Quantum Advantage for Two-Party Differential Privacy

Daniel Alabi[1], Emil T. Khabiboulline[2,3]
[1]Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[2]Joint Center for Quantum Information and Computer Science, NIST/University of Maryland, College Park, MD 20742, USA
[3]Joint Quantum Institute, NIST/University of Maryland, College Park, MD 20742, USA
alabid@illinois.edu, ekhabibo@umd.edu

***Abstract*—We introduce information-theoretically private quantum protocols for two-party Hamming distance when both parties must output the same estimate. Classically, for input length n, information-theoretic protocols require $\Omega(\sqrt{n})$ error under pure differential privacy and $\Omega(\sqrt{n}/\log n)$ error under strong approximate differential privacy, whereas computational security permits $O(1)$ error. In Klauck's honest, nonpreemptive, message-preserving model, we give an $O(n)$-communication quantum protocol with pure ε quantum differential privacy (QDP) and expected error at most $\frac{2}{\sinh \varepsilon} + \gamma$, for every $\gamma > 0$. For approximate (ε, δ) QDP, an exact finite-cycle hockey-stick calculation yields strictly smaller error, while preserving the $O(1)$-versus-$\Omega(\sqrt{n}/\log n)$ separation for $\delta = o(1/n)$. Thus, quantum communication achieves $O(1)$ information-theoretic error, matching the accuracy available classically only under computational assumptions.**

The main construction uses a guarded coherent round trip and an equal-Gram rigidity principle that prevents an honest player from retaining input-dependent complementary information. We also separate this model from weaker prescribed-channel privacy, which already admits an exact classical realization, and from fully retention-robust security, against which measurement-and-abort attacks remain possible. Therefore, we identify preservation of non-orthogonal quantum messages as a resource for privacy, but leave open whether a separation exists in the malicious setting.

***Index Terms*—differential privacy, quantum communication, two-party protocols, Hamming distance**

I. Introduction

Differential privacy (DP) is best understood in the trusted curator setting: a curator collects the data, evaluates a low-sensitivity statistic, and releases a noisy answer [1], [2]. For Hamming distance, whose sensitivity is one, this gives expected error $O(1/\varepsilon)$, independent of the input length n [3]. In many applications, however, no trusted curator exists [4]–[6]. Instead, two mutually distrustful parties hold separate inputs and wish to compute a common statistic [7], [8] while protecting each input from the other party. Requiring privacy simultaneously against both participants fundamentally changes the problem: the transcript needed to coordinate the computation may itself reveal information that a trusted curator would never release [9], [10].

McGregor, Mironov, Pitassi, Reingold, Talwar, and Vadhan initiated the systematic study of lower bounds in this two-party setting [3], [11]. For Hamming distance, they showed that classical information-theoretic privacy is dramatically less accurate than central differential privacy: for suitable constant privacy parameters, purely differentially private protocols require $\Omega(\sqrt{n})$ error, and strong approximate privacy requires $\Omega(\sqrt{n}/\log n)$ error [3], [11]. By contrast, secure function evaluation can emulate the trusted curator and achieve $O(1)$ error when privacy is required only against computationally bounded parties [12], [13].

This led McGregor et al. [3], [11] to pose the following question:

> *What is the largest gap in accuracy between information-theoretic and computationally differentially private two-party protocols?*

They showed that a linear separation is possible for certain functions. For natural functions such as Hamming distance, however, their results expose a particularly important gap:

$$\underbrace{O(1)}_{\substack{\text{central or computational}\\ \text{differential privacy}}} \quad \text{versus} \quad \underbrace{\widetilde{\Omega}(\sqrt{n})}_{\substack{\text{classical information-theoretic}\\ \text{two-party privacy}}} .$$

This raises a basic question left outside the classical theory: *is the loss in accuracy inherent to information-theoretic privacy itself, or does it arise from the classical, copyable nature of communication?*

The distinction matters both conceptually and cryptographically. If the classical lower bound extended to quantum protocols, then to obtain the accuracy of the central model in the two-party differential privacy would inherently require computational assumptions or trusted setup [14]. Conversely, an information-theoretic quantum protocol with $O(1)$ error would show that computational assumptions are not the only way to cross the classical accuracy barrier [4], [12], [13]. It would identify noncopyability [15] and coherent information flow as genuine resources for achieving differential privacy, not merely for reducing communication complexity or performing cryptographic key distribution [16]–[18]. Moreover, determining exactly which adversarial models permit such an advantage is essential: a separation obtained only by discarding prescribed workspace has a different meaning from one that remains secure when a party may retain arbitrary quantum information [19].

In this work, we show that the answer is model dependent. Alice holds $x \in \{0,1\}^n$, Bob holds $y \in \{0,1\}^n$, and both parties must produce the same random estimate $\widehat{d}$ of

$$d = \mathrm{HD}(x,y) = \big|\{i : x_i \neq y_i\}\big|.$$

Under Klauck's honest, nonpreemptive, message-preserving definition [20], we give the first information-theoretic quantum protocol that crosses the classical two-party accuracy barrier. For every $\gamma > 0$, the protocol is pure ε-quantum differentially private, communicates $O(n)$ qubits and bits, and satisfies

$$\mathbb{E}\Big[|\widehat{d} - d|\Big] < \frac{2}{\sinh \varepsilon} + \gamma.$$

Thus, for constant ε, quantum communication achieves $O(1)$ error, whereas every classical information-theoretic protocol under the same honesty model requires $\Omega(\sqrt{n})$ error. In this model, quantum communication therefore (information-theoretically) recovers the accuracy that is available classically only through computationally secure evaluation [21].

We also obtain a strict improvement under approximate differential privacy. Rather than using the pure-DP noise parameter ε, we compute the privacy using the hockey-stick divergence and select the minimum $\alpha_n^\star(\varepsilon,\delta) > \varepsilon$ for a particular δ-budget [22]. The resulting protocol is (ε,δ)-quantum differentially private and has expected error at most

$$\frac{2}{\sinh \alpha_n^\star(\varepsilon,\delta)} + \gamma < \frac{2}{\sinh \varepsilon} + \gamma.$$

For $\delta = o(1/n)$, this gives $O(1)$ quantum error compared with the classical lower bound $\Omega(\sqrt{n}/\log n)$.

The source of the separation is that classical messages may be copied without disturbance, so a classical honest observer may retain the complete transcript [19]. A quantum observer constrained to return the prescribed pure message cannot do the same. Our construction exploits an equal-Gram rigidity principle [23]–[25]: exact preservation of a family of non-orthogonal pure messages forces the observer's complementary state to be independent of the protected input. A small, input-independent coherent guard makes this rigidity exact, moving the approximation entirely into a rare fallback event that affects accuracy rather than privacy.

Finally, we delineate the boundary of this advantage. Under a fixed prescribed-channel view model, reversible classical computation already realizes the private functionality exactly, so that view model cannot support a meaningful quantum–classical separation [26]. At the other extreme, our protocol is not secure against arbitrary retention-capable quantum strategies: a receiver that measures the first message and aborts can distinguish adjacent inputs. We formalize the corresponding retention obstruction and provide a classical randomized-response baseline for this stronger model, which matches the accuracy guarantees of McGregor et al. [3], [11]. Thus, quantum communication resolves the Hamming distance accuracy gap under Klauck's exact message-preserving honesty condition, while the question of whether it improves accuracy under fully malicious security remains open.

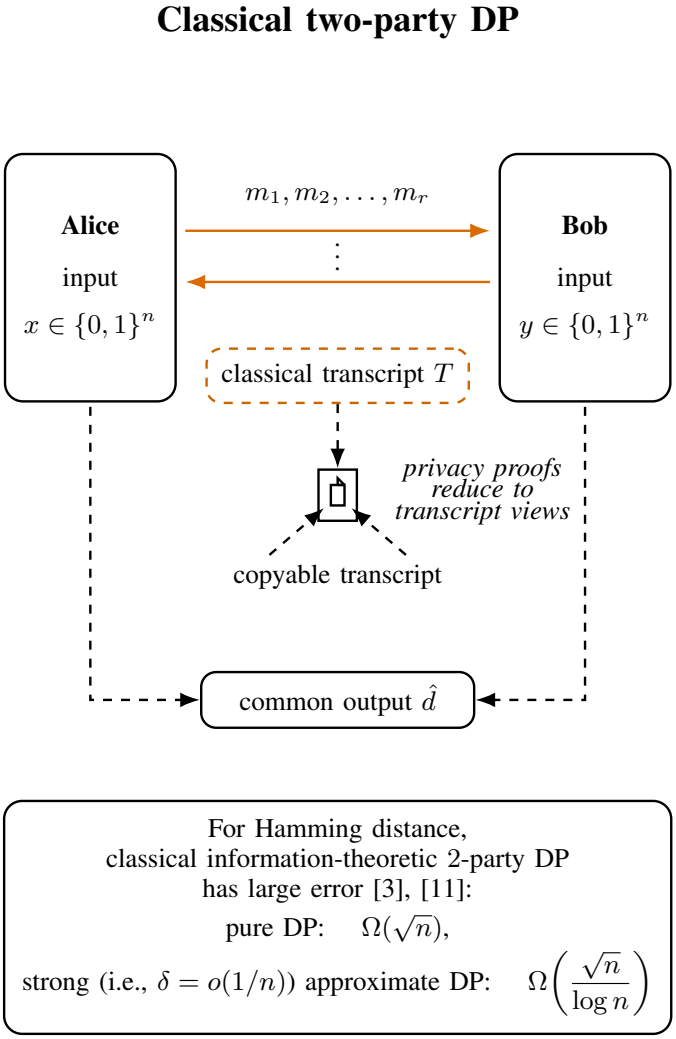


Fig. 1. Classical two-party differential privacy. Classical privacy proofs reduce to transcript views. The interaction transcript can be copied.

We consider a range of view models: PC, KHNP, and RR which stand for *prescribed-channel*, *Klauck's honest nonpreemptive view*, and *retention-robust* views, respectively. In the PC view model, the parties execute the fixed local channels specified by the protocol, including their prescribed discard operations. Conceptually, KHNP captures an honest-but-curious setting: a party may retain arbitrary private information, but must send the prescribed messages and complete the protocol. In contrast, our retention-robust (RR) model is malicious, allowing arbitrary CPTP deviations, including measurement, message modification, retention, and abort.

A. Contributions

In Table I, we summarize our main results and conclusions. We work in the standard idealized communication model allowing exact real-valued private randomness and arbitrary exact local unitaries; local computational complexity and finite-

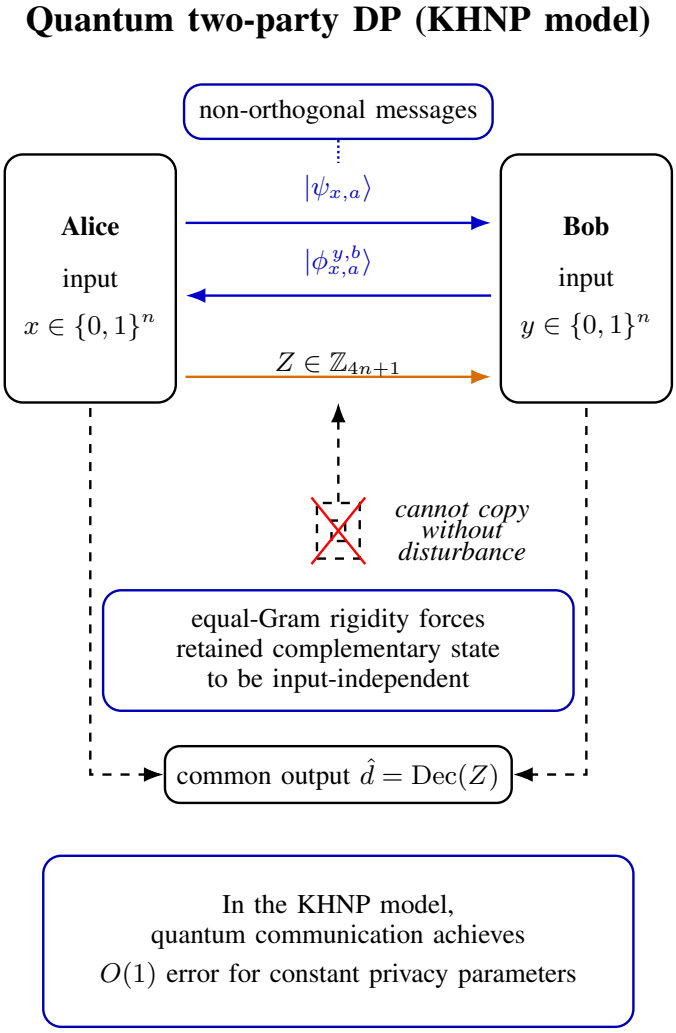


Fig. 2. Quantum two-party differential privacy in the KHNP model. Exact preservation of non-orthogonal messages prevents transcript copying and yields the quantum advantage.

precision synthesis are not accounted for. Our main results are as follows.

1) We show that a common-output functionality, via distributed noise generation, satisfies ε-DP and has error at most $2/\sinh\varepsilon$. The proof of privacy and accuracy follows from the construction via the use of geometric arguments and interval decoding.
2) Fixed prescribed channels (Definition 2) can realize this functionality exactly with a classical reversible protocol. In Klauck's stronger model, a round trip protocol gives pure ε-QDP, error at most $2/\sinh\varepsilon + \gamma$, and $O(n)$ communication; exact noise calibration gives strictly smaller noise for a protocol that satisfies (ε, δ)-QDP.
3) We define and analyze the Klauck-honest nonpreemptive view (KHNP). See Definition 3. The classical version of KHNP privacy includes the transcript-copying observer, so the classical pure-DP $\Omega(\sqrt{n})$ lower bound transfers and gives a same-model quantum accuracy separation.
4) We formalize retention-robust QDP (Definition 4), accounting for both adversarial strategies and side information.
5) Randomized response with geometric output noise gives pure ε-QDP in the retention-robust class (and hence (ε, δ)-QDP), with an explicit error bound stated in Theorem 10.

Figures 1 and 2 summarize the differences and similarities between the protocols that satisfy two-party classical DP versus two-party quantum DP.

B. Main Results and Implications

Our main result is an information-theoretic quantum advantage for the two-party computation of Hamming distance under Klauck's honest, nonpreemptive, message-preserving privacy model [20]. Alice holds $x \in \{0,1\}^n$, Bob holds $y \in \{0,1\}^n$, and both parties must output the same estimate of $d = \mathrm{HD}(x, y)$. The protocol uses quantum communication but requires no computational assumption, trusted setup, prior entanglement, bounded-storage assumption, or secure erasure [27].

