Doubly infinite separation of quantum information and communication

Zi-Wen Liu,[1,*] Christopher Perry,[2] Yechao Zhu,[1] Dax Enshan Koh,[3] and Scott Aaronson[4]

[1] *Center for Theoretical Physics and Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

[2] *Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom*

[3] *Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

[4] *Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

(Received 22 July 2015; revised manuscript received 30 November 2015; published 29 January 2016)

We prove the existence of (one-way) communication tasks with a subconstant versus superconstant asymptotic gap, which we call "doubly infinite," between their quantum information and communication complexities. We do so by studying the exclusion game [C. Perry *et al.*, Phys. Rev. Lett. **115**, 030504 (2015)] for which there exist instances where the quantum information complexity tends to zero as the size of the input n increases. By showing that the quantum communication complexity of these games scales at least logarithmically in n, we obtain our result. We further show that the established lower bounds and gaps still hold even if we allow a small probability of error. However in this case, the n-qubit quantum message of the zero-error strategy can be compressed polynomially.

DOI: 10.1103/PhysRevA.93.012347

I. INTRODUCTION

The field of communication complexity, originated by Yao's 1979 seminal work [1], aims to study the minimum amount of communication needed for multiple distributed parties to accomplish a given communication task. Such tasks are typically formalized as follows: Players are given private inputs and asked to solve some computational problems based on them. To do this, some communication will have to take place in the form of exchanging messages.

While such models were originally considered in the context of classical protocols, it has since been realized that quantum resources, e.g., quantum communication channels (players are allowed to exchange quantum states instead of classical messages), may offer significant advantage. For example, there exist tasks for which quantum strategies can consume exponentially less communication than any classical one, even without shared entanglement [2–7]. These results sparked interest in further characterizing which tasks exhibit distinctions between quantum and classical communication protocols, and what kind of distinctions there can be. The vast majority of previous work in this field was carried out in the constant bounded-error setting. Here, we shall focus on a scenario where the allowed probability of error is zero or vanishingly small.

Recently, a peculiar type of one-way communication task between two players Alice and Bob, namely the exclusion game, was introduced by Perry, Jain, and Oppenheim (PJO) [8]: Alice randomly draws an n-bit string x, and Bob randomly draws some subset $y \subseteq [n]$, where $|y| = m$, both from uniform distributions. Alice then sends a single message regarding her input to Bob. They win the game if Bob outputs a string z that is different from x restricted to the bits specified by y. A particular exclusion game can be denoted by $\mathrm{EXC}_{n,m,\gamma}$, where γ is the allowed probability of error. Comparing to the conventional bounded-error tasks of computing functions, exclusion games exhibit some special properties. They are relational tasks: multiple outputs can be accepted for one certain input; and they are extremely sensitive to error: if $\gamma \geqslant 2^{-m}$, then no communication is required as Bob can succeed at his task by guessing a string at random. In Ref. [8], PJO demonstrated a new kind of quantum-classical separation: They devised a zero-error quantum strategy that only reveals vanishingly small amount of information regarding Alice's input for the exclusion games with large m, while any zero-error classical strategy must reveal almost everything. More formally, there is an infinite gap between the quantum and classical information complexities (the minimum amount of information regarding Alice's private input that needs to be revealed) of these exclusion games.

In this paper, we further analyze the complexities of different exclusion games, and exhibit several features. The PJO strategy requires that exactly n qubits be sent from Alice to Bob, i.e., the communication cost is n. Since the amount of information actually revealed is vanishingly small, an interesting question that naturally arises is as to how much we can possibly reduce the communication cost. For zero-error exclusion games with m scaling strongly sublinearly in n, we show that any winning quantum strategy can be classically simulated with at most exponential overhead. Combining with the linear lower bound on the classical communication complexity, we establish a logarithmic lower bound on the quantum communication complexity of these games. As a result, there is an at least subconstant versus logarithmic gap between the quantum information and communication complexities of the exclusion games for which both sides hold simultaneously (they exist). That is, a vanishingly small amount of extractable information must be carried by a diverging amount of communication for these tasks. This gap is an example of doubly infinite gaps, which we shall motivate and define later.

Next, we extend our analysis to the cases where error may be allowed ($\gamma > 0$). By slightly different arguments, we show that the overhead of classically simulating a quantum strategy is still at most exponential for small γ. Furthermore, for $\gamma \leqslant (n+1)^{-m}$, we show that the classical communication complexity is still at least linear, thus the logarithmic lower bound on the quantum communication complexity and the doubly infinite gap between the quantum information and communication complexities of certain exclusion games hold.

[*] zwliu@mit.edu

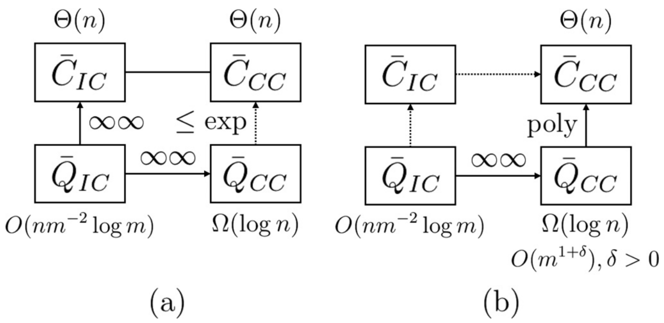

FIG. 1. Complexities of $\mathrm{EXC}_{n,m,\gamma}$ with (a) $m \in \omega(\sqrt{n\log n})$, $m \in \tilde{o}(n)$, $\gamma = 0$; (b) $m \in \omega(\sqrt{n\log n})$, $m \in \tilde{o}(n)$, $\gamma = (n+1)^{-m}$. Solid arrows denote established gaps (pointing towards the larger complexity), while the dashed ones denote unknown gaps. "$\infty\infty$" means doubly infinite.

The significance of the doubly infinite gap between quantum information and communication complexities may be compared with its classical counterpart. For constant nonzero probability of error, the gap between classical information and communication complexities is at most exponential for any communication task [9,10]. For zero and asymptotically vanishing probability of error, the largest known gap is constant versus superconstant ("singly infinite") and occurs for the equality function [10]. Our results may lead to a better understanding of the relation between information and communication complexities, which is a major objective of recent research in the field of communication complexity (in both classical and quantum settings).

We should mention that with regards to the gaps between quantum and classical communication complexities, it was shown in Ref. [11] that for computing functions in the bounded-error model (without shared entanglement or randomness), they can be at most exponential. As the exclusion game is a relational problem and the interesting separations occur only when the probability or error is zero or tends to zero asymptotically, the arguments of Ref. [11] cannot be directly applied here. Our results indicate that the conclusion holds for almost all exclusion games. However, it indeed remains open as to whether the gap can be superexponential for those games with m scaling linearly in n.

In addition, we show that $\gamma \geqslant (n+1)^{-m}$ allows the n-qubit quantum communication of the PJO strategy to be compressed at least polynomially. Most of our results are summarized in Fig. 1.

II. COMPLEXITIES OF COMMUNICATION TASKS

Two types of information-theoretic quantities associated with a certain communication task are of great interest and importance in our context, namely, the communication [1,12] and information [13] complexities. Here, we formally define them.

The communication cost of a λ protocol Π [where $\lambda = C$ (classical) or $= Q$ (quantum) in our context], denoted by $\lambda_{CC}(\Pi)$, is defined to be the maximum number of bits or qubits that are exchanged in any run of the protocol, where the maximum is taken over all inputs and the value of any randomness used.

The information cost of a λ protocol Π, denoted by $\lambda_{IC}(\Pi)$, aims to measure the amount of information regarding the players' inputs revealed by Π. Here, we consider one-way protocols, i.e., the communication is only from Alice to Bob. Suppose that X and Y are, respectively, Alice and Bob's inputs, distributed according to a joint distribution μ. Then, $\lambda^{\mu}_{IC}(\Pi) = I(X:\Pi|Y)$, where Π on the right-hand side essentially denotes the message exchanged during the protocol together with the public randomness used, and $I(S:T|U) = H(SU) + H(TU) - H(STU) - H(U)$ measures the mutual information between S and T given knowledge of U [14]. The distribution-independent information cost is then defined to be $\lambda_{IC}(\Pi) = \sup_{\mu}\lambda^{\mu}_{IC}(\Pi)$ [10].

Complexities measure the minimum possible amount of certain costs that need to take place to accomplish the task, where the minimization is taken over all winning protocols. The λ communication complexity of a task Ξ is then defined to be $\bar{\lambda}_{CC}(\Xi) = \inf_{\Pi_{\Xi}}\lambda_{CC}(\Pi_{\Xi})$, where Π_{Ξ} are all winning λ protocols for Ξ. The distribution-dependent and distribution-independent λ information complexities of one-way tasks are defined similarly.

We emphasize that these quantities of interest are only associated with the communication between players. Players themselves can have unlimited access to any kind of local resources.

III. INFINITE ASYMPTOTIC GAPS

We are interested in the limiting behaviors of complexities as the size of the task n tends to infinity. Throughout this paper, we adopt the standard notation to describe asymptotic complexities. In addition to the widely used O, o, Ω, ω (Bachmann-Landau) symbols (formal definitions can be found in, e.g., Ref. [15]), the following soft symbols are also used when needed. For example, $\tilde{O}(n)$ (soft-O) means $O(n \,\mathrm{polylog}\, n)$, i.e., $O(n\log^k n)$ for some k, while $\tilde{o}(n)$ (soft-o) means $o(n \,\mathrm{polylog}\, n)$, i.e., $o(n\log^k n)$ for any k. Soft-Ω and soft-ω are defined analogously.

