Limits of quantum one-way communication by matrix Hypercontractive Inequality

Yaoyun Shi[1] Xiaodi Wu[2] Wei Yu[3]

[1]Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI 48105, USA

[2]Center for Theoretical Physics
Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[3] Center for the Theory of Interactive Computations (CTIC)
and Center for Massive Data Algorithmics (MADALGO)
Aarhus University, 8200 Aarhus N, Denmark

[1]shiyy@umich.edu, [2]xiaodiwu@mit.edu, [3]yuwei@cs.au.dk

September 19, 2015

Abstract

An important discovery in quantum information processing is that quantum one-way communication protocols can be exponentially more efficient than classical protocols. Those extraordinary quantum advantages were demonstrated through the *Hidden Matching Problem* and its variants, where the underlying basic task is to determine the parity of some $k = 2$ bits of an n-bit string. We prove that for larger values of k, the quantum complexities of those problems increase exponentially from $O(\log n)$ to $\Omega(n^{1-2/k})$, which is almost tight and renders any super-polynomial quantum-classical gaps impossible. Our results also rule out a "quantum argument", in the sense of Kerenidis and de Wolf (*Journal of Computer and System Sciences*, 69(3)395–420, 2004), for proving any super-polynomial lower bound on Locally Decodable Codes for more than 2 queries. Our proofs are new applications of the matrix Hypercontractive Inequality developed by Ben-Aroya, Regev, and de Wolf (FOCS 2008).

1 Background and summary of the main results

The lower bound problems and motivations. A central question in quantum information processing is to identify its power and limitations in comparison with classical models. Because of the apparent difficulty of the question, researchers have been focusing on simple yet useful models. One-way communication is one such model: Alice and Bob wish to compute a function $f(x, y)$, for which the input x is known to Alice only and y known to Bob only. Alice sends a single message to Bob, who is required to output their best guess for $f(x, y)$. The *one-way communication complexity* of f is the smallest integer k such that the function can be computed using a length k message for all inputs. More generally, the function f may be partially defined or a relation. Despite its simplicity, the model has shown to be instrumental in tackling some important computational problems, such as streaming algorithms and locally decodable codes. The simplicity of the model is also deceiving: proving classical lower bounds may turn out to be highly nontrivial.

An important discovery by Bar-Yossef *et al.* [2] and Gavinsky *et al.* [8] is that quantum one-way communication protocols can be exponentially more efficient than classical protocols. Those extraordinary quantum advantages were demonstrated through a relational problem called "Hidden Matching Problem" (HM), and its partial function variants. The main challenges for establishing the results were to prove strong classical lower bounds.

However, techniques for proving strong *quantum lower bounds* are necessary to complete our understanding on quantum complexities and the quantum-classical boundaries. There is much work to be done in this direction. For example, the basic question that if any *asymptotic* quantum-classical gap is possible remains open for total functions. Therefore, advancing quantum lower bound techniques is of fundamental importance. To this end, our approach is to identify and study explicit problems that are not only of interest on their own, but are also at the forefront of challenging existing lower bound techniques. We therefore focus on the following straightforward generalizations of the Hidden Matching Problem and its functional variants mentioned above.

Definition 1 (Generalized Hidden Matching)**.** *Let $k, n \geq 2$ be integers. In* the k-Generalized Hidden Matching Problem of n subsets, *denoted k-HM$_n$, Alice is given $x \in \{0,1\}^{kn}$ and Bob is given a partition of $[nk]$ into n subsets of size k, both uniformly distributed. Bob is required to output a k-set G in its partition and the parity of the bits of x in G.*

The α-Partial Matching Problem of [8] is similarly generalized in Definition 2. For $k = 2$, the definitions coincide with the original problems, which were shown to have $O(\log n)$ quantum complexity but $\Omega(\sqrt{n})$ classical complexity. For a general constant k, the following quantum *upper bound* is known

Proposition 1.1 (Kerenidis and de Wolf [11])**.** *For any integer $k \geq 2$, there is a quantum protocol using $O(n^{1-1/\lceil k/2 \rceil} \log n)$ qubits for each of the generalized problems with parameters k and n.*

However, no strong, i.e., $n^{\Omega(1)}$, quantum lower bound is known for $k \geq 3$. This set of problems are appropriate for studying quantum lower bounds first for the challenges they pose to the existing techniques. For problems such as those whose two-way complexity is low, two-way lower bound methods necessarily fail. Thus useful techniques need to exploit the one-way nature of the model. Three works are representative on this regard: the quantum random access code lower bound by Nayak [13], the *trace distance method* by Aaronson [1], and the direct product theorems of Ben-Aroya, Regev, and de Wolf [3] using a matrix Hypercontractive Inequality.

Recall that a quantum random access code encodes a classical binary string in a quantum state from which each bit can be recovered. Thus the quantum message solving the Generalized Hidden Matching Problem appears to be a significantly weaker object. Therefore, it is not clear how the information theoretical argument for the quantum random access code would work. The trace distance method appears to be devised specifically for total functions, thus not directly applicable to our problems. The matrix Hypercontractive Inequality turns out to be a powerful tool for us yet the application of it requires additional insights.

Those lower bound problems are also interesting for their own reasons. First, as the case of $k = 2$ provides the important instances for exponential quantum-classical gap, it is natural to ask if a dramatic gap remains for general k. In particular, the super-efficient quantum protocols all make use of the quantum fingerprinting state, a simple yet powerful object discovered by Buhrman *et al.* [4] and widely studied subsequently (see, e.g. [18, 7]). Could one use or generalize the fingerprinting states to solve the larger k problems?

Secondly, the Generalized Hidden Matching Problem is closely related to *locally decodable codes* (LDC), an important object that finds many applications in complexity theory and cryptography. Recall that a LDC for n-bit strings is an encoding such that any bit can be recovered with high probability from a small number of queries, even a constant fraction of the code word is corrupted. A major question on LDC's is the tradeoff between the codeword length and the number of queries. In a seminal work, Kerenidis and de Wolf [11] proved an exponential lower bound on codeword length for 2-query LDC's. Their proof makes use of a *quantum* argument: an efficient 2-query LDC encoding can be turned into an efficient quantum random access code. This reduction is made through an efficient solution to the Hidden Matching Problem (though historically the latter problem was inspired by the reduction in [11]). The reduction holds for general k: a super-efficient quantum encoding for the k-Generalized Hidden Matching Problem would imply a strong lower bound for k-query LDC. Therefore, our quantum lower bound question is to ask if such a quantum proof for lower-bounding LDC's works for more than 2 queries.

The main results and their implications. Our main result is that none of the generalized problems admits a super-efficient quantum one-way protocol for $k \geq 3$.

Theorem 1.2 (Main Theorem)**.** *(Informally) For all constant $k \geq 2$, the k-Generalized Hidden Matching Problem and its function variants for n subsets require $\Omega(n^{1-2/k})$ quantum one-way communication.*

Comparing with the upper bound in Proposition 1.1, those bounds are up to a logarithmic factor (for even k) or close to (for odd k) optimal. Our results indicate that the exponential separation achieved by the Hidden Matching problem is unique for $k = 2$. One can also extend the classical lower and upper bounds for $k = 2$ in [8, 2] to a lower and upper bound of $\Theta(n^{1-1/k})$ for the generalized problems. Thus for general $k \geq 2$, the quantum advantage is at most an $O(n^{1/k})$ factor. This, in particular, rules out the possibility of a "quantum proof", in the sense of Kerenidis and de Wolf [11], for a super-polynomial lower bound on LDC's of more than 2 queries. We also note that in [6] Efremenko already provided 3-query locally decodable code of subexponential length using in part key ideas of Yekhanin [17], which renders exponential lower bounds impossible.

We do not know if for odd k, our lower bound can be improved to $\Omega(n^{1-1/\lceil k/2 \rceil})$ or the upper bound can be lowered by removing the ceiling function. (Note that the upper bound is achieved by having Alice send $O(n^{1-1/\lceil k/2 \rceil})$ copies of the fingerprint state [4]). We are able to prove that the class of protocols that we call *quantum fingerprint protocols*, where Alice sends multiple copies of the fingerprint state, $\Omega(n^{1-1/\lceil k/2 \rceil})$ copies of the fingerprint state are necessary.

