Exponential Quantum Communication Reductions from Generalizations of the Boolean Hidden Matching Problem

João F. Doriguello[1]
School of Mathematics, University of Bristol, United Kingdom
Quantum Engineering Centre for Doctoral Training, University of Bristol, United Kingdom
http://www.joaodoriguello.com
joao.doriguellodiniz@bristol.ac.uk

Ashley Montanaro
School of Mathematics, University of Bristol, United Kingdom
ashley.montanaro@bristol.ac.uk

Abstract

In this work we revisit the Boolean Hidden Matching communication problem, which was the first communication problem in the one-way model to demonstrate an exponential classical-quantum communication separation. In this problem, Alice's bits are matched into pairs according to a partition that Bob holds. These pairs are compressed using a Parity function and it is promised that the final bit-string is equal either to another bit-string Bob holds, or its complement. The problem is to decide which case is the correct one. Here we generalize the Boolean Hidden Matching problem by replacing the parity function with an arbitrary function f. Efficient communication protocols are presented depending on the sign-degree of f. If its sign-degree is less than or equal to 1, we show an efficient classical protocol. If its sign-degree is less than or equal to 2, we show an efficient quantum protocol. We then completely characterize the classical hardness of all symmetric functions f of sign-degree greater than or equal to 2, except for one family of specific cases. We also prove, via Fourier analysis, a classical lower bound for any function f whose pure high degree is greater than or equal to 2. Similarly, we prove, also via Fourier analysis, a quantum lower bound for any function f whose pure high degree is greater than or equal to 3. These results give a large family of new exponential classical-quantum communication separations.

2012 ACM Subject Classification Theory of computation → Communication complexity; Theory of computation → Quantum complexity theory

Keywords and phrases Communication Complexity, Quantum Communication Complexity, Boolean Hidden Matching Problem

Digital Object Identifier 10.4230/LIPIcs.TQC.2020.1

Related Version A full version of the paper is available at https://arxiv.org/abs/2001.05553.

Funding *João F. Doriguello*: EPSRC grant EP/L015730/1.
Ashley Montanaro: QuantERA ERA-NET Cofund in Quantum Technologies implemented within the European Union's Horizon 2020 Programme (QuantAlgo project); the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 817581); and EPSRC grants EP/L021005/1, and EP/R043957/1.

Acknowledgements We would like to thank Ronald de Wolf for pointing out Ref. [24], and Makrand Sinha for useful discussions about the hypercontractive inequality.

[1] Corresponding author

1 Introduction

One of the main aims of the field of quantum information and quantum computation is to establish the superiority of quantum computers and quantum resources over their classical counterparts. While in some areas this superiority is based on a belief in the impossibility of classical computers or classical resources solving particular tasks, e.g. the efficiency of Shor's algorithm [25] coming from the belief that there is no efficient classical factoring algorithm, in other areas like communication complexity one can establish unconditional exponential separations between classical and quantum performances.

Communication complexity is a model of computation first introduced by Yao [28]. In this model, two parties (normally called Alice and Bob) hold each a piece of data and want to solve some computational task that jointly depends on their data. More specifically, if Alice holds some information x and Bob holds some information y, they want to solve some function $f(x, y)$ or relational problem with several valid outputs for each x and y. In order to do so, they will need to communicate between themselves, and their goal is to solve the problem with minimal communication. The protocol that Alice and Bob employ could be *two-way*, where they take turns sending messages to each other; *one-way*, where Alice sends a single message to Bob who then outputs the answer; or *simultaneous*, where Alice and Bob each pass one message to a third party (the referee) who outputs the answer. Apart from these different types of communication settings, one is also interested in the error of a protocol when solving a communication problem: the zero-error *communication complexity* is the worst-case communication of the best protocol that gives a correct output with probability 1 for every input (x, y); the bounded-error *communication complexity* is the worst-case communication cost of the best protocol that gives a correct output with probability $1 - \epsilon$ for every input (x, y), with $\epsilon \in [0, 1/2)$.

An interesting extension of the original communication model is the model of *quantum* communication complexity [8], also introduced by Yao [29]. In this model, Alice and Bob each has a quantum computer and they exchange qubits instead of bits and/or make use of shared entanglement. The use of quantum resources can drastically reduce the amount of communication in solving some problems in comparison to the classical communication model.

Exponential quantum-classical separations are known in the two-way (e.g. [22]), one-way (e.g. [4, 15]) and simultaneous (e.g. [9, 12]) models. Indeed, it is even known that one-way quantum communication can be exponentially more efficient than two-way classical communication [14, 23]. However, surprisingly few examples of such exponential separations are known, compared (for example) with the model of query complexity in which Shor's algorithm operates.