Now, we introduce the notion of infinite asymptotic gaps and discuss different types of such gaps in an intuitive manner. Formal definitions are left to Appendix A. The gap between two asymptotic complexities is normally characterized by a type of increasing monotone function. For example, there is a quadratic gap between $O(\sqrt{n})$ and $\Omega(n)$, and an exponential gap between $O(\log n)$ and $\Omega(n)$. However, when one side is $\omega(1)$, i.e., superconstant [or $o(1)$, i.e., subconstant], while the other side is not, the gap between them grows faster than any such monotone function. We regard such gaps as infinite. In fact, all (positive) asymptotic complexities belong to one of the following three classes: $o(1)$, $\Theta(1)$, or $\omega(1)$. In the logarithmic scale, these three types of asymptotic complexities tend to negative infinity, constant, positive infinity, respectively. The gap between any two of them is infinite. In particular, an $o(1)$ versus $\omega(1)$ gap can be regarded as doubly infinite, whereas an $o(1)$ versus $\Theta(1)$ or $\Theta(1)$ versus $\omega(1)$ gap is only singly infinite. Evidently, the gap between any two asymptotic complexities cannot be larger than doubly infinite. The general behaviors and comparisons of infinite gaps are illustrated in Fig. 2.

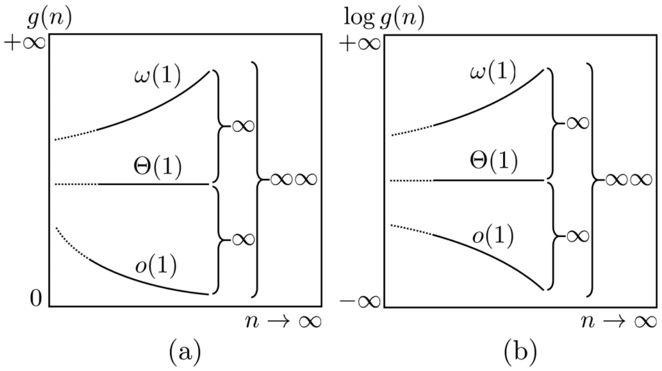

FIG. 2. Illustrations of infinite gaps in the (a) linear scale and (b) logarithmic scale. "∞" means singly infinite; "∞∞" means doubly infinite.

IV. EXCLUSION GAME

A communication task between two players is typically defined by a function $f : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^*$ ($\{0,1\}^*$ denotes the set of bit strings of arbitrary length): Alice and Bob are respectively given some string $x \in \{0,1\}^*$ and $y \in \{0,1\}^*$. They are allowed to exchange messages, and one of them outputs a string $z \in \{0,1\}^*$ in the end. They succeed at the task if $z = f(x,y)$.

To formally define the exclusion game, a more general framework of communication tasks is needed. The problem is now defined by a relation $R \subseteq \{0,1\}^* \times \{0,1\}^* \times \{0,1\}^*$. Furthermore, we restrict the protocol to be one way: Alice can send one message to Bob, and Bob outputs an answer. They succeed at the task if $(x,y,z) \in R$. It is evident that the general framework reduces to the original one if for all $x,y \in \{0,1\}^*$, there exists a unique $z \in \{0,1\}^*$ for which $(x,y,z) \in R$. Generically, relational tasks are those admitting multiple winning outputs for one certain input.

The exclusion game is a relational task defined by the relation $R_{\rm EXC} = \{(x,y,z) | z \neq \mathcal{M}_y(x)\}$: Alice's input $x \in \{0,1\}^n$ and Bob's input $y \subseteq [n], |y| = m$ (y can be encoded as a string to conform to the above general framework) are both drawn randomly from uniform distributions, and $\mathcal{M}_y(x)$ denotes the string given by x restricted to the bits specified by y. The winning condition is that Bob's output $z \neq \mathcal{M}_y(x)$, for given x and y.

PJO devised the following quantum strategy [8] that wins every exclusion game with certainty, i.e., works for any $\text{EXC}_{n,m,\gamma}$. Given the input $x = x_1 \ldots x_n$, Alice encodes each classical bit x_i as the qubit

$$|\phi(x_i;\theta_m)\rangle = \cos\left(\frac{\theta_m}{2}\right)|0\rangle + (-1)^{x_i} \sin\left(\frac{\theta_m}{2}\right)|1\rangle, \tag{1}$$

where $\theta_m = 2\tan^{-1}(2^{1/m} - 1)$. The n-bit string x is then encoded as the joint state

$$|\Phi(x;\theta_m)\rangle = \bigotimes_{i=1}^{n} |\phi(x_i;\theta_m)\rangle, \tag{2}$$

which she sends to Bob via the quantum channel. Upon receiving the state from Alice, Bob performs a global measurement across the m systems specified by y [denoted by $|\Phi(\mathcal{M}_y(x);\theta_m)\rangle$]. The measurement is given by

$$|\zeta(z)\rangle = \frac{1}{\sqrt{2^m}}\left[|0\rangle - \sum_{s\neq 0}(-1)^{z\cdot s}|s\rangle\right]. \tag{3}$$

As one can verify, $\langle\Phi(\mathcal{M}_y(x);\theta_m)|\zeta(\mathcal{M}_y(x))\rangle = 0$ [16]. That is, Bob always outputs some $z \neq \mathcal{M}_y(x)$ according to the measurement outcome. Therefore, they win the game with certainty. This measurement technique may be described as a conclusive-exclusion measurement. It was first introduced in Ref. [17], and was subsequently used to prove the Pusey-Barrett-Rudolph (PBR) theorem [18], a result in the field of quantum foundations that rules out a certain class of ψ-epistemic models of quantum mechanics.

The PJO strategy exhibits a striking property: The amount of information Alice actually reveals to Bob (the information cost) tends to zero as n increases, in a certain regime. More specifically, it can be calculated that $Q_{IC}(\text{PJO}) \leqslant 2S(M_Q) \in O(nm^{-2}\log m)$, where $S(M_Q)$ is the von Neumann entropy of the quantum message M_Q [the ensemble of $|\Phi(x;\theta_m)\rangle$ where x is an n-bit string with each of the 2^n possibilities being equally likely] that Alice sends to Bob. When $m \in \omega(\sqrt{n\log n})$, $\lim_{n\to\infty} Q_{IC}(\text{PJO}) = 0$. This directly indicates that $\lim_{n\to\infty} \bar{Q}_{IC}(\text{EXC}_{n,m,0}) = 0$ in the specified regime. Note that this actually holds for any prior distribution of inputs [8].

V. ZERO-ERROR QUANTUM COMMUNICATION COMPLEXITY

Here, we prove an $\Omega(\log n)$ lower bound on $\bar{Q}_{CC}(\text{EXC}_{n,m,0})$, when $m \in \tilde{o}(n)$. That is, there cannot exist any winning quantum strategy whose communication cost scales sublogarithmically in n in this regime.

The main idea of the proof is to approximately simulate any quantum protocol for $\text{EXC}_{n,m,0}$ by a classical protocol with exponential overhead, and show that the task can still be accomplished with zero probability of error. Because of the tiny possible probability of error associated with the exclusion game, the existence of such a simulation is nonobvious and the results of Ref. [11] cannot be applied, but we show that it can be made to work when $m \in \tilde{o}(n)$. Then, lower bounds on the classical communication complexities in this regime would directly imply exponentially smaller lower bounds on the corresponding quantum communication complexities. The following lemma sets a linear lower bound on the classical communication complexities of almost all exclusion games:

Lemma 1. For $m \leqslant \alpha n$ where $0 < \alpha < 1/2$ is a constant, $\bar{C}_{CC}(\text{EXC}_{n,m,0}) \in \Omega(n)$.

The proof of this lemma is given in Appendix B. Note that the applicable regime of this lemma covers $m \in \tilde{o}(n)$. This enables us to prove the following result:

Theorem 2. For $m \in \tilde{o}(n)$, $\bar{Q}_{CC}(\text{EXC}_{n,m,0}) \in \Omega(\log n)$.