Theorem 1.3. *A quantum fingerprint protocol for any of the generalized problems with parameters k and n requires $\Omega(n^{1-1/\lceil k/2 \rceil})$ copies of the fingerprint state.*

We note that there are several variants of quantum fingerprint state. Our result applies to a general notion as defined in Definition 3. Thus it shows an additional limitation of the often powerful fingerprint state in addition to that pointed out in [7].

Proof techniques and comparison with previous works. We now describe our proof techniques and compare our proofs to those for the classical lower bound proofs [8, 2] and the quantum lower bound in [3].

Our proofs make use of the Fourier analysis of matrix-valued functions and the proof for the main theorem relies on the matrix-valued Hypercontractive Inequality of Ben-Aroya, Regev, and de Wolf [3]. The standard line of attack using the Fourier technique is to reduce the problem to properties of the Fourier coefficients. The crux of applying the technique often lies in setting up a massive cancelation of the Fourier coefficients, and in exploiting the pattern of the surviving ones. Those are precisely the two places where we have new insights comparing with previous works, which we elaborate below.

Two works provide the main sources of inspirations for our proofs: the classical lower bound proofs [8, 2] and the quantum lower bound in [3]. Both works use Fourier analysis, the first uses the Kahn-Kalai-Linial Hypercontractive Inequality, while the latter developed its matrix-valued analogue.

It is not clear how to make the classical arguments work for the quantum cases for the following two reasons. The first is, the classical proofs are based on arguments conditioned on the classical messages. There is no clear quantum analogy for such conditioning. Consequently, while the Fourier analysis is applied in the classical cases to the pre-images of a message, we apply it to the encoding function. Secondly, the classical proof (for the partial function variants) makes use of Parseval's Identity. The Identity extends to the quantum (matrix) case, but is useless as it is an identity of the Frobenius norm, while for our purpose the trace norm is needed. The difference between our proofs and the classical proofs becomes apparent when one restricts the quantum arguments to classical, i.e., commuting encoding states. The resulting lower bounds are of the magnitude of the quantum bounds, weaker than the (almost) tight classical lower bounds.

The connections of our proofs with that of Ben-Aroya *et al.* [3] are more intimate. For the relation problem (i.e., k-HM$_n$), our proof shows that the quantum encoding in a constant-bias k-HM$_n$ protocol is by itself an average-$\Omega(1/n)$-bias k-XOR-quantum random access code (k-XOR-QRAC) as defined in [3]. Our reduction is accomplished in two steps. First, we show that an arbitrary protocol can be converted to one that outputs the subset identity first, and then outputs the XOR of bits inside that subset. This is due to a general conversion of POVMs that output (a, b) into two-stage POVMs that outputs a in the first stage and b in the second one, which might be of independent interest. The reduction then follows from a data-processing inequality when applied to an "artificially" defined physical realizable super-operator.[1]

Our proof of the quantum lower bound for the function problem is, to the authors' knowledge, is the first example that makes full use of the matrix Hyper-contractive Inequality. The previous application of this inequality to lower-bound k-XOR-QRAC is a direct consequence of the matrix Hyper-contractive inequality when restricted to the Fourier coefficients in low levels, while it is crucial in our application to apply this inequality to Fourier coefficients in all levels. Moreover, as mentioned earlier, there are two difficulties (i.e., no quantum analogy of the pre-image of a message and no counter-part of Parseval's Identity of the trace norm) in extending the classical argument for proving quantum lower bounds for the function problem, which we overcame by some novel technical manipulations. We feel that these technical contributions might be useful in other contexts when dealing with quantum messages.

When restricted to quantum fingerprint protocols, a better lower bound (when k is odd) can be obtained through the use of two crucial properties of the fingerprint state (of a more general version in Definition 3): the projection to each index i has equal length and the "bit states" are linear function of x_i's. Those properties enable the application of the generalized birthday paradox to show that most of the weight of the Fourier coefficients concentrates on low levels.

Organizations. We survey the necessary technical backgrounds and define formally the general-

[1] After the first appearance of our paper, we learned that the same result could be obtained by a few different but similar in spirit arguments [16].

ized partial function variant in Section 2. The quantum lower bounds for the relational problem and the function problem are proved in Section 3 and Section 4, respectively. The proof for Theorem 1.3 is sketched in Section 5. We include the classical and quantum upper bound results in Appendix A.1, as well as the classical lower bounds in Appendix A.2, as they are simple extensions of known results.

2 Preliminaries

Quantum States. The state space $\mathcal{A}$ of m-qubit is the complex Euclidean space $\mathbb{C}^{2^m}$. An m-qubit quantum state is represented by a density operator ρ, i.e., a positive semidefinite matrices with trace 1, over $\mathcal{A}$. The set of all quantum states in $\mathcal{A}$ is denoted by $\mathrm{D}(\mathcal{A})$. The Hilbert-Schmidt inner product on the operator space $\mathrm{L}(\mathcal{A})$ is defined by $\langle X, Y\rangle = \mathrm{tr}(X^*Y)$ for all $X, Y \in \mathrm{L}(\mathcal{A})$, where $*$ is the adjoint operator.

Let Σ be a finite nonempty set of *measurement outcomes*. A *positive-operator valued measure (POVM)* on the state space $\mathcal{A}$ with outcomes in Σ is a collection of positive semidefinite operators $\{P_a : a \in \Sigma\}$ such that $\sum_{a\in\Sigma} P_a = \mathbb{1}_{\mathcal{A}}$. If instead of equality, $\sum_{a\in\Sigma} P_a \leq \mathbb{1}_{\mathcal{A}}$, the collection is a *sub-normalized* POVM. When this POVM is applied to a quantum state ρ, the probability of each outcome $a \in \Sigma$ is $\langle \rho, P_a\rangle$. When outcome a is observed, the quantum state ρ becomes the state $\sqrt{P_a}\rho\sqrt{P_a}/\langle \rho, P_a\rangle$.

Norms. For any $X \in \mathrm{L}(\mathcal{A})$ with singular values $\sigma_1, \cdots, \sigma_d$, where $d = \dim(\mathcal{A})$, we define its (normalized Schatten) p-norm as $\|X\|_p = (\frac{1}{d}\sum_{i=1}^d \sigma_i^p)^{1/p}$. The *trace norm*, $\|X\|_{\mathrm{tr}}$, is $\|X\|_{\mathrm{tr}} = \sum_{i=1}^d \sigma_i$. Clearly, $\|X\|_{\mathrm{tr}} = d\|X\|_1$. The following well-known fact relates the trace distance with the optimal probability of distinguishing quantum states.

Fact 2.1 ([9]). *Let ρ_0, ρ_1 be two quantum states which appear with probability p and $1-p$ respectively, the optimal success probability of predicting which state it is by a POVM is*

$$\frac{1}{2} + \frac{1}{2}\|p\rho_0 - (1-p)\rho_1\|_{\mathrm{tr}}.$$

Matrix Hypercontractive Inequality. Consider a matrix-valued function on $\{0,1\}^n$, $f : \{0,1\}^n \to \mathrm{L}(\mathcal{A})$. For example, f may encode an n-bit string x by a quantum state in $\mathcal{A}$. The *Fourier transform* $\hat{f}$ is defined similarly as for scaler functions. Denote by $[n]$ the set of all indices $1, \cdots, n$. We identify a subset S of $[n]$ with the n-bit binary string consisting of 1's at the indices in S. For every subset $S \subseteq [n]$ and $x \in \{0,1\}^n$, let $\chi_S(x) = (-1)^{x\cdot S}$ be the (sign-represented) parity of the bits of x indexed by S. The *Fourier transform* of a matrix-valued function $f : \{0,1\}^n \to \mathrm{L}(\mathcal{A})$ is the function $\widehat{f} : \{0,1\}^n \to \mathrm{L}(\mathcal{A})$ defined by

$$\widehat{f}(S) = \frac{1}{2^n}\sum_{x\in\{0,1\}^n} f(x)\chi_S(x). \tag{1}$$

The values $\widehat{f}(S)$ are called the *Fourier coefficients* of f and now are matrices[2] over $\mathcal{A}$. An important property of the Fourier transform is that we can express f in terms of its Fourier coefficients as

$$f(x) = \sum_{S\subseteq[n]} \widehat{f}(S)\chi_S(x). \tag{2}$$

[2] An equivalent definition is by applying the standard Fourier transform to each (i,j)-entry separately: $\widehat{f}(S)_{i,j} = \widehat{f(\cdot)_{i,j}}(S)$.