The Hidden Matching problem [4] was the first problem to exhibit an exponential separation between the bounded-error classical communication complexity and the bounded-error quantum communication complexity in the one-way model. The problem can be efficiently solved by one quantum message of $\log n$ qubits, while any classical one-way protocol needs to send $O(\sqrt{n})$ bits to solve it. The hardness of the problem is essentially one-way: it could be efficiently solved by having Bob sent a classical message of $\log n$ bits to Alice. The Hidden Matching problem is a relational problem. In the same paper [4] the authors proposed a Boolean version of the problem, the Boolean Hidden Matching problem (which is a partial Boolean function), and conjectured that the same quantum-classical gap holds for it as well, which was later proven to be true by Gavinsky *et al.* [15]. Generalizing this separation is the focus of this work.

1.1 Hidden matching problems

Throughout the paper, $[n] = \{1, 2, \ldots, n\}$ and given $x, y \in \{-1, 1\}^n$, we denote by $x \circ y$ the Hadamard (elementwise) product of x and y, and by $\overline{x}$ the complement of x, such that $x \circ \overline{x} = 1^n$.

The Hidden Matching (HM_n^α) and Boolean Hidden Matching (BHM_n^α) problems are defined with respect to some $\alpha \in (0, 1]$. Alice is given a string $x \in \{-1, 1\}^n$[4] and Bob is given a sequence $M \in \mathcal{M}_{\alpha n/2}$ of $\alpha n/2$ disjoint pairs $(i_1, j_1), (i_2, j_2), \ldots, (i_{\alpha n/2}, j_{\alpha n/2}) \in [n]^2$. Such a sequence is called an α-matching, and $\mathcal{M}_{\alpha n/2}$ denotes the family of all α-matchings – i.e. partial matchings of a fixed size in the complete graph on n vertices. Together x and M induce a string $z \in \{-1, 1\}^{\alpha n/2}$ defined by the parities of the $\alpha n/2$ edges, i.e., $z_\ell = x_{i_\ell} x_{j_\ell}$ for $\ell = 1, \ldots, \alpha n/2$. Then the HM_n^α and BHM_n^α problems are defined as follows.

▶ **Definition 1** (**The Hidden Matching problem** (HM_n^α)). *Let $n \in \mathbb{N}$ be even and $\alpha \in (0, 1]$. Alice receives $x \in \{-1, 1\}^n$ and Bob receives $M \in \mathcal{M}_{\alpha n/2}$. Their goal is to output a tuple $\langle i, j, b\rangle$ such that $(i, j) \in M$ and $b = x_i x_j$.*

▶ **Definition 2** (**The Boolean Hidden Matching problem** (BHM_n^α)). *Let $n \in \mathbb{N}$ be even and $\alpha \in (0, 1]$. Alice receives $x \in \{-1, 1\}^n$ and Bob receives $M \in \mathcal{M}_{\alpha n/2}$ and $w \in \{-1, 1\}^{\alpha n/2}$. It is promised that $z \circ w = b^{\alpha n/2}$ for some $b \in \{-1, 1\}$. Their goal is to output b.*

Given inputs x and M, it is clear that there are many possible correct outputs for the HM_n^α problem ($\alpha n/2$ correct outputs, actually), making it a relational problem. On the other hand, the BHM_n^α is a partial Boolean function due to the promise statement.

Bar-Yossef *et al.* [4] gave a simple quantum protocol to solve the HM_n^1 problem with just $O(\log n)$ qubits of communication[5], while proving that any classical protocol needs to communicate at least $\Omega(\sqrt{n})$ bits in order to solve it. Similarly with the BHM_n^α problem, Gavinsky *et al.* [15] demonstrated the same exponential classical-quantum communication gap for any $\alpha \leq 1/2$ (note that the definition of α they use differs from ours by a factor of 2). As HM_n^α is at least as difficult as BHM_n^α, their result implies the same lower bound for HM_n^α. The approach taken by Gavinsky *et al.* in proving the classical lower bound is particularly interesting in that it uses the Fourier coefficients inequality of Kahn, Kalai, and Linial [17], which is proven via the Bonami-Beckner inequality [7, 5]. We also mention that Fourier analysis had been previously used in communication complexity by Raz [21] and Klauck [18].

A slightly weaker separation ($O(\log n)$ vs. $\Omega(n^{7/16})$) for a closely related problem was shown in [19] using similar techniques. The BHM_n^α problem was generalized by Verbin and Yu [26] to a problem that they named Boolean Hidden Hypermatching (BHH_n^t). In this problem, instead of having the bits from Alice matched in pairs, they are now matched in tuples of t elements. In other words, a bit from the final string z is obtained by XORing t bits from Alice's string. More precisely, Alice is given a string $x \in \{-1, 1\}^n$ and Bob is given a sequence $M \in \mathcal{M}_{n/t}$ of n/t disjoint tuples $(M_{1,1}, \ldots, M_{1,t}), \ldots, (M_{n/t,1}, \ldots, M_{n/t,t}) \in [n]^t$ called a hypermatching, where $\mathcal{M}_{n/t}$ denotes the family of all hypermatchings. Both x and M induce a string $z \in \{-1, 1\}^{n/t}$ defined by the parities of the n/t edges, i.e., $z_\ell = \prod_{j=1}^t x_{M_{\ell,j}}$ for $\ell = 1, \ldots, n/t$. The BHH_n^t problem is defined as follows.