Proof. Here, we only sketch the main steps of the proof. Details are given in Appendix C. Suppose that for $\text{EXC}_{n,m,0}$ with $m \in \tilde{o}(n)$, there exists a winning quantum strategy Π_Q such that $Q_{CC}(\Pi_Q) \equiv q \in o(\log n)$. Then based on Π_Q, we can devise a corresponding classical strategy Π_C such that $C_{CC}(\Pi_C) \in o(n)$ as follows. Given input x, the quantum message that Alice sends to Bob in Π_Q can be encoded as a $2q$-qubit pure state $|\psi(x)\rangle$. First, Alice prepares a classical message $C(|\psi(x)\rangle)$ that approximately encodes $|\psi(x)\rangle = \sum_{j=1}^{2^{2q}} \alpha_j |j\rangle = \sum_{j=1}^{2^{2q}} (b_j + i c_j)|j\rangle$, by registering the real (b_j) and imaginary parts (c_j) of all amplitudes (α_j) to accuracy $2^{-(m+q)}/20$ (the approximations are denoted by $\tilde{b}_j$ and $\tilde{c}_j$). It can be shown that $|C(|\psi(x)\rangle)| \in o(n)$, when $m \in \tilde{o}(n)$. Alice then sends $C(|\psi(x)\rangle)$ to Bob, whose local strategy can be considered as a positive-operator valued measure (POVM) $\{P_z\}$ with 2^m elements, each indicating an m-bit output string z. Bob first normalizes the amplitude vector encoded in $C(|\psi(x)\rangle)$, and then applies Born's rule to compute the approximate probability p_z of obtaining each z. Given the above accuracy of encoding, it can be shown that $p_{\mathcal{M}_y(x)} < 2^{-m}$. Therefore, Bob simply outputs a z such that $p_z > 2^{-m}$, which always exists. Since $C_{CC}(\Pi_C) \in o(n)$, we have reached a contradiction to Lemma 1. Therefore, no quantum strategies Π_Q such that $Q_{CC}(\Pi_Q) \in o(\log n)$ for $\text{EXC}_{n,m,0}$ with $m \in \tilde{o}(n)$ can exist: $\bar{Q}_{CC}(\text{EXC}_{n,m,0}) \in \Omega(\log n)$ in this regime. ■

This result directly implies the following gaps between complexities:

Corollary 3. For $\text{EXC}_{n,m,0}$ with $m \in \omega(\sqrt{n \log n})$ and $m \in \tilde{o}(n)$, we have $\bar{Q}_{IC} \in O(nm^{-2} \log m)$ (tends to zero as n increases), $\bar{Q}_{CC} \in \Omega(\log n)$ and $\bar{C}_{CC} \in \Theta(n)$. Thus, the gap between

(i) $\bar{Q}_{IC}$ and $\bar{Q}_{CC}$: doubly infinite;

(ii) $\bar{Q}_{CC}$ and $\bar{C}_{CC}$: at most exponential.

VI. ROBUSTNESS AGAINST ERROR

In the discussions above, Bob is required to output a right answer every single time. If error is allowed sometimes, do the properties of the zero-error instances still hold? Note that $\gamma \geqslant 2^{-m}$ is trivial since such probability of error can be achieved by randomly guessing without any communication. With a variant of the zero-error simulation protocol, we show the following general result for $\gamma < 2^{-m}$:

Theorem 4. Consider some $h(m)$ such that γ satisfies $-\log(2^{-m} - \gamma) \in O[h(m)]$. Suppose that for $\text{EXC}_{n,m,\gamma}$ with $\gamma < 2^{-m}$, there exists a winning quantum strategy Π_Q^{γ} such that $Q_{CC}(\Pi_Q^{\gamma}) \equiv s \in O[\xi(n)]$. Then, one can construct a classical strategy $\Pi_C^{0^+}$ such that $C_{CC}(\Pi_C^{0^+}) \in \{O[h(m)] + O[\xi(n)]\}2^{O[\xi(n)]}$, whose probability of error can be made arbitrarily small.

Proof. Here, we only sketch the main steps of the proof. Details are given in Appendix D. We revise Bob's local part of Π_C presented in Theorem 2 to devise $\Pi_C^{0^+}$ as follows. As for the zero-error case, given input x, Alice prepares an $\{O[h(m)] + O[\xi(n)]\}2^{O[\xi(n)]}$-bit classical message that encodes the real and imaginary parts of all amplitudes of the pure quantum message $|\psi^{\gamma}(x)\rangle$ in Π_Q^{γ} to accuracy $(2^{-m} - \gamma)2^{-s}/20$, and sends it to Bob, who then normalizes the amplitude vector. Instead of classically calculating the probability distribution of the output as in Π_C, Bob now resorts to local quantum resources. He simply prepares a new quantum state $|\tilde{\psi}^{\gamma}(x)\rangle$ according to the normalized amplitudes, and then feeds it into his original local quantum computation. It can be shown that the probability of outputting the wrong answer $p_{\mathcal{M}_y(x)}$ is always less than 2^{-m}, which guarantees that $\mathcal{M}_y(x)$ is not the winning output. Therefore, Bob can run $\Pi_C^{0^+}$ multiple times and take majority vote to suppress the probability of error to an arbitrarily small value by the Chernoff bound (amplitude amplification). ■

When $\gamma = 2^{-(m+1)}$, $m \geqslant \sqrt{n}$, it was shown in an early version of Ref. [8] that only one classical bit of communication is needed. For completeness we include the proof in Appendix D. Therefore, we consider only the regime of even smaller γ to be of interest. Since $m < -\log(2^{-m} - \gamma) < m + 1$ under this constraint, $h(m)$ can be replaced by m in the above discussions. Like the zero-error case, for $m \in \tilde{o}(n)$, this theorem indicates that the gap between $\bar{Q}_{CC}$ and $\bar{C}_{CC}$ for $\text{EXC}_{n,m,\gamma}$, when any nontrivial γ is allowed, is at most exponential. Consequently, the logarithmic lower bound on $\bar{Q}_{CC}$ and the gaps established for the zero-error case still hold even if some γ such that $\bar{C}_{CC} \in \Omega(n)$ is allowed. The permissible range of γ is identified by the following theorem:

Theorem 5. For $m \leqslant \alpha n$ where $0 < \alpha < 1/2$ is a constant and $\gamma \leqslant (n+1)^{-m}$, $\bar{C}_{CC}(\text{EXC}_{n,m,\gamma}) \in \Omega(n)$.

The proof of Theorem 5 is given in Appendix E. Combining Theorems 4 and 5, we obtain the following results:

Corollary 6. For $m \in \tilde{o}(n)$ and $\gamma \leqslant (n+1)^{-m}$, $\bar{Q}_{CC}(\text{EXC}_{n,m,\gamma}) \in \Omega(\log n)$. By further restricting $m \in \omega(\sqrt{n \log n})$, the gaps established in Corollary 3 still hold.

VII. COMPRESSING QUANTUM COMMUNICATION

Although the PJO strategy succeeds with vanishingly small amount of information cost, it requires exactly n qubits of communication, which is maximal. For $m \in \tilde{o}(n)$, the possibility of superexponential compression of quantum communication cost has been ruled out, but it remains unsettled if any compression is possible at all. In particular, one may wonder if quantum strategies can be more efficient than classical ones in communication cost. Here, we show that a polynomial reduction of quantum communication cost can be achieved by abandoning an insignificant part of the PJO message, while only causing a tiny probability of error such that $\bar{C}_{CC} \in \Omega(n)$ still holds:

Theorem 7. For $m \in \Theta(n^{\alpha})$, $1/2 < \alpha < 1$ and $\gamma \geqslant (n+1)^{-m}$, $\bar{Q}_{CC}(\text{EXC}_{n,m,\gamma}) \in O(m^{1+\delta})$ for any $\delta > 0$.

Proof. Here, we only sketch the main steps of the proof. Details are given in Appendix F. Instead of directly sending the n-qubit state given by Eq. (2), Alice now compresses the message by projecting it onto the subspace spanned by the computational basis vectors with Hamming weight (the number of ones) at most k. Upon receiving the message, Bob performs the same measurement on the quantum state as in the original PJO strategy. Obviously, this would lead to some probability of error ϵ_k. However, it can be shown that taking $k = m^{1+\eta}$ with any $\eta > 0$ is sufficient to guarantee that $\epsilon_k \leqslant (n+1)^{-m}$ asymptotically. It then follows that the compressed message scales as $O(m^{1+\delta})$ for any $\delta > 0$. ■

Combining Theorems 5 and 7, we obtain another quantum-classical separation:

Corollary 8. For $\text{EXC}_{n,m,\gamma}$ with $\gamma \geqslant (n+1)^{-m}$, $m \in \Theta(n^{\alpha})$, $1/2 < \alpha < 1$, there is a polynomial gap between $\bar{Q}_{CC}$ and $\bar{C}_{CC}$.

VIII. CONCLUDING REMARKS

In this paper, we obtained some new knowledge about quantum communication by studying different regimes of the exclusion game. The key result of this paper is a logarithmic lower bound on the quantum communication complexity of most exclusion games. This bound indicates the following results: (i) a doubly infinite gap between the quantum information and communication complexities; (ii) the gap between the quantum and classical communication complexities of certain relational tasks with exponentially small possible probability of error is at most exponential. In contrast, the largest known gap between classical information and communication complexities is singly infinite [10], and the known upper bound on the gap between quantum and classical communication complexities only applies to bounded-error function problems [11]. For exclusion games, we leave open the problems of whether the $\Omega(\log n)$ lower bound on the quantum communication complexity for $m \in \tilde{o}(n)$ is tight, and whether the gap between the quantum and classical complexities for $m \in \tilde{\Omega}(n)$ can be superexponential. (Interestingly, for a slight modification of the exclusion game, there exists a singly infinite gap between the entanglement-assisted communication complexity and the ordinary classical communication complexity [8].) Another set of important open problems is related to how large the gap between quantum information and communication complexities can be in different settings, e.g., bounded error, entanglement assisted, interactive. Answers to these problems will provide significant insight into the relation between these complexities, and the power of quantum resources in the communication model.

ACKNOWLEDGMENTS

The authors would like to thank R. Jain and J. Oppenheim for insightful discussions, and the anonymous referee for helpful comments. Z.W.L. is supported by an Army Research Office award (No. W911NF-11-1-0400). Y.Z. is supported by the ARO grant Contract No. W911NF-12-0486. D.E.K. is supported by the National Science Scholarship from the Agency for Science, Technology and Research (A*STAR). S.A. is supported by an Alan T. Waterman Award from the National Science Foundation, under Grant No. 1249349.