The main tool we are going to use is an extension of the hypercontractive inequality to matrix-valued functions [3] as follows.

Theorem 2.2 ([3])**.** *For every $f : \{0,1\}^n \to \mathrm{L}(\mathcal{A})$ and $1 \le p \le 2$,*

$$\left(\sum_{S \subseteq [n]} (p-1)^{|S|} \left\| \widehat{f}(S) \right\|_p^2 \right)^{1/2} \le \left(\frac{1}{2^n} \sum_{x \in \{0,1\}^n} \| f(x) \|_p^p \right)^{1/p}.$$

In particular, we consider the special case where the matrix-valued function f maps every $x \in \{0,1\}^n$ to a m-qubit density operator (i.e., an encoding of n-bit string to a m-qubit quantum state).

Corollary 2.3 ([3])**.** *Let $f : \{0,1\}^n \to \mathrm{D}\left(\mathbb{C}^{2^m}\right)$ be any mapping from n-bit strings to m-qubit density matrices. Then for any $0 \le \delta \le 1$, we have*

$$\sum_{S \subseteq [n]} \delta^{|S|} \| \widehat{f}(S) \|_{\mathrm{tr}}^2 \le 2^{2\delta m}.$$

A direct consequence is the following upper bound on the average trace norm of the Fourier coefficients $\widehat{f}(S)$ for S of constant size k. We note the original statement in [3] is slightly more general on the choice of k. Denote by $S \sim \binom{n}{k}$ that the random variable S is a uniformly distributed size-k subset of S.

Corollary 2.4 ([3])**.** *There exist constants $\eta, C > 0$ such that for any k and sufficiently large n, any encoding of n-bit strings to m-qubit density operators $f : \{0,1\}^n \to \mathrm{D}\left(\mathbb{C}^{2^m}\right)$ has the property that,*

$$\mathbb{E}_{S \sim \binom{n}{k}} \left[\| \widehat{f}(S) \|_{\mathrm{tr}} \right] \le C \left(\frac{\eta m}{n} \right)^{k/2}.$$

Note that Corollary 2.4 fails to give the correct bound for large non-constant k. We shall make use of Corollary 2.3 to handle the case of large k directly.

Generalized Hidden Matching problems. The Hidden Matching Problem was inspired by the quantum argument for locally decodable codes [11] and was proposed in [2] as the first *relation problem* to establish an exponential separation between quantum and bounded-error randomized one-way communication complexity. The exponential separation was later proved to be true even for a *function version* of the problem [8]. We give the definition of our generalized relation problem in Definition 1, and define the generalized partial function variant as follows. [3]

Let k, n be positive integers such that $k \ge 2$, and $0 \le \alpha \le 0.5$. Alice's input is a kn-bit string $x \in \{0,1\}^{kn}$. Part of Bob's input is an α-partition M that consists of αn disjoint subsets of indices from $[kn]$ each of size k, such as $G_1 = \left\{ i_1^1, \cdots, i_k^1 \right\}, \cdots, G_{\alpha n} = \left\{ i_1^{\alpha n}, \cdots i_k^{\alpha n} \right\}$. Let α-$\mathcal{M}_{k,n}$ be the set of all possible α-partitions over $[kn]$ (denote by $\mathcal{M}_{k,n}$ if $\alpha = 1$) . We may also regard an α-partition on $[kn]$ as an $(\alpha n \times kn)$ matrix M over GF(2), where the r-th row corresponds to the subset G_r and the entry (r, c) is 1 iff $c \in G_r$ and is 0 otherwise. In this way, the product Mx is an αn-bit string $z = z_1, \cdots, z_r, \cdots, z_{\alpha n}$ where $z_r = x_{i_1^r} \oplus \cdots \oplus x_{i_k^r}$.

Definition 2 (Generalized Partial Matching)**.** *The (k, α)-Partial Matching Problem, denoted (k, α)-PM$_n$, is a partial Boolean function problem, where Alice is given $x \in \{0,1\}^{kn}$ and Bob is given $M \in \alpha$-$\mathcal{M}_{k,n}$ and $w \in \{0,1\}^{\alpha n}$. The promise is that there is a bit $b \in \{0,1\}$ such that $Mx \oplus b^{\alpha n} = w$. Bob is required to output this b.*

[3] For readers who are familiar with the history of the partial function variants, we only include the generalization of the "α-Partial Matching" problem, whereas similar result can be obtained for the other variant, called the "Noisy Perfect Matching" problem.

3 Lower bound for Generalized HM

We prove the *quantum* lower bound for the Generalized Hidden Matching Problem (Definition 1).

Theorem 3.1. *Any quantum protocol $\mathcal{P}$ (with m-qubit encoding) for k-HM$_n$ problem has the successful probability no more than $\frac{1}{2} + O(n(\frac{\eta m}{nk})^{k/2})$ for some constant $\eta > 0$. In particular, in order to achieve a constant bias, any quantum protocol $\mathcal{P}$ needs to send an $\Omega(n^{1-2/k})$-qubit message.*

Fix a quantum protocol $\mathcal{P}$ for k-HM$_n$. Recall that in this problem, Alice is given $x \in \{0,1\}^{kn}$ and Bob is given $M \in \mathcal{M}_{k,n}$, both uniformly distributed. Let the matrix-valued function $\rho : \{0,1\}^{kn} \to \mathrm{D}\left(\mathbb{C}^{2^m}\right)$ be Alice's m-qubit encoding of her input x. The protocol succeeds if Bob outputs a subset $G \in M$ and the parity of the bits of x in that subset. The successful probability minus a half is the *bias* of the protocol. We can assume without loss of generality that Bob never outputs a subset not in M, since doing so would not increase the bias.

We start with an important observation that one can assume without loss of generality that Bob performs a two-stage POVM as follows to accomplish the above task and leave the proof in Appendix A.3.

Lemma 3.2. *Any protocol for k-HM$_n$ can be converted to one such that*

(a) Bob first performs a measurement to output a subset, then a second measurement to output the input-parity for that subset, and,

(b) for any input (x, M), the output distribution is unchanged.

Based on this lemma, we assume from now on that Bob's measurement is performed in the two stages described. Let $\left\{\Pi_M^G\right\}$ be Bob's first-stage POVM on input M with outcome G. Note we use G to index all subsets of $[kn]$ of size k. But it is clear that $\Pi_M^G = 0$ if $G \notin M$.

Our next step is to upper-bound the bias ϵ_{bias} in a two-stage protocol, in terms of the Fourier coefficients $\hat{\rho}(G)$ for k-sets G. Define $\rho_x \overset{\text{def}}{=} \rho(x)$, $p_x \overset{\text{def}}{=} 1/2^{kn}$, and $p_M \overset{\text{def}}{=} 1/|\mathcal{M}_{k,n}|$. From our notation, it is easy to see that the chance protocol $\mathcal{P}$ outputs subset G on input x, M is given by

$$\mathbf{Pr}[x, M, G] = p_x p_M \left\langle \rho_x, \Pi_M^G \right\rangle .$$

Lemma 3.3. *The bias ϵ_{bias} of any protocol $\mathcal{P}$ satisfies*

$$\epsilon_{bias} \le \frac{1}{2} \sum_{M,G} p_M \left\| \sqrt{\Pi_M^G} \hat{\rho}(G) \sqrt{\Pi_M^G} \right\|_{\text{tr}}. \tag{3}$$

Proof. Conditioning on Bob's outputting subset G on input M , which occurs with chance $\mathbf{Pr}[G|M]$, each Alice's message ρ_x collapses to an unnormalized state

$$\sqrt{\Pi_M^G} \rho_x \sqrt{\Pi_M^G}.$$

Let $G+$ (resp. $G-$) be the subset of $\{0,1\}^{kn}$ where the parity on subset G is 0 (resp. 1). Bob wishes to distinguish between the post-measurement states of $\sum_{x \in G+} p_x \rho_x$ and $\sum_{x \in G-} p_x \rho_x$. By Fact 2.1, the largest bias Bob can achieve conditioned on M and G is

$$\frac{1}{2\mathbf{Pr}[G|M]} \left\| \sqrt{\Pi_M^G} \sum_{x \in G+} p_x \rho_x \sqrt{\Pi_M^G} - \sqrt{\Pi_M^G} \sum_{x \in G-} p_x \rho_x \sqrt{\Pi_M^G} \right\|_{\text{tr}} . \tag{4}$$