▶ **Definition 3** (**The Boolean Hidden Hypermatching problem** (BHH_n^t)). *Let $n, t \in \mathbb{N}$ be such that $2t|n$. Alice receives $x \in \{-1, 1\}^n$ and Bob receives $M \in \mathcal{M}_{n/t}$ and $w \in \{-1, 1\}^{n/t}$. It is promised that $z \circ w = b^{n/t}$ for some $b \in \{-1, 1\}$. Their goal is to output b.*

[4] Throughout this paper we shall use $\{-1, 1\}$ instead of $\{0, 1\}$ for convenience.
[5] Their protocol extends easily to the more general HM_n^α problem.

Verbin and Yu proved a classical lower bound of $\Omega(n^{1-1/t})$ communication for every bounded-error one-way protocol, showing the increasing hardness of the problem with t, as one should expect since the BHH_n^t problem can be reduced from the BHM_n problem (we will show how this is done in detail later). The authors subsequently used this problem to prove various streaming lower bounds, i.e., lower bounds on the space required of streaming algorithms (algorithms that read the input from left to right, use a small amount of space, and approximate some function of the input). However, no efficient quantum protocol was proposed for solving the BHH_n^t problem for $t > 2$. It was only later that Shi, Wu and Yu [24] showed that such efficient quantum protocols do not exist. More specifically, they proved a quantum lower bound of $\Omega(n^{1-2/t})$ communication for every bounded-error one-way protocol for the BHH_n^t problem. Their proof is similar to the ones used in the classical lower bound, the difference lying in the use of Fourier analysis of *matrix-valued* functions and the matrix-valued Hypercontractive Inequality of Ben-Aroya, Regev, and de Wolf [6].

Note that the lower bound of Verbin and Yu does not use an α parameter, unlike the lower bound of [15]. However, their lower bound requires n/t to be even, otherwise Alice can just send the parity of her bit-string. (The result of [15] can be extended to hold for any $\alpha < 1$ fairly straightforwardly, but achieving a strong lower bound for $\alpha = 1$ requires some more work.)

1.2 Our Results

This paper focuses on the study of a broad generalization of the BHH_n^t problem. In the (Boolean) Hidden Matching and Boolean Hidden Hypermatching problems, the task Alice and Bob want to solve can be viewed as rearranging Alice's data according to some permutation that Bob holds, and "compressing" the data to a final bit-string by applying some Boolean function to the bits. Then Alice and Bob's goal is to determine some information about this final bit-string. The way this compression was originally done was via the Parity function, but, apart from the obvious reason that Parity gives the desired classical-quantum communication gap and, less obviously, leads to a clear proof, there is no particular need to restrict to this function in order to arrive at the final bit-string. This observation leads to a generalization of the Boolean Hidden Hypermatching problem, which we named the f-Boolean Hidden Partition (f-$\mathrm{BHP}_n^{\alpha,t}$) problem, where $f : \{-1,1\}^t \to \{-1,1\}$ is the Boolean function used to compress Alice's bits.

Given $y \in \{-1,1\}^n$, we define by $y^{(j;t)} = (y_{(j-1)t+1}, y_{(j-1)t+2}, \ldots, y_{jt}) \in \{-1,1\}^t$ the j-th block of size t from y, with $t|n$ and $j = 1, \ldots, n/t$. When the size of the block is clear from the context, we shall simply write $y^{(j)}$.

The f-Boolean Hidden Partition problem is defined as follows. Alice is given a bit-string $x \in \{-1,1\}^n$, and Bob is given a permutation $\sigma \in S_n$ and a bit-string $w \in \{-1,1\}^{\alpha n/t}$, where $\alpha \in (0,1]$ is fixed. Given a Boolean function $f : \{-1,1\}^t \to \{-1,1\}$, we can define the map $B_f : \{-1,1\}^n \to \{-1,1\}^{\alpha n/t}$ by $B_f(x) = \big(f(\sigma(x)^{(1)}), \ldots, f(\sigma(x)^{(\alpha n/t)})\big)$, where $\sigma(x)_i = x_{\sigma^{-1}(i)}$. Hence x and σ induce a bit-string given by $B_f(x)$, each of whose bits is obtained by applying f to a block of the permuted bit-string $\sigma(x)$. The f-$\mathrm{BHP}_n^{\alpha,t}$ problem can be defined as follows.