At the *ideal* functional level, Alice and Bob independently sample cyclic geometric noise shares $A, B \in \mathbb{Z}_{4n+1}$ and release

$$Z = d + A + B \pmod{4n+1}.$$

Both parties apply the same interval decoder to Z. Neither party knows the other party's noise share, so neither can remove all of the privacy noise from its own view. The resulting common-output functionality is pure ε-differentially private and has expected error at most

$$\frac{2}{\sinh\varepsilon}.$$

The factor of two reflects the common-output requirement for the mistrustful parties: privacy for both inputs requires both independently generated noise shares to appear in the same released value.

The quantum protocol implements this functionality through a *guarded coherent round trip* [20]. By *guarded*, we mean that every input encoding contains a small common fallback component; by *coherent*, we mean that Alice sends a superposition rather than a classical input message; and by *round trip*, we mean that Bob coherently computes the answer into the message and returns the entire quantum state to Alice. So, Alice prepares a superposition containing a dominant data-dependent branch and a small, input-independent *guard branch*. Bob coherently computes the noisy Hamming distance release, appends it to the message, and returns the entire prescribed pure state. Alice then measures the branch and output registers and communicates the resulting classical release to Bob. Except on the input-independent fallback branch, both parties obtain the output of the ideal distributed-noise functionality.

Theorem 1 (Pure QDP for Klauck-honest players. Informal Version of Theorem 7)**.** *For every $n \geq 1$, $\varepsilon > 0$, and $\gamma > 0$, there exists a protocol that provides information-theoretic pure*

TABLE I

SUMMARY OF THE ESTABLISHED PRIVACY AND ERROR BOUNDS. WE ANALYZE THESE BOUNDS UNDER DIFFERENT VIEW MODELS: PC IS THE PRESCRIBED CHANNEL VIEW (DEFINITION 2); KHNP IS THE KLAUCK-HONEST NONPREEMPTIVE VIEW (DEFINITION 3); RR IS THE RETENTION-ROBUST ADVERSARIAL VIEW (DEFINITION 4). CPTP STANDS FOR COMPLETELY POSITIVE TRACE-PRESERVING.

View model	Privacy	Status and utility gap (error)
PC	pure ε	Proposition 6: exact classical common output, $2/\sinh\varepsilon$.
KHNP	pure ε	Theorem 7: quantum common output, $< 2/\sinh\varepsilon + \gamma$, $O(n)$ communication.
KHNP	(ε, δ)	Theorem 8: $\leq 2/\sinh(\alpha_n^\star) + \gamma$ for some $\alpha_n^\star > \varepsilon$, $O(n)$ communication.
Classical KHNP	pure and approximate	Corollary 9: $\Omega(\sqrt{n})$ pure error and $\Omega(\sqrt{n}/\log n)$ error for $\delta = o(1/n)$, versus $O(1)$ quantum error.
RR, all CPTP	pure ε	Theorem 10: $\sqrt{n}/(2\sinh(\varepsilon/2)) + 1/(2\sinh^2(\varepsilon/2))$.
RR, all CPTP	(ε, δ)	The pure protocol of Theorem 10 gives the same bound for every $\delta \geq 0$; improvement is open.

ε-QDP for the complete terminal view of every Klauck-honest nonpreemptive observing strategy. In an honest execution both parties produce a common output,

$$\mathbb{E}|\widehat{d} - \mathrm{HD}(x,y)| < \frac{2}{\sinh\varepsilon} + \gamma,$$

and total communication is $2n + 2 + O(\log n) = O(n)$ qubits and bits in the standard (arbitrary) unitary communication model.

Theorem 7 establishes pure quantum differential privacy for this protocol. For every $n \geq 1$, $\varepsilon > 0$, and $\gamma > 0$, the protocol is pure ε-QDP for the complete terminal view of every Klauck-honest nonpreemptive observing party. It communicates

$$2n + 2 + O(\log n) = O(n)$$

qubits and classical bits and satisfies

$$\mathbb{E}\Big[\Big|\widehat{d} - \mathrm{HD}(x,y)\Big|\Big] < \frac{2}{\sinh\varepsilon} + \gamma.$$

For constant ε and arbitrarily small constant γ, the expected error is therefore $O(1)$, independent of the input length.

This gives a genuine same-model separation from classical information-theoretic protocols. A classical Klauck-honest party may copy every classical message while still sending exactly the prescribed outgoing messages and completing the protocol. Its terminal view can therefore contain the full classical transcript and private randomness. The lower bounds of McGregor et al. consequently apply under the same honesty requirement: classical pure-DP protocols for Hamming distance require $\Omega(\sqrt{n})$ error for suitable constant privacy parameters. Hence, under Klauck's message-preserving model, quantum communication improves the optimal information-theoretic accuracy from polynomial error to constant error:

$$\underbrace{O(1)}_{\text{quantum information-theoretic}} \quad \text{versus} \quad \underbrace{\Omega(\sqrt{n})}_{\text{classical information-theoretic}}.$$

In particular, the quantum protocol attains information-theoretically the accuracy (in the central model) that classical two-party protocols can obtain through computationally secure function evaluation [12], [13].

Theorem 2 (Noise-calibrated approximate KHNP-QDP. Informal Version of Theorem 8)**.** *Let $n \geq 1$, $\varepsilon > 0$, $0 < \delta < 1$, and $\gamma > 0$. There exists a protocol such that every Klauck-honest nonpreemptive terminal view satisfies (ε, δ)-QDP, both honest parties produce a common output, and*

$$\mathbb{E}|\widehat{d} - d| \leq \frac{2}{\sinh\alpha_n^\star} + \gamma < \frac{2}{\sinh\varepsilon} + \gamma,$$

for some $\alpha_n^\star > \varepsilon$. Communication is $O(n)$ in the (arbitrary) unitary model.

Theorem 8 shows that approximate privacy provides a further, strict accuracy improvement. A direct application of the pure-DP protocol would yield (ε, δ)-QDP with the same noise parameter $\alpha = \varepsilon$, leaving the δ-budget unused. Instead, we compute the exact quantum hockey-stick divergence of the finite cyclic geometric distribution. For every $0 < \delta < 1$, this identifies a unique parameter

$$\alpha_n^\star(\varepsilon, \delta) > \varepsilon$$

for which the ideal functionality is exactly (ε, δ)-QDP. Running the same guarded quantum protocol with noise shares sampled using $\alpha_n^\star(\varepsilon, \delta)$ yields

$$\mathbb{E}\Big[\Big|\widehat{d} - \mathrm{HD}(x,y)\Big|\Big] \leq \tag{1}$$

$$\leq \frac{2}{\sinh\alpha_n^\star(\varepsilon,\delta)} + \gamma \tag{2}$$

$$< \frac{2}{\sinh\varepsilon} + \gamma. \tag{3}$$

Thus approximate privacy does not merely inherit the pure-privacy guarantee: the allowed failure probability is converted directly into a more concentrated noise distribution and strictly better utility. For strong approximate privacy, $\delta = o(1/n)$, the quantum error remains $O(1)$, whereas the corresponding classical information-theoretic lower bound is $\Omega(\sqrt{n}/\log n)$.

These results give a model-dependent answer to the accuracy-gap question of McGregor et al [3], [11]. Under exact message-preserving honesty, quantum communication removes the polynomial accuracy loss suffered by classical information-theoretic protocols. The advantage does not, however, automatically extend to arbitrary malicious or retention-capable quantum

strategies. A receiver allowed to measure the first message and abort can distinguish adjacent inputs. Accordingly, our result should be interpreted as showing that noncopyable communication is an information-theoretic privacy resource under exact message-preservation [15], rather than as a simulation-secure [28] or maliciously secure two-party protocol [29]. Whether quantum communication improves accuracy under fully retention-robust security remains open.

C. Main Techniques

The proof of our main results (Theorem 7 and 8) combines four ingredients: distributed cyclic-geometric noise, coherent guarded communication, equal-Gram rigidity, and exact hockey-stick-divergence calibration.

a) Distributed noise for a common private output: The starting point is a common-output functionality in which the two parties independently sample

$$A, B \sim q_\alpha, \qquad q_\alpha(g) \propto \exp(-\alpha d_m(g,0)),$$

over the cycle $\mathbb{Z}_m$, where $m = 4n+1$ and d_m is circular distance. The functionality releases

$$Z = \mathrm{HD}(x,y) + A + B \pmod m.$$

Because changing one input bit changes the Hamming distance by at most one, the ratio between adjacent output probabilities is controlled by the one-step likelihood ratio of q_α.

Privacy of Alice's input from Bob is provided by Alice's noise share A, which is unknown to Bob. Likewise, privacy of Bob's input from Alice is provided by Bob's noise share B, which is unknown to Alice.

The choice $m = 4n+1$ ensures that the valid distance interval $\{0,\dots,n\}$ is separated from its wraparound copies by a sufficiently long complementary arc. A nearest-interval decoder $\mathrm{Dec}(\cdot)$ therefore satisfies

$$|\mathrm{Dec}(Z) - d| \le d_m(A+B,0) \le d_m(A,0) + d_m(B,0).$$

The expected magnitude of each cyclic-geometric share is at most $1/\sinh\alpha$, which gives the error bound $2/\sinh\alpha$.

b) A coherent guard that moves approximation into utility: A direct coherent encoding of the classical input would produce orthogonal states for different inputs. An observing party could then copy or measure the input label without disturbing the prescribed outgoing message. To prevent this, Alice adds a small common component to every encoded state. For $i = (x,a)$, she sends

$$|\psi_i\rangle = \sqrt{1-\kappa}\,|0\rangle\,|i\rangle + \sqrt{\kappa}\,|1\rangle\,|\perp\rangle,$$

where the second branch is independent of i. Consequently, for every $i \neq j$,

$$\langle\psi_i|\psi_j\rangle = \kappa > 0.$$

Bob coherently computes

$$f_{y,b}(x,a) = \mathrm{HD}(x,y) + a + b \pmod m$$

on the data branch and appends a fixed fallback output on the guard branch, returning

$$|\phi_i^{y,b}\rangle = \sqrt{1-\kappa}\,|0\rangle\,|i\rangle\,|f_{y,b}(i)\rangle + \sqrt{\kappa}\,|1\rangle\,|\perp\rangle\,|z_\perp\rangle.$$

The input and returned state families have exactly the same Gram matrix.

The guard probability is chosen as

$$\kappa = \frac{\gamma}{n+\gamma}.$$

On the dominant branch, the parties obtain the ideal noisy release. On the guard branch, they output a fixed fallback value, whose error is at most n. Thus the guard contributes less than γ to the expected error. The construction is therefore exactly private for every positive κ; small κ affects utility, not privacy. This avoids taking a trace-distance limit, which would generally be insufficient for pure QDP because pure privacy is sensitive to arbitrarily small support mismatches. [1]

c) Equal-Gram rigidity and the no-information principle: The central quantum argument is an equal-Gram rigidity lemma [20], [23], [24], [31]. This is the equal-Gram form of the no-information principle underlying Klauck's honest-player quantum privacy model [20]. Suppose a Completely Positive Trace-Preserving (CPTP) map sends each pure input state $|u_i\rangle$ to a prescribed pure output state $|v_i\rangle$, and suppose

$$\langle u_i|u_j\rangle = \langle v_i|v_j\rangle \neq 0$$

for every i, j. Let W be a Stinespring isometry for the map [32]. Purity of the prescribed output implies

$$W\,|u_i\rangle = e^{\mathrm{i}\theta_i}\,|v_i\rangle\,|e_i\rangle$$

for some environment state $|e_i\rangle$. Preservation of inner products gives

$$\langle u_i|u_j\rangle = e^{\mathrm{i}(\theta_j-\theta_i)}\,\langle v_i|v_j\rangle\,\langle e_i|e_j\rangle.$$

Since the two Gram matrices agree and their entries are nonzero, $\langle e_i|e_j\rangle$ must be a phase of unit magnitude. Hence all environment states agree up to phase, and their density operators are identical. The observer's retained complementary state is therefore independent of i.

Applied to the guarded protocol, this means that any Klauck-honest Bob who returns the prescribed pure state cannot retain information about Alice's pair (x,a). His complete terminal state is merely an input-independent quantum channel applied to the ideal view (y,B,Z). A symmetric postprocessing argument handles an observing Alice. Privacy then follows from the ideal functionality by data processing.

This is precisely where the quantum advantage comes into play. Classical messages can be copied while remaining unchanged, so exact preservation of a classical outgoing message places no comparable restriction on the observer's retained state. For non-orthogonal quantum states, by contrast,

[1] Pure differential privacy [1], [2] is equivalent to a bound on the max-divergence. It is known that the max-divergence between any two distributions P and Q becomes infinite if the support of the first probability distribution P is not contained within the support of the second probability distribution Q [30].

exact preservation of the entire pure-state family prevents the observer from creating an input-dependent copy in a complementary register.

d) Exact calibration for approximate privacy: For approximate privacy, we avoid generic composition or continuity bounds and calculate the hockey-stick divergence exactly. On the finite cycle, the likelihood ratio between neighboring shifts of q_α takes only three forms:

$$1, \qquad e^{\alpha}, \qquad e^{-\alpha}.$$

The region on which the ratio exceeds e^{ε} is exactly one half-cycle. Summing the excess mass on that region gives

$$\Delta_{n,\varepsilon}(\alpha) = \left(1 - e^{\varepsilon-\alpha}\right) \frac{1 - e^{-2n\alpha}}{1 + e^{-\alpha} - 2e^{-(2n+1)\alpha}}.$$

We prove that $\Delta_{n,\varepsilon}(\alpha)$ is continuous and strictly increasing from 0 at $\alpha = \varepsilon$ to 1 as $\alpha \to \infty$. Therefore, for every $0 < \delta < 1$, there is a unique solution

$$\Delta_{n,\varepsilon}(\alpha_n^\star(\varepsilon, \delta)) = \delta.$$

Since $\alpha_n^\star(\varepsilon, \delta) > \varepsilon$, the corresponding cyclic-geometric distribution is more concentrated than the pure-DP distribution. The equal-Gram argument is independent of α, so this improved ideal functionality lifts to the quantum protocol without consuming any additional δ-budget. This yields the strict approximate-privacy utility improvement in Theorem 8.

D. Comparison to Previous Lower Bounds

We now further discuss how previous lower bounds in the literature apply to our work.

a) The classical transcript model: A classical interactive protocol induces a random transcript

$$T = (M_1, \ldots, M_r)$$

consisting of all messages exchanged by the parties [33]–[36]. An honest-but-curious classical player can copy every message it receives or sends without changing the execution. Its terminal view may therefore contain the complete transcript, its private input, and its private random coins. Consequently, privacy of the player's complete view implies privacy of the classical transcript and coins [37].