Since by definition $\sum_{x \in G+} p_x \rho_x - \sum_{x \in G-} p_x \rho_x = \widehat{\rho}(G)$, **Eq.** (4) then becomes

$$\frac{1}{2\mathbf{Pr}[G|M]} \left\| \sqrt{\Pi_M^G} \widehat{\rho}(G) \sqrt{\Pi_M^G} \right\|_{\mathrm{tr}}.$$

Thus for the overall bias,

$$\epsilon_{\mathrm{bias}} \leq \sum_{M,G} p_M \mathbf{Pr}[G|M] \frac{1}{2\mathbf{Pr}[G|M]} \left\| \sqrt{\Pi_M^G} \widehat{\rho}(G) \sqrt{\Pi_M^G} \right\|_{\mathrm{tr}} = \frac{1}{2} \sum_{M,G} p_M \| \sqrt{\Pi_M^G} \widehat{\rho}(G) \sqrt{\Pi_M^G} \|_{\mathrm{tr}}.$$

□

Now we make another important observation in our proof. In order to upper bound the sum over M of trace norms in (5) for a fixed G, we shall treat $\{\Pi_M^G\}_M$ as a sub-normalized POVM, although it is not performed in the protocol. Denote by $p_k \overset{\mathrm{def}}{=} n/\binom{kn}{k}$.

Lemma 3.4. *For any G, the set $\{\frac{p_M}{p_k} \Pi_M^G\}_M$ is a sub-normalized POVM.*

Proof. Because $\Pi_M^G = 0$ if $G \notin M$, then we have

$$\sum_M p_M \Pi_M^G = \sum_{M: G \in M} p_M \Pi_M^G \leq \left(\sum_{M: G \in M} p_M \right) \mathbb{1} = \mathbf{Pr}_M[G \in M] \mathbb{1}.$$

Since $\mathbf{Pr}_M[G \in M] = p_k$, we have $\sum_M \frac{p_M}{p_k} \Pi_M^G \leq \mathbb{1}$. □

This observation allows us to upper-bound the sum of trace norms of transformed $\widehat{\rho}(G)$ to the sum of trace norms of $\widehat{\rho}(G)$ itself, through the following data-processing inequality[4].

Fact 3.5 ([12]). *For any Hermitian operator R and any sub-normalized POVM $\{P_a\}_a$, $\sum_a \| \sqrt{P_a} R \sqrt{P_a} \|_{\mathrm{tr}} \leq \|R\|_{\mathrm{tr}}$.*

Therefore by Lemma 3.4 and Fact 3.5, we have

$$\epsilon_{\mathrm{bias}} \leq \frac{1}{2} p_k \sum_G \| \widehat{\rho}(G) \|_{\mathrm{tr}} = \frac{1}{2} n \mathbb{E}_{G \sim \binom{kn}{k}} [\| \widehat{\rho}(G) \|_{\mathrm{tr}}] \leq O(n(\frac{\eta m}{nk})^{k/2}), \tag{5}$$

where the last inequality comes from Corollary 2.4. This directly implies our main theorem.

We note that when $k = 2$ our argument yields only a constant lower bound. However, a $\Omega(\log(n))$ lower bound of quantum protocols can be obtained by simulating the quantum protocol with classical messages in the most trivial way and making use of the classical lower bound $\Omega(\sqrt{n})$.

[4] Note the inequality essentially follows from the fact that the trace distance is non-increasing under admissible quantum operations and any Hermitian operator is just a re-scaled difference between two weighted density operators.

4 The function problem

In this section we prove the quantum lower bound for the partial function problem defined in Section 2. First, we note that a lower bound of quantum protocols with shared randomness can be obtained from a lower bound for quantum protocols under some 'hard' input distribution (i.e., distributional complexity). The distribution we choose is uniform on Alice's input $x \in \{0,1\}^{kn}$, Bob's input $M \in \alpha\text{-}\mathcal{M}_{k,n}$ and the function value b, which fixes Bob's second input $w = Mx \oplus b^{\alpha n}$.

Our proof generalizes several technical lemmas in their classical lower bound proof for $k = 2$. However, as summarized in the introduction there is also a significant difference between the two arguments. In particular, our Fourier analysis is performed directly on the encoding messages rather than the pre-images of a fixed encoding message, as a consequence of the fact that there is no clear quantum analogue of conditioning on a message. Moreover, we only make use of the matrix-version hypercontractive inequality but never use the Parseval Identity that is crucial in the classical argument. Comparing to [3], our use of the matrix-valued hypercontractive inequality is more involved and requires more careful analysis and additional inequality tricks.

Theorem 4.1. *For any integer k with $k \geq 2$ and $0 \leq \alpha \leq 0.5$, the quantum bounded-error one-way communication complexity of (k,α)-PM$_n$ is $\Omega(\log_2(1/\alpha)n^{1-2/k})$.*

Fix an arbitrary quantum protocol $\mathcal{P}$ with Alice's encoding function $\rho : \{0,1\}^{kn} \to \mathrm{D}\left(\mathbb{C}^{2^m}\right)$. Similar to the proof for the relation problem, we will derive the lower bound through an upper bound on the bias ϵ_{bias}. Our proof strategy is also similar in that we first upper-bound the bias by the trace norms $\|\widehat{\rho}(G)\|_{\text{tr}}$, then make use of matrix hyper-contractive inequality to derive the desired bound. However, the technical execution will be different.

Let $p_x \overset{\text{def}}{=} 1/2^{nk}$, $p_b \overset{\text{def}}{=} 1/2$, $p^\alpha_M \overset{\text{def}}{=} 1/|\alpha\text{-}\mathcal{M}_{k,n}|$ and $\delta_{a,b} \overset{\text{def}}{=} 1$ iff $a = b$, the hard distribution is given by

$$\mathbf{Pr}[x, b, M, w] = p_x p^\alpha_M p_b \delta_{Mx\oplus b^{\alpha n},w}. \tag{6}$$

To upper-bound the bias by the trace norms $\|\widehat{\rho}(G)\|_{\text{tr}}$, we define $t_w(Mx) \overset{\text{def}}{=} (\delta_{Mx,w} - \delta_{Mx\oplus 1^{\alpha n},w})/2$ and $u(M, w, S) \overset{\text{def}}{=} \sum_{x\in\{0,1\}^{kn}} p_x p^\alpha_M t_w(Mx)(-1)^{x\cdot S}$.

Lemma 4.2. *The bias ϵ_{bias} satisfies*

$$\epsilon_{bias} \leq \frac{1}{2} \sum_{S,M,w} |u(M, w, S)| \|\widehat{\rho}(S)\|_{\text{tr}}. \tag{7}$$

Proof. Conditioning on Bob's input M, w, in his eyes, Alice's message ρ_x appears with chance $\mathbf{Pr}[x|M, w]$. The best strategy for Bob to determine b conditioning on his input (M, w) is no more than the chance to distinguish between two subsets of ρ_x selected according to b. Namely, no more than the chance to distinguish between the following $\rho^{M,w}_0$ and $\rho^{M,w}_1$ each appearing with the chance $\mathbf{Pr}[b = 0|M, w]$ and $\mathbf{Pr}[b = 1|M, w]$, respectively,

$$\rho^{M,w}_0 = \frac{\sum_x \mathbf{Pr}[x, 0, M, w]\rho_x}{\mathbf{Pr}[0, M, w]} \text{ and } \rho^{M,w}_1 = \frac{\sum_x \mathbf{Pr}[x, 1, M, w]\rho_x}{\mathbf{Pr}[1, M, w]}.$$

Then by Fact 2.1, we have $\mathbf{Pr}[\mathcal{P} \text{ succeeds}|M, w]$ at most

$$\frac{1}{2} + \frac{1}{2}\|\mathbf{Pr}[0|M, w]\rho^{M,w}_0 - \mathbf{Pr}[1|M, w]\rho^{M,w}_1\|_{\text{tr}}.$$