▶ **Definition 4 (The f-Boolean Hidden Partition problem (f-$\mathrm{HM}_n^{\alpha,t}$)).** *Let $n, t \in \mathbb{N}$ be such that $t|n$ and $\alpha \in (0,1]$. Alice receives $x \in \{-1,1\}^n$ and Bob receives $\sigma \in S_n$ and $w \in \{-1,1\}^{\alpha n/t}$. It is promised that there exists $b \in \{-1,1\}$ such that $B_f(x) \circ w = b^{\alpha n/t}$. The problem is to output b.*

The adoption of the word "Partition" instead of "(Hyper)Matching" from previous works comes from our decision to view the problem in terms of a hidden partition that Bob holds, instead of an α-(Hyper)Matching. Bob shuffles Alice's data according to some permutation, and then just partitions the resulting data in adjacent blocks of size t and uses f to get the final bit-string. Obviously both views are equivalent, but we think that the permutation approach eases the analysis of the problem.

Our aim is to study the f-Boolean Hidden Partition problem in terms of the function f. It should be clear that for some functions the problem is hard to solve classically, e.g. when f is the Parity function and we recover the usual Boolean Hidden Hypermatching problem. On the other hand, for some functions it becomes easily solvable, e.g. when f is the AND function (Alice needs only to send the position of any 0 in her string). We would like to characterize for which functions the problem can be efficiently solved classically, i.e., with $O(\log n)$ bits of communication, and for which functions it is hard to solve classically, i.e., requires $\Omega(n^a)$ bits of communication for some $a \in (0, 1]$. And the same question applies to quantum communication complexity: we would like to determine for which functions the problem admits or not an efficient quantum communication protocol. Given this characterization, we can check for which functions there is an exponential classical-quantum communication gap.

We partially prove that the whole f-BHP$_n^{\alpha,t}$ problem can be fully characterized by just one quantity: the *sign-degree* of the function f. A polynomial $p : \{-1, 1\}^t \to \mathbb{R}$ is said to *sign-represent* f if $f(x) = \mathrm{sgn}(p(x))$. If $|p(x)| \leq 1$ for all x, we say that p is *normalized*. The *bias* of a normalized polynomial p is defined as $\beta = \min_x |p(x)|$. The sign-degree ($sdeg(f)$) of f is the minimum degree of polynomials that sign-represent it. In Appendix A we prove the following upper bounds on the classical and quantum communication complexity of the f-Boolean Hidden Partition problem based on the sign-degree:

▶ **Theorem 5.** *Let $f : \{-1, 1\}^t \to \{-1, 1\}$ be a Boolean function. If $sdeg(f) \leq 1$, then there exists a bounded-error classical protocol that solves the f-BHP$_n^{\alpha,t}$ problem with error probability ϵ and $O\left((\frac{t}{\alpha\beta})^2 \log \frac{1}{\epsilon} \log n\right)$ bits of communication, where β is the maximal bias of a polynomial of degree $sdeg(f)$ that sign-represents f.*

▶ **Theorem 6.** *Let $f : \{-1, 1\}^t \to \{-1, 1\}$ be a Boolean function. If $sdeg(f) \leq 2$, then there exists a bounded-error quantum protocol that solves the f-BHP$_n^{\alpha,t}$ problem with error probability ϵ and $O\left((\frac{t}{\alpha\beta})^2 \log \frac{1}{\epsilon} \log n\right)$ qubits of communication, where β is the maximal bias of a polynomial of degree $sdeg(f)$ that sign-represents f.*

Note that the bias β can be very small, but can also be lower-bounded in terms only of t: indeed, it is shown in [10] that β is lower-bounded by $t^{-O(t^{sdeg(f)})}$. In this work we will usually assume that $t = O(1)$, so $\beta = \Omega(1)$. We assume throughout that Alice and Bob do not have access to shared randomness or entanglement. The classical complexity in the above theorem can actually be improved to an additive dependence on $\log n$ via applying Newman's Theorem [20] to a protocol with shared randomness, but at the expense of making the protocol less intuitive.

The classical upper bound stated above comes from the observation that, if f has a sign-representing polynomial p of degree 1, it is possible to determine whether $f(z) = 1$ with probability $> 1/2$ by only evaluating f on one uniformly random bit of z, by writing down a probabilistic procedure whose expectation on z mimics $p(z)$. So Alice sends a few uniformly random bits to Bob, who matches them to blocks in his partition, and evaluates f on the corresponding blocks with success probability $> 1/2$ for each block. Only a few repetitions are required to determine whether $f(x) = w$ or $f(x) = \overline{w}$ with high probability.

