Electronic Colloquium on Computational Complexity, Report No. 83 (2023)

Trade-offs between Entanglement and Communication

Srinivasan Arunachalam
IBM Quantum, Almaden Research Center
Srinivasan.Arunachalam@ibm.com

Uma Girish
Princeton University
ugirish@cs.princeton.edu

Abstract

We study the advantages of quantum communication models over classical communication models that are equipped with a limited number of qubits of entanglement. In this direction, we give explicit partial functions on n bits for which reducing the entanglement increases the classical communication complexity exponentially. Our separations are as follows. For every $k \geq 1$:

$\mathsf{Q}\|^*$ versus $\mathsf{R2}^*$: We show that quantum simultaneous protocols with $\tilde{\Theta}(k^5 \log^3 n)$ qubits of entanglement can exponentially outperform two-way randomized protocols with $O(k)$ qubits of entanglement. This resolves an open problem from [Gav08] and improves the state-of-the-art separations between quantum simultaneous protocols with entanglement and two-way randomized protocols without entanglement [Gav19, GRT22].

$\mathsf{R}\|^*$ versus $\mathsf{Q}\|^*$: We show that classical simultaneous protocols with $\tilde{\Theta}(k \log n)$ qubits of entanglement can exponentially outperform quantum simultaneous protocols with $O(k)$ qubits of entanglement, resolving an open question from [GKRW06, Gav19]. The best result prior to our work was a relational separation against protocols without entanglement [GKRW06].

$\mathsf{R}\|^*$ versus $\mathsf{R1}^*$: We show that classical simultaneous protocols with $\tilde{\Theta}(k \log n)$ qubits of entanglement can exponentially outperform randomized one-way protocols with $O(k)$ qubits of entanglement. Prior to our work, only a relational separation was known [Gav08].

1 Introduction

One of the central goals in complexity theory is to understand the power of different computational resources. In the past four decades, communication complexity has provided a successful toolbox to establish various results in different areas of research in theoretical computer science such as circuit complexity [KW90, KRW95], streaming algorithms [KKS14], property testing [BBM12], extension complexity [FMP+15], data structures [MNSW95], proof complexity [HN12]. In the standard two-player model of communication complexity introduced by Yao [Yao79] there are two parties Alice and Bob whose goal is to compute a partial function $F : \mathcal{X} \times \mathcal{Y} \to \{-1, 1, \star\}$. Alice receives $x \in \mathcal{X}$ (unknown to Bob) and Bob receives $y \in \mathcal{Y}$ (unknown to Alice) and their goal is to compute $F(x, y)$ for all $(x, y) \in F^{-1}(1) \cup F^{-1}(-1)$, while minimizing the amount of communication. In this setting, there are three models of communication in increasing order of strength:

(*i*) Simultaneous message passing (SMP) model: Alice and Bob send a message to a referee Charlie, whose goal is to output $F(x, y)$.

(*ii*) One-way model: Alice sends a message to Bob, whose goal is to output $F(x, y)$.

(*iii*) Two-way model: Alice and Bob can exchange several rounds of messages and their goal is to output $F(x, y)$.

ISSN 1433-8092

In all these models, the complexity of the protocol is the total number of bits used to describe the message. It is not hard to see that the communication complexity in model (i) is at least the complexity in model (ii) which in turn is at least the complexity in model (iii).

One variant of these models is when the players are allowed to use *quantum resources*, for instance, the players could send quantum messages or share entanglement. Over the past two decades, several works have established the advantage of quantum over classical communication complexity in various settings. In a sequence of works [BCW98, BCWW01, Raz99, GKRW06, GKK$^+$07, KR11, Gav20], it has been shown that quantum communication can exponentially outperform classical communication. In particular, a few works [Gav09, Gav19, GRT22] have demonstrated communication tasks that are easy to solve in the SMP model if the players share entanglement, however, every interactive randomized protocol without entanglement has exponentially larger cost. This leads to a natural and fundamental question (which has been asked many times before [JKN07, CH19, Shi05, Gav08]): *How much entanglement do quantum protocols really need?* Given any small-cost quantum protocol, can we simulate it by a small-cost quantum protocol that uses only a small amount of entanglement? Answering this question is one of the central questions in quantum communication complexity; in fact giving *any* upper bound on the number of qubits in a potentially helpful shared state has been open for decades.

A similar question of how much shared *randomness* is necessary in classical communication complexity is well understood. In a famous result, Newman [New91] showed that to solve communication tasks on n-bit inputs, with an additive overhead of $O(\log n)$ bits in communication one can assume that the players only have private randomness. Jain et al. [JRS05] showed that blackbox arguments similar to the one in [New91] cannot be used to reduce the entanglement in a quantum protocol. Motivated by the question of how much entanglement protocols need, we study a fine-grained variant of this question, which will be the topic of this work.

> *Can we reduce the entanglement in a quantum communication protocol from k qubits to $k/\log n$ qubits using a classical protocol of only polynomially larger cost?*

In this direction, Shi [Shi05] showed that we can remove any amount of entanglement using a classical communication protocol of exponentially larger cost. Subsequently, [Gav08, JKN07] showed that this exponential blowup is inevitable, in particular they constructed a *relational* problem for which we cannot reduce the entanglement with just a polynomial overhead using one-way communication alone. Their works left open the question of reducing entanglement in a quantum protocol computing a *partial* function, using two-way classical communication between the players.[1]

1.1 Main Result

In this work, we provide a strong negative answer to this question. We give partial functions for which, reducing the entanglement by even a logarithmic factor, increases the communication cost by an exponential factor. To discuss our results, we set up some notation first. Let $\mathsf{R}\|^*$ (resp. $\mathsf{Q}\|^*$) denote the SMP communication model where Alice and Bob share entanglement and send classical (resp. quantum) messages to the referee. Let $\mathsf{R1}^*, \mathsf{R2}^*$ be the one-way and two-way models of classical communication where Alice and Bob share entanglement. The models $\mathsf{R1}$ and $\mathsf{R2}$ are similarly defined with the difference being that Alice and Bob don't share entanglement. The

[1] Relational separations are known as the "weakest" form of separations between communication models. A partial function separation immediately implies a relational separation, however, the converse is false [GKdW06].

model $\mathsf{Q}\|^{\mathrm{pub}}$ is also defined similarly to $\mathsf{Q}\|^*$ but without entanglement, additionally, the players are allowed public randomness. We first summarize our results informally below. All these results hold for every $k \geq 1$ which is any parameter that is allowed to depend on n.

Our first result shows that for simultaneous quantum protocols, more entanglement cannot be simulated by two-way classical communication with less entanglement (and a polynomial overhead).

Result 1. *There is a partial function on $\tilde{O}(kn)$ bits that can be computed in $\mathsf{Q}\|^*$ with $\tilde{O}(k^5 \log^3 n)$ qubits of communication and entanglement, but if the players only share $O(k)$ qubits of entanglement, requires $\Omega(n^{1/4})$ bits of communication in the $\mathsf{R2}^*$ model.*

There are two ways to view this result: (i) It shows that in the rather weak quantum SMP model, reducing the entanglement by a polylogarithmic factor increases the classical communication by an exponential factor, even if Alice and Bob are allowed to interact. This answers an open question in [Gav08]. (ii) This result can also be viewed in the context of quantum versus classical separations in communication complexity. As we mentioned earlier, numerous works [BCW98, Raz99, GKK$^+$07, KR11, Gav20] have shown that quantum provides exponential savings for partial functions in various settings. The state-of-the-art separations between quantum and classical communication complexity for partial functions are due to [Gav19, GRT22]; they show separations between $\mathsf{Q}\|^*$ and $\mathsf{R2}$. One drawback of the aforementioned works, in the context of our work, is that the lower bound can only be made to work for protocols where Alice and Bob share $\ll \log n$ qubits of entanglement. We improve upon this by showing separations between $\mathsf{Q}\|^*$ (with more entanglement) and $\mathsf{R2}^*$ (with less entanglement). Our result can thus be seen as the current best-known separation between quantum and classical communication complexity for partial functions. In particular, we give a lower bound technique against $\mathsf{R2}^*$ protocols with $O(\log^c n)$ qubits of entanglement for every $c \in \mathbb{N}$. To the best of our knowledge, there were no known lower bound techniques that distinguished $\mathsf{R2}^*$ (with more entanglement) and $\mathsf{R2}^*$ (with less entanglement) once the number of qubits of entanglement is $\gg \log n$, even for *relational problems*.

Our second result shows that for SMP protocols where the players share entanglement but only send classical messages, entanglement cannot be reduced even by quantum simultaneous protocols or by one-way classical protocols (with a polynomial overhead).

Result 2. *There is a partial function on $\tilde{O}(kn)$ bits that can be computed in $\mathsf{R}\|^*$ using $\tilde{O}(k \log n)$ bits of communication and $\tilde{O}(k \log n)$ qubits of entanglement, but if the players share $O(k)$ qubits of entanglement, requires $\Omega(n^{1/3})$ qubits of communication in the $\mathsf{Q}\|^*$ model and $\Omega(\sqrt{n})$ bits in the $\mathsf{R1}^*$ model.*

We remark that the trade-offs obtained in this result are more fine-grained in comparison to Result 1, i.e., our separations hold even if we reduce the entanglement by a $O(\log n)$-factor. Prior to our work, the best known separation between $\mathsf{R}\|^*$ (with more entanglement) and $\mathsf{Q}\|^*$ (with less entanglement) was a relational separation between $\mathsf{R}\|^*$ and $\mathsf{Q}\|^{\mathrm{pub}}$ [GKRW06]. Their work left open two questions: (i) Does there exist a *partial function* separating $\mathsf{R}\|^*$ and $\mathsf{Q}\|^{\mathrm{pub}}$? The weaker question of showing a functional separation between $\mathsf{Q}\|^*$ and $\mathsf{Q}\|^{\mathrm{pub}}$ was also open and recently asked by [Gav19]. (ii) Is there a *relational* separation between $\mathsf{Q}\|^*$ (with more entanglement) and $\mathsf{Q}\|^*$ (with less entanglement)? Our result answers both these questions. Firstly, we prove separations for partial functions improving upon the relational separations; secondly, we also show lower bounds for $\mathsf{Q}\|^*$ with limited entanglement. With regards to separations between $\mathsf{R}\|^*$ (with more entanglement) and $\mathsf{R1}^*$ (with less entanglement), prior to our work these were established in [Gav08, JKN07], again for relational problems. Gavinsky [Gav08] left open the question of showing a similar separation for *partial* functions and our work resolves this.

In the next two sections, we discuss the problems witnessing these separations followed by the proof sketches. Our first result is based on the Forrelation problem and the second result is based on the Boolean Hidden Matching problem.

1.2 Result 1: Separations based on the Forrelation problem

1.2.1 Problem Definition: The Forrelation Problem

The Forrelation problem was first introduced by Aaronson in the context of query complexity [Aar10] and subsequently has been studied again in the context of separating quantum and classical computation [RT22, AA15]. Variants of the Forrelation problem have been used to show various quantum versus classical separations in communication complexity [GRT22, BS21, SSW21, GRZ21]. The state-of-the-art separations for quantum versus classical communication complexity of partial functions are between $\mathsf{Q}\|^*$ and $\mathsf{R2}$; one such separation is due to [GRT22] and is based on the Forrelation problem, which we define now.

Definition 1.1 (Forrelation Function)**.** *Let $n \in \mathbb{N}, n \geq 2$ be a power of two. Let H_n be the (unitary) $n \times n$ Hadamard matrix. For $z_1, z_2 \in \{-1,1\}^{n/2}$, define the* forrelation *function as*

$$\mathrm{forr}(z_1, z_2) = \frac{1}{n}\langle z_2, H_n(z_1)\rangle.$$

Let $\varepsilon \in (0,1]$ be a parameter. We typically set $\varepsilon = \Theta\left(\frac{1}{\log n}\right)$ if it is not specified. We are interested in the communication complexity version of the Forrelation problem defined below.

Definition 1.2 (The Forrelation Problem)**.** *In the Forrelation problem, Alice is given $x \in \{-1,1\}^n$, Bob is given $y \in \{-1,1\}^n$. Their goal is to compute $\textsc{forr}(x,y)$ given by*

$$\textsc{forr}(x,y) = \begin{cases} -1 & \mathrm{forr}(x \odot y) \geq \varepsilon/4 \\ 1 & \mathrm{forr}(x \odot y) \leq \varepsilon/8. \end{cases}$$

Here, $\odot$ denotes the pointwise product. Let $k \in \mathbb{N}$ be a parameter satisfying $k = o(n^{1/50})$. We are interested in the XOR of k copies of the Forrelation problem. This problem was first studied in [GRZ21] in the context of XOR lemmas.

Definition 1.3 ($\oplus^k$-Forrelation Problem)**.** *This problem is the XOR of k independent instances of the Forrelation problem where $\varepsilon = \frac{1}{60k^2 \ln n}$. To be precise, Alice and Bob receive $x = (x^{(1)}, \ldots, x^{(k)})$ and $y = (y^{(1)}, \ldots, y^{(n)})$ where $x^{(i)}, y^{(i)} \in \{-1,1\}^n$ for all $i \in [k]$, and they need to compute*

$$\textsc{forr}^{(\oplus k)}(x,y) = \prod_{i=1}^{k} \textsc{forr}\left(x^{(i)}, y^{(i)}\right).$$

1.2.2 Main Theorem

We now state our main theorem. For $n \in \mathbb{N}$, let $k \in \mathbb{N}$ be a parameter satisfying $k = o(n^{1/50})$.

Theorem 1.1. *The $\oplus^k$-Forrelation problem can be solved with $\tilde{O}(k^5 \log^3 n)$ qubits of communication in the $\mathsf{Q}\|^*$ model if Alice and Bob share $\tilde{\Theta}(k^5 \log^3 n)$ EPR pairs. However, if they share $O(k)$ qubits of entanglement, then this problem requires $\Omega(n^{1/4})$ bits of communication even in the $\mathsf{R2}^*$ model.*

We make a few remarks. First, the upper bound holds provided Alice and Bob share $\tilde{\Theta}(k^5 \log^3 n)$ EPR pairs, however, the lower bound holds for all possible entangled states on $O(k)$ qubits, not necessarily EPR pairs. See Section 2.2 for a formal description of the models and EPR pairs. Second, although Theorem 1.1 is stated for bounded-error models, our lower bound also holds for protocols with advantage $2^{-o(k)}$. To prove our lower bound, our main technical contribution is to show a *Fourier growth bound for R2* protocols* with limited entanglement.[2] In the following lemma, $O_\ell(t)$ is a shorthand notation for $O(t \cdot 2^{O(\ell)})$.

Lemma 1.1. *Let $C : \{-1,1\}^n \times \{-1,1\}^n \to [-1,1]$ be an R2* protocol of cost c where Alice and Bob share an entangled state on at most $2d$ qubits for some parameter $d \in \mathbb{N}$. Let H be the XOR-fiber of C as in Definition 2.2. Then, for all $\ell \in \mathbb{N}$, we have*

$$L_{1,\ell}(H) \triangleq \sum_{|S|=\ell} \left| \widehat{H}(S) \right| \leq 2^{5d} \cdot O_\ell(c^\ell).$$

We also study lower bounds for the Forrelation problem in the quantum SMP model. Prior to our work, there was no partial function separating $\mathsf{Q}\|^*$ with more entanglement and $\mathsf{Q}\|^*$ with less entanglement. In particular, it was unknown whether the Forrelation problem can be solved in the $\mathsf{Q}\|^{\text{pub}}$ model with small cost and no entanglement. Our result shows that this is not the case, and that the Forrelation problem separates $\mathsf{Q}\|^*$ from $\mathsf{Q}\|^{\text{pub}}$, resolving an open problem from [Gav19].

Theorem 1.2. *The Forrelation problem requires $\Omega(n^{1/4})$ qubits of communication in the $\mathsf{Q}\|^{\text{pub}}$ model.*

This is also proved using a Fourier growth bound.

Lemma 1.2. *Let $C : \{-1,1\}^n \times \{-1,1\}^n \to [-1,1]$ be a $\mathsf{Q}\|^{\text{pub}}$ protocol of cost c and let H be its XOR-fiber as in Definition 2.2. Then, for all $\ell \in \mathbb{N}$, we have*

$$L_{1,\ell}(H) \leq O_\ell(c^\ell).$$

We remark that our techniques also implies that for the $\oplus^k$-Forrelation problem, if Alice and Bob only share $O(k)$ EPR pairs, they require $\Omega(n^{1/4})$ *qubits* of communication in the $\mathsf{Q}\|^*$ model. We don't show the details of this proof, instead, we prove a stronger result, namely Theorem 1.3. We give an example of a partial function such that with $\Theta(k \log n)$ EPR pairs, it is solvable in the $\mathsf{R}\|^*$ model with cost $O(k \log n)$, however, with only $O(k)$ qubits of entanglement requires cost $\Omega(n^{1/3})$ even in the $\mathsf{Q}\|^*$ model.

1.3 Result 2: Separations based on Boolean Hidden Matching

1.3.1 Problem Definition: The Boolean Hidden Matching Problem

We first define the Boolean Hidden Hatching problem. The (relational) Hidden Matching problem was first defined by Bar Yossef et al. [BJK08]. The Boolean Hidden Matching problem was defined by Gavinsky et al. [GKK+07] in the context of one-way communication complexity and was used to separate the $\mathsf{R}\|^*$ and $\mathsf{R1}$ models. Subsequently this problem and its variants have found several applications, especially in proving streaming lower bounds starting with the seminal work of Kapralov et al. [KKS14]. The Boolean Hidden Matching problem, denoted $\mathrm{BHM}_{m,n}$, is defined as follows. Let $n, m \in \mathbb{N}$ be parameters and $m = \alpha n$ for a small enough constant $\alpha \ll 1$.

[2] For the introduction, we will loosely say "Fourier growth of communication protocols", when strictly speaking, we are referring to Fourier growth of *XOR-fibers* and other functions associated with communication protocols.

Definition 1.4 (Boolean Hidden Matching)**.** *Alice gets $x \in \{-1,1\}^n$, Bob gets a matching on $[n]$ with m edges and a string $y \in \{-1,1\}^m$. Their goal is to compute $\mathrm{BHM}_{m,n}(x,y,M)$ given by*

$$\mathrm{BHM}_{m,n}(x,y,M) = \begin{cases} -1 & \text{if } Mx = \overline{y} \\ 1 & \text{if } Mx = y. \end{cases}$$

Here, we use $Mx \in \{-1,1\}^m$ to denote the vector whose k-th coordinate is $x_{i_k} \cdot x_{j_k}$ for $k \in [m]$, where the edges of M are $(i_1, j_1), \ldots, (i_m, j_m) \in [n]^2$. We also use $\overline{y}$ to denote $-y$. Below we will be concerned with computing the XOR of k independent copies of $\mathrm{BHM}_{m,n}$.