By averaging over different inputs M, w, we have the overall successful chance $\mathbf{Pr}[\mathcal{P}$ succeeds$]$ at most

$$\frac{1}{2}+\frac{1}{2}\sum_{M,w}\mathbf{Pr}[M,w]\|\mathbf{Pr}[0|M,w]\rho_0^{M,w}-\mathbf{Pr}[1|M,w]\rho_1^{M,w}\|_{\mathrm{tr}}.$$

Thus

$$\epsilon_{\text{bias}}\leq\frac{1}{2}\sum_{M,w}\|\sum_x\mathbf{Pr}[x,0,M,w]\rho_x-\sum_x\mathbf{Pr}[x,1,M,w]\rho_x\|_{\mathrm{tr}}. \tag{8}$$

Plugging in the definitions of the two probabilities in **Eq.** (6), we have

$$\epsilon_{\text{bias}}\leq\frac{1}{2}\sum_{M,w}p_M\|\sum_x p_x t_w(Mx)\rho_x\|_{\mathrm{tr}}, \tag{9}$$

which by the definition of $t_w(Mx)$ and **Eq.** (1) gives

$$\epsilon_{\text{bias}}\leq\frac{1}{2}\sum_{M,w}\|\sum_{S\subseteq[kn]}u(M,w,S)\widehat{\rho}(S)\|_{\mathrm{tr}}.$$

Thus (7) holds. □

Now we analyze $|u(M,w,S)|$ for different M,w,S. For any fixed partition $M=\{G_1,G_2,\cdots,G_n\}$, let $v(M)$ be the set of unions of an arbitrary number (including zero) of elements in M. We shall use a n-bit string $T_{|M}\in\{0,1\}^n$ to denote which subsets are included in T. With the above notation and similar to [8], the quantity $|u(M,w,S)|$ is given by the following lemma.

Lemma 4.3. *If $S\in v(M)$ with an odd hamming weight $l=h(S_{|M})$, $|u(M,w,S)|=p_M^{\alpha}/2^{\alpha n}$; otherwise $|u(M,w,S)|=0$.*

Proof. First notice that if $S\notin v(M)$ then for any $z\in\{0,1\}^n$, the expectation of $(-1)^{x\cdot S}$ over $x\in\{0,1\}^{kn}$ conditioned on $Mx=z$ is 0, i.e., $\mathbb{E}_x[(-1)^{x\cdot S}\mid Mx=z]=0$. Also note that $t_w(Mx)$ only depends on $z=Mx$. Thus we have,

$$u(M,w,S)=p_M\sum_{x\in\{0,1\}^{kn}}p_x t_w(Mx)(-1)^{x\cdot S}=p_M\sum_{z\in\{0,1\}^n}\mathbf{Pr}[z]t_w(z)\mathbb{E}_{Mx=z}[(-1)^{x\cdot S}]=0,$$

where $\mathbf{Pr}[z]=1/2^n$. Otherwise if $S\in v(M)$, it is easy to see $\mathbb{E}_x[(-1)^{x\cdot S}\mid Mx=z]=(-1)^{z\cdot S_{|M}}$. Thus,

$$\begin{aligned}|u(M,w,S)| &= p_M^{\alpha}|\sum_{z\in\{0,1\}^{\alpha n}}\frac{1}{2^{\alpha n}}t_w(z)(-1)^{z\cdot S_{|M}}| \\ &= \begin{cases}p_M^{\alpha}/2^{\alpha n}, & h(S_{|M})\text{ is odd;}\\ 0, & h(S_{|M})\text{ is even.}\end{cases}\end{aligned}$$

Hence the lemma follows. □

We further simplify the upper bound on ϵ_{bias} in the following lemma.

Lemma 4.4. *The bias ϵ_{bias} satisfies*

$$2\epsilon_{bias}\leq\sum_{l\in[\alpha n]\ odd}\alpha^l\sum_{h(S)=kl}\binom{n}{l}\binom{kn}{kl}^{-1}\|\widehat{\rho}(S)\|_{\mathrm{tr}}. \tag{10}$$

Proof. By Lemma 4.3 and **Eq.** (7),

$$\epsilon_{\text{bias}} \leq \frac{1}{2} \sum_{\substack{h(S)=kl \\ l \leq \alpha n \text{ odd}}} \sum_{\substack{M \in \alpha\text{-}\mathcal{M}_{k,n} \\ S \in v(M)}} p_M^\alpha \sum_{w \in \{0,1\}^{\alpha n}} \frac{1}{2^{\alpha n}} \|\widehat{\rho}(S)\|_{\text{tr}}.$$

This allows us to remove w from the summation, arriving at

$$\epsilon_{\text{bias}} \leq \frac{1}{2} \sum_{l \in [\alpha n] \text{ odd}} \sum_{\substack{S \subseteq [kn] \\ h(S)=kl}} \sum_{\substack{M \in \alpha\text{-}\mathcal{M}_{k,n} \\ S \in v(M)}} p_M^\alpha \|\widehat{\rho}(S)\|_{\text{tr}}. \tag{11}$$

The value $\sum_{M \text{ s.t. } S \in v(M)} p_M^\alpha$ is exactly $\mathbf{Pr}_M[S \in v(M)]$, for a random $M \in \alpha\text{-}\mathcal{M}_{k,n}$. Through a counting problem, we show in Appendix A.4 that

$$\mathbf{Pr}_{M \in \alpha\text{-}\mathcal{M}_{k,n}}[S \in v(M)] = \binom{\alpha n}{l} \cdot \binom{kn}{kl}^{-1}. \tag{12}$$

Hence,

$$\epsilon_{\text{bias}} \leq \frac{1}{2} \sum_{l \in [\alpha n] \text{ odd}} \binom{\alpha n}{l} \left[\sum_{h(S)=kl} \binom{kn}{kl}^{-1} \|\widehat{\rho}(S)\|_{\text{tr}} \right].$$

Noting that $\binom{\alpha n}{l} \leq \alpha^l \binom{n}{l}$ for $0 \leq \alpha \leq 1$ we have,

$$2\epsilon_{\text{bias}} \leq \sum_{l \in [\alpha n] \text{ odd}} \alpha^l \sum_{h(S)=kl} \binom{n}{l} \binom{kn}{kl}^{-1} \|\widehat{\rho}(S)\|_{\text{tr}}.$$

□

We are ready to apply the matrix hypercontractive inequality to prove Theorem 4.1.

PROOF OF THEOREM 4.1. Let $\gamma_l = \binom{n}{l}^2 \binom{kn}{kl}^{-1}$ and $\delta_l = \gamma_l^{1/kl}$. It follows from Lemma 4.4 that

$$\begin{aligned} 2\epsilon_{\text{bias}} &\leq \sum_{l \in [\alpha n] \text{ odd}} \alpha^l \binom{n}{l}^{-1} \left[\sum_{|S|=kl} \binom{n}{l}^2 \binom{kn}{kl}^{-1} \|\widehat{\rho}(S)\|_{\text{tr}} \right] \\ &= \sum_{l \in [\alpha n] \text{ odd}} \alpha^l \binom{n}{l}^{-1} \left[\sum_{|S|=kl} \delta_l^{kl} \|\widehat{\rho}(S)\|_{\text{tr}} \right]. \end{aligned} \tag{13}$$

By Cauchy-Schwarz,

$$\sum_{|S|=kl} \delta_l^{kl} \|\widehat{\rho}(S)\|_{\text{tr}} \leq \sqrt{\sum_{|S|=kl} \delta_l^{kl}} \sqrt{\sum_{|S|=kl} \delta_l^{kl} \|\widehat{\rho}(S)\|_{\text{tr}}^2}. \tag{14}$$

It is clear that $\sqrt{\sum_{|S|=kl} \delta_l^{kl}} = \sqrt{\gamma_l \binom{kn}{kl}} = \binom{n}{l}$. By expanding the definition of δ_l, we have

$$\delta_l \leq (l/n)^{1-2/k} \leq 1, \tag{15}$$

which we justify in Appendix A.5. Apply the matrix-valued hypercontractive inequality (Corollary 2.3), we have

$$\sum_{|S|=kl} \delta_l^{kl} \|\widehat{\rho}(S)\|_{\mathrm{tr}}^2 \leq \sum_{S\subseteq[kn]} \delta_l^{|S|} \|\widehat{\rho}(S)\|_{\mathrm{tr}}^2 \leq 2^{2\delta_l m}.$$

