knowledge property. In this conference paper, we use these notions as guiding design targets, while reserving full formal treatment for subsequent journal work.

III. Theoretical Framework

For classical ZKPs a commitment can be a hash with a salt. The quantum commitment is similar but has many open research questions. We observed low collision but have not explored its pre-image resistance (given an output difficult to find input that produced it) or its second-preimage (given one input find a different input with same output), or its avalanche (small change in input produces big change in output) although we did observe some evidence the quantum circuit may have avalanche properties. The quantum commitment compresses into a low-dimensional structure but with less-understood structure. The quantum commitment may leak patterns depending on protocol. More experiments are needed.

Definition 3 (Quantum Circuit Commitment Function)**.** *Let $U(\theta)$ be a parameterized quantum circuit on n qubits. Define $C_{\mathrm{QC}} : \{0,1\}^* \to \mathbb{R}^{2^n}$ as: (1) Encode input s to parameters $\theta = \mathrm{Encode}(s)$; (2) Prepare $|\psi(\theta)\rangle = U(\theta)|0\rangle^{\otimes n}$; (3) Measure M times to obtain frequencies f; (4) Output $C_{\mathrm{QC}}(s) = f$.*

In the present conference version, C_{QC} should be understood as a commitment-like protocol map rather than as a fully formal hiding-and-binding commitment scheme.

A working hypothesis of this paper is the **separation between quantum circuit properties and classical encoding properties**: the quantum circuit architecture appears to govern the protocol's scrambling and verification behavior. The classical encoding is expected to affect input-sensitivity metrics such as avalanche behavior and distinguishability, although a fuller study is deferred to future work.

Remark 4 (Heuristic One-Wayness Motivation)**.** *If $U(\theta)$ operates in a regime associated with approximate 2-design behavior and anticoncentration, then recovering secret-dependent input information from accepted response statistics may be classically difficult under standard sampling-hardness assumptions. We treat this as heuristic motivation in the present conference version; a formal reduction-based one-wayness argument is reserved for subsequent journal treatment.*

Classical hash pre-image resistance is a standard example of one-wayness.

IV. Protocol and Implementation

Before describing the protocol, we distinguish three related objects. First, the prover's private secret s is the witness-like input retained across sessions. Second, the classical encoding map $\theta = \mathrm{Encode}(s)$ converts that secret into the parameters of the quantum circuit $U(\theta)$. Third, repeated measurements of the circuit output produce an empirical frequency vector, which serves as the commitment-like protocol object. For a fixed secret and circuit, the underlying ideal output distribution is stable, while the finite-shot empirical frequency vector is only an approximation to it and therefore varies across repeated executions.

Circuit Architecture. We use $n = 4$ qubits, $L = 10$ layers with R_Y/R_Z rotations per qubit per layer and circular CNOT entanglement ($i \to i{+}1 \bmod 4$), yielding $p = 80$ parameters. This depth/topology is intended to place the circuit in a strongly scrambling regime; approximate 2-design results for local random circuits motivate this choice [2].

Setup Phase. Public parameters are the circuit architecture $\mathcal{C}$, the measurement count $M = 1000$, the acceptance tolerance $\delta = 0.10$, and the round count $k = 3$.

Registration. The prover chooses a private secret s, computes the secret-dependent public reference object $C = C_{\mathrm{QC}}(s)$, and registers C publicly.

Authentication Goal. The verifier seeks to check that the prover knows a hidden secret s consistent with the registered public reference object C, without requiring the prover to reveal s directly.

Authentication. In the present proof-of-concept implementation, the verifier issues a challenge specifying a subset of qubits to be checked, together with a fresh nonce carried in the challenge metadata. The prover then re-executes the secret-parameterized circuit, computes empirical measurement counts, and returns the marginal distribution over the challenged qubit subset. The verifier compares that marginal, within tolerance δ, to the corresponding marginal extracted from the registered public reference object C. Thus, C denotes the registered public reference object, r_i a fresh challenge nonce, and δ the tolerance used in approximate consistency checks. In the current codebase, the nonce serves as a freshness parameter and challenge identifier, while fuller nonce-conditioned response generation and branch-specific verification are deferred to future work.

V. Experimental Results

Implementation in Qiskit [4] with statevector simulation yields the results in Table I. The 4-qubit simulation is intended only as a proof-of-concept demonstration of protocol mechanics and metric behavior, not as a practical hardness benchmark. The current proof-of-concept codebase evaluates both challenged-marginal verification and a simplified direct distribution-comparison variant used for bulk authentication testing; accordingly, the reported metrics should be interpreted as prototype authentication indicators rather than as measurements of a fully formalized challenge–response proof system.

The average authentication time reflects the computational cost of the present proof-of-concept implementation on small simulated instances and should be interpreted as a prototype runtime indicator rather than as a deployment benchmark. The 95.0% success rate for correct passwords provides empirical evidence of completeness-like behavior under the chosen

TABLE I
SUMMARY OF PROOF-OF-CONCEPT EXPERIMENTAL INDICATORS REPORTED FOR THE CURRENT IMPLEMENTATION. AUTHENTICATION TIMING AND ACCURACY WERE OBTAINED FROM BULK TESTING IN A SIMPLIFIED DISTRIBUTION-COMPARISON CONFIGURATION.