Definition 1.5 ($\oplus^k$-Boolean Hidden Matching Problem)**.** *This problem is the XOR of k independent instances of the Boolean Hidden Matching problem. To be precise, Alice receives $x = (x^{(1)}, \ldots, x^{(k)})$ and Bob receives $y = (y^{(1)}, \ldots, y^{(k)})$ and $M_1, \ldots, M_k$ where $x^{(i)} \in \{-1,1\}^n$, $y^{(i)} \in \{-1,1\}^m$ and M_i is a matching on $[n]$ with m edges for all $i \in [k]$. They need to compute*

$$\mathrm{BHM}_{m,n}^{(\oplus k)}(x,y) = \prod_{i=1}^{k} \mathrm{BHM}_{m,n}\left(x^{(i)}, y^{(i)}, M_i\right).$$

1.3.2 Main Theorem

We now state our main theorem. Here, $\alpha \ll 1$ is some absolute constant and $k \in \mathbb{N}$ is a parameter, possibly depending on $n \in \mathbb{N}$.

Theorem 1.3. *The $\oplus^k$-Boolean Hidden Matching problem can be solved with $\tilde{O}(k \log n)$ bits of communication in the $\mathsf{R}\|^*$ model if Alice and Bob share $\tilde{\Theta}(k \log n)$ EPR pairs. However, if Alice and Bob only share $O(k)$ qubits of entanglement, then this problem requires*

- *$\Omega(k\sqrt{n})$ bits of communication in the $\mathsf{R1}^*$ model,*
- *$\Omega(kn^{1/3})$ qubits of communication in the $\mathsf{Q}\|^*$ model.*

Similar to the discussion below Theorem 1.1, our upper bound holds provided Alice and Bob share $\tilde{\Theta}(k \log n)$ EPR pairs, however, the lower bound holds for all possible entangled states on $O(k)$ qubits, and our lower bound also holds for protocols even with advantage $2^{-o(k)}$. The main technical contribution of this part is to argue that $\mathsf{R1}$ and $\mathsf{Q}\|^{\mathrm{pub}}$ protocols satisfy an XOR lemma with respect to computing the Boolean Hidden Matching problem.

XOR Lemmas. XOR lemmas study the relation between the computational resources of F and the k-fold XOR of F on k independent inputs. In particular, XOR lemmas for communication complexity are of the following format: If cost-t protocols have advantage at most $2/3$ in computing F, then cost-$o(tk)$ protocols have advantage at most $2^{-\Theta(k)}$ in computing the k-fold XOR of F. XOR lemmas provide a framework to construct hard objects in a black-box way and have applications to several areas in theoretical computer science such as one-way functions, pseudorandom generators and streaming algorithms. We prove an XOR lemma for $\mathsf{R1}$ and $\mathsf{Q}\|^{\mathrm{pub}}$ protocols with respect to computing the Boolean Hidden Matching problem.

Lemma 1.3. *Let C be any $\mathsf{Q}\|^{\mathrm{pub}}$ protocol of cost c. Then its advantage in computing the $\oplus^k$-Boolean Hidden Matching problem is at most $O_k\left(\frac{(c/k)^3}{n}\right)^{k/2} + O_k(n^{-k/2})$.*

Lemma 1.4. *Let C be any R1 protocol of cost c. Then its advantage in computing the $\oplus^k$-Boolean Hidden Matching problem is at most $O_k\left(\frac{(c/k)^2}{n}\right)^{k/2} + O_k(n^{-k/2})$.*

Until very recently [Yu22], we didn't have an XOR lemma for R1 and as far as we are aware we do not have *any* XOR lemmas for the quantum communication model. However we do have direct product and direct sum theorems for classical and quantum communication models (which are strictly weaker than XOR lemmas) and in fact this was used in the prior work of [GKRW06, JKN07]. Our main technical contribution here is an XOR lemma for R1 and $\mathsf{Q}\|^{\text{pub}}$ protocols for the Boolean Hidden Matching problem. Only during completion of this project, we were made aware of a recent work by Yu [Yu22] proving an XOR lemma for all constant round classical protocols (hence implying Lemma 1.4). Given the technicality of his proof, in our paper we present a simple proof for an XOR lemma for the R1 model for the Boolean Hidden Matching problem.

1.4 Proof Sketch

One of the difficulties of proving lower bounds against classical models equipped with entanglement is that these models are quite powerful; using the quantum teleportation protocol, any quantum protocol with q qubits of communication can be classically simulated using q EPR pairs. Thus, all known partial functions that separate quantum and classical communication complexity are easy to classically simulate in the presence of $O(\log^c n)$ EPR pairs for some small constant $c > 0$.

One approach to show a fine-grained separation between protocols with more entanglement and protocols with less entanglement is the following. Consider any communication task F that exhibits an exponential separation between quantum and classical communication complexity. We have many examples of such tasks that are easy with $O(\log n)$ EPR pairs but exponentially harder in the absence of entanglement. Consider the problem of solving k independent and parallel instances of F. Here, the players receive k pairs of inputs (x_i, y_i) and need to compute $F(x_i, y_i)$ *for every* $i \in [k]$. We denote this problem by $F^{(k)}$. The hope is that entanglement obeys a direct sum theorem of sorts, that is, if the players need at least $\Omega(\log n)$ qubits of entanglement to solve the original task, then to solve k independent and parallel instances, they need at least $\Omega(k \log n)$ qubits of entanglement. In particular, we might hope that protocols that compute $F^{(k)}$ using only $O(k)$ qubits of entanglement require exponentially larger cost. There is a way to make this idea work and this was done in [Gav08]. We describe this idea. Assume by contradiction that we have a small-cost protocol computing $F^{(k)}$ using only $O(k)$ qubits of entanglement.

Step 1: Remove entanglement. The first step is to remove all entanglement from this protocol. To do this, we replace the entangled state on $O(k)$ qubits by the maximally mixed state on $O(k)$ qubits. Since the maximally mixed state is unentangled, the resulting protocol effectively uses no entanglement. Furthermore, the mixed state can be viewed as a probability distribution over states, where the original entangled state occurs with probability $2^{-\Theta(k)}$. It follows that this protocol succeeds with probability at least $2^{-\Theta(k)}$.

Step 2: Direct Product Theorems. The second step is to prove a direct product theorem in the absence of entanglement. Direct product theorems in communication complexity are of the following form: If for cost-t protocols, the probability of solving one instance of F is at most $2/3$, then for cost-$o(tk)$ protocols, the probability of solving k parallel and independent instances of F is at most $2^{-\Theta(k)}$. Establishing such theorems is highly non-trivial and for one-way protocols, this was done by [Gav08, JKN07].

Following this framework, the work of [Gav08] gives examples of relational problems that are easy to solve with $\Theta(k \log n)$ EPR pairs but difficult with only $O(k)$ EPR pairs. One drawback of this approach is that the task $F^{(k)}$ has many output bits, regardless of whether F is a partial function or a relational problem. To get separations for functions with single-bit outputs, we need to modify this approach. We ask the players to solve the XOR of k independent instances of F. Here, the players receive k pairs of inputs (x_i, y_i) and they need to compute $\prod_{i\in[k]} F(x_i, y_i)$. We denote this problem by $F^{(\oplus k)}$. We will show that there is no small-cost protocol solving $F^{(\oplus k)}$ using only $O(k)$ qubits of entanglement. To do this, we assume by contradiction that there exists such a protocol.

Step 1: Remove entanglement. We produce a small-cost protocol for $F^{(\oplus k)}$ that uses no entanglement and has success probability at least $1/2+2^{-\Theta(k)}$, i.e., the advantage is at least $2^{-\Theta(k)}$.

Step 2: XOR Lemmas. We establish an XOR lemma for protocols without entanglement. We show that if for cost-t protocols, the probability of solving one instance is at most $2/3$, then for cost-$o(tk)$ protocols, the probability of solving the XOR of k independent instances is at most $1/2 + 2^{-\Theta(k)}$, i.e., the advantage is at most $2^{-\Theta(k)}$.

Together, this would establish the desired result. We now discuss some of the difficulties in executing these steps and present our solutions. We first present the details of step 2 and then step 1.

Details of Step 2. One difficulty with step 2 is that XOR lemmas are stronger than direct product theorems and are thus harder to establish. In this work, we present XOR lemmas that are tailored for particular functions. The functions we will be interested in are the Forrelation problem and the Boolean Hidden Matching problem. For the former problem, XOR lemmas for $\mathsf{R2}$ protocols were established in [GRZ21]. For the Boolean Hidden Matching problem, we show an XOR lemma for the $\mathsf{R1}$ and $\mathsf{Q}\|^{\text{pub}}$ models (Lemma 1.4 and Lemma 1.3). We now describe this part in more detail.

Let F be the Boolean Hidden Matching problem. One central ingredient in our XOR lemmas for F is the construction of hard distributions. The result of [GKK+07] shows hard distributions $\mathcal{Y}$ and $\mathcal{N}$ on the YES and NO instances of F respectively, such that no small-cost protocol can distinguish these distributions with $1/3$ advantage. We produce hard distributions $\mu_{-1}^{(k)}$ and $\mu_1^{(k)}$ for the $F^{(\oplus k)}$ problem such that no small-cost protocol can distinguish them with advantage $2^{-\Theta(k)}$. To get bounds of the form $2^{-\Theta(k)}$, it turns out that our distributions need to agree on moments of size at most $\Theta(k)$. Motivated by this, we define the following distributions.

$$\mu_1^{(k)} := \frac{1}{2^{k-1}} \sum_{\substack{K\subseteq[k] \\ |K| \text{ is even}}} \mathcal{Y}_K \mathcal{N}_{\overline{K}} \quad \text{and} \quad \mu_{-1}^{(k)} := \frac{1}{2^{k-1}} \sum_{\substack{K\subseteq[k] \\ |K| \text{ is odd}}} \mathcal{Y}_K \mathcal{N}_{\overline{K}},$$

Here, $\mathcal{Y}_K \mathcal{N}_{\overline{K}}$ is a product of k independent distributions, where the i-th distribution is $\mathcal{Y}$ if $i \in K$ and is $\mathcal{N}$. We show that the distributions $\mu_1^{(k)}$ and $\mu_{-1}^{(k)}$ are indeed distributions on the NO and YES instances respectively of the $F^{(\oplus k)}$ problem, furthermore, they agree on all moments of size at most k. To complete the argument, we need to show that small-cost protocols cannot distinguish these distributions with more than $2^{-\Theta(k)}$ advantage. This is fairly technical and involves the use of Fourier analysis. For $\mathsf{R1}$ protocols, the $k = 1$ version of this was proved in [GKK+07]. Their work in particular makes use of the level-k inequality. We build on their work for $\mathsf{R1}$ protocols and prove the desired XOR lemma for larger k (Lemma 1.4). For $\mathsf{Q}\|^{\text{pub}}$ protocols, we are not aware of any works that study the communication complexity of F or that analyze the Fourier spectrum of such protocols, which is a contribution in this paper. In particular we prove Fourier growth bounds for $\mathsf{Q}\|^{\text{pub}}$ protocols (Lemma 1.2) as well as an XOR lemma for the Boolean Hidden Matching problem (Lemma 1.3). For these, we make use of a matrix version of the level-k inequality [BRW08].

Details of Step 1. One difficulty with step 1 is that the trick of replacing an entangled state by the maximally mixed state no longer works. It is possible for a protocol to be correct when using a particular entangled state, but wrong for every orthogonal state. In this case, executing the protocol on the maximally mixed state would bias the output towards the wrong answer. Thus, carrying out step 1 is non-trivial and in particular, difficult to do for $\mathsf{R2}^*$ protocols. We take an alternate approach for $\mathsf{R2}^*$ protocols to sidestep this difficulty, which we will describe later. We are able to carry out step 1 for $\mathsf{R1}^*$ and $\mathsf{Q}\|^*$ protocols (Lemma 5.2 and Lemma 5.1). Given a cost c protocol for a function in the $\mathsf{R1}^*$ model or $\mathsf{Q}\|^*$ model using at most $2d$ qubits of entanglement, we produce a cost $c + O(d)$ protocol in the $\mathsf{R1}$ or $\mathsf{Q}\|$ model[3] respectively; these protocols use no entanglement and have advantage $2^{-\Theta(d)}$. We now give an illustrative example of this simulation.

Consider a simple $\mathsf{Q}\|^*$ protocol where the entangled state consists of d EPR pairs and Alice and Bob apply a unitary operator to their part of the entangled state and send all their qubits to Charlie. If Alice's and Bob's unitary operators map $|i\rangle$ to $|u_i(x)\rangle$ and $|v_i(y)\rangle$ respectively, then the state received by the referee is the pure state $\sum_{i\in\{-1,1\}^d} |u_i(x)\rangle\, |v_i(y)\rangle$ (ignoring the normalization). We now construct a $\mathsf{Q}\|$ protocol that produces the same state with probability $2^{-\Theta(d)}$, furthermore, Charlie is able to detect when this state was successfully produced. We have Alice and Bob send the pure states $\sum_i |i\rangle\, |u_i(x)\rangle$ and $\sum_j |j\rangle\, |v_j(y)\rangle$ respectively to Charlie. Charlie first projects onto states such that $i = j$ and obtains the pure state $\sum_{i\in\{-1,1\}^d} |i, i\rangle\, |u_i(x)\rangle\, |v_i(y)\rangle$ with probability 2^{-d}. He then applies Hadamard on the first $2d$ qubits and measures. He obtains the outcome $|0^{2d}\rangle$ with probability 2^{-2d} in which the resulting state is the pure state $\sum_i |u_i(x)\rangle\, |v_i(y)\rangle$ as in the original $\mathsf{Q}\|^*$ protocol. We use similar ideas to remove entanglement from arbitrary $\mathsf{Q}\|^*$ protocols. To remove entanglement from an $\mathsf{R1}^*$ protocol, we need to take a different approach which involves Alice sending Bob a random coordinate of a certain density matrix. We omit the details.

Alternate Approach to Step 1 for $\mathsf{R2}^*$ Protocols. We now present an alternative to step 1 for $\mathsf{R2}^*$ protocols. The idea is to prove Fourier growth bounds for $\mathsf{R2}^*$ protocols. The results of [GRZ21] imply that for protocols whose level-$2k$ Fourier growth is at most α, their advantage in solving the $\oplus^k$-Forrelation problem is at most $\alpha \cdot n^{-k/2} + o(n^{-k/2})$. We directly establish a Fourier growth bound on $\mathsf{R2}^*$ protocols. In particular, we show that for $\mathsf{R2}^*$ protocols of communication cost c that use d qubits of entanglement, their level-ℓ Fourier growth is at most $\text{poly}(2^d) \cdot O_\ell(c^\ell)$ (Lemma 1.1). Choosing $\ell = 2k$, $d = \Theta(k)$ and $c = \Theta(n^{1/4})$ for appropriate constants, we have that the advantage is at most $\text{poly}(2^d) \cdot O_\ell(c^\ell) \cdot n^{-k/2} \ll 1$. This would complete the proof. We now describe how we prove the Fourier growth bound on $\mathsf{R2}^*$ protocols (Lemma 1.1).

(1) We first show that if the players share a $2d$-qubit entangled state, then we can decompose the state into a small linear combination of $\text{poly}(2^d)$ many *two*-qubit quantum states that are either unentangled, or locally equivalent to $|\Phi^+\rangle\langle\Phi^+|$, the EPR state. (By locally equivalent we mean that the players can transform one state into the other using local unitaries and no communication.) This is formalized in Claim 3.2. This gives us the pre-factor of $\text{poly}(2^d)$ in Lemma 1.1.

(2) We analyze protocols where Alice and Bob share the EPR state and bound the Fourier growth of such protocols. Observe that if they share an unentangled state, the protocol is essentially an $\mathsf{R2}$ protocol and the work of [GRT22] showed Fourier growth for such protocols. We strengthen this result by proving similar Fourier growth for $\mathsf{R2}^*$ protocols where Alice and Bob share the EPR state. To do this, we study the structure of such protocols. We first show that the expected output

[3] We use $\mathsf{Q}\|$ to denote the private-coin version of the $\mathsf{Q}\|^{\text{pub}}$ model.

of any R2* protocol of cost c where Alice and Bob share the EPR state can be written as

$$C(x,y) = \sum_{z\in\{-1,1\}^c} \alpha_z \cdot \mathrm{Tr}((E_z(x)\otimes F_z(y))\cdot\rho).$$

where $E_z(x)$ and $F_z(y)$ are positive semidefinite matrices, $\sum_{z\in\{-1,1\}^c} E_z(x)\otimes F_z(y) = \mathbb{I}$, $\alpha_z \in \{-1,1\}$, and $\rho = |\Phi^+\rangle\langle\Phi^+| \otimes |0^{2m}\rangle\langle 0^{2m}|_{AB}$ for some $m\in\mathbb{N}$ that is possibly large (Claim 3.7). We give some intuition on this expression. The qubits $|0^m\rangle\langle 0^m|_A$ and $|0^m\rangle\langle 0^m|_B$ in ρ correspond to Alice and Bob's private memory respectively and ρ captures the initial state of all the qubits in the system. The matrices $E_z(x)\otimes F_z(y)$ arise out of Alice's and Bob's sequence of quantum operations (i.e., POVMs) in the R2* protocol and the quantity $\mathrm{Tr}(E_z(x)\otimes F_z(y)\cdot\rho)$ captures the probability of the transcript being $z\in\{-1,1\}^c$. The number $\alpha_z\in\{-1,1\}$ is 1 if and only if the transcript z results in the players outputting 1.

We now write out the Fourier expansion of the XOR fiber $H(z) = \mathbb{E}_{x\sim\{-1,1\}^n}[C(x,x+z)]$. Using the convolution property of Fourier coefficients, we can express the Fourier coefficients of $H(z)$ in terms of the Fourier coefficients of $E_z(x), F_z(y)$. In particular, we get

$$\sum_{|S|=\ell} |\widehat{H}(S)| = \sum_{|S|=\ell} \left| \sum_{z\in\{-1,1\}^c} \alpha_z \cdot \mathrm{Tr}\left(\left(\widehat{E_z}(S)\otimes\widehat{F_z}(S)\right)\cdot\rho\right)\right|.$$

We now use the fact that ρ has exactly four non-zero entries. This zeroes out all but four coordinates of $\widehat{E}_z(S)\otimes\widehat{F}_z(S)$ in the above expression. At this point, we use the level-k inequality by Lee [Lee19] to bound each of the coordinates separately in terms of the entries of $E_z(x)$ and $F_z(y)$. Using the fact that $\{E_z(x)\otimes F_z(y)\}_z$ forms a POVM, we can bound the coordinates of these matrices and combine them with the bounds we obtain from the level-k inequality in a nice way. Putting everything together involves some calculation and is done in Section 3.2.

Acknowledgements. We thank Vojtech Havlicek, Ran Raz and Penghui Yao for many discussions during this project. We also thank Ran Raz for feedback on the presentation.

2 Preliminaries

Sets. For $n\in\mathbb{N}$, let $[n]=\{1,\dots,n\}$. We use $\mathbb{1}$ to denote the indicator function, i.e., for a predicate E, $\mathbb{1}[E]$ is 1 if E is satisfied and 0 otherwise. For a subset $S\subseteq[n]$, we use $\overline{S} := [n]\setminus S$ to denote the complement of S. We denote the $n\times n$ identity matrix by $\mathbb{I}_n$, and we omit the subscript if it is implicit.