Therefore,

$$\begin{aligned} 2\epsilon_{\text{bias}} &= \sum_{l\in[\alpha n] \text{ odd}} \alpha^l \binom{n}{l}^{-1} \binom{n}{l} 2^{\delta_l m} \leq \alpha 2^{\delta_1 m} + \sum_{3\leq l\leq n \text{ odd}} \alpha^{l/2}\alpha^{l/2} 2^{\delta_l m} \\ &\leq \alpha 2^{\delta_1 m} + \sum_{3\leq l\leq n \text{ odd}} \alpha^{l/2} \max_{3\leq l\leq n \text{ odd}} \alpha^{l/2} 2^{\delta_l m} \leq \alpha 2^{\delta_1 m} + \frac{\alpha^{1.5}}{1-\alpha} \max_{3\leq l\leq n \text{ odd}} \alpha^{l/2} 2^{\delta_l m}, \end{aligned}$$

where $\delta_l = (\binom{n}{l}^2 \binom{kn}{kl}^{-1})^{1/kl} \leq (l/n)^{1-2/k}$. For sufficiently small distributional error (such that $2\epsilon_{\text{bias}} \geq 0.95$) and for $0 \leq \alpha \leq 0.5$, we have either $\frac{\alpha^{1.5}}{1-\alpha}\alpha^{l/2}2^{\delta_l m} \geq \alpha^{1.5}$ for some $l \geq 3$ or $\alpha 2^{\delta_1 m} \geq 0.95 - \alpha^{1.5} \geq \alpha^{0.9}$ when $0 \leq \alpha \leq 0.5$. In the first case, we have $2^{\delta_l m} \geq (1-\alpha)\alpha^{-l/2}$ and thus $m = \Omega(\log_2(1/\alpha) n^{1-2/k} l^{2/k})$ for $l \geq 3$. In the second case $2^{\delta_1 m} \geq \alpha^{-0.1} = 2^{0.1 \log_2(1/\alpha)}$ and thus $m = \Omega(\log_2(1/\alpha) n^{1-2/k})$, and we conclude the proof for Theorem 4.1. □

Remarks. We note that the rewriting of the right hand side in **Eq.** (10) as that in **Eq.** (13) allows the application of the matrix-valued hypercontractive inequality to obtain the desired bound. Different from the relation problem, we cannot directly use Corollary 2.4 to upper bound the summation of trace norms in **Eq.** (13), because kl can be arbitrarily close to kn. A more precise bound from Corollary 2.3 then becomes necessary. We remark this is, to the authors' knowledge, the first example that makes use of the full power of the matrix-valued hypercontractive inequality.

For constant α, the above theorem gives the desired lower bound $\Omega(n^{1-2/k})$. However, for very small α such as $\alpha = \Theta(1/n)$, the problem reduces to the XOR quantum random access code defined in [3] and has a tight bound $\Theta(n)$. Our lower bound is weaker in that case and only proves $\Omega(n^{1-2/k}\log_2(n))$. This is because our analysis for small α is loose and a more sophisticated analysis might provide a better lower bound.

Classical lower bounds. Extending the classical lower bound to the case of subset-size k is straightforward but tedious. We will just point out the change needed when adapting the lower bound proof for (k,α)-PM$_n$ in the journal version of [8]. For a general k, the probability in [8, Lemma 3.3] becomes $\binom{\alpha n}{l}/\binom{kn}{kl}$. [5] Then by choosing the classical message size $c = O(n^{1-1/k}/\alpha^{1/k})$, the rest argument implies a lower bound $\Omega(n^{1-1/k}/\alpha^{1/k})$ of the message size.

5 Fingerprint protocols

In this section, we sketch a proof for Theorem 1.3. There are several variants of fingerprint states. Some variant may make the Theorem trivial, e.g., for odd k, for each $x \in \{0,1\}^{kn}$, the fingerprint state $\frac{1}{\sqrt{kn}}\sum_{i\in[kn]}(-1)^{x_i}|i\rangle$ is the same as that for $\bar{x}$ (up to a global phase), thus is not useful at all. Our result applies to a more general definition of fingerprint state described below.

Definition 3 (Generalized quantum fingerprint state). *Given $x \in \{0,1\}^m$ for some positive integer m, we call a quantum state $|\phi_x\rangle$ a* generalized fingerprint state *if for some Hilbert space $\mathcal{H}$ of a finite*

[5] Note we use different notations from [8]. More precisely, our "kl"is "k" in [8]. Our "n" and "α" are the same as [8]

dimension and some quantum states $|\alpha_{b,i}\rangle \in \mathcal{H}$, $b \in \{0,1\}, i \in [m]$,

$$|\phi_x\rangle = \frac{1}{\sqrt{m}} \sum_{i\in[m]} |i\rangle \otimes |\alpha_{x_i,i}\rangle.$$

The properties of generalized quantum fingerprint states key to our proof are that the projection to each index i has equal length and that the "bit states" $|\alpha_{x_i,i}\rangle$ are linear function of x_i.

We now prove by contradiction Theorem 1.3 where "quantum fingerprint state" is demonstrated in Definition 3.

PROOF OF THEOREM 1.3. Let $\bar{k} \stackrel{\text{def}}{=} \lceil k/2 \rceil$. Suppose that r copies of the fingerprint state are sent, for an $r = o(n^{1-1/\bar{k}})$. Let $|\Psi_x\rangle \stackrel{\text{def}}{=} |\psi_x\rangle^{\otimes r}$ and $\rho \stackrel{\text{def}}{=} \rho_x = |\Psi_x\rangle\langle\Psi_x|$ be Alice's messages. For a fixed partition $M = \{G_1, \cdots, G_n\}$, a random index i induces a random variable that indicates which subset out of n subsets it is from. Then in a sequence of r random numbers uniformly and independently drawn from $[n]$, the probability of having a subset number occurring for $\bar{k}$ times is $o(1)$. Therefore, removing those base vectors $|i_1, i_2, ..., i_r\rangle$ where $\bar{k}$ indices from a same subset occur, we obtain a (un-normalized) vector $|\Psi'_{x,M}\rangle$, which is $o(1)$-close to $|\Psi_x\rangle$. Let

$$\rho' \stackrel{\text{def}}{=} \rho'_{x,M} \stackrel{\text{def}}{=} |\Psi'_{x,M}\rangle\langle\Psi'_{x,M}|.$$

We replace every Alice's message ρ_x in the case where Bob receives M by $\rho'_{x,M}$. Note that this is not a possible operation by Alice but rather an intermediate step for the analysis. Since $\|\rho_x - \rho'_{x,M}\|_{\text{tr}} = o(1)$ for all x, M, the new overall bias ϵ'_{bias} only differs $o(1)$ from ϵ_{bias} because the replacement only incurs a $o(1)$ change in terms of ℓ_1 norm on the output distribution.

A second important property of $\rho'_{x,M}$ is that its Fourier coefficients $\hat{\rho}'_M(S) = 0$ for all $S \in v(M)$ (or any S that has k indexes from a same subset in M). We prove this in Lemma 5.1 below.