For every fixed classical transcript t, the probability of observing t factors into an Alice-dependent and a Bob-dependent term:

$$\Pr[T = t \mid X = x, Y = y] = a_t(x) b_t(y).$$

Equivalently, a transcript induces a combinatorial rectangle in the input space. Conditional on a transcript, the parties' private random coins are independent. These factorization and conditional-independence properties are fundamental to classical communication complexity arguments [7], [8].

McGregor et al. exploit this classical structure in two complementary ways [3], [11]. Their first technique relates the likelihood-ratio constraints imposed by differential privacy to deterministic extraction from Santha–Vazirani sources [38]. Their second technique connects accurate two-party differentially private protocols to low-communication approximations of the underlying function [39]. Both arguments reason about classical transcript distributions, transcript rectangles, and the information that can be retained by copying the interaction.

b) Why these arguments do not directly extend to quantum protocols: A quantum protocol does not, in general, generate a simultaneously available classical transcript [40]. At round j, the communicated register may be in a non-orthogonal state

$$\rho_{M_j}^{x,y},$$

possibly entangled with private workspaces and with registers communicated in other rounds [41]. There need not exist a classical random variable T whose realizations describe the interaction while preserving all of its operational properties.

In particular, the classical step of copying every message is unavailable [40], [41]. An isometry of the form

$$|\psi_i\rangle |0\rangle \longmapsto |\psi_i\rangle |\psi_i\rangle$$

cannot exist for a family of distinct non-orthogonal states. Likewise, a measurement that records which input-dependent state was sent may disturb the message and prevent the observer from returning the exact state prescribed by the protocol. Such a measure-and-record strategy [42] is therefore not Klauck-honest.

This failure of copyability removes the classical reduction from a private interactive protocol to a private transcript distribution [15]. The observer's view is instead a density operator, and privacy is expressed through operator inequalities or quantum hockey-stick divergences [22]. There is no general pointwise factorization

$$p_{x,y}(t) = a_t(x) b_t(y)$$

over classical transcript values, and hence no immediate quantum analogue of the rectangle partition used in the classical lower bound [40], [41]. Measuring all messages to manufacture a classical transcript is also insufficient: the measurement is an additional channel that can alter interference, destroy coherence, and violate the exact outgoing-message requirement defining the KHNP model [20].

More concretely, our guarded messages are deliberately non-orthogonal [16], [20]. If Alice's encoded input and noise share is indexed by $i = (x, a)$, then the first-message states satisfy

$$\langle \psi_i | \psi_j \rangle = \kappa > 0 \qquad \text{for all } i \neq j.$$

Bob is required to return a prescribed family $\{|\phi_i^{y,b}\rangle\}_i$ with exactly the same Gram matrix. Equal-Gram rigidity then implies that any complementary state retained by a Klauck-honest Bob is independent of i [23]. Thus the action that is automatic classically (i.e., keeping a copy of the input-dependent message while forwarding the prescribed message) is physically impossible for these quantum states. This is because physical transformations between prescribed families of pure states can be characterized through their Gram matrices [23], [43]. The obstruction can also be seen directly from a Stinespring dilation [23], [32], as described earlier.

c) What the classical lower bounds still imply: The preceding discussion does not mean that the McGregor et al. [3], [11] bounds are irrelevant to our model. They continue to apply to every *classical* KHNP protocol. A classical KHNP observer may copy every message, retain its private randomness, send exactly the prescribed outgoing messages, and complete the protocol. Hence, privacy against all classical KHNP observers implies privacy of the transcript-and-coins views considered by McGregor et al [3], [11]. Moreover, when honest executions produce the same output with probability one, that common output is a function of the classical transcript on its support. Their lower-bound argument therefore yields the classical side of our separation:

$$\Omega(\sqrt{n}) \quad \text{under pure privacy,}$$

and

$$\Omega(\sqrt{n}/\log n) \quad \text{under strong approximate privacy.}$$

Theorems 7 and 8 show that these bounds cannot be extended unchanged from classical to quantum communication.

d) Dependence on the adversarial model: The quantum advantage is specific to the exact message-preserving KHNP condition. It is not a lower-bound violation in a model where the receiver may arbitrarily measure, modify, retain, or abort. For example, a receiver can measure the projector onto the data-dependent support of the first message and then abort, distinguishing adjacent inputs. That strategy is included in our retention-robust all-CPTP model but excluded from KHNP because it cannot subsequently return the prescribed coherent state.

Accordingly, prior classical lower bounds and our quantum upper bounds address different physical realizations, not inconsistent privacy guarantees [3], [11]–[13]. Classically, exact message preservation does not prevent transcript retention, so the classical lower bounds still apply even under KHNP. In the quantum regime, exact preservation of non-orthogonal pure messages can force the retained complementary state to contain no input information. Whether an analogue of the classical accuracy lower bound holds for quantum protocols against fully retention-capable adversaries remains open.

E. Organization of the Paper

The remainder of the paper is organized as follows. In Section II, we review related work on classical two-party differential privacy, quantum privacy models, and stronger cryptographic notions of two-party security. In Section III, we introduce the three view models considered in this work (fixed prescribed-channel privacy, Klauck-honest nonpreemptive privacy, and retention-robust adversarial privacy) and discuss the quantum differential privacy preliminaries used throughout the paper.

In Section IV, we develop the ideal common-output functionality for Hamming distance. We introduce the distributed cyclic-geometric noise mechanism, prove its pure privacy and accuracy guarantees, and derive the exact finite-cycle hockey-stick divergence needed for approximate privacy. The corresponding protocol realizations are given in Section V. We first show that the prescribed-channel model admits an exact classical reversible implementation, and then present the guarded coherent quantum protocol. Furthermore, we prove the equal-Gram no-information principle, establish the pure QDP guarantee of Theorem 7, derive the sharper approximate QDP guarantee of Theorem 8, and compare the resulting accuracy of the quantum protocol with the classical lower bounds of McGregor et al. [11].

In Section VI, we discuss the interpretation of the different privacy models, the distinction between pure and approximate privacy, and the possibility of verification of quantum states in the KHNP model. Finally, we conclude in Section VII.

II. Related Work

A. Differential privacy in two-party computation

In two-party differential privacy, Alice and Bob hold private inputs and interactively compute a statistic of their joint data [2]. Privacy must hold in both directions: Alice's view should be insensitive to a change in Bob's input, and Bob's view should be insensitive to a change in Alice's input [4]. This notion of privacy differs from the trusted curator model, in which a single trusted algorithm sees both inputs and releases only a noisy output. For a sensitivity-one statistic such as Hamming distance, the trusted curator model achieves expected error $O(1/\varepsilon)$, independent of the input length. In the two-party setting, by contrast, the interaction required to compute the statistic may itself leak information.

McGregor et al. [11] derive and discuss strong information-theoretic limitations for classical two-party differential privacy. For Hamming distance on n-bit inputs, their results imply an $\Omega(\sqrt{n})$ error lower bound for suitable constant pure-privacy parameters and an $\Omega(\sqrt{n}/\log n)$ lower bound under strong approximate privacy. These bounds separate the two-party setting from the trusted curator model and from computationally private protocols [4]. Indeed, standard secure computation can emulate the trusted curator under cryptographic assumptions, allowing the parties to recover accuracy, in the central model, against computationally bounded observers.

Subsequent work investigated the computational assumptions needed to cross this information-theoretic accuracy barrier. Haitner et al. [12] showed that sufficiently accurate computationally differentially private protocols for Hamming distance or integer inner product imply key agreement. Thus, improving substantially over the classical information-theoretic bounds is not merely an algorithmic improvement within the classical model: it entails nontrivial cryptographic structure. Several other related works have further studied the relationship between computational differential privacy, key agreement, oblivious transfer, and other public-key primitives [4], [12], [13].

Our result identifies quantum communication as a different resource for crossing the classical accuracy barrier. The protocol of Theorem 7 uses no computational assumption, trusted setup, prior entanglement, bounded-storage assumption, or secure-erasure assumption. Nevertheless, under Klauck's honest,

nonpreemptive, message-preserving model, it achieves $O(1)$ expected error for constant ε. This does not contradict the lower bounds of McGregor et al. [3], [11], because those bounds are lower bounds for *classical* interactive protocols. Their proofs use structural properties of classical transcripts that do not hold for coherent quantum communication [27], [40], [41].

B. Quantum privacy models

Klauck explicitly separated privacy at every round from privacy only in the final state of honest nonpreemptive parties [20]. The honest players are defined by exact preservation of every prescribed outgoing message, not by a fixed sequence of local channels; he also observes that classical messages may be copied and that final-state privacy is therefore no weaker classically [20]. Our guarded construction uses the exact no-information principle used in Klauck's work [20]. Kerenidis et al. further showed that privacy in quantum communication depends on whether inputs are classical, coherent, or externally purified [44]. In our work, the inputs are classical but quantum communication is allowed. The classical differential privacy definition has been extended to allow for more general measurements [22]. We rely on such quantum differential privacy definitions (which include the classical differential privacy definition as a special case) in our work.

We note that analyses for stronger adversaries exist in the literature [6]. Such adversaries may retain auxiliary work while remaining locally consistent with an honest execution [45]; even a private SWAP cannot be implemented in the bare model in that framework [45]. Salvail, Schaffner, and Sotakova study honest-but-curious quantum parties that purify their actions and postpone measurement [46]. General impossibility and leakage results for two-party quantum computation imply that message-preserving honesty *cannot*, without additional extensions, lead to malicious or simulation-based security [47], [48]. The retention barrier we discuss in our work is consistent with these observations.

III. Models and QDP Preliminaries

Write $x \sim x'$ when two strings differ in one coordinate. For density operators, $A \preceq B$ denotes the Loewner order and

$$E_\lambda(\rho\|\sigma) = \mathrm{Tr}(\rho - \lambda\sigma)_+$$

is the quantum hockey-stick divergence [49].

Motivated by the definition of classical differential privacy for two-party protocols (Definition 2.3 in [11]), we define quantum differential privacy for two-party protocols:

Definition 1 (QDP for a family of views)**.** *A family $\{\rho^B_{x,y}\}$ is pure ε-QDP for Alice's input if, for every $x \sim x'$ and fixed y,*

$$\rho^B_{x,y} \preceq e^\varepsilon \rho^B_{x',y}, \qquad \rho^B_{x',y} \preceq e^\varepsilon \rho^B_{x,y}.$$

It is (ε,δ)-QDP if the corresponding two hockey-stick divergences with $\lambda = e^\varepsilon$ are at most δ. The symmetric conditions protect Bob's input in Alice's view.

These inequalities are equivalent to the usual bounds for every quantum measurement $0 \preceq M \preceq I$ [22], [50]. They are preserved by quantum channels, partial trace, and tensoring input-independent side information.

Definition 2 (Fixed prescribed-channel view)**.** *In the* prescribed-channel *(PC) view model, the parties execute the fixed local channels specified by the protocol, including their prescribed discard operations. A party's view is the collection of registers designated by those channels at termination.*

Definition 3 (Klauck-honest nonpreemptive view)**.** *Fix a quantum protocol Π. A strategy is* Klauck-honest *if, for every private input and every sequence of pure message states it could have received, the density operator of its next outgoing message is exactly the one prescribed by Π. Subject to this constraint it may perform arbitrary local operations and retain arbitrary private systems. It is* nonpreemptive *if it completes every round. Its view is its complete terminal private state. Protocol Π is KHNP-QDP if Definition 1 holds for the terminal views of every such strategy for the observing party.*

This is Klauck's definition [20, Sec. 2.3]. Copying a prescribed classical message is allowed. By contrast, his no-information lemma shows that no unitary can extract a distinguishing record while leaving each of two non-orthogonal pure messages unchanged.

The pure KHNP result below uses classical input registers and classical private random tapes. The parties may keep their own input and sampled noise, and honesty is imposed conditional on every realization of that tape. The input is not supplied in coherent superposition or entangled with an external reference. Subject to these restrictions, the observer may still use arbitrary quantum private work allowed by Definition 3.

Definition 4 (Retention-robust adversarial QDP)**.** *For corrupted Bob, let $\mathfrak{A}^{\mathrm{all}}_B$ contain all finite-dimensional interactive CPTP strategies, including strategies that change messages, retain auxiliary systems, or abort. A strategy may start with an arbitrary side register R_B in state $\sigma^y_{R_B}$ that may depend on Bob's fixed input y, but is identical for adjacent choices x, x' of Alice's protected input and is independent of Alice's fresh private coins. The strategy itself is independent of x except through received messages. Let $\rho^{B^*,\sigma}_{x,y}$ be its complete state when it terminates. RR-QDP for Alice means that Definition 1 holds for every $B^* \in \mathfrak{A}^{\mathrm{all}}_B$, every y, σ^y, and every $x \sim x'$. The symmetric quantifiers, with side information allowed to depend on x but identical for adjacent y, y', protect Bob. Accuracy and common output are required only in the all-honest execution.*

Allowing initial side information that already changes with the protected input would make privacy impossible independently of the protocol. Definition 4 therefore fixes the same adversarial side state in the two adjacent experiments.

IV. The Ideal Common-Output Functionality

Fix the odd modulus $m = 4n + 1$. For an integer u, let $[u]_m \in \{0, \ldots, m-1\}$ be its least nonnegative residue. The circular metric on $\mathbb{Z}_m$ is

$$d_m(a,b) := \min\{[a-b]_m, [b-a]_m\} = \min_{k\in\mathbb{Z}} |\widetilde{a} - \widetilde{b} + km|, \quad (4)$$

where $\widetilde{a}, \widetilde{b}$ are arbitrary integer lifts. For $\alpha > 0$, define the full-support cyclic-geometric distribution

$$q_\alpha(g) = \frac{\exp[-\alpha d_m(g,0)]}{\sum_{h\in\mathbb{Z}_m} \exp[-\alpha d_m(h,0)]}.$$

Circular distance is 1-Lipschitz, so

$$e^{-\alpha} \le \frac{q_\alpha(g+1)}{q_\alpha(g)} \le e^{\alpha}.$$

Alice samples $A \sim q_\alpha$, Bob independently samples $B \sim q_\alpha$, and the ideal functionality releases

$$Z = d + A + B \pmod m.$$

Let $D = \{0, \dots, n\} \subset \mathbb{Z}_m$, and define

$$\mathrm{Dec}(z) := \min \arg\min_{r\in D} d_m(z, r), \tag{5}$$

so every tie is broken toward the smaller integer. Both parties receive the same Z and output the common estimate $\widehat{d} = \mathrm{Dec}(Z)$. Alice's ideal view is (x, A, Z), while Bob's is (y, B, Z).

Lemma 3 (Circular interval decoding)**.** *For every $d \in D$ and $z \in \mathbb{Z}_m$,*

$$|\,\mathrm{Dec}(z) - d| \le d_m(z, d).$$

Proof. If $z \in D$, then $\mathrm{Dec}(z) = z$, and $|z - d| = d_m(z, d)$ because $|z - d| \le n < m/2$.