APPENDIX A: FORMAL DEFINITIONS OF INFINITE GAPS

Here, we formally define and classify infinite gaps between two positive asymptotics $g_1(n)$ and $g_2(n)$ [without loss of generality, assume that $\lim_{n\to\infty} g_2(n)/g_1(n) \geqslant 1$], as n tends to infinity. The key idea of properly characterizing all possible gaps is to symmetrize the increasing and decreasing asymptotics by using the logarithmic scale. As discussed in the main text, finite gaps are characterized by a type of well-behaved increasing monotone function. However, there exist gaps that are larger than any finite one, in the sense that they grow faster than any monotone function asymptotically:

Definition 1 (Infinite gap). The gap between $g_1(n)$ and $g_2(n)$ is infinite, if there does not exist any strictly increasing monotone function g such that $\lim_{n\to\infty} \log g_2(n)/g[\log g_1(n)]=1$.

Infinite gaps can be further classified:

Definition 2 (Doubly infinite gap). The gap between $g_1(n)$ and $g_2(n)$ is doubly infinite, if there exists an intermediate asymptotic $g_m(n)$ satisfying $\lim_{n\to\infty} g_2(n)/g_m(n) \geqslant 1$, $\lim_{n\to\infty} g_m(n)/g_1(n) \geqslant 1$ such that $g_1(n)$ vs $g_m(n)$ and $g_m(n)$ vs $g_2(n)$ are both infinite gaps.

Definition 3 (Singly infinite gap). The gap between $g_1(n)$ and $g_2(n)$ is singly infinite, if $g_1(n)$ vs $g_2(n)$ is an infinite gap, but there does not exist any asymptotic $g_m(n)$ satisfying $\lim_{n\to\infty} g_2(n)/g_m(n) \geqslant 1$, $\lim_{n\to\infty} g_m(n)/g_1(n) \geqslant 1$ such that $g_1(n)$ vs $g_m(n)$ and $g_m(n)$ vs $g_2(n)$ are both infinite gaps.

It is evident that, if the gap between $g_1(n)$ and $g_2(n)$ is doubly infinite, the only possibility is that $g_m(n) \in \Theta(1)$, $g_1(n) \in o(1)$, $g_2(n) \in \omega(n)$. This is the largest type of gap between two positive asymptotics. Gaps that take the form $o(n)$ vs $\Theta(1)$ or $\Theta(1)$ vs $\omega(n)$ are singly infinite. Infinite gaps are either singly infinite or doubly infinite.

APPENDIX B: LEMMAS FOR THEOREMS 2 AND 4

Here, we present the detailed proofs of some lemmas that are useful for proving Theorems 2 and 4, including Lemma 1, which has already been stated in the main text.

Lemma 1 For $m \leqslant \alpha n$ where $0 < \alpha < 1/2$ is a constant, $\bar{C}_{CC}(\mathrm{EXC}_{n,m,0}) \in \Omega(n)$.

Proof. By Theorem 2 of Ref. [8], for any classical strategy Π that wins $\mathrm{EXC}_{n,m,0}$, $C_{IC}(\Pi) \geqslant n - \log_2 [\sum_{i=0}^{m-1} \binom{n}{i}]$. For $m \leqslant \alpha n$ where $0 < \alpha < 1/2$ is a constant, $C_{IC}(\Pi) \in \Omega(n)$ (see Appendix C of Ref. [8]). Since the amount of information revealed cannot exceed the amount of communication, i.e., $C_{IC} \leqslant C_{CC}$ for any communication protocol [13,19], it follows that $C_{CC}(\Pi) \in \Omega(n)$. Note that Alice can always send the whole string to Bob in order to win, thus in fact $C_{CC}(\Pi) \in \Theta(n)$. Therefore, $\bar{C}_{CC}(\mathrm{EXC}_{n,m,0}) \in \Omega(n)$ for the specified regime of m asymptotically. ■

Lemma 9. A t-qubit pure quantum state can be classically described by a set of real numbers encoding the real and imaginary parts of all amplitudes to accuracy ϵ using $O[2^t \log(1/\epsilon)]$ bits.

Proof. Generically, a t-qubit pure state $|\psi_t\rangle$ can be written as $|\psi_t\rangle = \sum_{i=1}^{2^t} \alpha_i |i\rangle$, where $\alpha_i \in \mathbb{C}$, and $\{|i\rangle\}$ is a complete orthonormal basis set containing 2^t elements. We express all complex amplitudes as $\alpha_i = b_i + i c_i$ where $b_i, c_i \in \mathbb{R}$, satisfying $\sum_{i=1}^{2^t} |\alpha_i|^2 = \sum_{i=1}^{2^t} (b_i^2 + c_i^2) = 1$. Thus, $0 \leqslant |b_i|, |c_i| \leqslant 1$. To approximate each of these real numbers to accuracy $\epsilon = 2^{-r}$, we keep the first r bits after the binary point, and use one extra bit to indicate its sign, i.e., we can find an $(r+1)$-bit classical string that encodes an approximation $\tilde{b}_i$ of each b_i such that for all i,

$$\begin{aligned} \Delta b_i &= |\tilde{b}_i - b_i| \leqslant \epsilon, \\ \Delta c_i &= |\tilde{c}_i - c_i| \leqslant \epsilon. \end{aligned} \tag{B1}$$

Notice that there are 2×2^t such numbers in total, thus only $2^{t+1}(r+1) \in O[2^t \log(1/\epsilon)]$ bits are needed to encode $|\psi_t\rangle$

such that we have specified the real and imaginary parts of all amplitudes to accuracy ϵ. ■

Lemma 10. Let $\mathcal{H}$ be a Hilbert space of dimension $|\mathcal{H}| = l$, with orthonormal basis $\{|1\rangle,\ldots,|l\rangle\}$. Let $|\psi\rangle \in \mathcal{H}$ with $|\psi\rangle = \sum_{j=1}^{l}\alpha_j|j\rangle = \sum_{j=1}^{l}(b_j + ic_j)|j\rangle$. Suppose that we have $\{\tilde{b}_j,\tilde{c}_j\}$ such that $\forall\ j$, $|b_j - \tilde{b}_j|$, $|c_j - \tilde{c}_j| \leqslant \epsilon < (6\sqrt{2l})^{-1}$. Let $|\tilde{\psi}\rangle = \sum_{j=1}^{l}\tilde{\alpha}_j|j\rangle = \sum_{j=1}^{l}(\tilde{b}_j + i\tilde{c}_j)|j\rangle/\nu$ where $\nu \equiv \sqrt{\sum_{k=1}^{l}(\tilde{b}_k^2 + \tilde{c}_k^2)}$ is the norm. Then, $D(|\psi\rangle,|\tilde{\psi}\rangle) < 10\sqrt{l}\epsilon$, where $D(\cdot,\cdot)$ is the trace distance.

Proof. We first consider the normalization factor

$$\begin{aligned}\nu^2 \equiv \sum_{j=1}^{l}\tilde{b}_j^2 + \tilde{c}_j^2 &\leqslant \sum_{j=1}^{l}(|b_j| + \epsilon)^2 + (|c_j| + \epsilon)^2\\ &= 1 + 2\sum_{j=1}^{l}(|b_j| + |c_j|)\epsilon + 2l\epsilon^2.\end{aligned} \tag{B2}$$

By the Cauchy-Schwarz inequality, we have

$$\sum_{j=1}^{l}|b_j| + |c_j| \leqslant \sqrt{2l} \tag{B3}$$

and

$$2l\epsilon^2 < 2l\frac{1}{\sqrt{2l}}\epsilon = \sqrt{2l}\epsilon. \tag{B4}$$

Therefore,

$$\nu^2 < 1 + 3\sqrt{2l}\epsilon. \tag{B5}$$

Similarly,

$$1 - 2\sqrt{2l}\epsilon < \nu^2. \tag{B6}$$

Since

$$\frac{1}{\sqrt{1 + 3\sqrt{2l}\epsilon}} > \sqrt{1 - 3\sqrt{2l}\epsilon} > 1 - 3\sqrt{2l}\epsilon \tag{B7}$$

and

$$\frac{1}{\sqrt{1 - 2\sqrt{2l}\epsilon}} < \sqrt{1 + 3\sqrt{2l}\epsilon} < 1 + 3\sqrt{2l}\epsilon, \tag{B8}$$

we have

$$1 - 3\sqrt{2l}\epsilon < \frac{1}{\nu} < 1 + 3\sqrt{2l}\epsilon. \tag{B9}$$

Assuming $b_j > 0$, then

$$(b_j - \epsilon)(1 - 3\sqrt{2l}\epsilon) < \frac{\tilde{b}_j}{\nu} < (b_j + \epsilon)(1 + 3\sqrt{2l}\epsilon)\quad \text{if } b_j - \epsilon > 0, \tag{B10}$$

$$(b_j - \epsilon)(1 + 3\sqrt{2l}\epsilon) < \frac{\tilde{b}_j}{\nu} < (b_j + \epsilon)(1 + 3\sqrt{2l}\epsilon)\quad \text{if } b_j - \epsilon < 0. \tag{B11}$$

For both cases,

$$\begin{aligned}(b_j + \epsilon)(1 + 3\sqrt{2l}\epsilon) &= b_j + (1 + 3\sqrt{2l}b_j)\epsilon + 3\sqrt{2l}\epsilon^2\\ &< b_j + (2 + 3\sqrt{2l}b_j)\epsilon.\end{aligned} \tag{B12}$$

For $b_j - \epsilon > 0$,

$$(b_j - \epsilon)(1 - 3\sqrt{2l}\epsilon) > b_j - (1 + 3\sqrt{2l}b_j)\epsilon. \tag{B13}$$