Metric	Value
Avg. authentication time	14.92 ± 0.71 ms
Success rate (correct password)	95.0%
False positive rate	2.0%
Collisions observed	0 / 1000

tolerance and shot count. The observed 2.0% false-positive rate should not be interpreted as a fundamental limit of the protocol. If it is driven primarily by finite-shot statistical variation, then it can in principle be reduced by increasing the number of measurements, tightening acceptance thresholds, and/or adding rounds. If instead it reflects structural overlap in the current commitment and response design, then protocol modifications will be required. Distinguishing statistical from structural error is part of the intended journal extension. Finally, the absence of observed collisions in the reported small-scale tests should be viewed only as a preliminary empirical indicator and not as evidence of formal collision resistance.

VI. SECURITY DISCUSSION

Completeness (empirical). If the prover knows s, the verifier accepts with high probability, up to finite-shot sampling variation and the selected tolerance parameters.

Soundness (heuristic). A cheating prover who does not know s must either guess responses or reproduce verifier-accepted response statistics without access to the secret-dependent commitment structure. We hypothesize that this task becomes more difficult in sufficiently scrambling regimes associated with anticoncentration and sampling hardness, but a formal reduction-based soundness proof is left to future work.

Injectivity and protocol refinement. The present conference version does not assume that the commitment-like map C_{QC} is injective. Distinct secrets may in principle induce identical or near-identical verifier-visible response statistics due to gate periodicity, circuit symmetries, measurement-only observation, or finite-dimensional compression of the observable output. In addition, the classical secret-to-parameter encoding may itself affect how strongly nearby secrets separate before entering the circuit, although a fuller study of encoding choices is deferred to the journal version. The encoding method affects how secrets are mapped into parameter space, but the dominant risks to injectivity and verifier-side distinguishability appear to arise from the structure of the observable map itself, especially measurement-only observation and compression of the verifier-visible statistics.

The current authentication protocol already incorporates a fresh nonce in the challenge metadata, so that each proof is tied to a specific issued challenge rather than reused across rounds unchanged. In the present proof-of-concept implementation, however, this nonce does not yet define a nonce-conditioned family of quantum response maps of the form $C_{\mathrm{QC}}(s\|r)$; instead, it serves as a freshness parameter and challenge identifier. The journal version will investigate whether a stronger nonce-conditioned construction can improve effective injective behavior by reducing collisions and replay-style overlap. It will further study how the classical input-encoding method affects avalanche behavior, distinguishability, and approximate collision structure, and whether the observed false-positive behavior is primarily statistical or structural, by varying shot count, tolerance, round count, and encoding choice, and by comparing exact statevector distributions against finite-shot empirical frequencies.

The present conference version has several important limitations. First, the reported experiments use only small classically simulable circuits and therefore do not by themselves demonstrate practical hardness. Second, the security discussion is heuristic and does not yet include a formal reduction from impersonation to a standard sampling problem. Third, the reported leakage and error metrics are empirical protocol indicators rather than substitutes for simulator-based zero-knowledge analysis. These gaps motivate the fuller journal treatment.

VII. CONCLUSION

Anticoncentration in deep quantum circuits appears to be a promising design ingredient for quantum authentication and proof-of-knowledge style protocols. The present conference version provides a proof-of-concept protocol and small-scale simulation results, while questions involving circuit-family choice, input-encoding effects, injectivity, and formal security analysis are deferred to the journal extension. More broadly, if such commitment-like constructions can be formalized successfully, they may prove relevant not only to proof-of-knowledge settings but also to a wider family of hash-like cryptographic and systems primitives. Future work includes sharper formal security definitions and larger-scale evaluation. It also includes exploring whether a strongly mixing or approximate-2-design-like circuit family can be realized on a programmable photonic frequency-bin architecture, building on recent photonic processor work by Thornton and collaborators [5].

DATA AVAILABILITY

A research codebase and supporting experimental materials exist for the proof-of-concept results reported in this paper. As part of the conference-to-journal development process, these materials may be shared by the authors upon reasonable request. A cleaned and better-documented public release is planned after further formalization and validation.

ACKNOWLEDGMENTS

We thank Dr. Marco Cerezo (Los Alamos National Laboratory) for clarifying the distinction between barren plateaus and anticoncentration.

REFERENCES

[1] D. Hangleiter, J. Bermejo-Vega, M. Schwarz, and J. Eisert, "Anticoncentration theorems for schemes showing a quantum speedup," *Quantum*, vol. 2, p. 65, May 2018. [Online]. Available: https://doi.org/10.22331/q-2018-05-22-65

[2] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, "Local random quantum circuits are approximate polynomial-designs," *Communications in Mathematical Physics*, vol. 346, no. 1, pp. 397–434, 2016. [Online]. Available: https://doi.org/10.1007/s00220-016-2706-8

[3] S. Goldwasser, S. Micali, and C. Rackoff, "The knowledge complexity of interactive proof systems," *SIAM Journal on Computing*, vol. 18, no. 1, pp. 186–208, 1989.

[4] Qiskit contributors, "Qiskit: An open-source framework for quantum computing," 2021. [Online]. Available: https://qiskit.org/

[5] D. L. MacFarlane, H. Shahoei, M. V. Paul, R. M. Tuller, and M. A. Thornton, "Parallel architecture of a frequency comb qudit quantum processor," *arXiv preprint arXiv:2508.15981*, 2025.