Big O Notation. For simplicity in notation, for every $f,g:\mathbb{R}_{\geq 0}\to\mathbb{R}_{\geq 0}$ and $\ell\in\mathbb{N}$, we say that $g=O_\ell(f)$ if for some constant $c>0$, we have $g=O\left(f\cdot 2^{c\cdot\ell}\right)$. We say that $f=\tilde{O}(g)$ (respectively $f=\tilde{\Omega}(g)$) if for some constant $c>0$, we have $f=O(g\cdot\log^c(g))$ (respectively $f=\Omega(g\cdot\log^{-c}(g))$). We say that $f=\tilde{\Theta}(g)$ if $f=\tilde{O}(g)$ and $f=\tilde{\Omega}(g)$.

Probability Distributions. Let Σ be an alphabet and $\mathcal{D}$ be a probability distribution over Σ. We use $x\sim\mathcal{D}$ to denote x sampled according to $\mathcal{D}$. We use $\mathrm{supp}(\mathcal{D})$ to denote the support of the distribution $\mathcal{D}$. We use $x\sim\Sigma$ to denote a uniformly random sample from Σ. For a function $G:\Sigma\to\mathbb{R}^n$, we use $\mathbb{E}_{x\sim\mathcal{D}}[G(x)]$ to denote the expected value of G when the inputs are drawn according to $\mathcal{D}$. Let $k\in\mathbb{N}$, $S\subseteq[k]$ and $\mathcal{D},\mathcal{D}'$ be two distributions over Σ. We use $\mathcal{D}_S\mathcal{D}'_{\overline{S}}$ to denote the distribution over Σ^k which is a product of k independent distributions over Σ, where the ith distribution is $\mathcal{D}$ if $i \in S$ and $\mathcal{D}'$ if $i \notin S$ for all $i \in [k]$. For distributions, $\mathcal{D}, \mathcal{D}'$, define the total variation distance as $\|\mathcal{D} - \mathcal{D}'\|_1 := \sum_i |\mathcal{D}(i) - \mathcal{D}'(i)|$.

Norms. Let $k \in \mathbb{N}$. For a vector $v \in \mathbb{R}^n$, we use $\|v\|_k := \left(\sum_{i\in[n]} |v_i|^k\right)^{1/k}$ to denote the ℓ_k-norm of v. For any matrix $M \in \mathbb{R}^{n\times n}$, we use $|M|$ to denote $\sqrt{MM^\dagger}$ and we denote by $\|M\|_1$ the trace norm of M, that is $\|M\|_1 := \mathrm{Tr}(\sqrt{MM^\dagger}) = \mathrm{Tr}(|M|)$. We use $\|M\|_{\mathrm{op}} := \max_{\|v\|_2=1}(v^T M v)$ to denote the operator norm of M.

2.1 Quantum information

Quantum States. Let $d \in \mathbb{N}$ and let $\mathcal{H}$ be a Hilbert space of dimension 2^d. This is a vector space defined by the $\mathbb{R}$-span of the orthonormal basis $\{|x\rangle : x \in \{0,1\}^d\}$. We also identify this basis with $\{|i\rangle : i \in [2^d]\}$ using the lexicographic ordering as the correspondence. We use $|0^d\rangle$ to denote the vector $|0,\ldots,0\rangle$ with d zeroes. Let $\mathcal{P}(\mathcal{H})$ be the set of positive semidefinite matrices in $\mathbb{R}^{2^d\times 2^d}$. Let $\mathcal{S}(\mathcal{H})$ be the set of density operators on $\mathcal{H}$, that is, matrices in $\mathcal{P}(\mathcal{H})$ with trace 1. We typically use ρ and σ to refer to elements of $\mathcal{S}(\mathcal{H})$. The state of a quantum system on d qubits is described by a density operator $\rho \in \mathcal{S}(\mathcal{H})$. For states that are shared between Alice and Bob, we use the subscript A and B on qubits to denote whether Alice or Bob own those qubits.

Quantum State Evolution. Let $\mathcal{H}, \mathcal{H}'$ be Hilbert spaces. The evolution of a quantum state is described by a map $E : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H}')$ which is CPTP (i.e., completely positive and trace preserving). We use the notation $E(\rho)$ to denote the image of ρ under E. In particular, we will be interested in measurement operators. Any quantum measurement with k outcomes is specified by k matrices $M_1, \ldots, M_k \in \mathcal{P}(\mathcal{H})$ such that $\sum_{i\in[k]} M_i^\dagger M_i = \mathbb{I}$. The probability of getting outcome $i \in [k]$ is precisely $\mathrm{Tr}(M_i \rho M_i^\dagger)$ and the post measurement state upon obtaining the outcome i is $\frac{M_i\rho M_i^\dagger}{\mathrm{Tr}(M_i\rho M_i^\dagger)}$. We have a correspondence between $\{0,1\}$ and $\{1,-1\}$ defined by $0 \to 1, 1 \to -1$, hence, we will sometimes refer to measurement outcomes "1" and "0" as "-1" and "1" respectively.

Distance between States. Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ be density operators. We define the trace distance between ρ and σ to be $\frac{\|\rho-\sigma\|_1}{2}$. We will use the following standard facts about the trace distance. Firstly, the trace distance satisfies triangle inequality. Secondly, the trace distance between ρ and σ is equal to the maximum probability with which these states can be distinguished using any single projective measurement. Thirdly, the following inequality holds as a consequence of the Von-Neumann Inequality.

Fact 2.1. *For any matrices $M, \rho \in \mathbb{R}^{n\times n}$, we have* $\mathrm{Tr}(M\rho) \le \|M\|_{\mathrm{op}} \cdot \|\rho\|_1$.

2.2 Communication Complexity

The goal in communication complexity is for Alice and Bob to compute a function $F : \mathcal{X} \times \mathcal{Y} \to \{-1, 1, \star\}$. We interpret -1 as a YES and 1 as a NO. We say F is a *total* function if $F(x,y) \in \{-1,1\}$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, otherwise F is a *partial* function. In this paper we will be mostly concerned with partial functions and denote $\mathsf{dom}(F) = F^{-1}(\{-1,1\})$. Here Alice receives an input $x \in \mathcal{X}$ (unknown to Bob) and Bob receives an input $y \in \mathcal{Y}$ (unknown to Alice) promised that $(x,y) \in \mathsf{dom}(F)$ and their goal is to compute $F(x,y)$ with high probability, i.e., probability at least $2/3$. More formally, for any protocol $\mathcal{P}$, we let $\mathrm{cost}(\mathcal{P}, x, y)$ be the communication cost of $\mathcal{P}$ when Alice and Bob are given x, y as inputs. We say that $\mathcal{P}$ computes F, if, for every $(x,y) \in \mathsf{dom}(F)$, the output of the protocol is $F(x,y)$ with probability at least $2/3$ (where the probability is taken over the randomness/measurements of the protocol). The communication complexity of F is defined as

$$\min_{\mathcal{P} \text{ computes } F} \max_{(x,y)\in \mathsf{dom}(F)} \mathrm{cost}(\mathcal{P}, x, y).$$

The messages sent are referred to as the transcript of the protocol. We discuss a few models of communication of interest to us.

Simultaneous Message Passing Model. This is a general model of communication called the *simultaneous message passing* (SMP) model. In this model, Alice and Bob each send a single (possibly quantum or randomized) message to a referee Charlie. The goal is for Charlie to output $F(x,y)$ with high probability, i.e., at least $2/3$ probability. We measure the cost of a communication protocol by the total number of bits (or qubits) received by Charlie. There are many types of simultaneous protocols.

Quantum versus Classical protocols. We use $\mathsf{R}\|$ to denote the SMP model where the players can only send classical messages to Charlie. We use $\mathsf{Q}\|$ to denote the SMP model where the players can send a quantum message to Charlie.

Public-coin versus Private-coin Protocols. If we allow the players to use public coins, we use the superscript "pub". For instance, $\mathsf{Q}\|^{\text{pub}}$ denotes the public-coin quantum SMP model and $\mathsf{Q}\|$ denotes the private-coin quantum SMP model. Unless otherwise specified, all protocols are private coin protocols.

One-Way Model. In this model, Alice sends a single message to Bob, who should output $F(x,y)$ with probability at least $2/3$. The cost of the protocol is the size of message Alice sends. By a classical result of Newman [New91], we can assume that all one-way protocols are private-coin protocols with an $O(\log n)$ additive overhead in the communication complexity. We use $\mathsf{R1}$ to denote the one-way model of communication where Alice sends a classical message to Bob.

Two-Way Model. Here Alice and Bob are allowed to exchange messages, and Alice should finally output $F(x,y)$ with probability at least $2/3$. The cost of the protocol is the total size of the transcript. As before, by a result of Newman [New91], we can assume that all two-way protocols are private-coin protocols with an $O(\log n)$ additive overhead in the communication complexity. We use $\mathsf{R2}$ to denote the model of two-way communication where Alice and Bob exchange classical messages.

We now discuss protocols where Alice and Bob can share an entangled state that is independent of their inputs. One important type of entangled state is the EPR pair: This is the state $|\Phi^+\rangle\langle\Phi^+|$ where $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B)$. Here, the subscript on the qubit denotes which player owns the qubit. The upper bound in all our theorems will be established using quantum protocols where the players share a certain number of EPR pairs. We typically specify the dimension of the state shared by Alice and Bob.

Protocols with entanglement. We use $\mathsf{R}\|^*$ to denote the simultaneous model of communication where Alice and Bob share an entangled state and send classical messages to the referee. The model $\mathsf{Q}\|^*$ is similarly defined, but Alice and Bob can send quantum messages to the referee. We use $\mathsf{R1}^*$ to denote the one-way model where Alice and Bob share entanglement and Alice sends a classical message to Bob. We use $\mathsf{R2}^*$ to denote the two-way model where Alice and Bob share entanglement and the messages are classical. In this model, by teleportation , the players can exchange a limited number of qubits. Conversely, if the players can exchange a limited number of qubits, then Alice can create EPR pairs by herself and send the corresponding qubits to Bob.

For ease of readability, we summarize all the communication models in the tables below.

	Private Coins	Public Coins	Entanglement
Classical Messages	$\mathsf{R}\|$	$\mathsf{R}\|^{\text{pub}}$	$\mathsf{R}\|^{*}$
Quantum Messages	$\mathsf{Q}\|$	$\mathsf{Q}\|^{\text{pub}}$	$\mathsf{Q}\|^{*}$

Table 1: Table of Simultaneous Communication Models

	One-way Private Coins ($\equiv$ Public Coins)	Two-way Private Coins ($\equiv$ Public Coins)
Classical Messages	$\mathsf{R1}$	$\mathsf{R2}$
Classical Messages & Entanglement	$\mathsf{R1}^{*}$	$\mathsf{R2}^{*}$

Table 2: Table of One-Way & Two-Way Communication Models

2.3 XOR-Fibers of Communication Protocols

Definition 2.2. *Given a communication protocol $C : \{-1,1\}^n \times \{-1,1\}^n \to [-1,1]$, the XOR-fiber of C is a function $H : \{-1,1\}^n \to [-1,1]$ defined at $z \in \{-1,1\}^n$ by $H(z) = \mathbb{E}_{x\sim\{-1,1\}^n}[C(x, x\odot z)]$, where $\odot$ denotes point-wise product.*

The communication complexity of XOR functions are well-studied and have connections to the log-rank conjecture, parity decision trees, lifting theorems and separations between quantum and classical communication complexity [MO10, HHL18, TWXZ13, SZ08, Zha14]. XOR-fibers of communication protocols naturally arise in the study of communication complexity of XOR functions. Although we are not aware of any published works defining the term "XOR-fiber", this concept has been studied in many works, most notably [GRT22] and [Raz95].

2.4 Fourier analysis

Fourier Analysis on the Boolean Hypercube. We discuss some of the basics of Fourier analysis. Let $f : \{-1,1\}^n \to \mathbb{R}$ be a function. The Fourier decomposition of f is

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

where $\chi_S(x) = \prod_{i\in S} x_i$. The *Fourier coefficients* of f are defined as $\widehat{f}(S) = \mathbb{E}_{x\sim\{-1,1\}^n}[f(x)\cdot\chi_S(x)]$. For $\ell \in \mathbb{N}$, the level-ℓ Fourier mass of f is denoted by $L_{1,\ell}(f)$ and is defined as follows.

$$L_{1,\ell}(f) = \sum_{|S|=\ell} \left|\widehat{f}(S)\right|$$

By Fourier growth bounds, we typically mean upper bounds on $L_{1,\ell}(f)$. We will need the following technical lemma, often called the level-k inequality.

Lemma 2.3 ([Lee19, Lemma 10])**.** *Let $f : \{-1,1\}^n \to [-1,1]$ be a function with $\mathbb{E}_x[|f(x)|] = \alpha$. Then for every $\ell \in \mathbb{N}$,*

$$\sum_{|S|=\ell} \widehat{f}(S)^2 \le 4\alpha^2 \cdot (2e \cdot \ln(e/\alpha^{1/\ell}))^{\ell}.$$

Although [Lee19, Lemma 10] is only stated for functions with range $[0, 1]$, the same proof also applies for functions with range $[-1, 1]$. We remark that the bound often invoked [O'D14, GKK+07] is with the upper bound of $O(\alpha^2 \cdot \ln^\ell(1/\alpha))$ (i.e., without the $1/\ell$ exponent on α) which only holds for $\ell \leq 2\ln(1/\alpha)$. However this improved upper bound, proven recently in [Lee19], holds for all $\ell \in \mathbb{N}$. This makes our proofs much simpler and saves some logarithmic factors.

Fourier analysis for Matrix-Valued Functions. The Fourier coefficients of a matrix-valued function $f : \{-1, 1\}^n \to \mathbb{R}^{m\times m}$ are defined by

$$\widehat{f}(S) := \mathop{\mathbb{E}}_{x\sim\{-1,1\}^n} [f(x) \cdot \chi_S(x)]$$

for all $S \subseteq [n]$. We also use a matrix version of the level-k inequality.

Lemma 2.4 ([BRW08, Theorem 7])**.** *Let $\mathcal{H}$ be a Hilbert space of dimension 2^c and let $f : \{-1, 1\}^n \to \mathcal{S}(\mathcal{H})$ be a density-matrix valued function. Then, for any $\ell \in \mathbb{N}$ such that $\ell \leq 2\ln(2)c$,*

$$\sum_{|S|=\ell} \mathrm{Tr}^2 \left(|\widehat{\rho}_S| \right) \leq \left((2e\ln 2) \cdot c/\ell\right)^\ell .$$

Let $\mathcal{H}'$ be a Hilbert space that contains $\mathcal{H}$ of dimension $c' \geq c$. We can view f as a function $f : \{-1, 1\}^n \to \mathcal{S}(\mathcal{H}')$ by simply appending zeroes to the output matrix. Given any $\ell \in \mathbb{N}$ that is possibly larger than c, set $c' := c + \lceil \frac{\ell}{2\ln 2} \rceil$ so that $\ell \leq 2\ln(2) \cdot c'$. Using this setting of parameters and Lemma 2.4, we have the following corollary.

Corollary 2.5. *Let $\mathcal{H}$ be a Hilbert space of dimension 2^c and let $f : \{-1, 1\}^n \to \mathcal{S}(\mathcal{H})$ be a density-matrix valued function. Then, for any $\ell \in \mathbb{N}$,*

$$\sum_{|S|=\ell} \mathrm{Tr}^2 \left(|\widehat{\rho}_S| \right) \leq O_\ell \left((c/\ell)^\ell \right) + O_\ell(1).$$

3 Proof of Theorem 1.1

In this section, we prove Lemma 1.1, which we restate for convenience.

Lemma 1.1. *Let $C : \{-1, 1\}^n \times \{-1, 1\}^n \to [-1, 1]$ be an R2* protocol of cost c where Alice and Bob share an entangled state on at most $2d$ qubits for some parameter $d \in \mathbb{N}$. Let H be the XOR-fiber of C as in Definition 2.2. Then, for all $\ell \in \mathbb{N}$, we have*

$$L_{1,\ell}(H) \triangleq \sum_{|S|=\ell} \left| \widehat{H}(S) \right| \leq 2^{5d} \cdot O_\ell(c^\ell).$$

The proof of Theorem 1.1 follows from Lemma 1.1 and the techniques of [GRZ21]. The quantum upper bound is presented in [GRZ21, Theorem 3.8]. For the lower bound, let $\mathcal{C}$ be the set of R2* protocols of cost at most c using at most d qubits of entanglement. Let $\mathcal{H}$ be the set of XOR fibers of protocols in $\mathcal{C}$. Applying [GRZ21, Theorem 3.1] to $\mathcal{H}$, it follows that the maximum advantage that protocols in $\mathcal{C}$ have in solving the $\oplus^k$-Forrelation problem is at most $O\left(L_{1,2k}(\mathcal{H}) \cdot n^{-k/2}\right) + o(n^{-k/2})$. By Lemma 1.1, this is at most $O\left(2^{5d} \cdot c^{2k} \cdot n^{-k/2}\right)$. Since $d \leq k$, setting $c = \tau \cdot n^{1/4}$ for a small constant $\tau > 0$ implies that this is at most $1/3$. The details of this calculation are deferred to Appendix A.

The rest of this section will be devoted to the proof of Lemma 1.1. As described in the proof overview, this will consist of two parts. First, in Section 3.1, we show how to decompose entangled states as a linear combination "simple" states and next, in Section 3.2, we prove a Fourier growth bound on protocols that use "simple" states. We put these together in Section 3.3 to complete the proof of Lemma 1.1.

3.1 Decomposing an Arbitrary State as a Linear Combination of Simple States

Definition 3.1. *Let ρ, σ be (possibly entangled) states in $\mathcal{S}(\mathcal{H}_A \otimes H_B)$. We say that ρ is locally equivalent to σ if there exist unitaries U_A on $\mathcal{H}_A$ and V_B on $\mathcal{H}_B$ such that $(U_A \otimes V_B)\rho(U_A \otimes V_B)^\dagger = \sigma$. If $\sigma = |+\rangle\langle+|$ where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B)$ corresponds to the EPR state, we say that ρ is locally equivalent to the EPR state and if $\sigma = |0\rangle\langle 0|_A \otimes |0\rangle\langle 0|_B$, we say that ρ is locally equivalent to the zero state.*

We refer to the zero state and EPR state as simple. The main result of this section is as follows.

Claim 3.2. *Let $d \in \mathbb{N}$. Let $\mathcal{H}_A, \mathcal{H}_B$ be 2^d-dimensional Hilbert spaces. Given a (possibly entangled) state ρ in $\mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we can express it as*

$$\rho = \sum_{i=1}^{2^{4d}} \alpha_i \rho_i$$

where each $|\alpha_i| \leq 2^d$ and $\rho_i \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is locally equivalent to the zero state or the EPR state.

This decomposition is useful as it simplifies the study of communication protocols with arbitrary entangled states. In particular, we have the following.