On the other hand, to obtain the quantum upper bound we use the idea of *block-multilinear* polynomials from [1, 2], and some auxiliary results also from [2]. The idea is that Alice sends a superposition of her bits, and Bob, after collapsing the state onto one of the blocks from his partition (say block j), applies a controlled unitary operator that describes a block-multilinear polynomial $\widetilde{p}$ of degree 2, which is produced from a sign-representing polynomial p for f of degree 2. A Hadamard test is used to return an output with probability depending (roughly speaking) on $\widetilde{p}(\sigma(x)^{(j)}, \sigma(x)^{(j)})$, which in turn is equal to $p(\sigma(x)^{(j)})$ according to a theorem from [2]. The Hadamard test then outputs 1 with probability greater than $1/2$ if $f(x^{(j)}) = 1$ and 0 with probability greater than $1/2$ if $f(x^{(j)}) = -1$.

We remark that both of these protocols actually solve a natural generalization of the Hidden Matching problem [4] (i.e. they output the result of evaluating $f(x^{(j)})$ for Bob's block j, where j is arbitrary), which is at least as hard as the f-Boolean Hidden Partition problem. However, unlike the Hidden Matching problem, the output is not correct with certainty, but only with probability strictly greater than $1/2$.

In Section 2 we reduce the f-Boolean Hidden Partition problem from the Boolean Hidden Matching problem and prove that for almost all symmetric Boolean function f with $sdeg(f) \geq 2$ the f-BHP$_n^{\alpha,t}$ problem require at least $\Omega(\sqrt{n})$ bits of communication. The only functions for which the reduction does not work are the Not All Equal functions on an odd number of bits, i.e., NAE : $\{-1,1\}^t \to \{-1,1\}$, defined by NAE$(x) = -1$ if $|x| \in \{0, t\}$ and NAE$(x) = 1$ otherwise, with t odd.

▶ **Theorem 7.** *Let $f : \{-1,1\}^t \to \{-1,1\}$ be a symmetric Boolean function with $sdeg(f) \geq 2$. If f is not the NAE function on an odd number of bits, then any bounded-error classical communication protocol for solving the f-BHP$_n^{\alpha,t}$ problem needs to communicate at least $\Omega(\sqrt{n/(\alpha t)})$ bits.*

Finally, we generalize the Fourier analysis methods from [15, 26, 24] to prove a partial result on the hardness of the f-BHP$_n^{\alpha,t}$ problem, both classically and quantumly. Ideally we would like to prove that any bounded-error classical and quantum protocols would need to communicate $\Omega(n^{1-1/d})$ bits and $\Omega(n^{1-2/d})$ qubits, respectively, where $sdeg(f) = d$. What we obtained is this result but with d being the *pure high degree* of f. A Boolean function f is said to have pure high degree ($phdeg(f)$) d if $\widehat{f}(S) = 0$ for all $|S| = 0, 1, \ldots, d-1$, where $\widehat{f}(S) = \langle f, \chi_S \rangle = \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} f(x)\chi_S(x)$ is the Fourier transform of f and $\chi_S(x) = \prod_{i \in S} x_i$, with $S \subseteq [n]$, is a character function. It is possible to prove that $phdeg(f) \leq sdeg(f)$, so our result is a step towards proving a lower bound for all functions with sign degree ≥ 2.

▶ **Theorem 8.** *Let $f : \{-1,1\}^t \to \{-1,1\}$ be a Boolean function. If $phdeg(f) = d \geq 2$, then, for sufficiently small $\alpha > 0$ that does not depend on n, any bounded-error classical communication protocol for solving the f-BHP$_n^{\alpha,t}$ problem needs to communicate at least $\Omega(n^{1-1/d})$ bits.*

▶ **Theorem 9.** *Let $f : \{-1,1\}^t \to \{-1,1\}$ be a Boolean function. If $phdeg(f) = d \geq 3$, then, for sufficiently small $\alpha > 0$ that does not depend on n, any bounded-error quantum communication protocol for solving the f-BHP$_n^{\alpha,t}$ problem needs to communicate at least $\Omega(n^{1-2/d})$ qubits.*

The above lower bounds are proved in [11]. The classical proof follows the general idea from [15, 26], but the technical execution was substantially changed by borrowing ideas from [24]. First, we apply Yao's minimax principle [27], which says that it suffices to prove a lower bound for a *deterministic* protocol under a hard probability distribution on Alice and

Bob's inputs. We choose Alice's input x and Bob's input σ independently and uniformly over $\{-1,1\}^n$ and $\mathbb{S}_n$ (the set of all permutations on $[n]$), respectively. The input distribution is completed by choosing $w = B_f(x)$ with probability $1/2$ and $w = \overline{B_f(x)}$ with probability $1/2$.