For $b_j - \epsilon < 0$,

$$\begin{aligned}(b_j - \epsilon)(1 + 3\sqrt{2l}\epsilon) &= b_j - (1 - 3\sqrt{2l}b_j)\epsilon - 3\sqrt{2l}\epsilon^2\\ &> b_j - (2 + 3\sqrt{2l}b_j)\epsilon.\end{aligned} \tag{B14}$$

So, if $b_j > 0$,

$$\left|b_j - \frac{\tilde{b}_j}{\nu}\right| < (2 + 3\sqrt{2l}b_j)\epsilon. \tag{B15}$$

Similarly, if $b_j < 0$,

$$\left|b_j - \frac{\tilde{b}_j}{\nu}\right| < (2 - 3\sqrt{2l}b_j)\epsilon. \tag{B16}$$

So, we obtain

$$\left|b_j - \frac{\tilde{b}_j}{\nu}\right| < (2 + 3\sqrt{2l}|b_j|)\epsilon. \tag{B17}$$

Similarly,

$$\left|c_j - \frac{\tilde{c}_j}{\nu}\right| < (2 + 3\sqrt{2l}|c_j|)\epsilon. \tag{B18}$$

Recall that $|\tilde{\psi}\rangle = \sum\tilde{\alpha}_j|j\rangle$, where $\tilde{\alpha_j} = (\tilde{b}_j + i\tilde{c}_j)/\nu$. Then,

$$\begin{aligned}|\alpha_j - \tilde{\alpha}_j| &= \left|b_j + ic_j - \frac{\tilde{b}_j}{\nu} - i\frac{\tilde{c}_j}{\nu}\right|\\ &\leqslant \left|b_j - \frac{\tilde{b}_j}{\nu}\right| + \left|c_j - \frac{\tilde{c}_j}{\nu}\right|\\ &< [4 + 3\sqrt{2l}(|b_j| + |c_j|)]\epsilon.\end{aligned} \tag{B19}$$

Therefore,

$$\begin{aligned}1 - |\langle\psi|\tilde{\psi}\rangle|^2 &= (1 - |\langle\psi|\tilde{\psi}\rangle|)(1 + |\langle\psi|\tilde{\psi}\rangle|)\\ &\leqslant 2(1 - |\langle\psi|\tilde{\psi}\rangle|)\\ &\leqslant 2 - \langle\psi|\tilde{\psi}\rangle - \langle\tilde{\psi}|\psi\rangle\\ &= \sum_{j=1}^{l}|\alpha_j|^2 + |\tilde{\alpha}_j|^2 - \alpha_j\tilde{\alpha}_j^* - \alpha_j^*\tilde{\alpha}_j\\ &= \sum_{j=1}^{l}|\alpha_j - \tilde{\alpha}_j|^2\\ &< \sum_{j=1}^{l}[4 + 3\sqrt{2l}(|b_j| + |c_j|)]^2\epsilon^2,\end{aligned} \tag{B20}$$

where $[4 + 3\sqrt{2l}(|b_j| + |c_j|)]^2 = 16 + 24\sqrt{2l}(|b_j| + |c_j|) + 18l(|b_j| + |c_j|)^2$. Using $(|b_j| + |c_j|)^2 \leqslant 2(|b_j|^2 + |c_j|^2)$ and Eq. (B3), we obtain

$$\begin{aligned}1 - |\langle\psi|\tilde{\psi}\rangle|^2 &< \sum_{j=1}^{l}[16 + 24\sqrt{2l}(|b_j| + |c_j|)\\ &\quad + 36l(|b_j|^2 + |c_j|^2)]\epsilon^2\\ &\leqslant (16l + 48l + 36l)\epsilon^2 = 100l\epsilon^2.\end{aligned} \tag{B21}$$

Then,

$$D(|\psi\rangle,|\tilde{\psi}\rangle) = \sqrt{1-|\langle\psi|\tilde{\psi}\rangle|^2} < 10\sqrt{l}\epsilon. \tag{B22}$$

Thus, $D(|\psi\rangle,|\tilde{\psi}\rangle) < 10\sqrt{l}\epsilon$. ■

Lemma 11. Let $\{P_k\}$ be a POVM, with $p_k = \langle\psi|P_k|\psi\rangle$, $\tilde{p}_k = \langle\tilde{\psi}|P_k|\tilde{\psi}\rangle$. Then, $|p_k - \tilde{p}_k| < 20\sqrt{l}\epsilon$.

Proof. By Theorem 9.1 in [20], we directly obtain $|p_k - \tilde{p}_k| \leqslant \sum_{k=1}^{l} |p_k - \tilde{p}_k| \leqslant 2D(|\psi\rangle,|\tilde{\psi}\rangle) < 20\sqrt{l}\epsilon$, where the last step comes from Lemma 10. ■

APPENDIX C: DETAILED PROOF OF THEOREM 2

Theorem 2. For $m \in \tilde{o}(n)$, $\bar{Q}_{CC}(\mathrm{EXC}_{n,m,0}) \in \Omega(\log n)$.

Proof. Suppose that for $\mathrm{EXC}_{n,m,0}$ where $m \in \tilde{o}(n)$, there exists a winning quantum strategy Π_Q such that $Q_{CC}(\Pi_Q) \equiv q \in o(\log n)$. By definition, $q = \log|\mathcal{H}|$, where $\mathcal{H}$ is the Hilbert space of the largest quantum message. Then, based on Π_Q, we can devise a corresponding classical strategy Π_C with $o(n)$ bits of communication, which contradicts Lemma 1, therefore negating the existence of Π_Q.

Most generally, Π_Q can be divided into three steps: (i) Alice prepares a quantum message (state) of size at most q, based on her n-bit string x; (ii) Alice sends the state to Bob; (iii) Bob feeds the state into his local quantum computation, and obtains an m-bit string z such that $z \neq \mathcal{M}_y(x)$ according to the output (measurement outcome). Note that the quantum messages can in general be mixed, but each of them can always be encoded as a $2q$-qubit pure state by purification using an ancilla space of q qubits (append maximally mixed qubits when the original state contains less than q qubits). Denote the pure message corresponding to x as $|\psi(x)\rangle$. In addition, both players agree on a fixed basis for the matrix representation of operators and amplitudes of state vectors beforehand.

The essence of constructing Π_C is to classically simulate all steps of Π_Q. The basic procedure goes as follows. First, Alice prepares a classical message $C(|\psi(x)\rangle)$ that approximately encodes $|\psi(x)\rangle = \sum_{j=1}^{2^{2q}} \alpha_j |j\rangle = \sum_{j=1}^{2^{2q}} (b_j + ic_j)|j\rangle$ ($\{|j\rangle\}$ is the predetermined basis) by registering the real (b_j) and imaginary parts (c_j) of all amplitudes (α_j) to some desired accuracy $\bar{\epsilon}$ (the approximations are denoted by $\tilde{b}_j$ and $\tilde{c}_j$), and then send it to Bob. Note that the size of $C(|\psi(x)\rangle)$, i.e., the communication cost of Π_C, depends on $\bar{\epsilon}$: it grows as higher precision is desired. In Π_Q, Bob's local strategy can always be modeled as a quantum circuit with $|\psi(x)\rangle$ being the input, i.e., quantum operations followed by a generalized measurement by the principle of deferred measurement [20], which is altogether equivalent to some POVM $\{P_i\}$. Although $\{P_i\}$ may contain an arbitrary number of elements in principle, there are only 2^m possible strings that Bob can eventually output: $g(P_i) = z$, where z is an m-bit string. Therefore, all P_i's corresponding to the same z can be combined as an element P'_z of a new POVM $\{P'_z\}$ by

$$P'_z = \sum_{g(P_i)=z} P_i, \tag{C1}$$

or in the continuum limit where the elements are labeled by a continuous variable μ,

$$P'_z = \int_{g(P(\mu))=z} d\mu\, P(\mu). \tag{C2}$$

Due to the convexity of the set of all non-negative Hermitian operators (valid POVM elements), $\{P'_z\}$ forms a discrete effective POVM with 2^m elements labeled by z. A subtlety here is that the amplitude vector encoded in $C(|\psi(x)\rangle)$ is not necessarily normalized. Bob first normalizes the amplitude vector by dividing each component with the 2-norm $\nu \equiv \sqrt{\sum_{j=1}^{2^q}(\tilde{b}_j^2 + \tilde{c}_j^2)}$, and then applies Born's rule to compute the approximate probability of obtaining each z:

$$p'_z = \frac{1}{\nu^2}\sum_{j,k=1}^{2^q}(\tilde{b}_j\tilde{b}_k + i\tilde{b}_j\tilde{c}_k - i\tilde{c}_j\tilde{b}_k + \tilde{c}_j\tilde{c}_k)P'_{z,jk}, \tag{C3}$$

where $P'_{z,jk}$ is the (j,k)th entry of P'_z. The requirement that Π_Q never fails indicates that the probability of outputting the POVM elements corresponding to a wrong answer is exactly zero. As indicated by Lemma 11, the approximate distribution $\{p'_z\}$ can be arbitrarily close to the true one (denoted by $\{p_z\}$) when $\bar{\epsilon}$ is sufficiently small, so the probability corresponding to the wrong answer $\mathcal{M}_y(x)$ calculated by Eq. (C3) in Π_C is well bounded. Therefore, Bob simply sets an appropriate threshold value $\bar{\delta}(\bar{\epsilon})$ that $p'_{\mathcal{M}_y(x)}$ cannot exceed, and refuses to output any z with $p'_z < \bar{\delta}(\bar{\epsilon})$. As long as there exists an answer above this threshold, this protocol is guaranteed to succeed.