Definition 3.3. *We say that two communication protocols are* equivalent *if on the same inputs, they produce the same distribution on transcripts.*

Fact 3.4. *Let ρ and σ be two locally equivalent states. Any communication protocol where Alice and Bob use ρ as the entangled state can be transformed into an equivalent communication protocol where Alice and Bob use σ as the entangled state.*

Proof. Let U_A and V_B be unitary operators on $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively so that $(U_A \otimes V_B)\rho(U_A \otimes V_B)^\dagger = \sigma$. We modify the protocol by first having Alice and Bob apply U_A and V_B respectively to their parts of the entangled state, then apply U_A^{-1} and V_B^{-1} and then continue as per the original protocol. Observe that this transformation does not change the distribution on the transcripts. The first step of the new protocol transforms ρ into σ. Thus, we may think of the new protocol as an equivalent communication protocol where the initial entangled state is σ. □

We now turn to proving Claim 3.7. We will use the following fact.

Fact 3.5. *Let $|\phi\rangle \in \mathcal{H}_A$ and $|\psi\rangle \in \mathcal{H}_B$ be unit vectors. Then, the state $\rho := |\phi\rangle\langle\phi|_A \otimes |\psi\rangle\langle\psi|_B$ is locally equivalent to the zero state.*

Proof. Consider unitaries U_A and V_B such that $U_A |\phi\rangle = |0\rangle$ and $V_B |\psi\rangle = |0\rangle$. Observe that $(U_A \otimes V_B)\rho(U_A \otimes V_B)^\dagger = |0\rangle\langle 0|_A \otimes |0\rangle\langle 0|_B$ and hence ρ is locally equivalent to the zero state. □

We now complete the proof of Claim 3.2.

Proof of Claim 3.2. Let $i, j \in [2^{2d}]$ be such that $i \neq j$. Define $\rho^{(i,j)} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $\alpha_{i,j} \in \mathbb{R}$ by

$$\rho^{(i,j)} := \tfrac{1}{2}\left(|i\rangle + |j\rangle\right)\left(\langle i| + \langle j|\right) \quad \text{and} \quad \alpha_{i,j} = \rho_{i,j},$$

where $\rho_{i,j}$ is the (i,j)th entry of ρ. Define $\rho^{(i,i)} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $\alpha_{i,i} \in \mathbb{R}$ by

$$\rho^{(i,i)} := |i\rangle\langle i| \quad \text{and} \quad \alpha_{i,i} = \rho_{i,i} - \sum_{j \in [2^{2d}], j \neq i} \rho_{i,j}.$$

Observe that $\sum_{i,j \in [2^{2d}]} \alpha_{i,j}\rho^{(i,j)} = \rho$. Furthermore, for $i \neq j \in [2^{2d}]$, we have $|\alpha_{i,j}| = |\rho_{i,j}| \leq 1$ and $|\alpha_{i,i}| \leq \sum_{j \in [2^{2d}]} |\rho_{i,j}| \leq \sqrt{2^{2d}} \cdot \sqrt{\sum_{j \in [2^{2d}]} \rho_{i,j}^2} \leq 2^d$ due to Cauchy-Schwarz and the fact that ρ is a quantum state. It suffices to argue that $\rho^{(i,j)}$ is locally equivalent to the zero state or the EPR state. Firstly, every $i \in [2^{2d}]$ can be uniquely identified with (a,b) where $a, b \in [2^d]$. Similarly, each $j \in [2^{2d}]$ can be uniquely identified with (p,q) where $p, q \in [2^d]$. Consider the following cases.

Case 1: Suppose $a = p$ and $b = q$ (or equivalently $i = j$), then $\rho^{(i,j)} = |a\rangle\langle a|_A \otimes |b\rangle\langle b|_B$. By Fact 3.5, $\rho^{(i,j)}$ is locally equivalent to the zero state.

Case 2: Suppose $a = p$ and $b \neq q$, then

$$\begin{aligned}\rho^{(i,j)} &\triangleq \tfrac{1}{2}\left(|a\rangle_A |b\rangle_B + |p\rangle_A |q\rangle_B\right)\left(\langle a|_A \langle b|_B + \langle p|_A \langle q|_B\right) \\ &= |a\rangle\langle a|_A \otimes \left(\tfrac{|b\rangle_B + |q\rangle_B}{\sqrt{2}}\right)\left(\tfrac{\langle b|_B + \langle q|_B}{\sqrt{2}}\right).\end{aligned}$$

Since $|b\rangle$ and $|q\rangle$ are orthogonal, $\frac{|b\rangle + |q\rangle}{\sqrt{2}}$ is a unit vector and by Fact 3.5, $\rho^{(i,j)}$ is locally equivalent to the zero state.

Case 3: Suppose $a \neq p$ and $b \neq q$. Let U_A be any unitary operator that maps $|a\rangle$ to $|0\rangle$ and $|p\rangle$ to $|1\rangle$. This is well defined since $a \neq p$. Similarly define V_B to be any unitary operator that maps $|b\rangle$ to $|0\rangle$ and $|q\rangle$ to $|1\rangle$. This is well defined since $b \neq q$. Observe that

$$(U_A \otimes V_B)\rho^{(i,j)}(U_A \otimes V_B)^\dagger = \tfrac{1}{2}(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B)(\langle 0|_A \langle 0|_B + \langle 1|_A \langle 1|_B).$$

Thus, $\rho^{(i,j)}$ is locally equivalent to the EPR state. This completes the proof of Claim 3.2. □

3.2 Fourier Growth Bounds on R2* protocols with the EPR State

We now prove a Fourier growth bound on R2* protocols where Alice and Bob share a single EPR pair. Note that such protocols can easily simulate protocols that share the zero state, since Alice and Bob can simply produce $|0\rangle\langle 0|_A$ and $|0\rangle\langle 0|_B$ without communication and ignore the EPR pair. The main technical contribution of this subsection is the following lemma.

Lemma 3.6. *Let $C : \{-1,1\}^n \times \{-1,1\}^n \to [-1,1]$ be a R2* protocol of cost c where the players share the EPR state. Let H be its XOR-fiber as in Definition 2.2. Then, for all $\ell \in \mathbb{N}$, we have*

$$L_{1,\ell}(H) \leq O_\ell(1) + O_\ell\left((c/\ell)^\ell\right).$$

The following claim helps understand the acceptance probability of R2* protocols.

Claim 3.7. *Let $C : \{-1,1\}^n \times \{-1,1\}^n \to [-1,1]$ be any R2* protocol of cost c where Alice and Bob share a state $\rho \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ where $\mathcal{H}_A$ and $\mathcal{H}_B$ are Hilbert spaces of dimension 2^d. Then, there exists $m \in \mathbb{N}$ and Hilbert spaces $\mathcal{H}'_A, \mathcal{H}'_B$ of dimension m such that the expected output of the protocol on inputs $x, y \in \{-1,1\}^n$ can be expressed as*

$$C(x,y) = \sum_{z \in \{-1,1\}^c} \mathrm{Tr}\left((E_z(x) \otimes F_z(y)) \cdot \rho'\right) \cdot (-1)^{\mathbb{1}[z \in A]},$$

where for all $x, y \in \{-1,1\}^n$ and $z \in \{-1,1\}^c$,

1. *$E_z(x)$ is a $2^{d+m} \times 2^{d+m}$ positive semidefinite matrix acting on $\mathcal{H}_A \otimes \mathcal{H}'_A$,*
2. *$F_z(y)$ is a $2^{d+m} \times 2^{d+m}$ positive semidefinite matrices acting on $\mathcal{H}_B \otimes \mathcal{H}'_B$,*
3. *$\sum_{z' \in \{-1,1\}^c} E_{z'}(x) \otimes F_{z'}(y) = \mathbb{I}$,*
4. *$\rho' = \rho \otimes |0^m\rangle\langle 0^m|_A \otimes |0^m\rangle\langle 0^m|_B$, and*
5. *$A \subseteq \{-1,1\}^c$ is some subset.*

We defer the proof of this claim to Appendix B. We now prove Lemma 3.6 using Claim 3.7.

Proof of Lemma 3.6 using Claim 3.7. We use $|\Phi^+\rangle\langle\Phi^+|$ to denote the single EPR state. We use Claim 3.7 which describes the structure of arbitrary R2* protocols of cost c. By Claim 3.7, the expected output of the protocol C on inputs $x, y \in \{-1,1\}^n$ can be expressed as

$$C(x,y) = \sum_{z \in \{-1,1\}^c} \mathrm{Tr}((E_z(x) \otimes F_z(y) \cdot \rho) \cdot (-1)^{\mathbb{1}[z \in A]}$$

where $A \subseteq \{-1,1\}^c$, and $E_z(x)$ and $F_z(y)$ are positive semidefinite operators such that

$$\sum_{z' \in \{-1,1\}^c} E_{z'}(x) \otimes F_{z'}(y) = \mathbb{I}, \text{ and } \rho = |\Phi^+\rangle\langle\Phi^+| \otimes |0^m\rangle\langle 0^m|_A \otimes |0^m\rangle\langle 0^m|_B \tag{1}$$

for some parameter $m \in \mathbb{N}$. Recall that by the definition of XOR-fiber, for all $w \in \{-1,1\}^n$, we have $H(w) = \mathbb{E}_{x \sim \{-1,1\}^n}[C(x, x \odot w)]$. Hence,

$$\begin{aligned}
&\widehat{H}(S) \\
&= \mathop{\mathbb{E}}_{w \sim \{-1,1\}^n}[H(w)\chi_S(w)] \\
&= \mathop{\mathbb{E}}_{w,x \sim \{-1,1\}^n}\left[\sum_{z \in \{-1,1\}^c} \mathrm{Tr}\left((E_z(x) \otimes F_z(x \odot w)) \cdot \rho\right) \cdot \chi_S(w) \cdot (-1)^{\mathbb{1}[z \in A]}\right] \\
&= \mathop{\mathbb{E}}_{w,x}\left[\sum_{z \in \{-1,1\}^c} \sum_{T,Q} \mathrm{Tr}\left(\left(\widehat{E_z}(T) \otimes \widehat{F_z}(Q)\right) \cdot \rho\right) \cdot \chi_{S+Q}(w) \cdot \chi_{T+Q}(x) \cdot (-1)^{\mathbb{1}[z \in A]}\right] \\
&= \sum_{z \in \{-1,1\}^c} \mathrm{Tr}\left(\left(\widehat{E_z}(S) \otimes \widehat{F_z}(S)\right) \cdot \rho\right) \cdot (-1)^{\mathbb{1}[z \in A]}.
\end{aligned}$$

Our goal is to upper bound

$$\begin{aligned} L_{1,\ell}(H) \triangleq \sum_{|S|=\ell} \left|\widehat{H}(S)\right| &= \sum_{|S|=\ell} \left| \sum_{z\in\{-1,1\}^c} \mathrm{Tr}\left(\left(\widehat{E_z}(S) \otimes \widehat{F_z}(S)\right)\cdot\rho\right)\cdot(-1)^{\mathbb{1}[z\in A]}\right| \\ &\leq \sum_{z\in\{-1,1\}^c} \sum_{|S|=\ell} \left|\mathrm{Tr}\left(\left(\widehat{E_z}(S) \otimes \widehat{F_z}(S)\right)\cdot\rho\right)\right|. \end{aligned} \tag{2}$$

By Eq. (1), the density matrix of the state ρ has exactly four non-zero coordinates. This zeroes all but four coordinates of $\widehat{E_z}(S)\otimes\widehat{F_z}(S)$ in the R.H.S. of Eq. (2). More precisely, we have

$$\mathrm{Tr}\left(\left(\widehat{E_z}(S) \otimes \widehat{F_z}(S)\right)\cdot\rho\right) = \frac{1}{2}\cdot \sum_{i,j\in\{1,2^m+1\}} \widehat{E_z}(S)[i,j]\cdot\widehat{F_z}(S)[i,j].$$

Substituting this in Eq. (2),we have

$$\begin{aligned} L_{1,\ell}(H) &\leq \sum_{z\in\{-1,1\}^c}\sum_{|S|=\ell}\left|\sum_{i,j\in\{1,2^m+1\}} \widehat{E_z}(S)[i,j]\cdot\widehat{F_z}(S)[i,j]\right| \\ &\leq \sum_{z\in\{-1,1\}^c}\sum_{i,j\in\{1,2^m+1\}}\sum_{|S|=\ell}\left|\widehat{E_z}(S)[i,j]\cdot\widehat{F_z}(S)[i,j]\right| \\ &\leq \sum_{z\in\{-1,1\}^c}\sum_{i,j\in\{1,2^m+1\}}\sqrt{\sum_{|S|=\ell}\widehat{E_z}(S)[i,j]^2}\cdot\sqrt{\sum_{|S|=\ell}\widehat{F_z}(S)[i,j]^2}. \end{aligned}$$

For $i\neq j\in\{1,2^m+1\}$, let $e_z[i,j] := \mathbb{E}_x\big[\,|E_z(x)[i,j]|\,\big]$ and $f_z[i,j] := \mathbb{E}_y\big[\,|F_z(y)[i,j]|\,\big]$ (where the expectation is over the uniform distribution on $\{-1,1\}^n$). Similarly, let $e_z[i,i] = \mathbb{E}_x\big[E_z(x)[i,i]\big]$ and $f_z[i,i] = \mathbb{E}_y\big[F_z(y)[i,i]\big]$. Since E_x and F_y are positive semidefinite, their diagonal entries are non-negative. Using the level-k inequality (Lemma 2.3) we get that for all $i,j\in\{1,2^m+1\}$,

$$\begin{aligned} \sum_{|S|=\ell}\widehat{E_z}(S)[i,j]^2 &\leq 4(\mathbb{E}_x\left[|E_z(x)[i,j]|\right])^2\cdot\left(2e\ln\left(e/(\mathbb{E}_x\left[|E_z(x)[i,j]|\right])^{1/\ell}\right)\right)^{\ell} \\ &= 4(e_z[i,j])^2\cdot\left(2e\ln\left(e/(e_z[i,j])^{1/\ell}\right)\right)^{\ell}. \end{aligned}$$

We can now upper bound $L_{1,\ell}(H)$ as follows:

$$\begin{aligned} &L_{1,\ell}(H) \\ &\leq 4\sum_{\substack{z\in\{-1,1\}^c\\ i,j\in\{1,2^m+1\}}} e_z[i,j]\cdot f_z[i,j]\cdot\left(2e\ln\left(\frac{e}{e_z[i,j]^{1/\ell}}\right)\right)^{\ell/2}\cdot\left(2e\ln\left(\frac{e}{f_z[i,j]^{1/\ell}}\right)\right)^{\ell/2} \\ &\leq 4\sum_{\substack{z\in\{-1,1\}^c\\ i,j\in\{1,2^m+1\}}} e_z[i,j]\cdot f_z[i,j]\cdot\left(2e\ln\left(\frac{e^2}{e_z[i,j]^{1/\ell}\cdot f_z[i,j]^{1/\ell}}\right)\right)^{\ell}. \end{aligned}$$

We now use the concavity of the function $h(\gamma) = \gamma\cdot\ln(e^2/\gamma^{1/\ell})^{\ell}$ for $\gamma\in[0,1]$.[4] Jensen's inequality implies that for any $p_z\in[0,1]$, we have that $\mathbb{E}_z[h(p_z)]\leq h(\mathbb{E}_z[p_z])$. Setting $p_z = e_z[i,j]\cdot f_z[i,j]$, we

[4] The concavity of this function is proved in [Lee19, Claim 16]. In more detail, observe that $h(\gamma) = \gamma\cdot\ln\left((e/\gamma^{1/(2\ell)})^2\right)^{\ell} = \gamma\cdot\ln\left(e/\gamma^{1/(2\ell)}\right)^{\ell}\cdot 2^{\ell}$ which by their notation, equals $2^{\ell}\cdot\phi_{2\ell}(\gamma)$. It is shown that for all positive ℓ, the function $\phi_{2\ell}(\gamma)$ is concave and increasing for $\gamma\in[0,1]$.

conclude that $L_{1,\ell}(H)$ is at most

$$2^{c+2} \sum_{i,j \in \{1,2^m+1\}} \left(\mathop{\mathbb{E}}_{z \in \{-1,1\}^c} [e_z[i,j] f_z[i,j]] \right) \left(2e \ln \left(\frac{e^2}{\left(\mathbb{E}_{z \in \{-1,1\}^c} [e_z[i,j] f_z[i,j]] \right)^{1/\ell}} \right) \right)^{\ell}$$

$$= 4 \sum_{i,j \in \{1,2^m+1\}} \left(\sum_{z \in \{-1,1\}^c} e_z[i,j] f_z[i,j] \right) \left(2e \ln \left(\frac{e^2 \cdot 2^{c/\ell}}{\left(\sum_{z \in \{-1,1\}^c} e_z[i,j] f_z[i,j] \right)^{1/\ell}} \right) \right)^{\ell} .$$

To simplify notation, for $i, j \in \{1, 2^m+1\}$, we define $\beta_{i,j} \in \mathbb{R}$ by $\beta_{i,j} := \sum_{z \in \{-1,1\}^c} e_z[i,j] \cdot f_z[i,j]$. With this notation, we have

$$\begin{aligned} L_{1,\ell}(H) &\leq 4 \sum_{i,j \in \{1,2^m+1\}} \beta_{i,j} \cdot \left(2e \ln \left(\frac{e^2 \cdot 2^{c/\ell}}{\beta_{i,j}^{1/\ell}} \right) \right)^{\ell} \\ &\leq 4 \sum_{i,j \in \{1,2^m+1\}} \beta_{i,j} \cdot \left(2e \ln \left(\frac{e^2}{\beta_{i,j}^{1/\ell}} \right) + \frac{2ec}{\ell} \right)^{\ell} \\ &\leq 4 \sum_{i,j \in \{1,2^m+1\}} \beta_{i,j} \cdot 2^{\ell} \cdot \left(2e \ln \left(\frac{e^2}{\beta_{i,j}^{1/\ell}} \right) \right)^{\ell} + 4 \sum_{i,j \in \{1,2^m+1\}} \beta_{i,j} \cdot 2^{\ell} \cdot \left(\frac{2ec}{\ell} \right)^{\ell} . \end{aligned}$$

For all $x, y \in \{-1,1\}^n$, the matrix $E_z(x) \otimes F_z(y)$ is positive semidefinite. Furthermore, for any positive semidefinite matrix M, we have $|M_{i,j}| \leq \frac{1}{2}(M_{i,i} + M_{j,j})$ for all $i \neq j$. This implies that for all $x, y \in \{-1,1\}^n$ and $i \neq j \in \{1, 2^m+1\}$,

$$|E_z(x)[i,j]| \cdot |F_z(y)[i,j]| \leq \tfrac{1}{2} \left(E_z(x)[i,i] \cdot F_z(y)[i,i] + E_z(x)[j,j] \cdot F_z(y)[j,j] \right)$$

Taking an expectation over $x, y \sim \{-1,1\}^n$ implies that for $i \neq j \in \{1, 2^m+1\}$,

$$e_z[i,j] \cdot f_z[i,j] \leq \tfrac{1}{2} \left(e_z[i,i] \cdot f_z[i,i] + e_z[j,j] \cdot f_z[j,j] \right).$$

By Eq. (1), since $\sum_{z \in \{-1,1\}^c} E_z(x) \otimes F_z(y) = \mathbb{I}$ for $x, y \in \{-1,1\}^n$, we have for all $i \neq j \in \{1, 2^m+1\}$

$$\beta_{i,j} \triangleq \sum_{z \in \{-1,1\}^c} e_z[i,j] \cdot f_z[i,j] \leq \sum_{z \in \{-1,1\}^c} \tfrac{1}{2} \left(e_z[i,i] \cdot f_z[i,i] + e_z[j,j] \cdot f_z[j,j] \right) = 1,$$

where in the final equality we used that $e_z[i,i] = \mathbb{E}_x \left[E_z(x)[i,i] \right]$. Thus, we have

$$L_{1,\ell}(H) \leq O_{\ell}(1) \cdot \sum_{i,j \in \{1,2^m+1\}} \beta_{i,j} \cdot \left(2e \ln \left(\frac{e^2}{\beta_{i,j}^{1/\ell}} \right) \right)^{\ell} + O_{\ell}\left((c/\ell)^{\ell} \right).$$

It follows from [Lee19, Claim 16] that the function $h(\gamma) = \gamma \ln(e^2/\gamma^{1/\ell})^{\ell}$ is increasing for $\gamma \in [0,1]$ and the value at $\gamma = 1$ is 2^{ℓ}. Thus, we have

$$\beta_{i,j} \cdot \ln \left(\frac{e^2}{\beta_{i,j}^{1/\ell}} \right)^{\ell} \leq 2^{\ell}.$$

Therefore,

$$L_{1,\ell}(H) \leq O_{\ell}(1) + O_{\ell}\big((c/\ell)^{\ell} \big),$$

proving Lemma 3.6.