Alice sends a message to Bob. If the length of the message sent is c, then the inputs for which Alice could have sent that specific message define a set A of about 2^{n-c} x's. From Bob's perspective, he knows that the random variable X corresponding to Alice's bit-string is uniformly distributed in a set A and he knows his permutation σ, hence his knowledge of the random variable $B_f(X)$ is described by the distributions

$$p_\sigma(z) = \frac{|\{x \in A | B_f(x) = z\}|}{|A|} \text{ and } q_\sigma(z) = \frac{|\{x \in A | B_f(x) = \overline{z}\}|}{|A|}.$$

It is well known that the best success probability for distinguishing two distributions q_1 and q_2 with one sample is $1/2 + \|q_1 - q_2\|_{\text{tvd}}/4$. Therefore the bias of the protocol, i.e., the protocol's successful probability minus a half, is equal to the total variation distance between p_σ and q_σ. Differently from the approach of [15, 26], and following [24], we directly upper bound the expectation of the bias over Bob's permutation. By demanding a small distributional error, we arrive at the desired communication lower bound. Upper bounding the bias is done via Fourier analysis, using the inequality of Kahn, Kalai, and Linial [17].

The quantum proof follows the same idea from [24]. Yao's minimax principle is still applied and the "hard" input distribution is still uniform on Alice's input $x \in \{-1,1\}^n$, Bob's input $\sigma \in \mathbb{S}_n$ and the function value $b \in \{-1,1\}$, which fixes Bob's second input $w = B_f(x) \circ b^{\alpha n/t}$. The best strategy for Bob in determining b conditioned on his input (σ, w) is no more than the chance to distinguish between two subsets of Alice's messages, where a message corresponds to a quantum state ρ_x, selected according to b. In other words, no more than the chance to distinguish between the following $\rho_0^{\sigma,w}$ and $\rho_1^{\sigma,w}$, each appearing with probability $\Pr[b=0|\sigma,w]$ and $\Pr[b=1|\sigma,w]$, respectively,

$$\rho_0^{\sigma,w} = \frac{\sum_{x\in\{-1,1\}^n} \Pr[x,0,\sigma,w]\rho_x}{\Pr[x,0,\sigma,w]} \text{ and } \rho_1^{\sigma,w} = \frac{\sum_{x\in\{-1,1\}^n} \Pr[x,1,\sigma,w]\rho_x}{\Pr[x,1,\sigma,w]}.$$

It is known that any protocol that tries to distinguish two quantum states ρ_0 and ρ_1 appearing with probability p and $1-p$, respectively, by a POVM has bias at most $\|p\rho_0 - (1-p)\rho_1\|_{\text{tr}}/2$ [16]. The bias is then upper bounded by using Fourier analysis of matrix-valued functions, in particular by the matrix-valued hypercontractive inequality of Ben-Aroya, Regev, and de Wolf [6].

The difference between the classical and quantum lower bound proofs was considerably reduced in our paper, e.g., the classical proof now relies less on the use of the Parseval's identity. Still some differences persist. Apart from the obvious generalization of Fourier analysis to matrix-valued functions, the Fourier analysis in the quantum lower bound proof is performed directly on the encoding messages and not on the pre-images of a fixed encoding message, since there is no clear quantum analogue of conditioning on a message. The main technical difficulty we faced compared to [15, 26] is that the Fourier coefficients of Bob's distributions $p_\sigma(z)$ and $q_\sigma(z)$ are not nicely related to just one Fourier coefficient of the characteristic function of A any more, but instead to a more complicated sum of many coefficients. This requires us to carefully bound various combinatorial terms occurring in the proof and to use our freedom to choose α fairly small.

In Section 3 we analyse the limitations of our techniques and show that under the uniform distribution, which was used as the "hard" distribution during the proof of Theorem 8, we cannot obtain a lower bound depending on the sign degree instead of the pure high degree.

We finally remark that the one-way communication complexity separations we found can easily be used to obtain corresponding separations in the streaming model, similarly to [15, 26].

2 Reductions from the Boolean Hidden Matching problem

As mentioned before, in [15] it was proved that the Boolean Hidden Partition problem using PARITY on 2 bits (aka the BHM problem) is hard to solve, i.e., $R^1(\text{BHM}) = \Omega(\sqrt{n/\alpha})$. With this result alone it is possible to prove that the f-Boolean Hidden Partition problem for almost any symmetric Boolean function with $sdeg(f) \geq 2$ is at least as hard to solve. This can be achieved via a simple reduction from the BHM problem to the f-BHP$_n^{\alpha,t}$ problem with symmetric functions, which we shall show in this section.