Finally, we determine the appropriate values of $\bar{\epsilon}$ and $\bar{\delta}$ in the above protocol Π_C. To guarantee the existence of at least one valid output, it is sufficient that the upper bound on perturbation on all p_z's, δ, satisfies

$$\delta \equiv \sup_z |p_z - p'_z| < 2^{-m}. \tag{C4}$$

Then, we can simply set the threshold value to

$$\bar{\delta} = 2^{-m}, \tag{C5}$$

i.e., Bob only outputs a z with $p'_z \geqslant 2^{-m}$, which always exists. By Lemma 11, $\delta < 20\bar{\epsilon}2^q$. Then, according to Eq. (C5), we can set

$$\bar{\epsilon} = \frac{1}{20}2^{-(m+q)}, \tag{C6}$$

so that $\delta < \bar{\delta}$. In summary, Π_C runs as introduced with $\bar{\epsilon}$ and $\bar{\delta}$, respectively, specified by Eqs. (C6) and (C5).

By Lemma 9, $C_{CC}(\Pi_C)$ with the above accuracy scales as $O[(m+q)2^{2q}]$. For $m \in \tilde{o}(n)$, $m + q \in O(n^\beta)$ holds for any $0 < \beta < 1$. Since $2^{2q} \in o(n^\kappa)$ for any $\kappa > 0$, we simply set $\kappa = 1 - \beta$, and it can be directly seen that $C_{CC}(\Pi_C) \in o(n^{\beta+\kappa}) \in o(n)$. Since $m \in \tilde{o}(n)$ is within the scope of application of Lemma 1, we have reached a contradiction. ■

APPENDIX D: DETAILED PROOF OF THEOREM 4

Before presenting the proof, we note that a key point of this theorem is that overhead in communication cost of a successful classical simulation is dependent on the scaling of $(2^{-m} - \gamma)$. It was shown in an early version of Ref. [8] that only one bit of classical communication is needed for $m \geqslant \sqrt{n}$, $\gamma = 2^{-(m+1)}$.

We now sketch the argument here. Suppose that Alice sends a single bit to Bob indicating whether x contains a majority of zeros or a majority of ones. If it is the former case, Bob answers with $\vec{1} \in \{0,1\}^m$ for all y, while if it is the latter case, he answers with $\vec{0} \in \{0,1\}^m$. Without loss of generality, assume that x contains a majority of zeros and Bob thus answers with $\vec{1}$. If we denote the number of ones in x by j, $0 \leqslant j \leqslant \frac{n}{2}$, the fraction of y for which Bob makes an error, $\mathcal{M}_y(x) = \vec{1}$, is given by

$$\text{Probability of error for given } x: \begin{cases} \frac{\binom{j}{m}}{\binom{n}{m}} & \text{for } m \leqslant j \leqslant \frac{n}{2}, \\ 0 & \text{for } 0 \leqslant j < m. \end{cases} \tag{D1}$$

Combining with the fact that the number of x with Hamming weight j is $\binom{n}{j}$, the total probability of error of the strategy ϵ_t is given by

$$\begin{aligned} \epsilon_t &= \frac{\sum_{i=m}^{n/2} \binom{n}{i}\binom{i}{m}}{2^{n-1}\binom{n}{m}} \\ &< \frac{\binom{n}{\frac{n}{2}} \sum_{i=m}^{n/2} \binom{i}{m}}{2^{n-1}\binom{n}{m}} \\ &= \frac{\binom{n}{\frac{n}{2}}\binom{\frac{n}{2}+1}{m+1}}{2^{n-1}\binom{n}{m}} \\ &= \frac{\frac{n}{2}+1}{m+1} \frac{\binom{n}{\frac{n}{2}}\binom{\frac{n}{2}}{m}}{2^{n-1}\binom{n}{m}}. \end{aligned} \tag{D2}$$

For large n and $m = \sqrt{n}$,

$$\begin{aligned} \epsilon_t &\sim \frac{1}{2}\sqrt{n} \frac{4^{n/2}}{\sqrt{\frac{\pi n}{2}}} \frac{1}{2^{\sqrt{n}}} \frac{1}{\sqrt{e}} \frac{1}{2^{n-1}} \\ &= \frac{1}{\sqrt{\frac{e\pi}{2}} 2^{\sqrt{n}}} \\ &< \frac{1}{2^{\sqrt{n}+1}}. \end{aligned} \tag{D3}$$

Note that in the approximation we used Stirling's approximation for the $\binom{n}{n/2}$ term, and that

$$\frac{\binom{n/2}{m}}{\binom{n}{m}} \sim \frac{1}{2^m} e^{-1/2}. \tag{D4}$$

Thus, for $m = \sqrt{n}$, there exists a strategy using one bit of classical communication, when the allowed probability of error is greater than $1/2^{\sqrt{n}+1}$. Therefore, it makes sense to pay attention to the regime of even smaller probability of error only, when $m \geqslant \sqrt{n}$. For this case, the conclusion reduces to a simpler form (Corollary 12). However for $m < \sqrt{n}$ (where more communication should be needed), it is unsettled whether a nontrivial probability of error can be achieved with constant amount of communication. For now, we conjecture that for $\gamma = 2^{-(m+1)}$ and $m \in \Omega[\text{poly}(n)]$, $C_{CC} \in O(1)$. However, we have numerical results which indicate that for $m \in o[\text{poly}(n)]$, an $O(1)$ size of classical communication cannot guarantee any probability of error that is smaller than 2^{-m} in the limit of large n.

The most general form of our rigorous conclusion about the classical simulation when error is allowed goes as follows:

Theorem 4 (Error-bounded variant of Theorem 2). Consider some $h(m)$ such that γ satisfies $-\log(2^{-m} - \gamma) \in O[h(m)]$. Suppose that for $\text{EXC}_{n,m,\gamma}$ with $\gamma < 2^{-m}$, there exists a winning quantum strategy Π_Q^{γ} such that $Q_{CC}(\Pi_Q^{\gamma}) \equiv s \in O[\xi(n)]$. Then, one can construct a classical strategy $\Pi_C^{0^+}$ such that $C_{CC}(\Pi_C^{0^+}) \in \{O[h(m)] + O[\xi(n)]\}2^{O[\xi(n)]}$, whose probability of error can be made arbitrarily small.

Proof. We revise Bob's local part of the protocol presented in Theorem 2 to devise this $\Pi_C^{0^+}$ as follows. As for the zero-error game, given input x, Alice prepares a classical message that encodes the real and imaginary parts of all amplitudes of the $2s$-qubit pure quantum message $|\psi^{\gamma}(x)\rangle$ in Π_Q^{γ} to accuracy $\bar{\epsilon}_{\gamma}$ using $O[2^{2s}\log(1/\bar{\epsilon}_{\gamma})]$ bits, and sends it to Bob, who then normalizes the amplitude vector. Instead of classically calculating the probability distribution of the output as in Π_C, Bob now resorts to local quantum resources. He simply prepares a new quantum state $|\tilde{\psi}^{\gamma}(x)\rangle$ according to the normalized state vector (by Lemma 10, this state remains close to the original one when $\bar{\epsilon}_{\gamma}$ is small), and then feeds it into his original local quantum computation. By Lemma 11, the probability of outputting the wrong answer satisfies

$$p'_{\mathcal{M}_y(x)} < \gamma + 20\bar{\epsilon}_{\gamma}2^s. \tag{D5}$$

As long as $\mathcal{M}_y(x)$ is not the output with the largest probability, i.e.,

$$p'_{\mathcal{M}_y(x)} < 2^{-m}, \tag{D6}$$

Bob can apply amplitude amplification to suppress the probability of error: he simply repeats his local protocol for t times (he can use the classical message to prepare as many copies of $|\tilde{\psi}^{\gamma}(x)\rangle$ as he wants), and outputs the string z that comes out for most times. We denote the probability of error after the whole procedure by γ'. Then, by the Chernoff bound, for any $\tau > 0$, there exists a $\bar{t}$ such that as long as $t > \bar{t}$, $\gamma' < \tau$. That is, γ' can be made arbitrarily small simply by increasing t. Combining Eqs. (D5) and (D6), we can set

$$\bar{\epsilon}_{\gamma} = \frac{2^{-m} - \gamma}{20} 2^{-s} \tag{D7}$$

in the protocol. Since $-\log(2^m - \gamma) \in O[h(m)]$ and $s \in O[\xi(n)]$, $\log(1/\bar{\epsilon}_{\gamma}) \in O[h(m)] + O[\xi(n)]$. Therefore, $C_{CC}(\Pi_C^{0^+}) \in \{O[h(m)] + O[\xi(n)]\}2^{O[\xi(n)]}\}$. Note that the no-cloning theorem is not violated since Bob does not need to copy quantum states, and we do not care about the scaling of t since local computational resource is not limited. ■

As argued earlier, by restricting $m \geqslant \sqrt{n}$, any $\gamma \geqslant 2^{-(m+1)}$ becomes trivial. Then, Theorem 4 takes a simpler form because $\log(2^{-m} - \gamma) \in O(m)$:

Corollary 12. Suppose that for $\text{EXC}_{n,m,\gamma}$ with $m \geqslant \sqrt{n}$ and $\gamma \leqslant 2^{-(m+1)}$, there exists a winning quantum strategy Π_Q^{γ} such that $Q_{CC}(\Pi_Q^{\gamma}) \equiv s \in O[\xi(n)]$. Then, one can construct a classical strategy $\Pi_C^{0^+}$ such that $C_{CC}(\Pi_C^{0^+}) \in \{O(m) + O[\xi(n)]\}2^{O[\xi(n)]}\}$, whose probability of error can be made arbitrarily small.