3.3 Putting Things Together

We now prove Lemma 1.1 using Claim 3.2 and Lemma 3.6. Consider any interactive randomized communication protocol C_ρ of cost c that uses ρ as the entangled state, where ρ is an arbitrary state on $2d$ qubits for a parameter $d \in \mathbb{N}$. Consider the decomposition

$$\rho = \sum_{i=1}^{2^{4d}} \alpha_i \rho_i$$

as given by Claim 3.2. Let $C_\rho(x, y)$ denote the expected output of the protocol on inputs x, y as before. Observe that $C_\rho(x, y)$ is linear in ρ due to Claim 3.7. Thus, the expected output of the protocol can be expressed as

$$C_\rho(x, y) = \sum_{i=1}^{2^{4d}} \alpha_i C_{\rho_i}(x, y).$$

By Fact 3.4, we have $C_{\rho_i}(x, y) = C^{(i)}_{\sigma_i}(x, y)$ where σ_i is either the zero state or the EPR state and $C^{(i)}$ is some $\mathsf{R2}^*$ protocol that is equivalent to C and uses σ_i as the shared state. Therefore, we have for all $z \in \{-1, 1\}^n$,

$$H(z) = \sum_{i=1}^{2^{4d}} \alpha_i H^{(i)}_{\sigma_i}(z)$$

where $H^{(i)}_{\sigma_i}$ is the XOR-fiber of $C^{(i)}_{\sigma_i}$. In particular,

$$L_{1,\ell}(H) \le \sum_{i=1}^{2^{4d}} |\alpha_i| \sum_{|S|=\ell} \left| \widehat{H^{(i)}_{\sigma_i}}(S) \right| \le 2^{5d} \cdot \max_{i \in [2^{4d}]} \left(L_{1,\ell}\left(H^{(i)}_{\sigma_i} \right) \right).$$

Here, we use the fact that $|\alpha_i| \le 2^d$. Since each σ_i is either the zero state or the EPR state and $C^{(i)}$ has cost at most c, we can apply Lemma 3.6 to conclude that

$$L_{1,\ell}\left(H^{(i)}_{\sigma_i} \right) \le O_\ell(1) + O_\ell\big((c/\ell)^\ell\big) \le O(c^\ell).$$

This proves Lemma 1.1. □

4 Proof of theorem 1.2

The main technical contribution in this section is a Fourier growth bound on $\mathsf{Q}\|^{\mathrm{pub}}$ protocols. We restate Lemma 1.2 for convenience.

Lemma 1.2. *Let $C : \{-1, 1\}^n \times \{-1, 1\}^n \to [-1, 1]$ be a $\mathsf{Q}\|^{\mathrm{pub}}$ protocol of cost c and let H be its XOR-fiber as in Definition 2.2. Then, for all $\ell \in \mathbb{N}$, we have*

$$L_{1,\ell}(H) \le O_\ell(c^\ell).$$

Proof of Theorem 1.2 using Lemma 1.2. Let $c = \tau \cdot n^{1/4}$ for a small enough constant $\tau > 0$. Let $\mathcal{H}$ be the family of XOR-fibers of $\mathsf{Q}\|^{\mathrm{pub}}$ protocols of cost at most c. This is a restriction-closed family of Boolean functions. The results of [GRT22] imply that the maximum advantage that functions in

$\mathcal{H}$ have in computing the Forrelation problem is at most $O\left(\frac{L_{1,2}(H)}{\sqrt{n}}\right) + O\left(\frac{1}{\sqrt{n}}\right)$. Using Lemma 1.2, this quantity is at most $\frac{c^2}{\sqrt{n}} \leq \tau^2 \ll 1$. This, along with the techniques of [GRT22] implies that $\mathsf{Q}\|^{\mathrm{pub}}$ protocols solving the Forrelation problem require communication cost $\Omega(n^{1/4})$. □

We now prove Lemma 1.2.

Proof of Lemma 1.2. Firstly, it suffices to prove the lemma for $\mathsf{Q}\|$ protocols, by the fact that $\mathsf{Q}\|^{\mathrm{pub}}$ protocols are defined by expectations over $\mathsf{Q}\|$ protocols and by triangle inequality. Let C be any $\mathsf{Q}\|$ protocol of cost c where Alice and Bob don't share entanglement. Without loss of generality, the protocol has the following form. Let $\mathcal{H}$ be a Hilbert space of dimension 2^c. Alice sends Charlie $\rho(x)$ and Bob sends Charlie $\sigma(y)$ where $\rho(x), \sigma(y) \in \mathcal{S}(\mathcal{H})$. Then Charlie applies a two-outcome POVM $\{M_1, M_{-1}\}$ and announces the outcome as the answer. It is not too hard to see that the expected output of the protocol is precisely $C(x,y) = \mathrm{Tr}(E \cdot (\rho(x) \otimes \sigma(y))$ where $E = M_1^\dagger M_1 - M_{-1}^\dagger M_{-1}$. Observe that for all $S \subseteq [n]$,

$$\widehat{H}(S) \triangleq \mathop{\mathbb{E}}_{x,y\sim\{-1,1\}^n} [\mathrm{Tr}(E \cdot (\rho(x) \otimes \sigma(y))) \cdot \chi_S(x \odot y)] = \mathrm{Tr}(E \cdot (\widehat{\rho}_S \otimes \widehat{\sigma}_S)).$$

Therefore, we have

$$\begin{aligned}\sum_{|S|=\ell} \left|\widehat{H}(S)\right| &= \sum_{|S|=\ell} \left|\mathrm{Tr}\left(E \cdot (\widehat{\rho}_S \otimes \widehat{\sigma}_S)\right)\right| \\ &\leq 2 \sum_{|S|=\ell} \mathrm{Tr}\left(|\widehat{\rho}_S|\right) \cdot \mathrm{Tr}\left(|\widehat{\sigma}_S|\right) \leq \sqrt{\sum_{|S|=\ell} \mathrm{Tr}\left(|\widehat{\rho}_S|\right)^2} \cdot \sqrt{\sum_{|S|=\ell} \mathrm{Tr}\left(|\widehat{\sigma}_S|\right)^2}.\end{aligned} \tag{3}$$

Here, we used Fact 2.1 on $M_1^\dagger M_1$ and $M_{-1}^\dagger M_{-1}$. We now apply Corollary 2.5 to the density-matrix valued functions $\rho(x), \sigma(y)$ to conclude that

$$\max\left\{\sum_{|S|=\ell} \mathrm{Tr}^2\left(|\widehat{\rho}_S|\right), \sum_{|S|=\ell} \mathrm{Tr}^2\left(|\widehat{\sigma}_S|\right)\right\} \leq O_\ell\left((c/\ell)^\ell\right) + O_\ell(1).$$

Substituting this in Eq. (3), we have

$$L_{1,\ell}(H) \leq O_\ell\left((c/\ell)^\ell\right) + O_\ell(1) \leq O_\ell(c^\ell).$$

This proves the desired upper bound. □

Remark 4.1. *Lemma 1.2, along with [GRZ21] implies that the advantage that $\mathsf{Q}\|^{\mathrm{pub}}$ protocols of cost $o(n^{1/4})$ have in computing the $\oplus^k$-Forrelation problem is at most $\exp(-\Omega(k))$. Using this, along with Lemma 5.1, one can show that $\mathsf{Q}\|^*$ protocols of cost $o(n^{1/4})$ where Alice and Bob only share $O(k)$ qubits of entanglement cannot solve the $\oplus^k$-Forrelation problem. On the other hand, [GRZ21] shows that if Alice and Bob share $\tilde{O}(k^5 \log^3 n)$ EPR pairs, then the $\oplus^k$-Forrelation problem can be solved by $\mathsf{Q}\|^*$ protocols of cost $\tilde{O}(k^5 \log^3 n)$. This gives a separation between $\mathsf{Q}\|^*$ with more entanglement and $\mathsf{Q}\|^*$ with less entanglement.*

5 Proof of Theorem 1.3

For convenience we restate Theorem 1.3 below.

Theorem 1.3. *The $\oplus^k$-Boolean Hidden Matching problem can be solved with $\tilde{O}(k \log n)$ bits of communication in the $\mathsf{R}\|^*$ model if Alice and Bob share $\tilde{\Theta}(k \log n)$ EPR pairs. However, if Alice and Bob only share $O(k)$ qubits of entanglement, then this problem requires*

- *$\Omega(k\sqrt{n})$ bits of communication in the $\mathsf{R1}^*$ model,*
- *$\Omega(kn^{1/3})$ qubits of communication in the $\mathsf{Q}\|^*$ model.*

Proof. The proof of this theorem proceeds in two steps. First, we assume by contradiction that there exists an $\mathsf{R1}^*$ or $\mathsf{Q}\|^*$ protocol of cost c that uses $2d$ qubits of entanglement and solves the $\oplus^k$-Boolean Hidden Matching problem with advantage at least $\frac{1}{3}$. Using this, we produce an $\mathsf{R1}$ or $\mathsf{Q}\|$ protocol respectively that solves the $\oplus^k$-Boolean Hidden Matching problem with cost at most $c+O(d)$; this protocol uses no entanglement but only succeeds with advantage at least $2^{-\Theta(d)}$, that is, the success probability is at least $\frac{1}{2}+2^{-\Theta(d)}$. This will be proved in Lemma 5.1 and Lemma 5.2. We then argue that $\mathsf{R1}$ and $\mathsf{Q}\|$ protocols satisfy an XOR lemma with respect to the Boolean Hidden Matching problem. That is, the advantage that cost c protocols have in solving the $\oplus^k$-Boolean Hidden Matching problem is at most $O_k\left(\frac{(c/k)^3}{n}\right)^{k/2}$ for the $\mathsf{Q}\|$ model and at most $O_k\left(\frac{(c/k)^2}{n}\right)^{k}$ for the $\mathsf{R1}$ model. This will be proved in Lemma 1.3 and Lemma 1.4 respectively. Combining the aforementioned lemma, for the $\mathsf{Q}\|^*$ case and $\mathsf{R1}^*$ case, we have

$$\frac{1}{3} \cdot 2^{-4d} \leq O_k\left(\frac{((c+2d)/k)^3}{n}\right)^{k/2} \quad \text{and} \quad \frac{1}{6} \cdot 2^{-4d} \leq O_k\left(\frac{((c+10d)/k)^2}{n}\right)^{k/2} \tag{4}$$

respectively. Let $\tau, \tau' > 0$ be sufficiently small constants and $d = \tau' \cdot k$. For the $\mathsf{Q}\|^*$ model, we can set $c = \tau \cdot kn^{1/3}$ and for the $\mathsf{R1}^*$ model set $c = \tau \cdot kn^{1/2}$ so that Eq. (4) is violated. This proves our communication lower bound. The quantum upper bound is presented in Appendix C. □

5.1 Removing Entanglement from Protocols

Our main contributions in this section are the following lemmas. Below, we assume d is large enough.

Lemma 5.1. *Given any communication protocol for computing F with cost c and advantage $\frac{1}{3}$ in the $\mathsf{Q}\|^*$ model where Alice and Bob share at most $2d$ qubits of entanglement, there is a protocol of cost $c+2d$ in the $\mathsf{Q}\|$ model that computes F with advantage $\frac{1}{3} \cdot 2^{-4d}$.*

Lemma 5.2. *Given any communication protocol for computing F with cost c and advantage $\frac{1}{3}$ in the $\mathsf{R1}^*$ model where Alice and Bob share at most $2d$ qubits of entanglement, there is a protocol of cost $c+10d$ in the $\mathsf{R1}$ model that computes F with advantage $\frac{1}{6} \cdot 2^{-4d}$.*

Proof of Lemma 5.1. Let C be a communication protocol of cost c in the $\mathsf{Q}\|^*$ model that uses $2d$ qubits of entanglement. Without loss of generality, the protocol has the following form. Let $\mathcal{H}'_A, \mathcal{H}'_B$ be Hilbert spaces of dimension 2^c each. Alice applies a quantum channel $U_x : \mathcal{S}(\mathcal{H}_A) \to \mathcal{S}(\mathcal{H}'_A)$ to her part of the shared state and Bob applies a quantum channel $V_y : \mathcal{S}(\mathcal{H}_B) \to \mathcal{S}(\mathcal{H}'_B)$ to his part of the shared state. The players then send the resulting states to the referee. The referee evaluates a two-outcome POVM on the received state and returns the outcome as the answer. In particular, if

$$\rho := \sum_{a,b,p,q \in [2^d]} \rho_{a,b,p,q} |a\rangle\langle p|_A \otimes |b\rangle\langle q|_B$$

is the state shared by Alice and Bob, then the state received by the referee is

$$\sum_{a,b,p,q \in [2^d]} \rho_{a,b,p,q} U_x(|a\rangle \langle p|) \otimes V_y(|b\rangle \langle q|). \tag{5}$$

We now produce a $\mathsf{Q}\|$ protocol where Charlie receives the state in Eq. (5) with probability 2^{-4d}, furthermore, he knows when this state is received. When Charlie receives this state, he continues as per the original protocol and when he does not receive this state, he returns a uniformly random bit. This would produce a $\mathsf{Q}\|$ protocol that computes F with advantage $\frac{1}{3} \cdot 2^{-4d}$. The protocol is as follows. Consider a $\mathsf{Q}\|$ protocol where Alice and Bob create the states

$$\sum_{a,b,p,q \in [2^d]} \rho_{a,b,p,q} |a\rangle\langle p|_A \otimes |b\rangle\langle q|_A \quad \text{and} \quad 2^{-2d} \sum_{b',q' \in [2^d]} |b'\rangle\langle q'|_B \otimes |b'\rangle\langle q'|_B \quad \text{respectively.}$$

Alice applies $U(x)$ to the *first half* of the qubits of her state and sends the entire state to Charlie. Bob applies $V(y)$ to the *second half* of the qubits of his state and sends the entire state to Charlie. The state received by Charlie is

$$2^{-2d} \sum_{\substack{a,b,b' \in [2^d] \\ p,q,q' \in [2^d]}} \rho_{a,b,p,q} U_x \left(|a\rangle \langle p|\right) \otimes |b\rangle \langle q| \otimes |b'\rangle \langle q'| \otimes V_y \left(|b'\rangle \langle q'|\right).$$

Charlie projects onto states such that $b = b'$ and $q = q'$. More precisely, Charlie considers the measurement operator on the qubits from $d+1$ to $3d$ defined by projection onto $\{|b\rangle \langle q| \otimes |b\rangle \langle q| : b, q \in [2^d]\}$. This measurement operator applied to the above state produces the state

$$\sum_{a,b,p,q, \in [2^d]} \rho_{a,b,p,q} U_x \left(|a\rangle \langle p|\right) \otimes |b\rangle \langle q| \otimes |b\rangle \langle q| \otimes V_y \left(|b\rangle \langle q|\right).$$

with probability $1/2^{2d}$, furthermore, Charlie can tell when this state is produced using the measurement outcome. Charlie then applies the Hadamard operator to the qubits $d+1, \ldots, 3d$ and measures those qubits in the standard basis. With probability $1/2^{2d}$, Charlie obtains the state

$$\sum_{a,b,p,q, \in [2^d]} \rho_{a,b,p,q} U_x \left(|a\rangle \langle p|\right) \otimes |0^d\rangle \langle 0^d| \otimes |0^d\rangle \langle 0^d| \otimes V_y \left(|b\rangle \langle q|\right).$$

As before, Charlie can tell when he obtained this state. Ignoring the zero qubits from $d+1$ to $3d$, this state is precisely the state Charlie had received in the original $\mathsf{Q}\|^*$ protocol as in Eq. (5). This completes the proof of Lemma 5.1. □

Proof of Lemma 5.2. Let $\mathcal{H}_A, \mathcal{H}_B$ be Hilbert spaces of dimension 2^d. Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ and let $\rho \in \mathcal{S}(\mathcal{H})$ be the state shared by Alice and Bob. We can assume without loss of generality that an $\mathsf{R1}^*$ protocol using ρ as the entanglement has the following form. Suppose Alice and Bob got x, y respectively. Alice first measures her half of the shared state using a 2^c-outcome measurement operator $\{M_z(x) : z \in \{-1,1\}^c\}$. She obtains an outcome $z \in \{-1,1\}^c$ and sends z to Bob. Based on this message z and his input y, Bob applies the two-outcome POVM $\{N_1(y,z), N_{-1}(y,z)\}$ on his part of the shared state and outputs the measurement outcome as the answer. Let $\sigma(x,z)$ be the post-measurement state of ρ after Alice applies her measurement $M_z(x)$ and obtains the outcome z. The expected output of Bob is precisely

$$\mathrm{Tr}\left((\mathbb{I} \otimes F(y,z)) \cdot \sigma(x,z)\right)$$

where $F(y,z) = N_1^\dagger(y,z)N_1(y,z) - N_{-1}^\dagger(y,z)N_{-1}(y,z)$ and $\mathbb{I} \otimes F(y,z)$ is an operator that acts as identity on the first d qubits and acts as $F(y,z)$ on the last d qubits.