For this section, in a slight abuse of notation we define $|x| = |\{i : x_i = -1\}|$ to be the "Hamming weight" of x. Let $s, t \in \mathbb{N}$, with $s \leq t$. Consider a symmetric Boolean function $f_s : \{-1,1\}^t \to \{-1,1\}$ such that (without loss of generality) $f_s(1^n) = 1$ and

$$f_s(x) = \begin{cases} +1 \text{ if } 0 \leq |x| \leq \theta_1 \text{ or } & \theta_{2i} < |x| \leq \theta_{2i+1}, i = 1, 2 \ldots, \lfloor s/2 \rfloor, \\ -1 \text{ if } & \theta_{2j-1} < |x| \leq \theta_{2j}, j = 1, 2, \ldots, \lfloor (s+1)/2 \rfloor, \end{cases} \tag{1}$$

where $\theta_k \in \mathbb{N}$ for $k = 1, \ldots, s+1$ and $0 \leq \theta_1 < \cdots < \theta_s < \theta_{s+1} = t$ and $\theta_{k+1} - \theta_k \geq 1$ for all $k = 1, \ldots, s$. The following result from [3] tells us that $sdeg(f_s) = s$.

▶ **Lemma 10.** *(Lemma 2.6 from [3]) If f is a symmetric function, then $sdeg(f)$ is equal to the number of times f changes sign when expressed as a univariate function in $\sum_i x_i$.*

In order to reduce f_s-BHP$_n^{\alpha,t}$ from BHM we first need to reduce the function f_s from PARITY, i.e., we want that $\forall x' \in \{-1,1\}^2$, $\exists x \in \{-1,1\}^t$ such that $f_s(x) = \text{PARITY}(x')$. The key combinatorial step to achieve this is shown in the next Lemma.

▶ **Lemma 11.** *Let $f_s : \{-1,1\}^t \to \{-1,1\}$ be the symmetric Boolean function from Eq. 1 with $s \geq 2$ such that either $2|t$ or $\theta_2 - \theta_1 < t - 1$. Then there exists $a, b \in \mathbb{N}$ such that $\forall x' \in \{-1,1\}^2$, $\exists x \in \{-1,1\}^t$ such that $f_s(x) = \text{PARITY}(x')$ and $|x| = a|x'| + b$.*

Proof. The condition that $\forall x' \in \{-1,1\}^2$, $\exists x \in \{-1,1\}^t$ such that $f_s(x) = \text{PARITY}(x')$ and $|x| = a|x'| + b$ is equivalent to

$$\begin{cases} |x'| = 0 \implies f_s(b) = 1, \\ |x'| = 1 \implies f_s(a+b) = -1, \\ |x'| = 2 \implies f_s(2a+b) = 1. \end{cases} \tag{2}$$

We divide the proof into two cases: either there exists $k^* \in \{1, \ldots, s-1\}$ such that $\theta_{k^*+1} - \theta_{k^*}$ is odd or there does not exist such a k^*. Suppose first that such k^* exists. Without loss of generality we can assume that $f_s(x) = -1$ for $\theta_{k^*} < |x| \leq \theta_{k^*+1}$, otherwise we just flip the values of f_s. Then we set

$$\begin{cases} a = (\theta_{k^*+1} - \theta_{k^*} + 1)/2, \\ b = \theta_{k^*}. \end{cases}$$

First, $a, b \in \mathbb{N}$. Second, $a + b = (\theta_{k^*+1} + \theta_{k^*} + 1)/2$, hence $\theta_{k^*} < a + b \leq \theta_{k^*+1}$, since $\theta_{k^*+1} - \theta_{k^*} \geq 1$. And third, $2a + b = \theta_{k^*+1} + 1 \leq \theta_{k^*+2}$. Therefore all conditions from Eqs. 2 are satisfied.

Now suppose that for all $k = 1, \ldots, s-1$ we have $2|(\theta_{k+1} - \theta_k)$. Define the bit $\delta = [\theta_1 \neq 0]$ and set

$$\begin{cases} a = (\theta_2 - \theta_1 + 2)/2, \\ b = \theta_1 - \delta. \end{cases}$$

First, $a, b \in \mathbb{N}$ (note that $\delta = 1 \implies \theta_1 > 0$). Second, $a + b = (\theta_2 + \theta_1 + 2 - 2\delta)/2$, hence $\theta_1 < a + b \leq \theta_2$, since $\theta_2 - \theta_1 \geq 2$ by hypothesis. And third, $2a + b = \theta_2 + 2 - \delta \leq t$ since $\theta_2 - \theta_1 < t - 1$ and $\theta_2 < t$ (so that $\theta_2 = t - 1 \implies \delta = 1$). Therefore all conditions from Eqs. 2 are satisfied. ◀

If $2 \nmid t$ and $\theta_2 - \theta_1 = t - 1$, then our conditions give us

$$\begin{cases} b = 0, \\ 0 < a < t, \\ 2a = t, \end{cases}$$

and we see that the condition $2a = t$ cannot be fulfilled by $a \in \mathbb{N}$. This case corresponds to the symmetric Boolean function Not All Equal (NAE), defined by $\mathrm{NAE}(x) = 1$ if $|x| \in \{0, t\}$ and $\mathrm{NAE}(x) = -1$ otherwise, with t odd.