Suppose $z \notin D$. The complementary arc from n to 0 that avoids the interior of D has length $L = m - n = 3n + 1$. Let $t \in (0, L)$ be the distance along this arc from n to z. For every $r \in D$, the two circular paths from z to r have lengths

$$t + n - r \quad \text{and} \quad L - t + r,$$

and hence

$$d_m(z, r) = \min\{t + n - r, L - t + r\}.$$

Minimizing over $r \in D$ shows that the decoder selects n when $t < L/2$, selects 0 when $t > L/2$, and selects 0 at equality by (5).

If $\mathrm{Dec}(z) = n$, then

$$t + n - d \ge n - d, \qquad L - t + d > L/2 + d \ge n - d.$$

If $\mathrm{Dec}(z) = 0$, then

$$L - t + d \ge d, \qquad t + n - d \ge L/2 + n - d \ge d,$$

where the last inequality uses $d \le n$ and $L/2 = (3n + 1)/2$. Taking the minimum of the two path lengths proves the result in both cases. □

We provide the proof of privacy and accuracy for the use of geometric noise to satisfy privacy. Our work extends the analysis of Ghosh et al. [51] to circular interval decoding.

Theorem 4 (Ideal privacy and accuracy)**.** *The ideal views are pure ε-QDP and*

$$\mathbb{E}|\widehat{d} - d| \le \frac{2}{\sinh \varepsilon} \le \frac{2}{\varepsilon}.$$

Proof. For Alice,

$$\Pr[A = a, Z = z \mid x, y] = q_\varepsilon(a) q_\varepsilon(z - d - a).$$

Changing one bit of y changes d by at most one. The likelihood ratio is therefore between $e^{-\varepsilon}$ and e^{ε}. Exchanging the parties proves Bob's condition.

Set $E = A + B \in \mathbb{Z}_m$. Lemma 3, translation invariance of (4), and the triangle inequality give

$$|\widehat{d} - d| \le d_m(d + E, d) = d_m(E, 0) \le d_m(A, 0) + d_m(B, 0).$$

The random variable $d_m(A, 0)$ follows the distribution of the absolute value of a two-sided geometric random variable, except that it cannot exceed $2n$. Removing values larger than $2n$ can only decrease its expectation. That is, conditioning on this lower-tail event cannot increase its expectation. Consequently,

$$\mathbb{E} d_m(A, 0) \le \frac{2e^{-\varepsilon}}{1 - e^{-2\varepsilon}} = \frac{1}{\sinh \varepsilon},$$

and similarly for B. □

Proposition 5 (Exact finite-cycle approximate loss)**.** *For $\alpha \ge \varepsilon > 0$, define*

$$\Delta_{n,\varepsilon}(\alpha) := (1 - e^{\varepsilon - \alpha}) \frac{1 - e^{-2n\alpha}}{1 + e^{-\alpha} - 2e^{-(2n+1)\alpha}} \tag{6}$$

and

$$U_n(\alpha) := \frac{4 \sum_{j=1}^{2n} j e^{-\alpha j}}{1 + 2 \sum_{j=1}^{2n} e^{-\alpha j}}. \tag{7}$$

The ideal views are exactly $(\varepsilon, \Delta_{n,\varepsilon}(\alpha))$-QDP and satisfy

$$\mathbb{E}|\widehat{d} - d| \le U_n(\alpha) \le \frac{2}{\sinh \alpha}.$$

Moreover, $\Delta_{n,\varepsilon}$ is continuous and strictly increasing from 0 at $\alpha = \varepsilon$ to 1 as $\alpha \to \infty$. Thus every $0 < \delta < 1$ determines a unique $\alpha_n^\star(\varepsilon, \delta) > \varepsilon$ satisfying

$$\Delta_{n,\varepsilon}(\alpha_n^\star) = \delta.$$

Proof. Represent $\mathbb{Z}_m$ by $\{-L, \dots, L\}$, where $L = 2n$, put $r = e^{-\alpha}$, and write

$$q_\alpha(g) = \frac{r^{|g|}}{1 + 2 \sum_{j=1}^{L} r^j}.$$

For adjacent distances $d, d + 1$, Alice's probability masses are

$$P_d(a, z) = q_\alpha(a) q_\alpha(g), \qquad P_{d+1}(a, z) = q_\alpha(a) q_\alpha(g - 1),$$

where $g = z - d - a$ and subtraction is cyclic. The relevant ratios are

$$\frac{q_\alpha(g)}{q_\alpha(g-1)} = \begin{cases} 1, & g = -L, \\ e^{\alpha}, & -L + 1 \le g \le 0, \\ e^{-\alpha}, & 1 \le g \le L. \end{cases}$$

Consequently,

$$E_{e^\varepsilon}(P_d \| P_{d+1}) = (1 - e^{\varepsilon - \alpha}) \frac{\sum_{j=0}^{L-1} r^j}{1 + 2 \sum_{j=1}^{L} r^j} = \Delta_{n,\varepsilon}(\alpha).$$

The reverse divergence has the same value, supported on the opposite half-cycle. Averaging over the observer's own share and exchanging the parties proves the QDP statement.

Lemma 3 and the triangle inequality bound error by the sum of the two share magnitudes. Evaluating their finite expectations gives (7); comparison with the untruncated two-sided geometric law gives $U_n(\alpha) \le 2/\sinh\alpha$.

For monotonicity, let

$$A_L(r) = \sum_{j=0}^{L-1} r^j, \qquad W_L(r) = \frac{A_L(r)}{1+2rA_L(r)}.$$

Then

$$W_L'(r) = \frac{A_L'(r) - 2A_L(r)^2}{(1+2rA_L(r))^2} < 0,$$

because $A_L'(r) \le A_L(r)^2$ coefficientwise for $0 < r < 1$. Both $W_L(e^{-\alpha})$ and $1 - e^{\varepsilon-\alpha}$ therefore increase with α, strictly for the latter. Formula (6) gives the endpoint limits, proving existence and uniqueness. □

The other party's unknown share is essential: a party that sampled all the noise could subtract it from its own view.

Remark 1 (Why the factor two is present)**.** *The main problem requires a common output, so both independent shares enter the same released Z, giving $2/\sinh\alpha$. Without a common-output requirement, the ideal functionality could instead give Alice $d+B$ and Bob $d+A$ modulo m. Each party would then decode using only the other party's share and obtain expected error at most $1/\sinh\alpha$. We do not use that separate-output view model.*

V. Exact PC and Guarded KHNP Realizations

Proposition 6 (Exact classical realization in the PC view model)**.** *In the PC view model of Definition 2, there is an exact classical reversible protocol with $O(n)$ communication whose common terminal output and views equal the ideal functionality of Theorem 4. It is therefore pure ε-QDP and has expected error at most $2/\sinh\varepsilon$.*

Proof. Condition on the sampled shares a, b. Reversibly run a deterministic computational-basis protocol for

$$F_m((x,a),(y,b)) = \mathrm{HD}(x,y) + a + b \pmod m.$$

Copy its value Z into a permanent output register, copy the same classical Z into the other party's output register, and reverse the computation to clear the message history and work registers. The fixed channels of the PC view model perform this reversal and retain no additional transcript. The terminal views are exactly (x, A, Z) and (y, B, Z), and both parties apply the same decoder (5). Theorem 4 completes the proof. □

Thus PC security alone cannot demonstrate a quantum improvement: it already permits an exact classical pure-DP realization.

We proceed to the KHNP model where we introduce Algorithm 1.

A. Guarded coherent round trip

Set $\ell = \lceil \log_2 m \rceil$, choose a fixed fallback $z_\perp \in \mathbb{Z}_m$, and let

$$\kappa = \frac{\gamma}{n+\gamma} \in (0,1).$$

For $i = (x, a)$, Alice prepares and sends

$$|\psi_i\rangle = \sqrt{1-\kappa}\,|0\rangle\,|i\rangle + \sqrt{\kappa}\,|1\rangle\,|\perp\rangle\,. \tag{8}$$

For fixed y, b, define

$$f_{y,b}(x,a) = \mathrm{HD}(x,y) + a + b \pmod m.$$

Bob coherently appends an output register and returns the entire message:

$$|\phi_i^{y,b}\rangle = \sqrt{1-\kappa}\,|0\rangle\,|i\rangle\,|f_{y,b}(i)\rangle + \sqrt{\kappa}\,|1\rangle\,|\perp\rangle\,|z_\perp\rangle\,. \tag{9}$$

Alice measures the branch and residue registers, sends the measured residue to Bob, and both apply Dec.

B. Privacy

Theorem 7 (Pure QDP for Klauck-honest players)**.** *For every $n \ge 1$, $\varepsilon > 0$, and $\gamma > 0$, the guarded protocol with $A, B \sim q_\varepsilon$ is information-theoretic pure ε-QDP for the complete terminal view of every Klauck-honest nonpreemptive observing strategy. In an honest execution both parties produce a common output,*

$$\mathbb{E}|\widehat{d} - \mathrm{HD}(x,y)| < \frac{2}{\sinh\varepsilon} + \gamma,$$

and total communication is $2n + O(\log n) = O(n)$ qubits and bits in the standard (arbitrary) unitary communication model.

Proof. First fix y, b. For $i \ne j$,

$$\langle\psi_i|\psi_j\rangle = \kappa = \langle\phi_i^{y,b}|\phi_j^{y,b}\rangle,$$

while both inner products are 1 for $i = j$. Any observing Bob must return exactly the pure state (9). Claim 1 therefore makes his retained complementary state independent of $i = (x, a)$, conditional on his own y, b. He subsequently receives only the classical residue obtained from the ideal Z by the channel

$$Z \longmapsto \begin{cases} Z, & \text{with probability } 1-\kappa, \\ z_\perp, & \text{with probability } \kappa. \end{cases}$$

All later local work is postprocessing. Hence, for every such strategy of Bob, his complete terminal state is a channel output independent of x applied to the ideal state (y, B, Z). Theorem 4 and data processing prove pure ε-QDP for Alice's input, after averaging over B.

Now fix x and an observing Alice strategy. Conditional on $A = a$, the state returned by honest Bob depends on y, B only through

$$Z = \mathrm{HD}(x,y) + a + B \pmod m.$$

For fixed x, a, it is the fixed encoding

$$z \longmapsto \sqrt{1-\kappa}\,|0,x,a,z\rangle + \sqrt{\kappa}\,|1,\perp,z_\perp\rangle\,.$$

Alice's retained preparation registers, arbitrary subsequent instrument, prescribed final message, and terminal work together

form a channel independent of y applied to the classical ideal state (x, A, Z). Data processing of Theorem 4 proves pure ε-QDP for Bob's input. The two observing strategies were arbitrary, which falls under the KHNP model.

The branch measurement is input-independent, with guard probability κ. Conditional on the real branch, Theorem 4 gives error at most $2/\sinh\varepsilon$; on the fallback branch the error is at most n. Therefore

$$\mathbb{E}|\widehat{d} - d| \le (1-\kappa)\frac{2}{\sinh\varepsilon} + \kappa n < \frac{2}{\sinh\varepsilon} + \gamma.$$

The first message has $1+n+\ell$ qubits, its return has $1+n+2\ell$, and the final residue has ℓ classical bits. □

Claim 1 (Equal-Gram no-information principle)**.** *Let $\{|u_i\rangle\}$ and $\{|v_i\rangle\}$ be pure-state families with*

$$\langle u_i|u_j\rangle = \langle v_i|v_j\rangle \neq 0 \quad \textit{for all } i, j.$$

If a CPTP map (e.g., induced by either Alice or Bob) sends each $|u_i\rangle\langle u_i|$ to the pure outgoing state $|v_i\rangle\langle v_i|$, then its complementary output state is independent of i.

Proof. Let W be a Stinespring isometry [32]. Purity of the outgoing marginal implies

$$W|u_i\rangle = e^{\mathrm{i}\theta_i}|v_i\rangle|e_i\rangle$$

for some unit vectors $|e_i\rangle$. Preservation of inner products and the nonzero equal-Gram hypothesis give

$$\langle e_i|e_j\rangle = e^{\mathrm{i}(\theta_i-\theta_j)}.$$

Thus all $|e_i\rangle$ agree up to phase, and hence all complementary density operators $|e_i\rangle\langle e_i|$ are identical. □

C. Variations

The use of an exact $\sqrt{\kappa}$ rotation is possible in the (arbitrary) unitary model. A discrete variant chooses a power of two $K \ge \max\{2, n/\gamma\}$ and replaces (8) by

$$\frac{1}{\sqrt{K}}\left(|0,\perp\rangle + \sum_{r=1}^{K-1}|r,i\rangle\right).$$

The returned state analogously appends $z_\perp$ on branch 0 and $f_{y,b}(i)$ on every branch $r \ge 1$. The distinct-state overlap and fallback probability are $1/K$, the utility penalty is at most $n/K \le \gamma$, and Hadamards prepare the branch register exactly.[2] Its communication cost is $O(n + \log(1/\gamma))$.

Pure QDP already implies (ε, δ)-QDP with the same error. Approximate privacy permits strictly sharper calibration.

[2]When $K = 2^k$, the uniform branch register

$$\frac{1}{\sqrt{K}}\sum_{r=0}^{K-1}|r\rangle$$

is prepared exactly by applying $H^{\otimes k}$ to $|0^k\rangle$. Controlled basis-state preparation then produces

$$\frac{1}{\sqrt{K}}\left(|0\rangle|\perp\rangle + \sum_{r=1}^{K-1}|r\rangle|x,A\rangle\right)$$

exactly.

Algorithm 1 Guarded Coherent Protocol for Pure KHNP-QDP

Require: Alice has $x \in \{0,1\}^n$; Bob has $y \in \{0,1\}^n$; privacy parameter $\varepsilon > 0$; utility parameter $\gamma > 0$.

Ensure: A common output with pure ε-QDP in the Klauck-honest nonpreemptive model.