However, for k-HM$_n$, Lemma 3.3 with ρ_x replaced by $\rho'_{x,M}$ says

$$\epsilon'_{\text{bias}} \leq \frac{1}{2} \sum_{M,G} p_M \| \sqrt{\Pi_M^G} \widehat{\rho'_M}(G) \sqrt{\Pi_M^G} \|_{\text{tr}} = 0,$$

thus $\epsilon_{\text{bias}} = o(1)$, a contradiction. Similar arguments can be applied to **Eq.** (7), and therefore we can show $\epsilon'_{\text{bias}} = 0$ for both k-NPM$_n$ and (k, α)-PM$_n$ problems, getting the same contradiction that $\epsilon_{\text{bias}} = o(1)$. □

Lemma 5.1. *For any $S \in v(M)$ (defined in Section 4) with $S \neq \emptyset$, we have the Fourier coefficient of $\rho'_{x,M}$ denoted by $\widehat{\rho'_M}(S)$ is zero.*

Proof. Because the "bit states" $|\alpha_{x_i,i}\rangle$ are linear function of x_i, the Fourier coefficient $\widehat{\rho'_M}(S)$ can be formulated as a linear combination of the Fourier coefficient of the function

$$P_T(x) \stackrel{\text{def}}{=} \prod_{i\in T} x_i$$

for some $T \subseteq [kn]$ that is the union of any two r random indices $\{i_1, i_2, \cdots, i_r\}$ and $\{j_1, j_2, \cdots, j_r\}$. Since there is no $\bar{k}$-collision in any r random indices appearing in $\rho'_{x,M}$, the maximum number of elements in T that are in the same subset is at most $2(\bar{k} - 1) < k$. By definition for any $S \in v(M)$ with $S \neq \emptyset$, we have $S \nsubseteq T$. A directly calculation shows $\widehat{P_T}(S) = 0$ if $S \nsubseteq T$, thus the lemma follows. □

Acknowledgment

We thank Kai-Min Chung, Xiaoming Sun, and Shengyu Zhang for helpful discussions. We are also grateful to Oded Regev and Ronald de Wolf for their helpful feedback on our manuscript. Part of this work was done when YS was visiting Institute for Theoretical Computer Science, Tsinghua University. He thanks members of the institute for their hospitality, and the financial support of China Basic Research Grant 2011CBA00300 (sub-project 2011CBA00301). This work was completed when XW was a student at Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48105, USA. This work was also supported in part by US NSF Awards 1017335 and 1216729.

References

[1] S. Aaronson. Limitations of quantum advice and one-way communication. *Theory of Computing*, 1:1–28, 2005.

[2] Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis. Exponential separation of quantum and classical one-way communication complexity. *SIAM Journal on Computing*, 38(1):366–384, 2008.

[3] A. Ben-Aroya, O. Regev, and R. de Wolf. A hypercontractive inequality for matrix-valued functions with applications to quantum computing and ldc's. In *Proceedings of IEEE FOCS*, 2008.

[4] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf. Quantum fingerprinting. *Physical Review Letters*, 87:167902, 2001.

[5] H. Buhrman, O. Regev, G. Scarpa, and R. de Wolf. Near-optimal and explicit bell inequality violations. In *Proceedings of the 2011 IEEE 26th Annual Conference on Computational Complexity*, pages 157–166, 2011.

[6] K. Efremenko. 3-query locally decodable codes of subexponential length. In *Proceedings of the 2009 ACM International Symposium on Theory of Computing*, pages 39–44, 2009.

[7] D. Gavinsky, J. Kempe, and R. de Wolf. Strengths and weaknesses of quantum fingerprinting. In *Proc. of the 21st Conf. on Computational Complexity (CCC)*, pages 288–298, 2006.

[8] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf. Exponential separation for one-way quantum communication complexity, with applications to cryptography. *SIAM Journal on Computing*, 38(5):1695–1708, 2008.

[9] C. W. Helstrom. *Quantum Detection and Estimation Theory*. Academic Press Inc., New York, NY, USA, 1976.

[10] J. Kahn, G. Kalai, and N. Linial. The influence of variables on boolean functions. In *Proceedings of 29th IEEE FOCS*, pages 68–80, 1988.

[11] I. Kerenidis and R. de Wolf. Exponential lower bound for 2-query locally decodable codes via quantum argument. *Journal of Computer and System Sciences*, 69(3):395–420, 2004.

[12] A. Kitaev, A. Shen, and M. Vyalyi. *Classical and Quantum Computation*. American Mathematical Society, 2002.

[13] A. Nayak. Optimal lower bounds for quantum automata and random access codes. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 124–133, 1999.

[14] I. Newman. Private vs. common random bits in communication complexity. *Information Processing Letters*, 39:67–71, 1991.

[15] R. de Wolf. A brief introduction to fourier analysis on the boolean cube. *Theory of Computing*, Toc Library:Graduate Surveys 1, 2008.

[16] R. de Wolf, O. Regev. Personal communications. 2012.

[17] S. Yekhanin. Towards 3-query locally decodable codes of subexponential length. *Journal of the ACM*, 55(1):Art. 1, 16, 2008.

[18] A. C.-C. Yao. On the power of quantum fingerprinting. In *Proceedings of the 35th annual ACM symposium on Theory of computing*, pages 77–81, 2003.

A.1 Quantum and classical upper bounds

Since the upper bounds of the function problems follow from the upper bound of the relation problem, we will focus on describing the upper bounds of the relation problem and sketch how they extend to the function case. Note that the quantum upper bounds we describe are implicitly in [11, Section 6] and the classical upper bounds are extensions of the $k = 2$ case.

Quantum Upper Bounds. We sketch an $O(n^{1-1/\lceil k/2 \rceil} \log(n))$ quantum protocol for the relation problem. For any Alice's input $x \in \{0,1\}^{kn}$, let $|\psi_x\rangle$ be a uniform superposition of her bits $x = x_1 \cdots x_{kn}$ (i.e., a fingerprint state of x [4]):

$$|\psi_x\rangle = \frac{1}{\sqrt{kn}} \sum_{i=1}^{kn} (-1)^{x_i} |i\rangle .$$

Let us consider the case where k is even. Then Bob's input M partitions $[kn]$ into n subsets each consisting of $k/2$ pairs. Let $\{(i_t, j_t)\}_{t=1}^{kn/2}$ denote the set of these pairs. Thus by measuring $|\psi_x\rangle$ in the basis $\{\frac{1}{\sqrt{2}}(|i_t\rangle \pm |j_t\rangle) : t = 1 \cdots kn/2\}$, we can determine the parity of a uniformly picked pair. In order to determine the parity of some subset, we need to know the parities of the $k/2$ pairs in some subset. By the birthday paradox, we need $O(n^{1-2/k})$ copies of $|\psi_x\rangle$ to have a $k/2$-collision with high probability. Furthermore, the $k/2$ pairs are distinct with constant probability. Thus, with constant bias, Bob can determine the parity of a uniformly selected subset with $O(n^{1-2/k} \log(n))$ size messages. In the case when k is odd, it is trivial to reduce to the $k + 1$ case by filling dummy 0s in the input. Thus, we have an $O(n^{1-1/\lceil k/2 \rceil} \log(n))$ quantum upper bound.

In the case of (k, α)-PM$_n$, Bob might get the parity of some subset that does not lie in the desired α-fraction. This can be resolved by repeating the protocol $O(1/\alpha)$ times. Also by repeating the protocol $O(\log(1/\epsilon))$ times we can boost the correctness to $1 - \epsilon$ for any constant $\epsilon > 0$. Thus we obtain a upper bound of $O(n^{1-1/\lceil k/2 \rceil} \log(n)/\alpha)$.

Classical Upper Bounds. We first sketch an $O(n^{1-1/k})$ classical upper bound for the relation problem. Let Alice uniformly select a subset of $d \approx n^{1-1/k}$ bits of $x \in \{0,1\}^{kn}$ to send to Bob. By the birthday paradox, with high probability Bob will have k bits information that lie in the same subset (i.e., a k-collision instead of a 2-collision). Thus he can output the parity of that subset. By

Newman's Theorem [14], Alice only needs to send $O(n^{1-/k}+\log(n)) = O(n^{1-1/k})$ size classical messages to accomplish the task.

To extend the result to the function problem we observe that the subset, the parity of which Bob knows with high probability, is uniformly random. In the case of (k,α)-PM$_n$ problem, however, the subset uniformly picked might not lie in the desired α-fraction of subsets. A more careful analysis shows by selecting a subset of $O(n^{1-1/k}/\alpha^{1/k})$ bits randomly, with high probability Bob can recover the parity of some subset in the desired fraction.

We note all the classical upper bounds are also tight for the generalized Hidden Matching problems.

A.2 Classical lower bound for k-HM$_n$

We prove here an asymptotically tight classical lower bound for k-HM$_n$.