Consider an $\mathsf{R1}$ protocol where Alice obtains the measurement outcome $z \in \{-1,1\}^c$ and the post-measurement state $\sigma(x,z) \in \mathcal{S}(\mathcal{H})$. She knows the precise classical description of the state $\sigma(x,z)$. She will send the classical message z as before. She samples uniformly random $i,j \in [2^{2d}]$ and sends $(i,j,\widetilde{\sigma}(x,z)[i,j])$, where, $\widetilde{\sigma}(x,z)[i,j]$ is the (i,j)th coordinate of $\sigma(x,z)$ specified up to $\Theta(d)$ bits of precision. We will show that using this message, Bob can estimate $\mathrm{Tr}((\mathbb{I} \otimes F(y,z)) \cdot \sigma(x,z))$ with advantage 2^{-4d}. Let $F'(y,z) = \mathbb{I} \otimes F(y,z)$. First, consider an ideal situation where Alice sends exactly $\sigma(x,z)[i,j]$. Observe that

$$\begin{aligned} \forall i,j \in [2^{2d}], \quad & \left|F'(y,z)[i,j] \cdot \sigma(x,z)[i,j]\right| \leq 1, \\ & \mathop{\mathbb{E}}_{i,j \sim [2^{2d}]} \left[F'(y,z)[i,j] \cdot \sigma(x,z)[i,j]\right] = \tfrac{1}{2^{4d}} \mathrm{Tr}(F'(y,z) \cdot \sigma(x,z)). \end{aligned} \tag{6}$$

Consider the protocol where Bob computes $F'(y,z)[i,j] \cdot \sigma(x,z)[i,j]$ for uniformly randomly $i,j \sim [2^{2d}]$. He then returns a random $\{-1,1\}$-bit whose expectation is $F'(y,z)[i,j] \cdot \sigma(x,z)[i,j]$. This is well-defined due to the first line in Eq. (6). The assumption is that $\mathrm{Tr}(F'(y,z) \cdot \sigma(x,z))$ is at least $1/3$ for NO instances and at most $-1/3$ for YES instances. This, along with the second line of Eq. (6) implies that Bob's expected output is at least $\frac{1}{3} \cdot 2^{-4d}$ for NO instances and at most $-\frac{1}{3} \cdot 2^{-4d}$ for YES instances. Thus, Bob solves the problem with advantage $\frac{1}{3} \cdot 2^{-4d}$. Suppose Alice specifies $\widetilde{\sigma}(x,z)[i,j] \in [-1,1]$ up to $5d$ bits of precision, then we have

$$\begin{aligned} & \left| \mathop{\mathbb{E}}_{i,j \sim 2^{[2d]}} \left[F'(y,z)[i,j] \cdot \widetilde{\sigma}(x,z)[i,j]\right] - 2^{-4d} \cdot \mathrm{Tr}\left(F'(y,z) \cdot \sigma(x,z)\right) \right| \\ & \leq 2^{-4d} \cdot \left|\mathrm{Tr}\left(F'(y,z) \cdot (\sigma(x,z) - \widetilde{\sigma}(x,z))\right)\right| \\ & \leq 2^{-4d} \cdot 2^{-5d} \cdot 2^{4d} \leq 2^{-5d} \ll \frac{1}{6} \cdot 2^{-4d}. \end{aligned}$$

Following the same calculations as before, it follows that Bob solves the problem with advantage at least $\frac{1}{6} \cdot 2^{-4d}$. □

5.2 XOR Lemma for $\mathsf{Q}\|^{\mathrm{pub}}$ for the Boolean Hidden Matching Problem

Our main technical result in this section is an XOR lemma for $\mathsf{Q}\|^{\mathrm{pub}}$ protocols with regards to the Boolean Hidden Matching problem.

Lemma 1.3. *Let C be any $\mathsf{Q}\|^{\mathrm{pub}}$ protocol of cost c. Then its advantage in computing the $\oplus^k$-Boolean Hidden Matching problem is at most $O_k\left(\frac{(c/k)^3}{n}\right)^{k/2} + O_k(n^{-k/2})$.*

To prove this lemma, we will make use of some properties which are very similar to those proved in [GKK+07]. The proofs of the facts are deferred to the appendix. Let $\mathcal{M}$ be the uniform distribution on matchings on $[n]$ of size $m = \alpha n$ and $\mathcal{U}$ be the uniform distribution on $\{-1,1\}^n$.

Definition 5.3 (M matches S)**.** *Let $S \subseteq [nk]$ and $M \in \mathrm{supp}(\mathcal{M}^{\otimes k})$. We say that M matches S if M is an induced perfect matching on S. If M matches S, we use $M(S) \subseteq [mk]$ to denote the subset of edges of this induced perfect matching.*

Observe that the map $T = M(S)$ defines a bijection between sets S that are matched by M and subsets $T \subseteq [mk]$. Furthermore, $|T| = |S|/2$ and for any $i \in [k]$, $|T_i|$ is odd if and only if $|S_i|/2$ is odd. We now define some sets that will be important in the proof.

Definition 5.4. *Let $\mathcal{S}_{n,k} := \{S \subseteq [nk] : \forall i \in [k], |S_i|/2$ is an odd integer$\}$ and $\mathcal{T}_{n,k} := \{T \subseteq [mk] : \forall i \in [k], |T_i|$ is an odd integer$\}$. Define $\mathcal{S}^{\ell}_{n,k} := \{S \in \mathcal{S}_{n,k} : |S| = 2\ell\}$ and $\mathcal{T}^{\ell}_{n,k} := \{T \in \mathcal{T}_{n,k} : |T| = \ell\}$ for all $\ell \in [mk]$.*

The aforementioned map $T = M(S)$ provides a bijection between sets $S \in \mathcal{S}^{\ell}_{n,k}$ that are matched by M and sets $T \in \mathcal{T}^{\ell}_{n,k}$. The following facts can be proved using techniques in [GKK+07].

Fact 5.5. *Let $S \subseteq [nk]$ and $M \in \mathrm{supp}(\mathcal{M}^{\otimes k})$. Then, for any $w \in \{-1,1\}^{mk}$, the quantity*

$$\mathop{\mathbb{E}}_{x \sim \mathcal{U}^{\otimes k}} [\mathbb{1}[Mx = w] \cdot \chi_S(x)]$$

is non-zero if and only if M matches S. Furthermore, if it is non-zero, it equals $2^{-mk} \cdot \chi_{M(S)}(w)$.

Fact 5.6. *Let $S \subseteq [nk]$ with $|S| = 2\ell$. Then,*

$$\Pr_{M \sim \mathcal{M}^{\otimes k}}[M \text{ matches } S] \le O_\ell\left(\frac{\ell^\ell}{(nk)^\ell}\right).$$

We now prove our main lemma of this subsection.

Proof of Lemma 1.3. We assume that $(c/k)^{3/2} \le \tau \cdot n^{1/2}$ for some small constant $\tau > 0$, otherwise the statement of the lemma is vacuously true. We will construct distributions on the YES and NO instances of the $\oplus^k$-Boolean Hidden Matching problem such that no small cost $\mathsf{Q}\|$ protocol can distinguish them with considerable advantage. Consider the following two distributions.

- $\mathcal{N}$ is a distribution on NO-instances of $\mathrm{BHM}_{m,n}$: A random sample of $\mathcal{N}$ is of the form (x, M, y) where $x \sim \mathcal{U}$, $M \sim \mathcal{M}$ and $y := Mx$.
- $\mathcal{Y}$ is a distribution on YES-instances of $\mathrm{BHM}_{m,n}$ defined similarly to $\mathcal{N}$ except that $y := \overline{Mx}$.

Define two distribution $\mu^{(k)}_1, \mu^{(k)}_{-1}$ on inputs to the $\oplus^k$-Boolean Hidden Matching problem as follows.

$$\mu^{(k)}_1 := \frac{1}{2^{k-1}} \sum_{\substack{K \subseteq [k] \\ |K| \text{ is even}}} \mathcal{Y}_K \mathcal{N}_{\overline{K}} \quad \text{and} \quad \mu^{(k)}_{-1} := \frac{1}{2^{k-1}} \sum_{\substack{K \subseteq [k] \\ |K| \text{ is odd}}} \mathcal{Y}_K \mathcal{N}_{\overline{K}}. \tag{7}$$

Recall that $\mathcal{Y}_K \mathcal{N}_{\overline{K}}$ is a product of k independent distributions, where the i-th distribution is $\mathcal{Y}$ if $i \in K$ and is $\mathcal{N}$ if $i \notin K$. Clearly $\mu^{(k)}_{-1}$ and $\mu^{(k)}_1$ are distributions on the YES and NO instances respectively of the $\oplus^k$-Boolean Hidden Matching problem.

Consider any $\mathsf{Q}\|$ protocol with c qubits of communication and let $\mathcal{H}$ be a Hilbert space of dimension 2^c. Such a protocol can be described by density matrices $\rho(x) \in \mathcal{S}(\mathcal{H})$ and $\sigma_M(y) \in \mathcal{S}(\mathcal{H})$

for every $x \in \{-1,1\}^{nk}, y \in \{-1,1\}^{mk}$ and $M \in \mathrm{supp}(\mathcal{M}^{\otimes k})$. The state received by Charlie on these inputs is precisely $\rho(x) \otimes \sigma_M(y)$. We will show that the trace distance between the states $\mathbb{E}_{(x,M,y)\sim\mu_1^{(k)}}\left[\rho(x) \otimes \sigma_M(y)\right]$ and $\mathbb{E}_{(x,M,y)\sim\mu_{-1}^{(k)}}\left[\rho(x) \otimes \sigma_M(y)\right]$ is at most $O_k\left(\frac{(c/k)^{3k/2}}{n^{k/2}}\right)$. Since the trace distance measures the maximal distinguishing probability between the two states, this, along with Yao's principle would complete the proof. Towards this, define

$$\Delta := \mathop{\mathbb{E}}_{(x,M,y)\sim\mu_1^{(k)}}\left[\rho(x) \otimes \sigma_M(y)\right] - \mathop{\mathbb{E}}_{(x,M,y)\sim\mu_{-1}^{(k)}}\left[\rho(x) \otimes \sigma_M(y)\right].$$

Using the definition of $\mu_1^{(k)}$ and $\mu_{-1}^{(k)}$ in Eq. (7), we have

$$\Delta \triangleq \sum_{K\subseteq[k]} \frac{(-1)^{|K|}}{2^{k-1}} \cdot \mathop{\mathbb{E}}_{\substack{x\sim\mathcal{U}^{\otimes k}\\ M\sim\mathcal{M}^{\otimes k}}}\left[\rho(x) \otimes \sigma_M\left(\overline{(Mx)}_K (Mx)_{\overline{K}}\right)\right].$$

We introduce a variable $w \in \{-1,1\}^{mk}$ to represent Mx so that

$$\Delta = \sum_{\substack{w\in\{-1,1\}^{mk}\\ K\subseteq[k]}} \frac{(-1)^{|K|}}{2^{k-1}} \cdot \mathop{\mathbb{E}}_{\substack{x\sim\mathcal{U}^{\otimes k}\\ M\sim\mathcal{M}^{\otimes k}}}\left[\rho(x) \otimes \sigma_M\left(\overline{w}_K w_{\overline{K}}\right) \cdot \mathbb{1}[Mx = w]\right].$$

We expand $\rho(x)$ in the Fourier Basis to obtain

$$\Delta = \sum_{\substack{w\in\{-1,1\}^{mk}\\ K\subseteq[k]\\ S\subseteq[nk]}} \frac{(-1)^{|K|}}{2^{k-1}} \cdot \mathop{\mathbb{E}}_{\substack{x\sim\mathcal{U}^{\otimes k}\\ M\sim\mathcal{M}^{\otimes k}}}\left[\widehat{\rho}(S) \otimes \sigma_M\left(\overline{w}_K w_{\overline{K}}\right) \cdot \left[\mathbb{1}[Mx = w] \cdot \chi_S(x)\right]\right].$$

Consider the term $\mathop{\mathbb{E}}_{x\sim\mathcal{U}^{\otimes k}}\left[\mathbb{1}[Mx = w] \cdot \chi_S(x)\right]$. By Fact 5.5, this term is non-zero if and only if M matches S, in which case the term evaluates to $2^{-mk} \cdot \chi_{M(S)}(w)$. Substituting this in the equation above, we have that Δ equals

$$\sum_{\substack{w\in\{-1,1\}^{mk}\\ K\subseteq[k]\\ S\subseteq[nk]}} \frac{(-1)^{|K|}}{2^{k-1}} \mathop{\mathbb{E}}_{M\sim\mathcal{M}^{\otimes k}}\left[\widehat{\rho}(S) \otimes \sigma_M\left(\overline{w}_K w_{\overline{K}}\right) \cdot 2^{-mk} \cdot \chi_{M(S)}(w) \cdot \mathbb{1}[M \text{ matches } S]\right]. \tag{8}$$

We now expand $\sigma_M\left(\overline{w}_K w_{\overline{K}}\right)$ in the Fourier basis with respect to w. Consider

$$\begin{aligned}\sum_{K\subseteq[k]} (-1)^{|K|} \cdot \sigma_M\left(\overline{w}_K w_{\overline{K}}\right) &= \sum_{K\subseteq[k]} (-1)^{|K|} \cdot \sum_{T\subseteq[mk]} \widehat{\sigma_M}(T) \cdot \chi_T(\overline{w}_K, w_{\overline{K}})\\ &= \sum_{\substack{K\subseteq[k]\\ T\subseteq[mk]}} (-1)^{|K|} \cdot \widehat{\sigma_M}(T) \cdot \chi_T(w) \cdot (-1)^{\sum_{i\in K}|T_i|}\\ &= \sum_{T\subseteq[mk]} \widehat{\sigma_M}(T) \cdot \chi_T(w) \cdot \sum_{K\subseteq[k]} \left[(-1)^{|K|+\sum_{i\in K}|T_i|}\right].\end{aligned}$$

For $i \in [k]$, let $t_i = -1$ if $|T_i|$ is odd and $t_i = 1$ if $|T_i|$ is even. Observe that

$$\mathop{\mathbb{E}}_{K\subseteq[k]}\left[(-1)^{|K|+\sum_{i\in K}|T_i|}\right] = \mathop{\mathbb{E}}_{K\subseteq[k]}\left[\chi_K(-t_1,\ldots,-t_k)\right] = \begin{cases}1 & \text{if } \forall i \in [k], t_i = -1,\\ 0 & \text{otherwise.}\end{cases}$$

Hence, the quantity $\sum_{K\subseteq[k]}\left[(-1)^{|K|+\sum_{i\in K}|T_i|}\right]$ is non-zero if and only if $|T_i|$ is odd for all $i \in [k]$. Furthermore, if it is non-zero, then it equals 2^k. Recall that we defined $\mathcal{T}_{nk} := \{T \subseteq [mk] : \forall i \in [k], |T_i| \text{ is odd}\}$ in Definition 5.4. Using this, we have

$$\sum_{K\subseteq[k]} (-1)^{|K|} \cdot \sigma_M\left(\overline{w}_K w_{\overline{K}}\right) = 2^k \cdot \sum_{T\in\mathcal{T}_{n,k}} \widehat{\sigma_M}(T) \cdot \chi_T(w). \tag{9}$$

Substituting this in Eq. (8), we have that Δ equals

$$\begin{aligned}
&2 \sum_{\substack{w\in\{-1,1\}^{mk}\\ S\subseteq[nk]}} \mathop{\mathbb{E}}_{M\sim\mathcal{M}^{\otimes k}} \left[\widehat{\rho}(S) \otimes \sum_{T\in\mathcal{T}_{n,k}} \widehat{\sigma_M}(T)\cdot 2^{-mk}\cdot \chi_{M(S)}(w)\cdot\chi_T(w)\cdot \mathbb{1}[M \text{ matches } S]\right]\\
&= 2\sum_{S\subseteq[nk]} \mathop{\mathbb{E}}_{M\sim\mathcal{M}^{\otimes k}} \left[\widehat{\rho}(S) \otimes \sum_{T\in\mathcal{T}_{n,k}} \widehat{\sigma_M}(T)\cdot \mathop{\mathbb{E}}_{w\sim\{-1,1\}^{mk}}[\chi_{M(S)+T}(w)]\cdot \mathbb{1}[M \text{ matches } S]\right].
\end{aligned}$$

Observe that if M matches S, then $\mathbb{E}_{w\sim\{-1,1\}^{mk}}\left[\chi_{M(S)+T}(w)\right]$ equals 1 if $T = M(S)$ and 0 otherwise. Recall that the sets $S \subseteq [nk]$ such that M matches S and $M(S) \in \mathcal{T}_{n,k}$ are precisely those sets in $\mathcal{S}_{n,k}$ that are matched by M. Hence,

$$\Delta = \sum_{S\in\mathcal{S}_{n,k}} \widehat{\rho}(S) \otimes \mathbb{E}_M\left[\widehat{\sigma_M}(M(S))\cdot \mathbb{1}[M \text{ matches } S]\right].$$

We now upper bound $\|\Delta\|_1$ by triangle inequality as follows.

$$\|\Delta\|_1 \le \sum_{S\in\mathcal{S}_{n,k}} \|\widehat{\rho}(S)\|_1 \otimes \mathbb{E}_M\left[\|\widehat{\sigma_M}(M(S))\|_1 \cdot \mathbb{1}[M \text{ matches } S]\right].$$

We partition $\mathcal{S}_{n,k}$ and $\mathcal{T}_{n,k}$ into $\sqcup_\ell \mathcal{S}^\ell_{nk}$ and $\sqcup_\ell \mathcal{T}^\ell_{n,k}$ based on the size of the sets as in Definition 5.4. Observe that every set in $\mathcal{S}_{n,k}$ has size at least $2k$ and every set in $\mathcal{T}_{n,k}$ has size at least k. Thus,

$$\|\Delta\|_1 \le \sum_{\ell=k}^{mk} \sum_{S\in\mathcal{S}^\ell_{k,n}} \|\widehat{\rho}(S)\|_1 \otimes \mathbb{E}_M\big[\, \|\widehat{\sigma_M}(M(S))\|_1 \cdot 1[M \text{ matches } S]\big]. \tag{10}$$

We now apply Cauchy-Schwarz to conclude that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{M\sim\mathcal{M}^{\otimes k}}\big[\, \|\widehat{\sigma_M}(M(S))\|_1 \cdot \mathbb{1}[M \text{ matches } S]\big]\\
&\le \sqrt{\mathop{\mathbb{E}}_{M\sim\mathcal{M}^{\otimes k}}\big[\, \|\widehat{\sigma_M}(M(S))\|_1^2 \cdot \mathbb{1}[M \text{ matches } S]\big]} \cdot \sqrt{\mathop{\Pr}_{M\sim\mathcal{M}^{\otimes k}}[M \text{ matches S}]}.
\end{aligned}$$

Fact 5.6 implies that for any $S \in \mathcal{T}^\ell_{n,k}$, we have $\mathop{\Pr}_{M\sim\mathcal{M}^{\otimes k}}[M \text{ matches } S] \le O_\ell\left(\frac{\ell^\ell}{(nk)^\ell}\right)$. Substituting this in Eq. (10) implies that

$$\|\Delta\|_1 \le \sum_{\ell=k}^{mk} \sum_{S\in\mathcal{S}^\ell_{nk}} \|\widehat{\rho}(S)\|_1 \cdot \sqrt{\mathbb{E}_M\left[\|\widehat{\sigma_M}(M(S))\|_1^2 \cdot \mathbb{1}[M \text{ matches } S]\right]} \cdot O_\ell\left(\frac{\ell^{\ell/2}}{(nk)^{\ell/2}}\right).$$