1: Set
$$m \leftarrow 4n+1, \qquad \ell \leftarrow \lceil \log_2 m \rceil, \qquad \kappa \leftarrow \frac{\gamma}{n+\gamma}.$$
2: Fix a public fallback residue $z_\perp \in \mathbb{Z}_m$.
3: **Alice:** Sample
$$A \sim q_\varepsilon.$$
4: **Bob:** Sample
$$B \sim q_\varepsilon.$$
5: Alice prepares a one-qubit branch register C and an $(n+\ell)$-qubit data register M in the pure state
$$|\psi_{x,A}\rangle_{CM} = \sqrt{1-\kappa}\,|0\rangle_C|x,A\rangle_M + \sqrt{\kappa}\,|1\rangle_C|\perp\rangle_M,$$
where $|\perp\rangle$ is a fixed public computational-basis state.
6: **Alice → Bob:** Send the quantum message
$$M_{A\to B} = |\psi_{x,A}\rangle_{CM}.$$
7: Bob initializes an ℓ-qubit residue register R to $|0\rangle$.
8: Bob applies a controlled reversible operation satisfying
$$|0\rangle|x,A\rangle|0\rangle \longmapsto |0\rangle|x,A\rangle|\mathrm{HD}(x,y) + A + B \bmod m\rangle$$
and
$$|1\rangle|\perp\rangle|0\rangle \longmapsto |1\rangle|\perp\rangle|z_\perp\rangle.$$
9: The resulting prescribed pure return message is
$$|\phi^{y,B}_{x,A}\rangle_{CMR} = \sqrt{1-\kappa}\,|0\rangle_C|x,A\rangle_M|\mathrm{HD}(x,y)+A+B \bmod m\rangle_R + \sqrt{\kappa}\,|1\rangle_C|\perp\rangle_M|z_\perp\rangle_R.$$
10: **Bob → Alice:** Return the entire quantum message
$$M_{B\to A} = |\phi^{y,B}_{x,A}\rangle_{CMR}.$$
11: Alice measures C and R in the computational basis, obtaining outcomes
$$(c,z) \in \{0,1\}\times\mathbb{Z}_m.$$
12: Set the public released residue to
$$Z \leftarrow z.$$
▷ When $c = 1$, $z = z_\perp$.
13: **Alice → Bob:** Send the classical message
$$M^{\mathrm{final}}_{A\to B} = Z \in \mathbb{Z}_m.$$
14: **Alice:** Output
$$\widehat{d}_A \leftarrow \mathrm{Dec}(Z).$$
15: **Bob:** Output
$$\widehat{d}_B \leftarrow \mathrm{Dec}(Z).$$
16: **return**
$$\widehat{d}_A = \widehat{d}_B.$$

Theorem 8 (Noise-calibrated approximate KHNP-QDP)**.** *Let $n \geq 1$, $\varepsilon > 0$, $0 < \delta < 1$, and $\gamma > 0$. Run the guarded protocol with*

$$A, B \sim q_{\alpha_n^\star(\varepsilon,\delta)}.$$

Every Klauck-honest nonpreemptive terminal view satisfies (ε, δ)-QDP, both honest parties produce a common output, and

$$\mathbb{E}|\widehat{d} - d| \leq U_n(\alpha_n^\star) + \gamma \leq \frac{2}{\sinh \alpha_n^\star} + \gamma < \frac{2}{\sinh \varepsilon} + \gamma.$$

Communication is $O(n)$ in the (arbitrary) unitary model.

Proof. Proposition 5 makes the ideal views at $\alpha = \alpha_n^\star$ exactly (ε, δ)-QDP. The equal-Gram argument in Theorem 7 does not depend on the value of α: every complete KHNP terminal view is a quantum channel, independent of the protected input, applied to its corresponding ideal view. Data processing of the hockey-stick divergence proves privacy. The real branch has expected error at most $U_n(\alpha_n^\star)$; the input-independent guard branch contributes at most $\kappa n < \gamma$. The communication calculation is unchanged. □

A convenient explicit choice is

$$\bar{\alpha}(\varepsilon, \delta) := \log \frac{e^\varepsilon + \delta}{1 - \delta}. \tag{10}$$

Indeed, the finite-cycle factor in (6) is at most $(1 + e^{-\alpha})^{-1}$, since for $r = e^{-\alpha}$,

$$\frac{1 - r^{2n}}{1 + r - 2r^{2n+1}} \leq \frac{1}{1 + r}.$$

Therefore

$$\Delta_{n,\varepsilon}(\alpha) \leq \frac{e^\alpha - e^\varepsilon}{e^\alpha + 1}.$$

Thus $\alpha_n^\star \geq \bar{\alpha}$, and Theorem 8 implies the fully explicit bound

$$\mathbb{E}|\widehat{d} - d| \leq \frac{4(1 - \delta)(e^\varepsilon + \delta)}{(e^\varepsilon + 1)(e^\varepsilon - 1 + 2\delta)} + \gamma. \tag{11}$$

The exact finite-n calibration is at least as accurate.

D. Advantage

Corollary 9 (Same-model quantum advantages)**.** *For every fixed sufficiently small constant $\varepsilon > 0$ and fixed $\gamma > 0$, common-output Hamming distance has $O(1)$ expected error and $O(n)$ quantum communication in the KHNP model, whereas every classical information-theoretic pure-ε-DP KHNP protocol has $\Omega(\sqrt{n})$ error. For strong approximate privacy $\delta = o(1/n)$, the quantum error remains $O(1)$, whereas the classical KHNP error is $\Omega(\sqrt{n}/\log n)$.*

Proof. A classical KHNP observer may copy every received and sent classical message, retain its private coins, and still send exactly the prescribed messages and complete the protocol [20, Sec. 2.3]. Thus privacy against every classical KHNP strategy implies privacy of the standard transcript-and-coins views used by McGregor et al. [11, Def. 2.3]. Moreover, because the parties' private coins are independent conditional on a classical transcript, probability-one agreement makes the common output a function of that transcript on its support. The pure-DP partition-bound consequence of McGregor et al. then gives the $\Omega(\sqrt{n})$ Hamming distance error lower bound for suitable constant ε [11, Sec. 4.2]. A hypothetical $o(\sqrt{n})$ expected error would, by Markov's inequality, give $o(\sqrt{n})$ error with constant success probability and contradict that bound. The quantum upper bound follows from Theorem 7. The same transcript-view reduction and the approximate lower bound of McGregor et al. give $\Omega(\sqrt{n}/\log n)$ for $\delta = o(1/n)$, while Theorem 8 is $O(1)$ for fixed ε, γ. □

The separation is specific to exact KHNP honesty. For the averaged first message and adjacent x, x', the projector

$$P_x = \sum_{a \in \mathbb{Z}_m} |0, x, a\rangle\langle 0, x, a|$$

has probability $1 - \kappa$ on input x and probability 0 on input x'. A receiver allowed to measure P_x and abort therefore violates pure DP. Such a measurement cannot be followed by the exact coherent return (9); KHNP excludes it, whereas the general-CPTP RR model includes it.

VI. Discussion

A. Interpreting the Three View Models

The prescribed-channel (PC), Klauck-honest nonpreemptive (KHNP), and retention-robust (RR) models impose genuinely different requirements on a party's behavior and retained information. The distinction between these models is essential for interpreting both the positive results and their limitations.

In the PC model (Definition 2), the parties execute the fixed local channels specified by the protocol, including prescribed uncomputation and discard operations, and only the designated terminal registers are included in their views. This view model describes the terminal state of a fixed honest implementation. It does not allow a party to make additional copies of intermediate classical messages before the prescribed uncomputation. Indeed, Proposition 6 gives an exact classical reversible realization of the ideal functionality in this model: the protocol computes the noisy residue, copies it into the permanent output registers, reverses the remaining computation, and retains no transcript. Note that PC privacy alone cannot support a quantum–classical separation.

The KHNP model (Definition 3) allows substantially more local freedom. A party may perform arbitrary local quantum operations and retain arbitrary private systems, provided that every outgoing message density operator is exactly the one prescribed by the protocol and that the party completes the interaction. In the classical setting, this condition permits a party to copy every received message while continuing to send the prescribed messages. Thus a classical KHNP observer may retain the standard view (of the transcript and private coins) appearing in classical two-party differential-privacy lower bounds [3], [11].

Our guarded coherent construction exploits the fact that the same conclusion need not hold in the quantum setting. Its incoming messages are non-orthogonal, and the prescribed returned pure-state family has exactly the same nonzero Gram

matrix. If an observing party returns the exact prescribed state, Stinespring dilation and preservation of inner products force the complementary state retained by that party to be independent of the protected input. The resulting quantum–classical separation is therefore obtained under the same exact message-preserving KHNP model: classical messages can be copied while preserving the protocol, whereas the guarded non-orthogonal quantum messages cannot be copied in this manner.

The RR model (Definition 4) is stronger. It permits arbitrary interactive CPTP strategies, including strategies that retain extra systems, change outgoing messages, measure intermediate states, or abort. The guarded protocol is not RR secure. For example, a receiver may measure a projector that distinguishes the real branch for one candidate input and then terminate the protocol. Such a strategy is excluded by KHNP because it cannot subsequently return the exact coherent state, but it is included in the all-CPTP RR class of strategies.

The three models therefore can be used to answer *different questions*:

1) **PC:** What information remains in one prescribed honest implementation?
2) **KHNP:** What can a party retain while preserving every prescribed message exactly and completing the protocol?
3) **RR:** What privacy remains against arbitrary interactive deviations, retention, and early termination?

The quantum advantage established in our work is via the KHNP model. The advantage is stronger than a comparison based only on prescribed terminal registers, but it does not imply malicious or simulation-based security.

B. From KHNP Honesty to Verifiable Message Preservation

Although KHNP provides a clean same-model separation, its practical motivation might be less immediate. A user may reasonably object that a genuinely malicious counterparty has no obligation to preserve the prescribed message or complete the protocol. From that perspective, proving privacy only for KHNP strategies does not by itself guarantee protection against the adversaries of practical concern [20].

A potentially more compelling interpretation is to supplement the guarded protocol with a mechanism that allows the sender to *test* whether the returned quantum state is close to the state that should have been returned. KHNP would then describe the detection limit of a broader verifiable protocol: an exactly message-preserving party is accepted with probability one, while a party that significantly changes the prescribed return risks being detected.

Suppose Alice expects the pure returned state

$$|\phi_i^{y,b}\rangle .$$

If Alice can construct the corresponding verification measurement, she may apply the two-outcome projective test

$$\left\{\Pi_i^{y,b}, I - \Pi_i^{y,b}\right\}, \qquad \Pi_i^{y,b} := |\phi_i^{y,b}\rangle\langle\phi_i^{y,b}| .$$

For an honestly returned state, the acceptance probability is one. If Bob instead returns a state ρ_i, the acceptance probability is

$$p_{\mathrm{acc}}(i) = \mathrm{Tr}\Big(\Pi_i^{y,b}\rho_i\Big) = \langle\phi_i^{y,b}|\,\rho_i\,|\phi_i^{y,b}\rangle ,$$

and the rejection probability is

$$p_{\mathrm{rej}}(i) = 1 - \langle\phi_i^{y,b}|\,\rho_i\,|\phi_i^{y,b}\rangle .$$

Thus, deviations that significantly reduce the fidelity of the returned state, relative to the prescribed state, are detected with high probability.

This suggests an operational interpretation of the equal-Gram result. If Bob passes the ideal test with probability one for every possible input, then he returns the exact prescribed pure-state family. The equal-Gram no-information principle then implies that his retained complementary state is independent of Alice's protected input. At the opposite extreme, a strategy that extracts substantial distinguishing information is expected to disturb the returned state and thereby create a nonzero probability of rejection. Extracting information about a non-orthogonal quantum-state ensemble creates disturbance, and quantitative information and disturbance tradeoffs are well studied [52], [53]. Quantitative mathematical statements relating retained information, fidelity of the return, and detection probability could turn this intuition into an approximate, verifiable analogue of KHNP privacy. We believe this is important future work. Verification of a known pure target state can be formulated as a two-outcome test whose accepting operator contains the target-state projector [54].

Such verification could provide practical value even when it does not prevent every privacy violation. First, it could inform a user that a counterparty may have deviated from the privacy-preserving interaction. Second, repeated tests, trap rounds, or random audits could discourage malicious behavior by creating a risk of public detection. This may be relevant when service providers, laboratories, or network operators care about reputation, contractual compliance, or continued access to clients. Hidden trap states and randomly selected test computations are standard tools in verifiable delegated quantum computation [55].

The quantum setting is important here. Classically, a receiver can copy a message perfectly and then return or forward an unchanged copy. A test of the returned classical message cannot reveal that the receiver retained an additional copy. In contrast, the guarded quantum states are non-orthogonal. Extracting an input-dependent record while preserving the exact return is forbidden by the equal-Gram argument. Verification of the returned quantum state can therefore certify a property that has no direct classical analogue: consistency between low disturbance and the absence of an input-dependent retained complement [54].

VII. Conclusion

We have shown that classical information-theoretic accuracy lower bounds for two-party differential privacy do not automatically extend to models where quantum communication

is allowed. Under Klauck's honest, nonpreemptive, message-preserving model, two-party Hamming distance admits an information-theoretic quantum protocol with $O(n)$ communication and $O(1)$ expected error for constant privacy parameters. In the same model, classical protocols require $\Omega(\sqrt{n})$ error under pure privacy and $\Omega(\sqrt{n}/\log n)$ error under strong approximate privacy.

The separation is obtained through a distributed cyclic-geometric noise mechanism and a guarded coherent round trip. Exact preservation of the Gram matrix of the transmitted pure states forces any Klauck-honest observer's retained complementary state to be independent of the protected input. For approximate differential privacy, an exact finite-cycle hockey-stick calculation uses the δ-budget to choose a more concentrated noise distribution and strictly improve accuracy.

The result is inherently model dependent. Fixed prescribed channels already admit an exact classical reversible realization, whereas the guarded quantum protocol is not secure against arbitrary retention-capable CPTP strategies, such as measure-and-abort attacks. We provide a classical randomized-response baseline for this stronger model in the appendix. Whether quantum communication can improve retention-robust accuracy remains open.

VIII. Acknowledgements

We are grateful for helpful related discussions with Noam Mazor, Kaiyan Shi, Manasi Shingane, and Salil Vadhan. Daniel Alabi was supported by the Simons Foundation (965342, D.A.) as part of the Junior Fellowship from the Simons Society of Fellows. Emil T. Khabiboulline acknowledges support from the NRC Research Associateship Program at the National Institute of Standards and Technology (NIST), administered by the Fellowships Office of the National Academies of Sciences, Engineering, and Medicine, and was supported in part by ONR MURI (award No. N000142612102).

References

[1] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006*, ser. Lecture Notes in Computer Science, vol. 3876. Springer Berlin Heidelberg, 2006, pp. 265–284.

[2] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, "Our data, ourselves: Privacy via distributed noise generation," in *EUROCRYPT*, 2006, pp. 486–503.