Theorem A.2. *Any deterministic classical protocol $\mathcal{C}$ with m-bit messages for k-HM$_n$ problem has the best successful probability no more than $\frac{1}{2}+O((\frac{m^k}{n^{k-1}})^{1/2})$. To achieve constant bias, we have $m=\Omega(n^{1-1/k})$.*

Proof. The proof is a simple extension of that for 2-HM$_n$ in [5]. Fix a classical deterministic protocol $\mathcal{C}$. For any Alice's message $c\in\{0,1\}^m$, let $X_c\subseteq\{0,1\}^{kn}$ be the set of corresponding inputs. Define $p_c=|X_c|/2^{kn}$ and note that $\{p_c\}$ is a probability distribution over the 2^m messages c. Conditioned on receiving message c, Bob shall choose a subset G of size k from his input $M\in\mathcal{M}_{k,n}$ according to some distribution denoted by

$$q_c(G)=\mathbf{Pr}_{M\in\mathcal{M}_{k,n}}[\text{Bob outputs } G\mid \text{Bob received } c].$$

Clearly for any fixed G we have $q_c(G)\leq \mathbf{Pr}_{M\in\mathcal{M}_{k,n}}[G\in M]=1/\binom{kn-1}{k-1}$. Upon receiving c and outputting subset G, Bob's best strategy is then to output the parity of G that occurs most often among the $x\in X_c$. Define $\beta^c_G=\mathbb{E}_{x\in X_c}[(-1)^{x\cdot G}]$. Hence Bob's optimal success probability when guessing $(-1)^{x\cdot G}$ is $1/2+|\beta^c_G|/2$. This implies, for fixed c, Bob's successful chance is

$$\mathbf{Pr}_{\substack{x\in X_c\\ M\in\mathcal{M}_{k,n}}}[\text{Bob outputs the parity of some } G]=\frac{1}{2}+\epsilon^c_{\text{bias}}\leq \mathbb{E}_{G\sim q_c}\left[\frac{1}{2}+\frac{|\beta^c_G|}{2}\right],$$

where the expectation is taken over the distribution $q_c(G)$. As explained in [15, Section 4.2], it follows from the KKL inequality [10] that

$$\sum_{\substack{G\subseteq[kn]\\ |G|=k}}(\beta^c_G)^2\leq O\left(\log\frac{1}{p_c}\right)^k. \tag{16}$$

Thus we can upper bound ϵ^c_{bias} as follows:

$$2\epsilon^c_{\text{bias}}\leq \mathbb{E}_{G\sim q_c}[|\beta^c_G|]=\sum_G q_c(G)|\beta^c_G|\overset{(*)}{\leq}\sqrt{\sum_G q_c(G)^2}\cdot\sqrt{\sum_G(\beta^c_G)^2}\overset{(**)}{\leq}\binom{kn-1}{k-1}^{-1/2}\cdot O\left(\log\frac{1}{p_m}\right)^{k/2},$$

where $(*)$ is from Cauchy-Schwarz and $(**)$ follows from $\sum_G q_c(G)^2\leq\max_G q_c(G)\cdot\sum_G q_c(G)\leq\max_G q_c(G)\leq 1/\binom{kn-1}{k-1}$ and **Eq.** (16). Then the overall bias ϵ_{bias} is given by

$$\epsilon_{\text{bias}}=\sum_c p_c\epsilon^c_{\text{bias}}\leq\binom{kn-1}{k-1}^{-1/2}\sum_c p_c O(\log(1/p_c))^{k/2}\leq O\left((\frac{m^k}{(k-1)n^{k-1}})^{1/2}\right),$$

where the last inequality follows from $\binom{kn-1}{k-1} \geq ((k-1)n)^{k-1}$ and $\sum_c p_c \log^{k/2}(1/p_c) \leq m^{k/2}$ for $k \geq 2$.[6] □

A.3 Proof of Lemma 3.2

Proof. We prove a slightly more general result. Fix an arbitrary POVM $\Pi = \{\Pi_{a,b}\}$ indexed by $a \in \mathcal{A}, b \in \mathcal{B}$. We shall construct (sub-normalized) POVMs

$$P = \{P_a\}_{a\in\mathcal{A}}, \quad \text{and,} \quad Q_a = \{Q_{a,b}\}_{b\in\mathcal{B}}, \text{ for } a \in \mathcal{A},$$

such that for all state ρ that Π acts on, applying P, followed by Q_a, where a is the outcome of P, gives the same output distribution.

For an operator R, denote by R^+ the *Moore-Penrose pseudo-inverse* of R. We set

$$P_a = \sum_{b\in\mathcal{B}} \Pi_{a,b}, \quad \text{and,} \quad Q_{a,b} = \sqrt{P_a^+}\Pi_{a,b}\sqrt{P_a^+}.$$

By direct computation, one can verify that P is a POVM, while for each a, Q_a is a sub-normalized POVM. For an arbitrary quantum state ρ, the probability of observing (a, b) in the two-stage measurement is

$$\langle \rho, P_a\rangle \left\langle \frac{\sqrt{P_a}\rho\sqrt{P_a}}{\langle\rho, P_a\rangle}, Q_{a,b} \right\rangle = \left\langle \sqrt{P_a}\rho\sqrt{P_a}, \sqrt{P_a^+}\Pi_{a,b}\sqrt{P_a^+} \right\rangle,$$

which is $\langle \rho, \Pi_{a,b}\rangle$ since $\sqrt{P_a}\sqrt{P_a^+}$ is the projection on to the support of P_a, and $\Pi_{a,b}$'s support lies in that subspace. □

We notice that the post-measurement states after the two-stage measurement are in general different from those after the original measurement, but this difference bears no consequence for our proof.

A.4 Equation (12)

By a simple counting argument,

$$|\mathcal{M}_{k,n}| = (kn)! \cdot ((k!)^n n!)^{-1}$$

Thus for $S \subseteq [kn]$ with $|S| = kl$,

$$\mathbf{Pr}_{M\in\mathcal{M}_{k,n}}[S \in v(M)] = |\mathcal{M}_{k,l}| \cdot |\mathcal{M}_{k,n-l}|/|\mathcal{M}_{k,n}| = \binom{n}{l}\binom{kn}{kl}^{-1}.$$

Note that by definition,

$$|\alpha\text{-}\mathcal{M}_{k,n}| = \binom{kn}{\alpha kn}|\mathcal{M}_{k,\alpha n}|.$$

Thus for $S \subseteq [kn]$ with $|S| = kl$, setting $\beta = \frac{\alpha n - l}{n-l}$,

$$\mathbf{Pr}_{M\in\alpha\text{-}\mathcal{M}_{k,n}}[S \in v(M)] = |\mathcal{M}_{k,l}| \cdot |\beta\text{-}\mathcal{M}_{k,n-l}|/|\alpha\text{-}\mathcal{M}_{k,n}| = \binom{\alpha n}{l}\binom{kn}{kl}^{-1}.$$

[6] The proof of this inequality is subtle but straightforward by noticing the function is concave in certain interval.

A.5 Inequality (15)

Recall that $\gamma_l = \binom{n}{l}^2 \binom{kn}{kl}^{-1}$. We will make use of the inequality that

$$\frac{a+s}{b+s} \leq \frac{a}{b}, \text{ for all } a, b, s > 0 \text{ with } a \geq b.$$

For all integers i, j with $0 \leq i \leq l-1$,

$$\frac{n-i}{l-i} = \frac{k(n-i)}{k(l-i)} \leq \frac{kn-ik-1}{kl-ik-1}.$$

Thus

$$\left(\frac{n-i}{l-i}\right)^2 \leq \frac{kn-ik}{kl-ik} \frac{kn-ik-1}{kl-ik-1}.$$

For each integer j, $2 \leq j \leq k-1$,

$$\frac{kn-ik-j}{kl-ik-j} \geq \frac{n}{l}.$$

Therefore, for each i,

$$\frac{\left(\frac{n-i}{l-i}\right)^2}{\Pi_{j=0}^{k-1} \frac{kn-ik-j}{kl-ik-j}} \leq \left(\frac{l}{n}\right)^{k-2}.$$

Since γ_l is simply the product of the left hand side for $i = 0, .., l-1$, we have

$$\gamma_l \leq \left(\frac{l}{n}\right)^{kl-2l}.$$

Thus

$$\delta_l = \gamma_l^{1/kl} \leq \left(\frac{l}{n}\right)^{1-2/k}.$$