APPENDIX E: DETAILED PROOF OF THEOREM 5

Suppose that Bob is allowed to make an error with probability γ. In other words, for each pair of inputs (x,y), with probability less than or equal to γ, Bob is allowed to output an m-bit string z such that $z = \mathcal{M}_y(x)$. How much classical communication is required from Alice so that Bob does not err with probability more than γ? To answer this question, the following definitions and results will be useful. First, we formally define the one-way, public-coin randomized communication complexity:

Definition 4 (One-way, public-coin randomized communication complexity). For a relation $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, let $R_\epsilon^{1,\mathrm{pub}}(f)$ denote the communication complexity of the best one-way, public-coin randomized protocol that computes f with error at most ϵ on all inputs. When referring specifically to the exclusion game, we will replace this by $\bar{C}_{CC}(\mathrm{EXC}_{n,m,\epsilon})$.

A useful tool for obtaining bounds on the communication complexity is that of rectangle bounds. To define these, we first define (for one-way protocols) *rectangles* and *ϵ-monochromatic functions*.

Definition 5 (One-way rectangles). A one-way rectangle R is defined to be a set $S \times \mathcal{Y}$, where $S \subseteq \mathcal{X}$. For a distribution μ over $\mathcal{X} \times \mathcal{Y}$, let μ_R be the distribution formed from μ by conditioning on R. Let $\mu(R)$ be the probability of the event R under the distribution μ.

Definition 6 (One-way ϵ monochromatic). Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. A distribution λ on $\mathcal{X} \times \mathcal{Y}$ is one-way ϵ monochromatic for f if there exists a function $g : \mathcal{Y} \to \mathcal{Z}$, such that

$$P_{XY\sim\lambda}\{[X,Y,g(Y)] \in f\} \geqslant 1-\epsilon. \tag{E1}$$

With these in place, we now define *rectangle bounds* as follows:

Definition 7 (Rectangle bound). Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. For a distribution μ on $\mathcal{X} \times \mathcal{Y}$, the one-way rectangle bound is

$$\mathrm{rec}_\epsilon^{1,\mu}(f) = \min_R \left\{ \log_2 \frac{1}{\mu(R)} : R \text{ is one-way rectangle and } \mu_R \text{ is one-way } \epsilon \text{ monochromatic} \right\}. \tag{E2}$$

The one-way rectangle bound for f is

$$\mathrm{rec}_\epsilon^{1}(f) = \max_\mu \mathrm{rec}_\epsilon^{1,\mu}(f). \tag{E3}$$

If the above maximization is restricted to product distributions, we can also define

$$\mathrm{rec}_\epsilon^{1,[]}(f) = \max_{\mu:\text{product}} \mathrm{rec}_\epsilon^{1,\mu}(f). \tag{E4}$$

The utility of rectangle bounds to the problem at hand is given by the following result obtained from [21]:

Theorem 13 ([21]). Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and let $\epsilon \in [0,1/6]$. Then,

$$R_\epsilon^{1,\mathrm{pub}}(f) \in \Omega\big[\mathrm{rec}_\epsilon^{1,[]}(f)\big]. \tag{E5}$$

This theorem implies the following useful characterization for the communication complexity of the exclusion game for nonzero error γ:

Lemma 14. To show a lower bound of c for $\bar{C}_{CC}(\mathrm{EXC}_{n,m,\gamma})$, it is sufficient to show the following. Let S be any subset of $\{0,1\}^n$ of size 2^{n-c}. Let $A_M = \{z(y) \in \{0,1\}^m : y \text{ subset of } [n] \text{ of size } m\}$ be any set of answers for Bob. Then, for at least γ fraction of $\{(x,y) : x \in S, y \text{ a subset of } [n] \text{ of size } m\}$, $z(y)$ is an incorrect answer for x.

Proof. By Theorem 13 and the definition of rectangle bounds, we have

$$\bar{C}_{CC}(\mathrm{EXC}_{n,m,\gamma}) \in \Omega\big[\mathrm{rec}_\gamma^{1,\mathrm{unif}}(\mathrm{EXC}_{n,m,\gamma})\big], \tag{E6}$$

where "unif" is the product, uniform distribution over X and Y. For $R = S \times \mathcal{Y}$,

$$\mathrm{unif}(R) = \frac{1}{2^c}. \tag{E7}$$

Thus, if we can not find a set of answers for Bob, A_M (in the language of Definition 6, a function g) such that unif_R is one-way ϵ monochromatic, then

$$\mathrm{rec}_\gamma^{1,\mathrm{unif}}(\mathrm{EXC}_{n,m,\gamma}) > c, \tag{E8}$$

and $\bar{C}_{CC}(\mathrm{EXC}_{n,m,\gamma}) \in \Omega(c)$. ■

The following fact regarding sums of binomial coefficients will also be used:

Lemma 15. For $m \in \Theta(n^\alpha)$, $1/2 < \alpha < 1$,

$$n - \log_2 \left[\sum_{i=0}^{m} \binom{n}{i} \right] \geqslant n - o(n). \tag{E9}$$

For $m = \beta n$, $0 < \beta < 1/2$,

$$n - \log_2 \left[\sum_{i=0}^{m} \binom{n}{i} \right] \in \Omega(n). \tag{E10}$$

Proof. See Appendix C.2 of Ref. [8]. ■

Using these lemmas, we can now prove the following result:

Theorem 5. For $m \leqslant \alpha n$ where $0 < \alpha < 1/2$ is a constant and $\gamma \leqslant (n+1)^{-m}$, $\bar{C}_{CC}(\mathrm{EXC}_{n,m,\gamma}) \in \Omega(n)$.

Proof. First, let $\epsilon = 1/(\sum_{i=0}^m \binom{n}{i})$ and note that

$$\frac{1}{\sum_{i=0}^m \binom{n}{i}} \geqslant \frac{1}{(n+1)^m}. \tag{E11}$$

Our goal is to determine how large S can be taken to be in Lemma 14 subject to nonzero error ϵ. Note that from the proof of Theorem 2 in [8], we know that, for any choice of A_M, at most $\sum_{i=0}^{m-1} \binom{n}{i}$ strings can be contained in S without introducing any error. An example of when this occurs is when A_M is such that $z(y) = 0$ (the m-bit string of all zeros) for all y and S consists of all strings with strictly less than m zeros. What strings can be added into this S while keeping the error below ϵ?

There are $\binom{n}{m}$ strings such that $\mathcal{M}_y(x) = 0$ for precisely one value of y. These are the strings with precisely m zeros. If we define S as

$$S = \left\{ x : x \in \{0,1\}^n, \sum_{i=1}^n x_i \geqslant n - m \right\}, \tag{E12}$$

then the fraction of $\{(x,y) : x \in S, y$ subset of $[n]$ of size $m\}$ such that $z(y) = 0$ is an incorrect answer for x is given by

$$\frac{\binom{n}{m}}{\binom{n}{m}\sum_{i=0}^{m}\binom{n}{i}} = \epsilon. \tag{E13}$$

As S consists of the maximum number of strings that produce no error and strings that produce only one error, it is clear that this is the largest S can be taken to be for error given by ϵ. Thus, by Lemma 14,

$$\begin{aligned}\bar{C}_{CC}(\text{EXC}_{n,m,\epsilon}) &\in \Omega(n - \log_2|S|)\\ &= \Omega\left\{n - \log_2\left[\sum_{i=0}^{m}\binom{n}{i}\right]\right\}.\end{aligned} \tag{E14}$$

By Lemma 15, for $m \in \Theta(n^{\alpha}), 1/2 < \alpha < 1$, we obtain:

$$\bar{C}_{CC}(\text{EXC}_{n,m,\epsilon}) \in \Omega(n). \tag{E15}$$

Finally, as $\epsilon \geqslant (n+1)^{-m}$, the scaling holds for error parametrized by γ as given in the statement of the theorem. ■

APPENDIX F: DETAILED PROOF OF THEOREM 7

In the PJO quantum strategy [8], upon receiving x, Alice sends the state

$$\begin{aligned}|\Phi(x)\rangle &= \bigotimes_{i=1}^{n}\left[\cos\left(\frac{\theta_m}{2}\right)|0\rangle + (-1)^{x_i}\sin\left(\frac{\theta_m}{2}\right)|1\rangle\right]\\ &= \sum_{r\in\{0,1\}^n}(-1)^{x\cdot r}\left[\cos\left(\frac{\theta_m}{2}\right)\right]^{n-|r|}\left[\sin\left(\frac{\theta_m}{2}\right)\right]^{|r|}|r\rangle,\end{aligned} \tag{F1}$$

where $\theta_m = 2\tan^{-1}(2^{1/m} - 1)$.