[3] A. McGregor, I. Mironov, T. Pitassi, O. Reingold, K. Talwar, and S. P. Vadhan, "The limits of two-party differential privacy," in *51st Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, Las Vegas, Nevada, USA, October 23-26, 2010*. IEEE Computer Society, 2010, pp. 81–90. [Online]. Available: https://doi.org/10.1109/FOCS.2010.14

[4] A. Beimel, K. Nissim, and E. Omri, "Distributed private data analysis: Simultaneously solving how and what," in *Advances in Cryptology – CRYPTO 2008*, ser. Lecture Notes in Computer Science, D. Wagner, Ed., vol. 5157. Berlin, Heidelberg: Springer, 2008, pp. 451–468. [Online]. Available: https://doi.org/10.1007/978-3-540-85174-5_25

[5] E. Ben-Sasson, I. Bentov, Y. Horesh, and M. Riabzev, "Scalable zero knowledge with no trusted setup," in *Advances in Cryptology - CRYPTO 2019 - 39th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2019, Proceedings, Part III*, ser. Lecture Notes in Computer Science, vol. 11694. Springer, 2019, pp. 701–732.

[6] J. Katz and Y. Lindell, *Introduction to Modern Cryptography, Second Edition*, 2nd ed. Chapman & Hall/CRC, 2014.

[7] E. Kushilevitz and N. Nisan, *Communication Complexity*. New York: Cambridge University Press, 1997.

[8] A. Rao and A. Yehudayoff, *Communication Complexity: and Applications*. Cambridge University Press, 2020. [Online]. Available: https://books.google.com/books?id=PorPDwAAQBAJ

[9] C. Dwork and K. Nissim, "Privacy-preserving datamining on vertically partitioned databases," in *Advances in Cryptology - CRYPTO 2004, 24th Annual International Cryptology Conference*, ser. Lecture Notes in Computer Science, vol. 3152. Springer, 2004, pp. 528–544.

[10] C. Dwork, A. Smith, T. Steinke, and J. Ullman, "Exposed! a survey of attacks on private data," *Annual Review of Statistics and Its Application*, vol. 4, no. 1, pp. 61–84, 2017.

[11] A. McGregor, I. Mironov, T. Pitassi, O. Reingold, K. Talwar, and S. Vadhan, "The limits of two-party differential privacy," Electronic Colloquium on Computational Complexity, Tech. Rep. TR11-106, 2011. [Online]. Available: https://eccc.weizmann.ac.il/report/2011/106/

[12] I. Haitner, N. Mazor, J. Silbak, and E. Tsfadia, "On the complexity of two-party differential privacy," in *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*. New York, NY, USA: Association for Computing Machinery, 2022, pp. 1392–1405. [Online]. Available: https://doi.org/10.1145/3519935.3519982

[13] I. Haitner, N. Mazor, J. Silbak, E. Tsfadia, and C. Yan, "Computationally differentially private inner-product protocols imply oblivious transfer," in *Advances in Cryptology – CRYPTO 2025*, ser. Lecture Notes in Computer Science, Y. Tauman Kalai and S. F. Kamara, Eds., vol. 16003. Cham: Springer Nature Switzerland, 2025, pp. 174–206. [Online]. Available: https://doi.org/10.1007/978-3-032-01884-7_6

[14] I. Mironov, O. Pandey, O. Reingold, and S. Vadhan, "Computational differential privacy," in *Advances in Cryptology – CRYPTO 2009*, ser. Lecture Notes in Computer Science, S. Halevi, Ed., vol. 5677. Berlin, Heidelberg: Springer, 2009, pp. 126–142. [Online]. Available: https://doi.org/10.1007/978-3-642-03356-8_8

[15] V. Bužek and M. Hillery, "Quantum copying: Beyond the no-cloning theorem," *Phys. Rev. A*, vol. 54, pp. 1844–1852, Sep 1996. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.54.1844

[16] H. Klauck, "Lower bounds for quantum communication complexity," *SIAM J. Comput.*, vol. 37, no. 1, p. 20–46, Apr. 2007.

[17] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," *Theoretical Computer Science*, vol. 560, pp. 7–11, 2014.

[18] U. Vazirani and T. Vidick, "Fully device-independent quantum key distribution," *Phys. Rev. Lett.*, vol. 113, p. 140501, Sep. 2014. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.113.140501

[19] T. Vidick and S. Wehner, *Introduction to Quantum Cryptography*. Cambridge University Press, 2023.

[20] H. Klauck, "Quantum and approximate privacy," *Theory of Computing Systems*, vol. 37, no. 1, pp. 221–246, 2004.

[21] M. Naor and K. Nissim, "Communication preserving protocols for secure function evaluation," in *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, ser. STOC '01. New York, NY, USA: Association for Computing Machinery, 2001, p. 590–599. [Online]. Available: https://doi.org/10.1145/380752.380855

[22] C. Hirche, C. Rouzé, and D. S. França, "Quantum differential privacy: An information theory perspective," *IEEE Transactions on Information Theory*, vol. 69, no. 9, pp. 5771–5787, 2023.

[23] A. CHEFLES, R. JOZSA, and A. WINTER, "On the existence of physical transformations between sets of quantum states," *International Journal of Quantum Information*, vol. 02, no. 01, pp. 11–21, 2004. [Online]. Available: https://doi.org/10.1142/S0219749904000031

[24] M. D. Cock, M. Fannes, and P. Spincemaille, "Quantum dynamics and gram's matrix," *Europhysics Letters*, vol. 49, no. 4, p. 403, feb 2000. [Online]. Available: https://doi.org/10.1209/epl/i2000-00163-6

[25] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," *Nature*, vol. 299, no. 5886, pp. 802–803, 1982.

[26] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation*. USA: American Mathematical Society, 2002.

[27] H. Buhrman, R. Cleve, and A. Wigderson, "Quantum vs. classical communication and computation," in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, ser. STOC '98. New York, NY, USA: Association for Computing Machinery, 1998, p. 63–68. [Online]. Available: https://doi.org/10.1145/276698.276713

[28] Y. Lindell, *How to Simulate It – A Tutorial on the Simulation Proof Technique*. Cham: Springer International Publishing, 2017, pp. 277–346. [Online]. Available: https://doi.org/10.1007/978-3-319-57048-8_6

[29] Y. Lindell and B. Pinkas, "An efficient protocol for secure two-party computation in the presence of malicious adversaries," in *Advances in Cryptology - EUROCRYPT 2007, 26th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Barcelona, Spain, May 20-24, 2007, Proceedings*, ser. Lecture Notes in Computer Science, vol. 4515. Springer, 2007, pp. 52–78.

[30] I. Sason and S. Verdú, "Upper bounds on the relative entropy and rényi divergence as a function of total variation distance for finite alphabets," in *2015 IEEE Information Theory Workshop - Fall (ITW)*, 2015, pp. 214–218.

[31] A. Chefles, "Deterministic quantum state transformations," *Physics Letters A*, vol. 270, no. 1, pp. 14–19, 2000.

[32] W. F. Stinespring, "Positive functions on c^*-algebras," *Proceedings of the American Mathematical Society*, vol. 6, no. 2, pp. 211–216, 1955. [Online]. Available: https://ams.org

[33] S. Goldwasser, S. Micali, and C. Rackoff, "The knowledge complexity of interactive proof-systems," in *Proceedings of the Seventeenth Annual ACM Symposium on Theory of Computing*, ser. STOC '85. New York, NY, USA: Association for Computing Machinery, 1985, p. 291–304. [Online]. Available: https://doi.org/10.1145/22145.22178

[34] O. Goldreich, S. Micali, and A. Wigderson, "How to prove all np statements in zero-knowledge and a methodology of cryptographic protocol design (extended abstract)," in *Advances in Cryptology — CRYPTO' 86*, A. M. Odlyzko, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 1987, pp. 171–185.

[35] D. Beaver, "Secure multiparty protocols and zero-knowledge proof systems tolerating a faulty minority," *Journal of Cryptology*, vol. 4, no. 2, pp. 75–122, 1991. [Online]. Available: https://doi.org/10.1007/BF00196771

[36] ——, "Foundations of Secure Interactive Computing," in *Advances in Cryptology — CRYPTO '91*, J. Feigenbaum, Ed. Springer, 1992, pp. 377–391.

[37] 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. [Online]. Available: https://doi.org/10.1137/0218012

[38] M. Santha and U. V. Vazirani, "Generating quasi-random sequences from semi-random sources," *Journal of Computer and System Sciences*, vol. 33, no. 1, pp. 75–87, 1986. [Online]. Available: https://doi.org/10.1016/0022-0000(86)90044-9

[39] Z. Bar-Yossef, T. Jayram, R. Kumar, and D. Sivakumar, "An information statistics approach to data stream and communication complexity," *Journal of Computer and System Sciences*, vol. 68, no. 4, pp. 702–732, 2004, special Issue on FOCS 2002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022000003001855

[40] H. Buhrman, R. Cleve, and W. van Dam, "Quantum entanglement and communication complexity," *SIAM Journal on Computing*, vol. 30, no. 6, pp. 1829–1841, 2001, also available at arXiv:quant-ph/9705033. [Online]. Available: https://arxiv.org/abs/quant-ph/9705033

[41] G. Brassard, "Quantum communication complexity (a survey)," *Foundations of Physics*, vol. 33, no. 10, pp. 1593–1616, 2003, also available at arXiv:quant-ph/0101005. [Online]. Available: https://arxiv.org/abs/quant-ph/0101005

[42] C.-F. Li, Y.-S. Zhang, Y.-F. Huang, and G.-C. Guo, "Quantum strategies of quantum measurement," *Phys. Rev. A*, vol. 63, p. 054302, 2001. [Online]. Available: https://arxiv.org/abs/quant-ph/0007120

[43] J. Watrous, *The Theory of Quantum Information*. Cambridge University Press, 2018.

[44] I. Kerenidis, M. Laurière, F. Le Gall, and M. Rennela, "Information cost of quantum communication protocols," *Quantum Information and Computation*, vol. 16, no. 3–4, pp. 181–196, 2016.

[45] F. Dupuis, J. B. Nielsen, and L. Salvail, "Secure two-party quantum evaluation of unitaries against specious adversaries," in *Advances in Cryptology—CRYPTO 2010*, ser. Lecture Notes in Computer Science, vol. 6223. Springer, 2010, pp. 685–706.

[46] L. Salvail, C. Schaffner, and M. Sotáková, "Quantifying the leakage of quantum protocols for classical two-party cryptography," *International Journal of Quantum Information*, vol. 13, no. 4, p. 1450041, 2015.

[47] H. Buhrman, M. Christandl, and C. Schaffner, "Complete insecurity of quantum protocols for classical two-party computation," *Physical Review Letters*, vol. 109, no. 16, p. 160501, 2012.

[48] S. A. Osborn and J. Sikora, "A constant lower bound for any quantum protocol for secure function evaluation," in *17th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2022)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), vol. 232. Schloss Dagstuhl—Leibniz-Zentrum für Informatik, 2022, pp. 8:1–8:14.

[49] N. Sharma and N. A. Warsi, "Fundamental bound on the reliability of quantum information transmission," *Physical Review Letters*, vol. 110, no. 8, Feb. 2013. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.110.080501

[50] L. Zhou and M. Ying, "Differential privacy in quantum computation," in *30th IEEE Computer Security Foundations Symposium, CSF 2017, Santa Barbara, CA, USA, August 21-25, 2017*. IEEE Computer Society, 2017, pp. 249–262. [Online]. Available: https://doi.org/10.1109/CSF.2017.23

[51] A. Ghosh, T. Roughgarden, and M. Sundararajan, "Universally utility-maximizing privacy mechanisms," in *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, ser. STOC '09. New York, NY, USA: Association for Computing Machinery, 2009, p. 351–360. [Online]. Available: https://doi.org/10.1145/1536414.1536464

[52] C. A. Fuchs and A. Peres, "Quantum-state disturbance versus information gain: Uncertainty relations for quantum information," *Physical Review A*, vol. 53, no. 4, pp. 2038–2045, 1996.

[53] F. Buscemi and M. F. Sacchi, "Information-disturbance tradeoff in quantum-state discrimination," *Physical Review A*, vol. 74, no. 5, p. 052320, 2006.

[54] S. Pallister, N. Linden, and A. Montanaro, "Optimal verification of entangled states with local measurements," *Physical Review Letters*, vol. 120, no. 17, p. 170502, 2018.

[55] J. F. Fitzsimons and E. Kashefi, "Unconditionally verifiable blind quantum computation," *Physical Review A*, vol. 96, no. 1, p. 012303, 2017.

APPENDIX

The appendix provides expanded constructions and additional details supporting the main results.

We first give a complete description of the ideal common-output functionality for Hamming distance. In this functionality, Alice and Bob independently sample cyclic-geometric noise shares and receive the same noisy release. We explain why two independent shares are needed to protect both parties while producing a common output, and we give the corresponding privacy and accuracy guarantees. We then present an explicit reversible classical realization of this functionality in the prescribed-channel model. This construction clarifies why the prescribed-channel view is too weak to yield a quantum–classical separation: once intermediate computation is uncomputed, the designated terminal views coincide exactly with those of the ideal functionality. We next give the full noise-calibrated guarded protocol for approximate QDP. The construction uses the same coherent communication procedure as the pure QDP protocol, but chooses the cyclic-geometric parameter using the exact finite-cycle hockey-stick divergence. This allows the available δ-budget to be converted directly into a more concentrated noise distribution, yielding strictly smaller error without changing the communication complexity or the equal-Gram privacy argument. We also provide numerical examples illustrating the utility improvement over the pure QDP calibration. Finally, we give a common-output protocol for the stronger retention-robust model. The protocol combines randomized response for Alice's input with geometric output noise for Bob's input and remains private against arbitrary retention-capable CPTP strategies. Its expected error is $O(\sqrt{n})$ for constant privacy parameters, providing a classical retention-robust baseline against which future quantum protocols can be compared.

A. Implementing the Ideal Functionality

In general, cryptographic protocols are designed to be *realizations* of ideal functionalities, often conceptualized through *ideal views*. In modern cryptography, security is proven by showing that a real-world protocol emulates an ideal world where a trusted third party safely handles the task [6].

Theorem 4 gives the privacy and accuracy guarantees for the ideal views for solving the Hamming distance problem. We now describe, in more detail, how to implement the common-output distributed-noise ideal functionality. Throughout, we denote this functionality as $\mathcal{F}$.

a) Technical Explanation of Algorithm 2: The ideal functionality separates the privacy noise into two independently generated shares. Alice knows A, Bob knows B, and both parties receive only the common release

$$Z = \mathrm{HD}(x, y) + A + B \pmod{m}, \qquad m = 4n + 1.$$

The use of two noise shares is essential for a common-output protocol. If Alice generated all of the noise, she could subtract it from Z and recover the exact Hamming distance; the analogous problem would arise if Bob generated all of the noise. Instead,

Algorithm 2 Ideal Common-Output Distributed-Noise Functionality

Require: Alice has $x \in \{0,1\}^n$; Bob has $y \in \{0,1\}^n$; parameter $\varepsilon > 0$.

Ensure: Alice and Bob receive the same estimate of $d = \mathrm{HD}(x, y)$.