Again, by Cauchy-Schwarz, we have

$$\|\Delta\|_1 \le \sum_{\ell=k}^{mk} \sqrt{\sum_{S\in\mathcal{S}^\ell_{n,k}} \|\widehat{\rho}(S)\|_1^2} \cdot \sqrt{\sum_{S\in\mathcal{S}^\ell_{n,k}} \mathbb{E}_M\left[\|\widehat{\sigma_M}(M(S))\|_1^2 \cdot \mathbb{1}[M \text{ matches } S\right] \cdot O_\ell\left(\frac{\ell^{\ell/2}}{(nk)^{\ell/2}}\right).$$

By the aforementioned correspondence between sets $S \in \mathcal{S}^\ell_{n,k}$ such that M matches S and sets $T \in \mathcal{T}^\ell_{n,k}$, we have

$$\|\Delta\|_1 \le \sum_{\ell=k}^{mk} \sqrt{\sum_{S\in\mathcal{S}^\ell_{n,k}} \|\widehat{\rho}(S)\|_1^2} \cdot \sqrt{\sum_{T\in\mathcal{T}^\ell_{n,k}} \mathbb{E}_M\left[\|\widehat{\sigma_M}(T)\|_1^2\right]} \cdot O_\ell\left(\frac{\ell^{\ell/2}}{(nk)^{\ell/2}}\right). \tag{11}$$

We now apply the Matrix level-k inequality in Corollary 2.5 to the functions $\rho : \{-1,1\}^n \to \mathcal{S}(\mathcal{H})$ and $\sigma_M : \{-1,1\}^m \to \mathcal{S}(\mathcal{H})$ where $\mathcal{H}$ is a Hilbert space of dimension 2^c. Corollary 2.5 implies that

$$\sum_{|S|=2\ell} \|\widehat{\rho}(S)\|_1^2 \le O_\ell\left((c/\ell)^{2\ell}\right) + O_\ell(1) \quad \text{and} \quad \sum_{|T|=\ell} \|\widehat{\sigma_M}(T)\|_1^2 \le O_\ell\left((c/\ell)^{\ell}\right) + O_\ell(1).$$

Substituting this in Eq. (11), we get

$$\begin{aligned}\|\Delta\|_1 &\le \sum_{\ell=k}^{mk} \sqrt{\sum_{S:|S|=2\ell} \|\widehat{\rho}(S)\|_1^2} \cdot \sqrt{\sum_{T:|T|=\ell} \mathbb{E}_M\left[\|\widehat{\sigma_M}(T)\|_1^2\right]} \cdot O_\ell\left(\frac{\ell^{\ell/2}}{(nk)^{\ell/2}}\right)\\ &\le \sum_{\ell=k}^{mk} \max\left(O_\ell\left(\frac{c^{3\ell/2}}{\ell^{3\ell/2}}\right), O_\ell(1)\right) \cdot O_\ell\left(\frac{\ell^{\ell/2}}{(nk)^{\ell/2}}\right)\\ &\le \sum_{\ell=k}^{mk} O_\ell\left(\frac{c^{3\ell/2}}{\ell^\ell(nk)^{\ell/2}}\right) + \sum_{\ell=k}^{mk} O_\ell\left(\frac{\ell^{\ell/2}}{(nk)^{\ell/2}}\right).\end{aligned}$$

Since $\ell \le mk = \alpha \cdot nk$ for a sufficiently small constant $\alpha > 0$, the function $\ell^{\ell/2}/(nk)^{\ell/2}$ is exponentially decaying for $\ell \in [k, mk]$ and hence the second term is at most $O_k\left(n^{-k/2}\right)$. Our assumption that $(c/k)^{3/2} \le \tau \cdot n^{1/2}$ for a sufficiently small constant $\tau > 0$ implies that the function $c^{3\ell/2}/(\ell^\ell(nk)^{\ell/2})$ is exponentially decaying for $\ell \in [k, mk]$ and hence, the first term above is at most $O_k\left(\frac{(c/k)^{3k/2}}{n^{k/2}}\right)$. Together, we have

$$\|\Delta\|_1 \le O_k\left(\frac{(c/k)^{3k/2}}{n^{k/2}}\right) + O_k(n^{-k/2}).$$

This completes the proof of Lemma 1.3. □

5.3 XOR Lemma for R1 for the Boolean Hidden Matching Problem

In this subsection, we prove Lemma 1.4 which we restate here for convenience.

Lemma 1.4. *Let C be any R1 protocol of cost c. Then its advantage in computing the $\oplus^k$-Boolean Hidden Matching problem is at most $O_k\left(\frac{(c/k)^2}{n}\right)^{k/2} + O_k(n^{-k/2})$.*

Proof of Lemma 1.4. The proof of this lemma will be similar to the proof of Lemma 1.3 and hence we will follow similar notation. Let $z \in \{-1,1\}^c$ be any c-bit message sent by Alice and let $A_z \subseteq \{-1,1\}^{nk}$ be the set of Alice's inputs for which Alice would have sent z to Bob. Let $g(x) = \mathbb{1}[x \in A_z]$. Fix any $M \in \mathrm{supp}(\mathcal{M}^{\otimes k})$. Similar to Lemma 1.3, let $\mathcal{N}^M(y)$ be the distribution on $y \in \{-1,1\}^{mk}$ induced by sampling $x \sim A_z$ and letting $y = Mx$. Let $\mathcal{Y}^M(y)$ be similarly defined with $y := \overline{Mx}$. So we have that $\mathcal{N}^M(y) = \frac{|\{x \in A_z | Mx = y\}|}{|A_z|}$ and $\mathcal{Y}^M(y) = \frac{|\{x \in A_z | \overline{Mx} = y\}|}{|A_z|}$ for all $y \in \{-1,1\}^{mk}$. Define

$$\mu_1^{(k)} := \frac{1}{2^{k-1}} \sum_{\substack{K \subseteq [k] \\ |K| \text{ is even}}} \mathcal{Y}_K^M \mathcal{N}_{\overline{K}}^M \quad \text{and} \quad \mu_{-1}^{(k)} := \frac{1}{2^{k-1}} \sum_{\substack{K \subseteq [k] \\ |K| \text{ is odd}}} \mathcal{Y}_K \mathcal{N}_{\overline{K}}^M. \tag{12}$$

Below we show that for a typical $M \sim \mathcal{M}^{\otimes k}$, these two distributions are close in total variational distance. By arguments similar to [GKK+07], this would complete the proof. To this end, let

$$\Delta_{A_z} := \mathop{\mathbb{E}}_{M \sim \mathcal{M}^{\otimes k}} \left[\left\| \mu_1^{(k)} - \mu_{-1}^{(k)} \right\|_1 \right].$$

By Eq. (12), we have $\mu_1^{(k)} - \mu_{-1}^{(k)} = 2^{1-k} \cdot \sum_{K \subseteq [k]} (-1)^{|K|} \mathcal{Y}_K^M \mathcal{N}_{\overline{K}}^M$. Hence

$$\Delta_{A_z}^2 \leq 2^{mk} \cdot \mathbb{E}_M \left[\left\| \mu_1^{(k)} - \mu_{-1}^{(k)} \right\|_2^2 \right] = 2^{mk} \cdot \mathbb{E}_M \left[2^{-2k+2} \cdot \left\| \sum_{K \subseteq [k]} \mathcal{Y}_K^M \mathcal{N}_{\overline{K}}^M (-1)^{|K|} \right\|_2^2 \right],$$

where the first inequality is by the Cauchy-Schwarz inequality. By Parseval's theorem, we have

$$\Delta_{A_z}^2 \leq 2^{2mk-2k+2} \cdot \mathbb{E}_M \left[\sum_{\substack{T \subseteq [mk] \\ T \neq \emptyset}} \left(\sum_{K \subseteq [k]} \widehat{\mathcal{Y}_K^M \mathcal{N}_{\overline{K}}^M}(T) (-1)^{|K|} \right)^2 \right]. \tag{13}$$

Observe that

$$\begin{aligned}
\widehat{\mathcal{Y}_K^M \mathcal{N}_{\overline{K}}^M}(T) &= \frac{1}{2^{mk}} \sum_{y \in \{-1,1\}^{mk}} \left(\mathcal{Y}_K^M \mathcal{N}_{\overline{K}}^M \right)(y) \cdot \chi_T(y) \\
&= \frac{1}{2^{mk} \cdot |A_z|} \Big(\left| \{ x \in A_z | \chi_T \left((Mx)_{\overline{K}} (\overline{Mx})_K \right) = 1 \} \right| \\
&- \left| \{ x \in A_z \,|\, \chi_T \left((Mx)_{\overline{K}} (\overline{Mx})_K \right) = -1 \} \right| \Big) \\
&= \frac{1}{2^{mk} \cdot |A_z|} \Big(\left| \left\{ x \in A_z \,\middle|\, \chi_T(Mx) = (-1)^{\sum_{i \in K} |T_i|} \right\} \right| \\
&- \left| \left\{ x \in A_z \,\middle|\, \chi_T(Mx) \neq (-1)^{\sum_{i \in K} |T_i|} \right\} \right| \Big) \\
&= \frac{1}{2^{mk}} \sum_{y \in \{-1,1\}^{mk}} \mathcal{N}^{\otimes k}(y) \cdot \chi_T(y) \cdot (-1)^{\sum_{i \in K} |T_i|} \\
&= \widehat{\mathcal{N}^{\otimes k}}(T) \cdot (-1)^{\sum_{i \in K} |T_i|}.
\end{aligned}$$

By an argument analogous to [GKK+07, Eq. (3)], we have $\widehat{\mathcal{N}^{\otimes k}}(T) = \frac{2^{nk}}{|A_z| \cdot 2^{mk}} \cdot \widehat{g}(M^\dagger T)$. Hence

$$\sum_{K \subseteq [k]} \widehat{\mathcal{Y}_K^M \mathcal{N}_{\overline{K}}^M}(T) \cdot (-1)^{|K|} = \frac{2^{nk}}{2^{mk} \cdot |A_z|} \cdot \widehat{g}(M^\dagger T) \cdot \sum_{K \subseteq [k]} (-1)^{|K| + \sum_{i \in K} |T_i|}. \tag{14}$$

As we saw in Eq. (9), the term $\sum_{K\subseteq[k]}(-1)^{|K|+\sum_{i\in K}|T_i|}$ is 2^k if $|T_i|$ is odd for all $i \in [k]$ and zero otherwise. Hence, the R.H.S. of Eq. (14) is non-zero only if $T \in \mathcal{T}_{n,k}$ (defined in Definition 5.4), and in this case equals $\frac{2^{nk}}{2^{mk}\cdot|A_z|}\cdot\widehat{g}(M^\dagger T)\cdot 2^k$. Substituting this in Eq. (13), we have that $\Delta^2_{A_z}$ equals

$$2^{2mk-2k+2}\cdot\mathbb{E}_M\left[\sum_{T\in\mathcal{T}_{n,k}}\frac{2^{2nk+2k}}{2^{2mk}\cdot|A_z|^2}\widehat{g}(M^\dagger T)^2\right] = 4\cdot\mathbb{E}_M\left[\sum_{T\in\mathcal{T}_{n,k}}\frac{2^{2n}}{|A_z|^2}\widehat{g}(M^\dagger T)^2\right].$$

Recall the correspondence between $\mathcal{T}_{n,k}$ and $\mathcal{S}_{n,k}$ as in Definition 5.4. For every $S \in \mathcal{S}_{n,k}$, there is at most one $T \in \mathcal{T}_{n,k}$ such that $M^\dagger T = S$, furthermore, such a T exists if and only if M matches S. Hence we have that

$$\begin{aligned}\Delta^2_{A_z} &\le 4\cdot\mathbb{E}_M\left[\sum_{S\in\mathcal{S}_{n,k}}\frac{2^{2n}}{|A_z|^2}\widehat{g}(S)^2\cdot\mathbb{1}[M\text{ matches }S]\right]\\ &= 4\cdot\sum_{\ell=k}^{mk}\sum_{S\in\mathcal{S}^{\ell}_{n,k}}\frac{2^{2n}}{|A_z|^2}\widehat{g}(S)^2\cdot\Pr_M[M\text{ matches }S] \le \sum_{\ell=k}^{mk}\left(\sum_{|S|=2\ell}\frac{2^{2n}}{|A_z|^2}\widehat{g}(S)^2\right)\cdot O_\ell\left(\frac{\ell^\ell}{(nk)^\ell}\right),\end{aligned}$$

where we used Fact 5.6. Let $\mu(A_z) = \frac{|A_z|}{2^n}$. Applying Lemma 2.3, we have

$$\Delta^2_{A_z} \le \sum_{\ell=k}^{mk}\left(2e\cdot\ln\left(\tfrac{e}{\mu(A_z)^{1/(2\ell)}}\right)\right)^{2\ell}\cdot O_\ell\left(\tfrac{\ell^\ell}{(nk)^\ell}\right).$$

We now take square root on both sides (and use concavity of the square root function) to get

$$\Delta_{A_z} \le \sum_{\ell=k}^{mk}\left(2e\cdot\ln\left(\tfrac{e}{\mu(A_z)^{1/(2\ell)}}\right)\right)^{\ell}\cdot O_\ell\left(\tfrac{\ell^{\ell/2}}{(nk)^{\ell/2}}\right).$$

We now multiply both sides by $\mu(A_z)$ and add over all 2^c possibilities for the transcript $z \in \{-1,1\}^c$.

$$\Delta := \sum_{z\in\{-1,1\}^c}\Delta_{A_z} \le \sum_{z\in\{-1,1\}^c}\mu(A_z)\cdot\sum_{\ell=k}^{mk}\left(2e\cdot\ln\left(\tfrac{e}{\mu(A_z)^{1/(2\ell)}}\right)\right)^{\ell}\cdot O_\ell\left(\tfrac{\ell^{\ell/2}}{(nk)^{\ell/2}}\right).$$

We now use the concavity of the function $h'(\gamma) = \gamma\cdot\ln(e/\gamma^{1/2\ell})^\ell$ for $\gamma \in [0,1]$ and all $\ell \in \mathbb{N}$, (similarly to the proof in Section 3.2) to conclude that

$$\Delta \le \sum_{\ell=k}^{mk}\left(\sum_{z\in\{-1,1\}^c}\mu(A_z)\right)\cdot\left(2e\cdot\ln\left(\frac{e\cdot 2^{c/(2\ell)}}{\left(\sum_{z\in\{-1,1\}^c}\mu(A_z)\right)^{1/(2\ell)}}\right)\right)^{\ell}\cdot O_\ell\left(\frac{\ell^{\ell/2}}{(nk)^{\ell/2}}\right).$$

We use the fact that $\sum_{z\in\{-1,1\}^c}\mu(A_z) = 1$ to conclude that

$$\Delta \le \sum_{\ell=k}^{mk}\left(2e\cdot\ln\left(e\cdot 2^{c/(2\ell)}\right)\right)^{\ell}\cdot O_\ell\left(\frac{\ell^{\ell/2}}{(nk)^{\ell/2}}\right) \le \sum_{\ell=k}^{mk}O_\ell\left(\frac{\ell^{\ell/2}}{(nk)^{\ell/2}}\right)+\sum_{\ell=k}^{mk}O_\ell\left(\frac{c^\ell}{(\ell nk)^{\ell/2}}\right).$$

As before, the first term is at most $O(n^{-k/2})$. The assumption that $c\cdot k \le \tau\cdot n^{1/2}$ for a small enough constant $\tau > 0$ implies that the function $\frac{c^\ell}{(\ell nk)^{\ell/2}}$ is exponentially decaying for $\ell \in [k, mk]$. Hence, the second term is at most $O_\ell\left(\frac{(c/k)^k}{n^{k/2}}\right)$. This, along with the techniques of [GKK+07] completes the proof of Lemma 1.4. □

References

[AA15] Scott Aaronson and Andris Ambainis. Forrelation: A problem that optimally separates quantum from classical computing. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 307–316, 2015.

[Aar10] Scott Aaronson. BQP and the polynomial hierarchy. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 141–150, 2010.

[BBM12] Eric Blais, Joshua Brody, and Kevin Matulef. Property testing lower bounds via communication complexity. *computational complexity*, 21(2):311–358, 2012.

[BCW98] Harry Buhrman, Richard Cleve, and Avi Wigderson. Quantum vs. classical communication and computation. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 63–68, 1998.

[BCWW01] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. *Physical Review Letters*, 87(16):167902, 2001.

[BJK08] Ziv Bar-Yossef, Thathachar S Jayram, and Iordanis Kerenidis. Exponential separation of quantum and classical one-way communication complexity. *SIAM Journal on Computing*, 38(1):366–384, 2008.

[BRSW11] Harry Buhrman, Oded Regev, Giannicola Scarpa, and Ronald de Wolf. Near-optimal and explicit bell inequality violations. In *2011 IEEE 26th Annual Conference on Computational Complexity*, pages 157–166, 2011.

[BRW08] Avraham Ben-Aroya, Oded Regev, and Ronald de Wolf. A hypercontractive inequality for matrix-valued functions with applications to quantum computing and ldcs. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 477–486. IEEE, 2008.

[BS21] Nikhil Bansal and Makrand Sinha. k-forrelation optimally separates quantum and classical query complexity. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1303–1316, 2021.

[CH19] Matthew Coudron and Aram W. Harrow. Universality of EPR pairs in entanglement-assisted communication complexity, and the communication cost of state conversion. In *34th Computational Complexity Conference, CCC 2019, July 18-20, 2019, New Brunswick, NJ, USA*, volume 137 of *LIPIcs*, pages 20:1–20:25, 2019.

[FMP$^+$15] Samuel Fiorini, Serge Massar, Sebastian Pokutta, Hans Raj Tiwary, and Ronald de Wolf. Exponential lower bounds for polytopes in combinatorial optimization. *Journal of the ACM (JACM)*, 62(2):1–23, 2015.

[Gav08] Dmitry Gavinsky. On the role of shared entanglement. *Quantum Inf. Comput.*, 8(1):82–95, 2008.

[Gav09] Dmytro Gavinsky. Classical interaction cannot replace quantum nonlocality, 2009.

[Gav19] Dmitry Gavinsky. Quantum versus classical simultaneity in communication complexity. *IEEE Transactions on Information Theory*, 65(10):6466–6483, 2019.

[Gav20] Dmitry Gavinsky. Bare quantum simultaneity versus classical interactivity in communication complexity. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 401–411, 2020.

[GKdW06] Dmytro Gavinsky, Julia Kempe, and Ronald de Wolf. Strengths and weaknesses of quantum fingerprinting. 2006.

[GKK$^+$07] Dmitry Gavinsky, Julia Kempe, Iordanis Kerenidis, Ran Raz, and Ronald de Wolf. Exponential separations for one-way quantum communication complexity, with applications to cryptography. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 516–525, 2007.

[GKRW06] Dmitry Gavinsky, Julia Kempe, Oded Regev, and Ronald de Wolf. Bounded-error quantum state identification and exponential separations in communication complexity. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of Computing*, pages 594–603, 2006.

[GRT22] Uma Girish, Ran Raz, and Avishay Tal. Quantum versus randomized communication complexity, with efficient players. *computational complexity*, 31(2):17, 2022.

[GRZ21] Uma Girish, Ran Raz, and Wei Zhan. Lower bounds for XOR of forrelations. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM*, volume 207 of *LIPIcs*, pages 52:1–52:14, 2021.