Suppose that instead of directly sending $|\Phi(x)\rangle$, Alice compresses the message by projecting the state onto the space spanned by the computational basis vectors with with Hamming weight (the number of ones) at most k. The compressed quantum message reads as

$$\begin{aligned}|\Phi^{(k)}(x)\rangle = \frac{1}{\sqrt{A_k}}\sum_{\substack{r\in\{0,1\}^n\\ |r|\leqslant k}}(-1)^{x\cdot r}\left[\cos\left(\frac{\theta_m}{2}\right)\right]^{n-|r|}\\ \times\left[\sin\left(\frac{\theta_m}{2}\right)\right]^{|r|}|r\rangle,\end{aligned} \tag{F2}$$

where

$$A_k = \sum_{i=0}^{k}\binom{n}{i}\left[\cos\left(\frac{\theta_m}{2}\right)\right]^{2(n-i)}\left[\sin\left(\frac{\theta_m}{2}\right)\right]^{2i}. \tag{F3}$$

This compression reduces the number of qubits Alice sends to $\log\left[\sum_{i=0}^{k}\binom{n}{i}\right]$. Assuming that Bob performs the same measurement on the qubits specified by y as he would without the compression

$$|\zeta(z)\rangle = \frac{1}{\sqrt{2^m}}\left[|0\rangle - \sum_{s\neq 0}(-1)^{z\cdot s}|s\rangle\right], \tag{F4}$$

this would lead to some probability of error ϵ_k. If $\rho_{x,y}^{k} = \text{Tr}_{\backslash y}[|\Phi^{(k)}(x)\rangle\langle\Phi^{(k)}(x)|]$ denotes the state sent by Alice restricted to the locations specified by y, then

$$\epsilon_k = \langle\zeta(\mathcal{M}_y(x))|\rho_{x,y}^{k}|\zeta(\mathcal{M}_y(x))\rangle. \tag{F5}$$

To bound ϵ_k, we make use of the following lemma:

Lemma 16. For $|\Phi(x)\rangle$, $|\Phi^{(k)}(x)\rangle$, and ϵ_k, respectively, defined in Eqs. (F1), (F2), and (F5):

$$\sqrt{1 - |\langle\Phi(x)|\Phi^{(k)}(x)\rangle|^2} \geqslant \epsilon_k. \tag{F6}$$

Note that $\langle\Phi(x)|\Phi^{(k)}(x)\rangle$ is independent of x.

Proof. Recall that the trace distance between two density matrices ρ and σ is given by

$$D(\rho,\sigma) = \tfrac{1}{2}\text{Tr}\left[\sqrt{(\rho-\sigma)^{\dagger}(\rho-\sigma)}\right]. \tag{F7}$$

For pure states $|\psi\rangle$ and $|\phi\rangle$, this reduces to

$$D(|\psi\rangle,|\phi\rangle) = \sqrt{1 - |\langle\psi|\phi\rangle|^2}. \tag{F8}$$

We will also need the following facts. First, as the trace distance never increases under local operations, for bipartite states ρ_{AB} and σ_{AB},

$$D(\rho_{AB},\sigma_{AB}) \geqslant D(\rho_A,\sigma_A). \tag{F9}$$

Second, by Eq. (9.22) in [20],

$$D(\rho,\sigma) = \max_{P}\text{Tr}[P(\rho-\sigma)], \tag{F10}$$

where the maximization is taken over all projectors P. Combining these facts, we obtain

$$\begin{aligned}\epsilon_k &= \langle\zeta(\mathcal{M}_y(x))|\rho_{x,y}^{k}|\zeta(\mathcal{M}_y(x))\rangle\\ &= \langle\zeta(\mathcal{M}_y(x))|\rho_{x,y}^{k}|\zeta(\mathcal{M}_y(x))\rangle - \langle\zeta(\mathcal{M}_y(x))|\rho_{x,y}^{n}|\zeta(\mathcal{M}_y(x))\rangle\\ &\leqslant D\left(\rho_{x,y}^{k},\rho_{x,y}^{n}\right)\\ &\leqslant D(|\Phi^{(k)}(x)\rangle,|\Phi(x)\rangle)\\ &= \sqrt{1 - |\langle\Phi(x)|\Phi^{(k)}(x)\rangle|^2},\end{aligned} \tag{F11}$$

as required. ■

Lemma 16 enables us to prove the following theorem:

Theorem 7. For $m \in \Theta(n^{\alpha})$, $1/2 < \alpha < 1$, and $\gamma \geqslant (n+1)^{-m}$, $\bar{Q}_{CC}(\text{EXC}_{n,m,\gamma}) \in O(m^{1+\delta})$ for any $\delta > 0$.

Proof.

$$\begin{aligned}&\sqrt{1 - |\langle\Phi(x)|\Phi^{(k)}(x)\rangle|^2}\\ &= \sqrt{1 - \sum_{i=0}^{k}\binom{n}{i}\left[\cos\left(\frac{\theta_m}{2}\right)\right]^{2n-2i}\left[\sin\left(\frac{\theta_m}{2}\right)\right]^{2i}}\\ &= \sqrt{\sum_{i=k+1}^{n}\binom{n}{i}\left[\cos\left(\frac{\theta_m}{2}\right)\right]^{2n-2i}\left[\sin\left(\frac{\theta_m}{2}\right)\right]^{2i}}.\end{aligned} \tag{F12}$$

Now,

$$\binom{n}{i} \leqslant \left(\frac{ne}{i}\right)^{i}, \tag{F13}$$

$$\cos^2\left(\frac{\theta_m}{2}\right) \leqslant 1, \tag{F14}$$

$$\sin^2\left(\frac{\theta_m}{2}\right) < \frac{1}{m^2}, \quad \text{for large } m; \tag{F15}$$

so, for large m,

$$\begin{aligned} 1 - |\langle\Phi(x)|\Phi^{(k)}(x)\rangle|^2 &< \sum_{i=k+1}^{n}\left(\frac{ne}{i}\right)^i\left(\frac{1}{m}\right)^{2i} \\ &\leqslant (n+1)\left(\frac{ne}{m^2k}\right)^k, \end{aligned} \tag{F16}$$

as the $i = k+1$ term decays slowest for $m \in \omega(\sqrt{n})$. For this bound to be less than $\gamma^2 = (n+1)^{-2m}$, we require

$$\left(\frac{m^2k}{ne}\right)^k > (n+1)^{2m+1}, \tag{F17}$$

$$k\log\left(\frac{m^2k}{ne}\right) > (2m+1)\log(n+1). \tag{F18}$$

To satisfy this asymptotically, it suffices to take $k = m^{1+\eta}$ with any $\eta > 0$. The number of qubits sent [which, by Lemma 16, achieves a probability of error $\leqslant (n+1)^{-m}$] is then

$$\begin{aligned} \log\left[\sum_{i=0}^{m^{1+\eta}}\binom{n}{i}\right] &\leqslant \log[(n+1)^{m^{1+\eta}}] \\ &= m^{1+\eta}\log(n+1). \end{aligned} \tag{F19}$$

This can scale as $O(m^{1+\delta})$ for any $\delta > 0$, by choosing some $\eta < \delta$. Thus, $\bar{Q}_{CC}(\text{EXC}_{n,m,\gamma}) \in O(m^{1+\delta})$ for any $\delta > 0$. ■

[1] A. C.-C. Yao, in *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing*, STOC '79 (ACM, New York, 1979), pp. 209–213.

[2] R. Raz, in *Proceedings of the Thirty-first Annual ACM Symposium on Theory of Computing*, STOC '99 (ACM, New York, 1999), pp. 358–367.

[3] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Phys. Rev. Lett. **87**, 167902 (2001).

[4] Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis, in *Proceedings of the Thirty-sixth Annual ACM Symposium on Theory of Computing*, STOC '04 (ACM, New York, 2004), pp. 128–137.

[5] H. Buhrman, R. Cleve, and A. Wigderson, in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98 (ACM, New York, 1998), pp. 63–68.

[6] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf, in *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing*, STOC '07 (ACM, New York, 2007), pp. 516–525.

[7] A. Montanaro, Quantum Inf. Comput. **11**, 574 (2011).

[8] C. Perry, R. Jain, and J. Oppenheim, Phys. Rev. Lett. **115**, 030504 (2015).

[9] A. Ganor, G. Kol, and R. Raz, in *Proceedings of 55th IEEE Annual Symposium on Foundations of Computer Science*, FOCS 2014 (IEEE Computer Society, Los Alamitos, CA, 2014), pp. 176–185.

[10] M. Braverman, in *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing*, STOC '12 (ACM, New York, 2012), pp. 505–524.

[11] I. Kremer, Quantum Communication, Master's thesis, The Hebrew University, 1995.

[12] E. Kushilevitz and N. Nisan, *Communication Complexity* (Cambridge University Press, New York, 1997).

[13] A. Chakrabarti, Y. Shi, A. Wirth, and A. Yao, in *Proceedings of 42th IEEE Annual Symposium on Foundations of Computer Science*, FOCS 2001 (IEEE Computer Society, Los Alamitos, CA, 2001), pp. 270–278.

[14] B. Barak, M. Braverman, X. Chen, and A. Rao, SIAM J. Comput. **42**, 1327 (2013).

[15] D. E. Knuth, SIGACT News **8**, 18 (1976).

[16] S. Bandyopadhyay, R. Jain, J. Oppenheim, and C. Perry, Phys. Rev. A **89**, 022336 (2014).

[17] C. M. Caves, C. A. Fuchs, and R. Schack, Phys. Rev. A **66**, 062111 (2002).

[18] M. F. Pusey, J. Barrett, and T. Rudolph, Nat. Phys. **8**, 475 (2012).

[19] Z. Bar-Yossef, T. Jayram, R. Kumar, and D. Sivakumar, in *Proceedings of 43rd IEEE Annual Symposium on Foundations of Computer Science*, FOCS 2002 (IEEE Computer Society, Los Alamitos, CA, 2002), pp. 209–218.

[20] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, New York, 2000).

[21] R. Jain, H. Klauck, and A. Nayak, in *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08 (ACM, New York, 2008), pp. 599–608.