1: Set $m \leftarrow 4n + 1$.

2: Define the cyclic distance

$$d_m(u, v) := \min\{[u - v]_m, [v - u]_m\}.$$

3: Define the cyclic-geometric distribution

$$q_\varepsilon(g) := \frac{\exp(-\varepsilon d_m(g, 0))}{\sum_{h \in \mathbb{Z}_m} \exp(-\varepsilon d_m(h, 0))} \qquad (g \in \mathbb{Z}_m).$$

4: **Alice:** Sample $A \sim q_\varepsilon$.

5: **Bob:** Sample $B \sim q_\varepsilon$.

6: **Alice** $\to \mathcal{F}$**:** Send the classical pair

$$M_{A \to \mathcal{F}} = (x, A)$$

over Alice's ideal private interface.

7: **Bob** $\to \mathcal{F}$**:** Send the classical pair

$$M_{B \to \mathcal{F}} = (y, B)$$

over Bob's ideal private interface.

8: **Functionality** $\mathcal{F}$**:** Compute

$$Z \leftarrow \mathrm{HD}(x, y) + A + B \pmod{m}.$$

9: $\mathcal{F} \to$ **Alice:** Send the classical residue

$$M_{\mathcal{F} \to A} = Z \in \mathbb{Z}_m.$$

10: $\mathcal{F} \to$ **Bob:** Send the same classical residue

$$M_{\mathcal{F} \to B} = Z \in \mathbb{Z}_m.$$

11: Define

$$\mathrm{Dec}(z) := \min \arg \min_{r \in \{0, \ldots, n\}} d_m(z, r),$$

where ties are broken toward the smaller integer.

12: **Alice:** Output

$$\widehat{d}_A \leftarrow \mathrm{Dec}(Z).$$

13: **Bob:** Output

$$\widehat{d}_B \leftarrow \mathrm{Dec}(Z).$$

14: **return**

$$\widehat{d}_A = \widehat{d}_B = \mathrm{Dec}(Z).$$

each party knows only one share and must rely on the other party's unknown share for privacy.

For example, Alice's ideal view is

$$V_A^{\text{ideal}} = (x, A, Z).$$

Conditioned on $A = a$, its probability mass is

$$\Pr[A = a, Z = z \mid x, y] = q_\varepsilon(a)\, q_\varepsilon(z - \mathrm{HD}(x, y) - a)\,.$$

Changing one bit of y changes $\mathrm{HD}(x, y)$ by at most one. Since the cyclic-geometric distribution satisfies

$$e^{-\varepsilon} \le \frac{q_\varepsilon(g+1)}{q_\varepsilon(g)} \le e^{\varepsilon},$$

the likelihood ratio between Alice's views on adjacent values of y is at most e^{ε}. The symmetric argument protects Alice's input in Bob's view. Notice that the privacy parameter is ε, rather than 2ε: for either observer, only the other party's unknown noise share must absorb the unit change in Hamming distance.

The modulus $m = 4n+1$ is chosen so that the valid distance interval

$$D = \{0, \ldots, n\}$$

occupies a sufficiently short arc of the cycle. The nearest-point decoder therefore satisfies

$$|\mathrm{Dec}(Z) - \mathrm{HD}(x, y)| \le d_m(A+B, 0) \le d_m(A, 0) + d_m(B, 0).$$

Each cyclic-geometric share has expected circular magnitude at most $1/\sinh\varepsilon$, and hence

$$\mathbb{E}\left[|\mathrm{Dec}(Z) - \mathrm{HD}(x, y)|\right] \le \frac{2}{\sinh\varepsilon}.$$

The factor of two is a utility cost of requiring the parties to output the same private release: both independently generated shares appear in the common value Z.

B. Exact Classical Reversible Protocol

The prescribed-channel (PC) realization shows that the ideal common-output functionality can already be implemented exactly using classical reversible computation. Alice samples her noise share A, Bob samples his noise share B, and the parties reversibly compute

$$Z = \mathrm{HD}(x, y) + A + B \pmod m.$$

Once Z has been computed, it is copied into permanent output registers for both parties. The parties then run the reversible computation backward, clearing the temporary message and workspace registers while leaving the output registers unchanged. At termination, Alice's view is therefore exactly (x, A, Z), while Bob's view is exactly (y, B, Z), matching the views of the ideal functionality.

The important point is that reversibility does not make the intermediate messages private. While the computation is being carried out, the computational-basis messages may reveal substantial information about the parties' inputs. The PC model simply does not include copies of these intermediate messages in the final view unless the prescribed protocol explicitly retains them. Because the protocol uncomputes its temporary state, only the intended noisy output remains. Thus, in the PC model, the classical protocol inherits exactly the privacy and accuracy guarantees of the ideal functionality.

This construction also explains why the PC model cannot by itself demonstrate a quantum advantage. Classically, the parties can compute the functionality, copy out the noisy result, and erase the remaining transcript by reversible uncomputation. The distinction with the KHNP model is that a KHNP observer is allowed to retain additional information while still sending the prescribed messages. A classical KHNP observer can therefore copy the interaction transcript before it is uncomputed, whereas the PC model only evaluates the registers left by the prescribed implementation. This difference is why the classical lower bounds apply in the KHNP model but not directly to the weaker PC realization.

a) Technical Explanation of Algorithm 3: Algorithm 3 is an exact realization of the ideal functionality in the prescribed-channel model. The parties first evaluate

$$F_m((x, A), (y, B)) = \mathrm{HD}(x, y) + A + B \pmod m$$

by reversible classical computation. Once the residue Z has been computed, it is copied into permanent output registers Z_A and Z_B. The parties then reverse the computation, returning every temporary message and workspace register to its fixed initial state.

The role of reversibility is not to make the intermediate messages private. Indeed, intermediate computational-basis messages may reveal substantial information while they are present. Rather, reversibility ensures that, under the prescribed-channel view model, the designated terminal views contain no message history:

$$V_A^{\text{PC}} = (x, A, Z_A), \qquad V_B^{\text{PC}} = (y, B, Z_B).$$

These are exactly the views supplied by the ideal functionality. The pure ε-DP and $2/\sinh\varepsilon$ error guarantees therefore follow directly from the ideal analysis.

This construction also explains why the prescribed-channel model cannot support a meaningful quantum–classical separation. The model requires the parties to apply the specified uncomputation and discard operations and scores only the resulting designated terminal registers. An observer is not allowed to retain an additional copy of the classical transcript unless that copy is part of the prescribed channel. Consequently, the standard classical two-party lower bounds, which score transcript-and-coins views, do not apply directly to this terminal-view model.

This point is specific to the prescribed-channel model. Under Klauck's honesty model, a classical player may copy every message, retain the copy, and still send the prescribed outgoing messages. Thus, the same classical lower bounds do apply to classical KHNP protocols. Proposition 6 shows that it is necessary to distinguish these two notions of an honest view before claiming a quantum advantage.

Algorithm 3 Exact Classical Reversible Realization in the Prescribed-Channel Model

Require: Alice has $x \in \{0,1\}^n$; Bob has $y \in \{0,1\}^n$; parameter $\varepsilon > 0$.

Ensure: Both parties output the same estimate of $\mathrm{HD}(x, y)$.

1: Set
$$m \leftarrow 4n + 1, \qquad \ell \leftarrow \lceil \log_2 m \rceil .$$
2: **Alice:** Sample $A \sim q_\varepsilon$.
3: **Bob:** Sample $B \sim q_\varepsilon$.
4: Alice initializes an $(n + \ell)$-bit message register M_{XA} to 0.
5: Alice reversibly copies her classical input and noise share:
$$|x, A\rangle_{XA} |0\rangle_{M_{XA}} \longmapsto |x, A\rangle_{XA} |x, A\rangle_{M_{XA}} .$$
6: **Alice** $\rightarrow$ **Bob:** Send the computational-basis message
$$M_{XA} = |x, A\rangle .$$
7: Bob initializes an ℓ-bit temporary register T to $|0\rangle$.
8: Bob reversibly computes
$$|x, A\rangle_{M_{XA}} |y, B\rangle_{YB} |0\rangle_T \longmapsto |x, A\rangle_{M_{XA}} |y, B\rangle_{YB} |Z\rangle_T ,$$
where
$$Z = \mathrm{HD}(x, y) + A + B \pmod m.$$
9: Bob initializes a permanent output register Z_B to 0 and copies the result:
$$|Z\rangle_T |0\rangle_{Z_B} \longmapsto |Z\rangle_T |Z\rangle_{Z_B} .$$
10: Bob initializes an ℓ-bit message register M_Z to 0 and copies
$$|Z\rangle_T |0\rangle_{M_Z} \longmapsto |Z\rangle_T |Z\rangle_{M_Z} .$$
11: **Bob** $\rightarrow$ **Alice:** Send the computational-basis message
$$M_Z = |Z\rangle .$$
12: Alice initializes a permanent output register Z_A to 0 and copies
$$|Z\rangle_{M_Z} |0\rangle_{Z_A} \longmapsto |Z\rangle_{M_Z} |Z\rangle_{Z_A} .$$
13: Alice reversibly clears the received message register:
$$|Z\rangle_{M_Z} |Z\rangle_{Z_A} \longmapsto |0\rangle_{M_Z} |Z\rangle_{Z_A} .$$
14: Bob runs his computation of Z backward, clearing T and every temporary work register while leaving Z_B unchanged.
15: **Bob** $\rightarrow$ **Alice:** Return the computational-basis register
$$M_{XA} = |x, A\rangle .$$
16: Alice reversibly clears this returned copy using her retained register XA:
$$|x, A\rangle_{XA} |x, A\rangle_{M_{XA}} \longmapsto |x, A\rangle_{XA} |0\rangle_{M_{XA}} .$$
17: **Alice:** Output
$$\widehat{d}_A \leftarrow \mathrm{Dec}(Z_A).$$
18: **Bob:** Output
$$\widehat{d}_B \leftarrow \mathrm{Dec}(Z_B).$$
19: **return**
$$\widehat{d}_A = \widehat{d}_B.$$

C. Noise-Calibrated Guarded Protocol for Approximate QDP

The approximate QDP protocol uses the same guarded coherent communication structure as the pure QDP construction in the main body; the key difference is how the cyclic-geometric noise is calibrated. If the pure protocol were run with noise parameter $\alpha = \varepsilon$, it would automatically satisfy (ε, δ)-QDP for any $\delta > 0$, but the additional δ-budget would not be used to improve accuracy. So to use up this budget, the protocol computes the exact finite-cycle hockey-stick divergence between neighboring shifts of the cyclic-geometric distribution and chooses a larger parameter $\alpha_n^\star(\varepsilon, \delta) > \varepsilon$ that exactly meets the desired (ε, δ)-privacy level. Because larger α makes the cyclic-geometric distribution more concentrated around zero, Alice and Bob can add less noise while preserving the target privacy guarantee.

The guarded quantum realization itself is unchanged. Alice and Bob sample independent noise shares from the more concentrated distribution $q_{\alpha_n^\star}$, and Alice sends the same guarded non-orthogonal encoding used in the pure QDP protocol. Bob coherently appends the noisy Hamming distance residue and returns the prescribed pure state. The equal-Gram argument continues to apply because it does not depend on the value of the noise parameter. Consequently, each party's complete KHNP terminal view is still obtained by input-independent post-processing of the corresponding ideal view, so the hockey-stick divergence can only decrease under the quantum realization. In particular, no additional δ-loss is introduced by implementing the ideal functionality in the quantum setting.

The resulting protocol therefore converts the approximate-privacy budget directly into improved utility. Its expected error is bounded by

$$U_n\left(\alpha_n^\star(\varepsilon, \delta)\right) + \gamma \le \frac{2}{\sinh\left(\alpha_n^\star(\varepsilon, \delta)\right)} + \gamma < \frac{2}{\sinh \varepsilon} + \gamma,$$

while retaining $O(n)$ communication and the same exact message-preserving privacy argument as in the pure QDP construction. Thus, approximate privacy provides a genuine accuracy improvement (rather than merely inheriting the pure QDP guarantee).

a) Numerical savings from approximate privacy: The approximate QDP calibration permits a larger cyclic-geometric parameter

$$\overline{\alpha}(\varepsilon, \delta) = \log\left(\frac{e^\varepsilon + \delta}{1 - \delta}\right) > \varepsilon,$$

and therefore a smaller noise bound.

For convenience, denote the pure and approximate QDP expected error bounds as $E_{\mathrm{pure}}(\varepsilon)$ and $E_{\mathrm{approx}}(n, \varepsilon, \delta)$. Ignoring the arbitrarily small additive guard penalty γ, the approximate QDP calibration strictly improves on the pure-QDP bound:

$$E_{\mathrm{approx}}(n, \varepsilon, \delta) < E_{\mathrm{pure}}(\varepsilon),$$

since

$$\alpha_n^\star(\varepsilon, \delta) > \varepsilon.$$

Algorithm 4 Noise-Calibrated Guarded Protocol for Approximate KHNP-QDP

Require: Alice has $x \in \{0,1\}^n$; Bob has $y \in \{0,1\}^n$; privacy parameters $\varepsilon > 0$ and $0 < \delta < 1$; utility parameter $\gamma > 0$.

Ensure: A common output with (ε, δ)-QDP in the Klauck-honest nonpreemptive model.

Set

$$m \leftarrow 4n+1, \qquad \ell \leftarrow \lceil \log_2 m \rceil, \qquad \kappa \leftarrow \frac{\gamma}{n+\gamma}.$$

Fix a public fallback residue $z_\perp \in \mathbb{Z}_m$.

Define, for $\alpha \geq \varepsilon$,

$$\Delta_{n,\varepsilon}(\alpha) := \left(1 - e^{\varepsilon - \alpha}\right) \frac{1 - e^{-2n\alpha}}{1 + e^{-\alpha} - 2e^{-(2n+1)\alpha}}.$$

Let $\alpha_n^\star(\varepsilon, \delta)$ be the unique solution to

$$\Delta_{n,\varepsilon}\left(\alpha_n^\star(\varepsilon,\delta)\right) = \delta.$$

Set

$$\alpha \leftarrow \alpha_n^\star(\varepsilon, \delta).$$

Alice: Sample $A \sim q_\alpha$.

Bob: Sample $B \sim q_\alpha$.

Alice prepares

$$|\psi_{x,A}\rangle_{CM} = \sqrt{1-\kappa}\,|0\rangle_C\,|x,A\rangle_M + \sqrt{\kappa}\,|1\rangle_C\,|\perp\rangle_M.$$

Alice $\rightarrow$ **Bob:** Send the quantum message

$$M_{A\to B} = |\psi_{x,A}\rangle_{CM}.$$

Bob initializes an ℓ-qubit residue register R to $|0\rangle$.