[HHL18] Hamed Hatami, Kaave Hosseini, and Shachar Lovett. Structure of protocols for XOR functions. *SIAM J. Comput.*, 47(1):208–217, 2018.

[HN12] Trinh Huynh and Jakob Nordstrom. On the virtue of succinct proofs: Amplifying communication complexity hardness to time-space trade-offs in proof complexity. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 233–248, 2012.

[JKN07] Rahul Jain, Hartmut Klauck, and Ashwin Nayak. Direct product theorems for communication complexity via subdistribution bounds. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, pages 599–608, 2007.

[JRS05] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. Prior entanglement, message compression and privacy in quantum communication. In *20th Annual IEEE Conference on Computational Complexity (CCC'05)*, pages 285–296, 2005.

[KKS14] Michael Kapralov, Sanjeev Khanna, and Madhu Sudan. Streaming lower bounds for approximating max-cut. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1263–1282, 2014.

[KR11] Bo'az Klartag and Oded Regev. Quantum one-way communication can be exponentially stronger than classical communication. In *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*, page 31–40, 2011.

[KRW95] Mauricio Karchmer, Ran Raz, and Avi Wigderson. Super-logarithmic depth lower bounds via the direct sum in communication complexity. *Computational Complexity*, 5(3):191–204, 1995.

[KW90] Mauricio Karchmer and Avi Wigderson. Monotone circuits for connectivity require super-logarithmic depth. *SIAM J. Discret. Math.*, 3(2):255–265, 1990.

[Lee19] Chin Ho Lee. Fourier bounds and pseudorandom generators for product tests. In *34th Computational Complexity Conference, CCC*, volume 137 of *LIPIcs*, pages 7:1–7:25, 2019.

[MNSW95] Peter Bro Miltersen, Noam Nisan, Shmuel Safra, and Avi Wigderson. On data structures and asymmetric communication complexity. In *Proceedings of the twenty-seventh annual ACM symposium on Theory of computing*, pages 103–111, 1995.

[MO10] Ashley Montanaro and Tobias Osborne. On the communication complexity of xor functions, 2010.

[New91] Ilan Newman. Private vs. common random bits in communication complexity. *Information processing letters*, 39(2):67–71, 1991.

[O'D14] Ryan O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, 2014.

[Raz95] Ran Raz. Fourier analysis for probabilistic communication complexity. *Comput. Complex.*, 5(3/4):205–221, 1995.

[Raz99] Ran Raz. Exponential separation of quantum and classical communication complexity. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 358–367, 1999.

[RT22] Ran Raz and Avishay Tal. Oracle separation of BQP and PH. *ACM Journal of the ACM (JACM)*, 69(4):1–21, 2022.

[Shi05] Yaoyun Shi. Tensor norms and the classical communication complexity of nonlocal quantum measurement. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 460–467, 2005.

[SSW21] Alexander A. Sherstov, Andrey A. Storozhenko, and Pei Wu. An optimal separation of randomized and quantum query complexity. In Samir Khuller and Virginia Vassilevska Williams, editors, *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1289–1302, 2021.

[SZ08] Yaoyun Shi and Zhiqiang Zhang. Communication complexities of xor functions. *arXiv preprint arXiv:0808.1762*, 2008.

[TWXZ13] Hing Yin Tsang, Chung Hoi Wong, Ning Xie, and Shengyu Zhang. Fourier sparsity, spectral norm, and the log-rank conjecture. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 658–667, 2013.

[Yao79] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 209–213, 1979.

[Yu22] Huacheng Yu. Strong XOR lemma for communication with bounded rounds : (extended abstract). In *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS*, pages 1186–1192, 2022.

[Zha14] Shengyu Zhang. Efficient quantum protocols for xor functions. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pages 1878–1885. SIAM, 2014.

A Proofs in Section 3

Proof of Theorem 1.1 using Lemma 1.1. The quantum upper bound is presented in [GRZ21, Theorem 3.8]. We describe it below for completeness. Let $t = \Theta(k^5 \log^3 n \log k)$. For every $x^{(i)}, y^{(i)} \in \{-1,1\}^n$ given as input to Alice and Bob, Alice sends $|x^{(i)}\rangle = \frac{1}{\sqrt{n}} \sum_{j\in[n]} x_j^{(i)} |j\rangle$, and Bob sends $|y^{(i)}\rangle = \frac{1}{\sqrt{n}} \sum_{j\in[n]} y_j^{(i)} |j\rangle$ to the referee. The referee performs a swap test between $|x^{(i)}\rangle$ and $H_n |y^{(i)}\rangle$ and the bias of the swap test is precisely $\textsc{forr}(x^{(i)}, y^{(i)})$, which we are promised is either at least $\varepsilon/2$ or at most $\varepsilon/4$ for every $i \in [k]$ where $\varepsilon = \Theta\left(\frac{1}{k^2 \ln n}\right)$. The referee takes the threshold of $t = \Theta\left(k \cdot \log k \cdot k^4 \ln^2 n\right)$ many swap tests and a simple calculation similar to [GRZ21] shows that the referee can decide $\textsc{forr}^{(\oplus k)}$ with probability at least $2/3$.

The classical lower bound uses Lemma 1.1 and [GRZ21]. In more detail, [GRZ21] define two distributions $\tilde{\mu}_1^{(k)}$ and $\tilde{\mu}_{-1}^{(k)}$ and prove that these distribution put considerable mass (at least $1 - 1/\mathrm{poly}(n)$) on the YES and NO instances of the $\oplus^k$-Forrelation problem respectively [GRZ21, Lemma 2.11]. They also show [GRZ21, Theorem 3.1] that for any restriction-closed family $\mathcal{H}$ of Boolean functions on $2kn$ variables with outputs in $[-1,1]$, the maximum advantage that functions in $\mathcal{H}$ have in distinguishing $\tilde{\mu}_1^{(k)}$ and $\tilde{\mu}_{-1}^{(k)}$ is at most $O\left(L_{1,2k}(\mathcal{H}) \cdot n^{-k/2}\right) + o\left(n^{-k/2}\right)$.

Let $c = \tau \cdot n^{1/4}$ for a small enough constant $\tau > 0$. Let $\mathcal{H}$ be the set of all XOR-fibers of $\mathsf{R2}^*$ protocols of cost at most c that use ρ as the entangled state. It is not too hard to show that this family is closed under restrictions. Using the aforementioned results, as well as Lemma 1.1, we conclude that for all $H \in \mathcal{H}$,

$$\left| \mathop{\mathbb{E}}_{z\sim\tilde{\mu}_1^{(k)}} [H(z)] - \mathop{\mathbb{E}}_{z\sim\tilde{\mu}_{-1}^{(k)}} [H(z)] \right| \le O_k\left(2^{5d} \cdot c^{2k} \cdot n^{-k/2}\right) + o(n^{-k/2}) \le O_k(2^{5k} \cdot \tau^{2k}).$$

In the last step, we used the fact that $d \le k$ and $c = \tau \cdot n^{1/4}$. Setting $\tau \ll 1$ to be a sufficiently small constant, the R.H.S. of the above equation is at most $1/5$. This completes the proof. □

B Proofs in Section 3

We prove Claim 3.7 in this section. We begin by describing the structure of $\mathsf{R2}^*$ protocols that share a state $\rho \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with communication cost c. Without loss of generality, the protocol can be described as follows. Alice and Bob could each have private memory consisting of m qubits. Say $\mathcal{H}'_A$ and $\mathcal{H}'_B$ are Hilbert spaces of dimension 2^m. Note that the dimension m could potentially be very large. Consider a k-round protocol. Suppose Alice and Bob got x, y respectively. Below we let $|\Phi_0\rangle = |\Phi\rangle \otimes |0^m\rangle\langle 0^m|_A \otimes |0^m\rangle\langle 0^m|_B$, $z_0 = \emptyset$ and $t = 0$. Here, the subscript A, B on a qubit denotes which player has that qubit.

Consider the $(t+1)$th round of the protocol. Suppose Alice and Bob had exchanged messages $z_1, \ldots, z_{2t} \in \{-1, 1\}$ in the first t rounds. Without loss of generality, we can assume that each player sends at most one bit in each round. This assumption can increase the communication cost by a factor of at most two. Alice first applies a two-outcome POVM $\{M^{t+1}_{z_{2t+1}}(x, z_1, \ldots, z_{2t}) : z_{2t+1} \in \{-1, 1\}\}$ on the registers that she owns (i.e., her part of the shared state as well her memory). Alice sends the POVM outcome $z_{2t+1} \in \{-1, 1\}$ to Bob. Based on this message z_{2t+1}, his input y and the transcript $z_1, \ldots, z_{2t} \in \{-1, 1\}$, Bob applies applies a two-outcome POVM $\{N^{t+1}_{z_{2t+2}}(y, z_1, \ldots, z_{2t+1}) : z_{2t+2} \in \{-1, 1\}\}$ on the qubits that he owns (his part of the shared state and his memory). He sends the POVM outcome $z_{2t+2} \in \{-1, 1\}$ to Alice. Let the resulting state of all the qubits after the $(t + 1)$th round be $|\Phi_{t+1}\rangle$. They repeat this for k rounds after which Alice evaluates some predicate A on z and returns the answer as the output. We say that a protocol has cost c, if the size of the transcript is at most c bits. We can assume that there are $\lceil \frac{c}{2} \rceil$ rounds and that the players in fact communicate for exactly $\lceil \frac{c}{2} \rceil$ rounds, where c is the communication complexity of the protocol on the worst case inputs.

Proof of Claim 3.7. Let $A \subseteq \{-1, 1\}^n$ denote the set of $z \in \{-1, 1\}^n$ that satisfy the final predicate and let $k = \lceil \frac{c}{2} \rceil$. Let $z \in \{-1, 1\}^c$. Since $x, y \in \{-1, 1\}^n$ are fixed throughout this proof, for simplicity of notation, for any $j \in [k]$, let

$$M^j_z := M^j_{z_{2j-1}}(x, z_1, \ldots, z_{2j-2}) \quad \text{and} \quad N^j_z := N^j_{z_{2j}}(x, z_1, \ldots, z_{2j-1}),$$

$$M^{\leq j}_z := \prod_{j'=j}^{1} M^{j'}_z \quad \text{and } N^{\leq j}_z := \prod_{j'=j}^{1} N^{j'}_z.$$

Let $M_z = M^{\leq k}_z$ and $N_z = N^{\leq k}_z$. Note that the M^j_z are functions of x, z and N^j_z are functions of y, z. Finally, let $E_z(x) := {M_z}^\dagger M_z$ and $F_z(y) := {N_z}^\dagger N_z$. It is clear that properties 1 and 2 are satisfied by E_z and F_z. It is straightforward to see from the definition of $\mathsf{R2}^*$ protocols that the expected output of the protocol is precisely

$$\begin{aligned} &C(x, y) \\ &= (-1) \cdot \sum_{z \in A} \mathrm{Tr}\left((M_z \otimes N_z)\, \rho' \left(M^\dagger_z \otimes N^\dagger_z\right)\right) + 1 \cdot \sum_{z \notin A} \mathrm{Tr}\left((M_z \otimes N_z)\, \rho' \left(M^\dagger_z \otimes N^\dagger_z\right)\right) \\ &= \sum_{z \in \{-1,1\}^c} \mathrm{Tr}\left((E_z(x) \otimes F_z(y))\, \rho'\right) \cdot (-1)^{\mathbb{1}[z \in A]}. \end{aligned}$$

It only remains to prove property 3. Consider:

$$\begin{aligned} (*) &:= \sum_{z \in \{-1,1\}^c} E_z(x) \otimes F_z(y) \\ &\triangleq \sum_{z \in \{-1,1\}^{2k}} \left[\left(M^{\leq k}_z\right)^\dagger \cdot \left(M^{\leq k}_z\right)\right] \otimes \left[\left(N^{\leq k}_z\right)^\dagger \cdot N^{\leq k}_z\right] \\ &= \sum_{z \in \{-1,1\}^{2k-1}} \left[\left(M^{\leq k}_z\right)^\dagger \cdot \left(M^{\leq k}_z\right)\right] \otimes \left[\left(N^{\leq k-1}_z\right)^\dagger \cdot \left(\sum_{z_{2k} \in \{-1,1\}} \left(N^k_z\right)^\dagger \cdot N^k_z\right) \cdot \left(N^{\leq k-1}_z\right)\right] \end{aligned}$$

The last equality used the fact that the operators $N^{\leq k-1}_z$ and $M^{\leq k}_z$ do not depend on z_{2k}. For all $z \in \{-1, 1\}^{2k-1}$, $\left\{N^k_z : z_{2k} \in \{-1, 1\}\right\}$ is a two-outcome POVM, thus,

$$\sum_{z_{2k} \in \{-1,1\}} \left(N^k_z\right)^\dagger \cdot N^k_z = \mathbb{I}.$$

Let $N = \left(N_z^{\leq k-1}\right)^\dagger \cdot N_z^{\leq k-1}$. Substituting this above, we have

$$
\begin{aligned}
(*) &= \sum_{z \in \{-1,1\}^{2k-1}} \left[\left(M_z^{\leq k}\right)^\dagger \cdot M_z^{\leq k} \right] \otimes N \\
&= \sum_{z \in \{-1,1\}^{2k-2}} \left[\left(M_z^{\leq k-1}\right)^\dagger \cdot \left(\sum_{z_{2k-1} \in \{-1,1\}} \left(M_z^k\right)^\dagger \cdot M_z^k \right) \cdot \left(M_z^{\leq k-1}\right) \right] \otimes N.
\end{aligned}
$$

The last equality used the fact that the operators $M_z^{\leq k-1}$ and $N = \left(N_z^{\leq k-1}\right)^\dagger \cdot N_z^{\leq k-1}$ don't depend on z_{2k-1}. For all $z \in \{-1,1\}^{2k-2}$, $\left\{M_z^j : z_{2k-1} \in \{-1,1\}\right\}$ is a two-outcome POVM, thus,

$$
\sum_{z_{2k-1} \in \{-1,1\}} \left(M_z^k\right)^\dagger \cdot M_z^k = \mathbb{I}.
$$

Substituting this above, we have

$$
\begin{aligned}
(*) &= \sum_{z_1,\ldots,z_{2k-2} \in \{-1,1\}} \left[\left(M_z^{\leq k-1}\right)^\dagger \cdot \left(M_z^{\leq k-1}\right) \right] \otimes \left[\left(N_z^{\leq k-1}\right)^\dagger \cdot \left(N_z^{\leq k-1}\right) \right] \\
&= \ldots = \mathbb{I} \quad \text{by induction on } k.
\end{aligned}
$$

This proves property 3 and completes the proof of Claim 3.7. □

C Quantum upper bound in Theorem 1.3.

We discuss the quantum upper bound for a single instance of the Boolean Hidden Matching problem (this is similar to the protocol in [BRSW11]). Here Alice and Bob share the quantum state $\frac{1}{\sqrt{n}} \sum_{i=1}^n |i\rangle_A |i\rangle_B$. Alice applies the transformation $|i\rangle_A \rightarrow (-1)^{x_i} |i\rangle_A$ and transforms the state to $\frac{1}{\sqrt{n}} \sum_{i \in [n]} (-1)^{x_i} |i\rangle_A |i\rangle_B$. Bob now measures his register in the matching basis, in particular, he completes his α-partial matching arbitrarily to a complete matching. Let the resulting complete matching be $\{(e_i, e_j)\}$ wherein α-fraction of the edges belong to $\mathcal{E}$. Now Bob measures in the basis $\{|e_i\rangle\langle e_i| + |e_j\rangle\langle e_j|\}$. Now, Bob obtains a uniformly random edge in the matching (known to him). Furthermore, with probability $1/\alpha$, the obtained edge was in $\mathcal{E}$. Say he obtained $(e_i, e_j) \in \mathcal{E}$. The state then collapses to

$$
\frac{1}{\sqrt{2}} \Big((-1)^{x_i} |i\rangle |i\rangle + (-1)^{x_j} |j\rangle |j\rangle \Big). \tag{15}
$$

Note that Bob knows the edge (i, j). Now, both Alice and Bob apply the $(\log n)$-qubit Hadamard gate on their respective registers, the resulting state is given by

$$
\frac{1}{\sqrt{2}n} \sum_{a,b \in \{0,1\}^{\log n}} \left((-1)^{x_i + (a+b)\cdot i} + (-1)^{x_j + (a+b)\cdot j} \right) |a, b\rangle .
$$

Now observe that if Alice and Bob measure their respective registers, Alice obtains a uniformly random, Bob obtains b satisfying $(i \oplus j) \cdot (a \oplus b) = x_i + x_j$. Alice sends a and Bob sends (i, j), y_{ij} as well as b to the referee. The referee now can now compute $(i \oplus j) \cdot (a \oplus b)$ and learn $x_i \oplus x_j$.

The referee returns NO if $y_{ij} = x_i \oplus x_j$, and returns YES if $\overline{y}_{ij} = x_i \oplus x_j$. This solves the Boolean Hidden Matching Problem. Observe that this protocol succeeds with probability α (which is the probability that Bob's measurement gives an edge in his matching input).

In order to compute $\mathrm{BHM}_{m,n}^{(\oplus k)}$ Alice and Bob perform the following: for every $i \in [k]$, they carry out the protocol $O((\log k)/\alpha)$ many times and send all their outcomes to the referee. With probability at least 9/10, the measurement collapses to an edge in the matching, which Bob knows and can communicate to the referee. For this edge, the referee checks the predicate if $y_{ij} = x_i \oplus x_j$ is satisfied or not and hence knows the value of $\mathrm{BHM}_{m,n}(x^i, y^i)$. Hence after $O((\log k)/\alpha)$ bits of communication, the referee knows $\mathrm{BHM}_{m,n}(x^i, y^i)$ for *all* $i \in [k]$ and hence $\mathrm{BHM}_{m,n}^{(\oplus k)}$.

C.1 Proofs in Section 5.2

Proof of Fact 5.6. Let $|S_i| = 2\ell_i$ for $i \in [k]$ such that $\sum_{i\in[k]} \ell_i = \ell$. It is argued in [GKK+07] that the probability that a random matching on $[n]$ of size $m = \alpha n$ matches any given set of size $2\ell_i$ is precisely $\frac{\binom{\alpha n}{\ell_i}}{\binom{n}{2\ell_i}}$. Furthermore they showed that $\frac{\binom{\alpha n}{\ell_i}}{\binom{n}{2\ell_i}}$ is a decreasing function of ℓ_i, and is at most $O_{\ell_i}\left((\ell_i/n)^{\ell_i}\right)$. Thus, the probability that M matches S is

$$g(\ell_1, \ldots, \ell_k) = \prod_{i\in[k]} \frac{\binom{\alpha n}{\ell_i}}{\binom{n}{2\ell_i}} \leq \left(\max_{i\in[k]} \frac{\binom{\alpha n}{\ell_i}}{\binom{n}{2\ell_i}}\right)^k \leq \left(\frac{\binom{\alpha n}{\ell/k}}{\binom{n}{2\ell/k}}\right)^k = O_\ell\left(\frac{\ell^\ell}{(nk)^\ell}\right).$$

□