Bob applies a controlled reversible operation satisfying

$$|0\rangle\,|x,A\rangle\,|0\rangle \longmapsto |0\rangle\,|x,A\rangle\,|\mathrm{HD}(x,y) + A + B \bmod m\rangle$$

and

$$|1\rangle\,|\perp\rangle\,|0\rangle \longmapsto |1\rangle\,|\perp\rangle\,|z_\perp\rangle.$$

The resulting prescribed return state is

$$|\phi_{x,A}^{y,B}\rangle_{CMR} = \sqrt{1-\kappa}\,|0\rangle_C\,|x,A\rangle_M\,|\mathrm{HD}(x,y) + A + B \bmod m\rangle_R + \sqrt{\kappa}\,|1\rangle_C\,|\perp\rangle_M\,|z_\perp\rangle_R.$$

Bob $\rightarrow$ **Alice:** Return the entire quantum message

$$M_{B\to A} = |\phi_{x,A}^{y,B}\rangle_{CMR}.$$

Alice measures C and R in the computational basis, obtaining (c, z).

Set

$$Z \leftarrow z.$$

▷ On the fallback branch, $Z = z_\perp$.

Alice $\rightarrow$ **Bob:** Send the classical residue

$$M_{A\to B}^{\mathrm{final}} = Z \in \mathbb{Z}_m.$$

Alice: Output

$$\widehat{d}_A \leftarrow \mathrm{Dec}(Z).$$

Bob: Output

$$\widehat{d}_B \leftarrow \mathrm{Dec}(Z).$$

return

$$\widehat{d}_A = \widehat{d}_B.$$

Ignoring the common arbitrarily small additive guard penalty γ, the pure and approximate bounds are

$$E_{\mathrm{pure}} \leq \frac{2}{\sinh \varepsilon}, \qquad E_{\mathrm{approx}} \leq \frac{2}{\sinh(\overline{\alpha}(\varepsilon,\delta))}.$$

The exact finite-n calibration $\alpha_n^\star(\varepsilon,\delta)$ is at least as accurate as this explicit bound.

For example, let $\varepsilon = 0.5$ and $\delta = 0.01$. Pure privacy gives

$$E_{\mathrm{pure}} \leq \frac{2}{\sinh(0.5)} \approx 3.838.$$

In the approximate regime,

$$\overline{\alpha}(0.5, 0.01) = \log\left(\frac{e^{0.5} + 0.01}{0.99}\right) \approx 0.5161,$$

and hence

$$E_{\mathrm{approx}} \leq \frac{2}{\sinh(0.5161)} \approx 3.708.$$

Thus the explicit approximate-DP calibration lowers the expected-error upper bound by approximately

$$3.838 - 3.708 \approx 0.130,$$

corresponding to a reduction of about 3.4%.

The savings can be more pronounced in a stronger-privacy regime. For $\varepsilon = 0.1$ and $\delta = 0.01$, pure privacy gives

$$E_{\mathrm{pure}} \leq \frac{2}{\sinh(0.1)} \approx 19.967.$$

Approximate privacy permits

$$\overline{\alpha}(0.1, 0.01) = \log\left(\frac{e^{0.1} + 0.01}{0.99}\right) \approx 0.1191,$$

which gives

$$E_{\mathrm{approx}} \leq \frac{2}{\sinh(0.1191)} \approx 16.759.$$

The reduction is therefore approximately

$$19.967 - 16.759 \approx 3.208,$$

or about 16.1%.

For small ε and δ,

$$\overline{\alpha}(\varepsilon,\delta) \approx \varepsilon + 2\delta,$$

so the leading comparison is approximately

$$E_{\mathrm{pure}} \approx \frac{2}{\varepsilon}, \qquad E_{\mathrm{approx}} \approx \frac{2}{\varepsilon + 2\delta}.$$

Thus approximate privacy can provide a meaningful constant-factor reduction when δ is nonnegligible relative to ε. In the strong approximate-privacy regime $\delta = o(1/n)$, however, this numerical saving is lower order, and both quantum bounds remain $O(1)$ for constant ε.

D. A Common-Output Retention-Robust Baseline

Quantum communication can always be used to lift a classical protocol into the quantum setting. Slight extensions of previous works (e.g., [3], [11]) can be used to obtain the guarantees we outline in this section, since every DP algorithm is a QDP algorithm. However, previous works do not use the exact constructions we use in this work. Thus, for the sake of completeness and ease of comparison, we present our classical protocol here.

We define a clipping operator that will be used in this section. For $a \le b$, define the *clipping* operator

$$\mathrm{clip}_{[a,b]}(z) := \min\{b, \max\{a, z\}\}.$$

Equivalently,

$$\mathrm{clip}_{[a,b]}(z) = \begin{cases} a, & z < a, \\ z, & a \le z \le b, \\ b, & z > b. \end{cases}$$

The following asymmetric protocol uses randomized response for Alice's input and geometric output noise for Bob's input; its released estimate is common to both parties.

Theorem 10 (Common-output pure RR-QDP baseline)**.** *For every $n \ge 1$ and $\varepsilon > 0$, there is a two-message protocol using $n + O(\log n)$ classical bits (or computational-basis qubits) whose honest outputs agree with probability one and which is pure ε-RR-QDP under all quantifiers of Definition 4. Its expected absolute error is at most*

$$\frac{\sqrt{n}}{2\sinh(\varepsilon/2)} + \frac{1}{2\sinh^2(\varepsilon/2)}.$$

Proof. Let $p = (1 + e^{\varepsilon})^{-1}$ and $s = 1 - 2p = \tanh(\varepsilon/2)$. Independently for each coordinate, Alice sends $X'_i = x_i$ with probability $1 - p$ and $1 - x_i$ with probability p. Bob computes

$$C = \sum_i \mathbf{1}\{X'_i \ne y_i\}, \qquad \widetilde{d} = \frac{C - np}{s}.$$

He samples an independent two-sided geometric random variable

$$\Pr[G = g] = \frac{1 - e^{-\varepsilon}}{1 + e^{-\varepsilon}} e^{-\varepsilon|g|} \quad (g \in \mathbb{Z}),$$

sets $\widehat{d}$ to the clipping of $\widetilde{d} + G/s$ to $[0, n]$, and sends a finite encoding of this clipped value to Alice. Both parties output that value.

Fix any corrupted-Bob strategy and initial side state permitted by Definition 4. Alice's only input-dependent transmission is the randomized-response vector. A one-bit change of x changes its likelihood by at most e^{ε}; the adversary's arbitrary CPTP interaction, retention, or early termination is postprocessing. Hence Bob's complete state satisfies pure ε-QDP.

Now fix any corrupted-Alice strategy. Bob first measures the purported vector in the computational basis, obtaining some X'. The induced classical–quantum state of X' and Alice's retained register is independent of y. Conditional on every X', changing one bit of y changes C by at most one. The geometric mechanism therefore obeys the likelihood-ratio bound e^{ε} pointwise in X'. Multiplying each pointwise inequality by the same conditional retained state and summing proves the operator inequality for Alice's joint final view. This proves pure RR-QDP in both directions, for every allowed initial side state.

In an honest run, $\widetilde{d}$ is unbiased and

$$\mathrm{Var}(\widetilde{d}) = \frac{np(1-p)}{s^2} = \frac{n}{4\sinh^2(\varepsilon/2)}.$$

Moreover, $\mathbb{E}|G| = 1/\sinh\varepsilon$ and

$$\frac{\mathbb{E}|G|}{s} = \frac{1}{\tanh(\varepsilon/2)\sinh\varepsilon} = \frac{1}{2\sinh^2(\varepsilon/2)}.$$

Since $d = \mathrm{HD}(x, y) \in [0, n]$ and $\widehat{d} = \mathrm{clip}_{[0,n]}(z)$, projection onto the interval $[0, n]$ cannot increase the distance to d. Hence

$$|\widehat{d} - d| \le |z - d|.$$

Thus, clipping cannot increase distance from d. The triangle inequality and Cauchy–Schwarz give the stated error bound. After clipping there are only $O(n)$ possible interior values plus two endpoints, so the common output uses $O(\log n)$ bits. □

a) Technical Explanation of Algorithm 5: Algorithm 5 is designed for the stronger retention-robust model, in which a corrupted party may apply arbitrary CPTP maps, retain auxiliary systems, change its prescribed messages, or abort. Unlike the KHNP construction, its security does not rely on forcing a party to return a non-orthogonal quantum state. Instead, every input-dependent message is differentially private before it reaches the recipient.

Alice first applies independent randomized response to every input bit. With

$$p = \frac{1}{1 + e^{\varepsilon}},$$

she sends

$$X'_i = \begin{cases} x_i, & \text{with probability } 1 - p, \\ 1 - x_i, & \text{with probability } p. \end{cases}$$

For adjacent inputs x and x', only one coordinate distribution changes, and the likelihood ratio of the complete vector X' is at most e^{ε}. Therefore X' is already pure ε-DP for Alice's input. Any operation subsequently performed by a corrupted Bob (including coherent processing, retention, measurement, or early termination) is postprocessing of this private message and cannot weaken its privacy guarantee.

In an honest execution, Bob computes

$$C = \sum_{i=1}^{n} \mathbf{1}\{X'_i \ne y_i\}.$$

Let

$$d = \mathrm{HD}(x, y), \qquad s = 1 - 2p = \tanh(\varepsilon/2).$$

Algorithm 5 Common-Output Retention-Robust Randomized-Response Protocol

Require: Alice has $x \in \{0,1\}^n$; Bob has $y \in \{0,1\}^n$; privacy parameter $\varepsilon > 0$.

Ensure: A common output satisfying pure ε-retention-robust QDP.

1: Set
$$p \leftarrow \frac{1}{1+e^{\varepsilon}}, \qquad s \leftarrow 1-2p = \tanh(\varepsilon/2).$$
2: **for** $i = 1, \ldots, n$ **do**
3: Alice independently samples
$$N_i \sim \text{Bernoulli}(p).$$
4: Alice sets
$$X_i' \leftarrow x_i \oplus N_i.$$
5: **end for**
6: Let
$$X' = (X_1', \ldots, X_n') \in \{0,1\}^n.$$
7: **Alice** $\rightarrow$ **Bob:** Send the classical message
$$M_{A\to B} = X'.$$
$\triangleright$ Equivalently, send the basis state $|X'\rangle$.
8: If the message was sent as qubits, honest Bob measures it in the computational basis.
9: Bob computes
$$C \leftarrow \sum_{i=1}^{n} \mathbf{1}\{X_i' \neq y_i\}.$$
10: Bob forms the debiased estimator
$$\widetilde{d} \leftarrow \frac{C-np}{s}.$$
11: Bob independently samples a two-sided geometric random variable $G \in \mathbb{Z}$ with
$$\Pr[G=g] = \frac{1-e^{-\varepsilon}}{1+e^{-\varepsilon}} e^{-\varepsilon|g|} \qquad (g \in \mathbb{Z}).$$
12: Bob computes
$$\widehat{d} \leftarrow \text{clip}_{[0,n]}\left(\widetilde{d} + \frac{G}{s}\right).$$
13: Bob encodes the result using a public finite encoding:
$$W \leftarrow \text{Enc}_{\varepsilon}(\widehat{d}) \in \{0,1\}^{O(\log n)}.$$
14: **Bob** $\rightarrow$ **Alice:** Send the classical message
$$M_{B\to A} = W.$$
15: Alice decodes
$$\widehat{d}_A \leftarrow \text{Enc}_{\varepsilon}^{-1}(W).$$
16: Bob sets
$$\widehat{d}_B \leftarrow \widehat{d}.$$
17: **return**
$$\widehat{d}_A = \widehat{d}_B.$$

Each coordinate that originally agrees contributes a mismatch with probability p, whereas each coordinate that originally disagrees contributes a mismatch with probability $1-p$. Hence

$$\mathbb{E}[C \mid x, y] = np + sd,$$

so

$$\widetilde{d} = \frac{C-np}{s}$$

is an unbiased estimate of d.

The first message does not protect Bob's input from Alice, since Alice generated it without using y. Bob therefore adds an independent two-sided geometric variable G before sending the common estimate back:

$$\widehat{d} = \text{clip}_{[0,n]}\left(\frac{C-np+G}{s}\right).$$

Conditional on every possible received value X', changing one bit of y changes C by at most one. The geometric distribution satisfies the corresponding e^{ε} likelihood-ratio bound, so $C+G$, and hence $\widehat{d}$, is pure ε-DP for Bob's input.

This privacy argument remains valid even when Alice deviates from the protocol. Honest Bob first measures Alice's purported message in the computational basis, obtaining a classical string X'. Conditional on that outcome, Alice's retained quantum register is independent of y. The pointwise geometric-mechanism inequality can therefore be multiplied by the same conditional retained state and summed over X', yielding the required operator inequality for Alice's complete final view.

The two uses of ε do not compose to 2ε. Privacy is required separately in the two directions. To protect x from Bob, the only x-dependent transmission is Alice's randomized-response vector. To protect y from Alice, Alice's initial message is independent of y, and only Bob's geometric release consumes privacy. Each protected input therefore incurs a single ε-privacy loss.

For utility, independence across coordinates gives

$$\text{Var}(\widetilde{d}) = \frac{np(1-p)}{s^2} = \frac{n}{4\sinh^2(\varepsilon/2)}.$$

Thus,

$$\mathbb{E}\left[\left|\widetilde{d}-d\right|\right] \le \sqrt{\text{Var}(\widetilde{d})} = \frac{\sqrt{n}}{2\sinh(\varepsilon/2)}.$$

Moreover,

$$\frac{\mathbb{E}|G|}{s} = \frac{1}{2\sinh^2(\varepsilon/2)}.$$

Since clipping to $[0,n]$ cannot increase the distance from $d \in [0,n]$, the triangle inequality yields

$$\mathbb{E}\left[\left|\widehat{d}-d\right|\right] \le \frac{\sqrt{n}}{2\sinh(\varepsilon/2)} + \frac{1}{2\sinh^2(\varepsilon/2)}.$$

The protocol uses $n + O(\log n)$ classical bits, or equivalently computational-basis qubits. It establishes a fully retention-robust baseline but does not give a quantum advantage: its leading error remains $O(\sqrt{n})$ for constant ε. Whether genuinely quantum communication can improve this accuracy while

remaining secure against all retention-capable CPTP strategies is left open.

E. Statement on AI Use

The authors generated the main ideas, algorithms, and protocols. The pure DP protocol is also an approximate DP protocol (for $\delta = 0$). However, using GPT 5.6 Sol, the authors optimized the noise mechanism to obtain a more accurate approximate DP protocol. Furthermore, the images were generated in LaTeX via the use of GPT 5.6 Sol. The authors independently checked the mathematical expressions and figures.