Given the reduction above from PARITY to f_s, we can construct our reduction from the BHM problem to the f_s-$\mathrm{BHP}_n^{\alpha,t}$ problem.

▶ **Theorem 7.** *Let $f_s : \{-1,1\}^t \to \{-1,1\}$ be the symmetric Boolean function from Eq. 1 with $s \geq 2$ such that either $2|t$ or $\theta_2 - \theta_1 < t - 1$. Then $R^1(f_s\text{-}\mathrm{BHP}_n^{\alpha,t}) = \Omega(\sqrt{n/(\alpha t)})$.*

Proof. Suppose by contradiction that $R^1(f_s\text{-}\mathrm{BHP}_n^{\alpha,t}) = o(\sqrt{n/(\alpha t)})$, i.e., there exists a protocol Π that solves f_s-$\mathrm{BHP}_n^{\alpha,t}$ with $o(\sqrt{n/(\alpha t)})$ bits of communication. We are going to show that such protocol would allow Alice and Bob to solve the BHM problem with $o(\sqrt{n/\alpha})$ bits of communication, which leads to a contradiction.

Let $a, b \in \mathbb{N}$ be the numbers used in reducing f_s from PARITY in Lemma 11. Alice increases her bit string $x \in \{-1,1\}^n$ as follows: she makes a copies of x, obtaining $x^a \in \{-1,1\}^{an}$, where $x^a = xx\cdots x$ represents x repeated a times. She then adds $bn/2$ times the bit 1, obtaining $x^a 1^{bn/2}$. Finally, she adds $(t - 2a - b)n/2$ times the bit -1, to finally obtain $x_f = x^a 1^{bn/2} \text{-} 1^{(t-2a-b)n/2}$. Note that $x_f \in \{-1,1\}^{nt/2}$.

Bob, on the other hand, increases his permutation $\sigma \in S_n$ to a new permutation $\sigma_f \in S_{nt/2}$. In order to describe how he does this, we ease the notation by referring to the j-th block $(\pi^{-1}((j-1)t+1), \ldots, \pi^{-1}(jt))$ of a given permutation π as $(B_{j,1}, \ldots, B_{j,t})$. With this notation, the j-th block $(B_{j,1}, B_{j,2})$ of the permutation σ is mapped to the j-th block

$$\Big(B_{j,1}, B_{j,2}, n + B_{j,1}, n + B_{j,2}, \ldots, (a-1)n + B_{j,1}, (a-1)n + B_{j,2}, \\ an + j, an + j + \frac{n}{2}, \ldots, an + j + (t - 2a - 1)\frac{n}{2}\Big)$$

of the new permutation σ_f. Note that the new block has t elements, as expected.

Consider the block strings $\sigma_f(x_f)^{(j;t)} \in \{-1,1\}^t$ and $\sigma(x)^{(j;2)} \in \{-1,1\}^2$, with $j = 1, \ldots, n/2$. By construction we have that $|\sigma_f(x_f)^{(j;t)}| = a|\sigma(x)^{(j;2)}| + b$ and, according to Lemma 11, we get $f_s(\sigma_f(x_f)^{(j;t)}) = \mathrm{PARITY}(\sigma(x)^{(j;2)})$ for all $j = 1, \ldots, n/2$. Hence we see that every instance of the problem $\mathrm{BHM} : \{-1,1\}^n \to \{-1,1\}$ is mapped to an instance of the problem f_s-$\mathrm{BHP}_n^{\alpha,t} : \{-1,1\}^{nt/2} \to \{-1,1\}$. Therefore we could map the BHM problem into the f_s-$\mathrm{BHP}_n^{\alpha,t}$ problem and use the protocol Π in order to solve it with $o(\sqrt{n/(\alpha t)})$ bits of communication, which is impossible. Thus $R^1(f_s\text{-}\mathrm{BHP}_n^{\alpha,t}) = \Omega(\sqrt{n/(\alpha t)})$. ◀

3 Limitations of proof technique

Theorem 8 guarantees the classical hardness of the f-BHP$_n^{\alpha,t}$ problem if f has pure high degree ≥ 2, and not sign degree ≥ 2, which would be a stronger result. To arrive at this result, we used the uniform distribution as a "hard" distribution for Yao's principle. In this section we shall prove that under the uniform distribution we cannot obtain a better result. More specifically, we shall prove that under the uniform distribution there is an efficient bounded-error classical protocol for solving the f-BHP$_n^{\alpha,t}$ problem if $phdeg(f) \leq 1$.

▶ **Theorem 12.** *Under the uniform distribution for Alice and Bob's inputs, if* $phdeg(f) \leq 1$ *then* $R^1(f\text{-BHP}_n^{\alpha,t}) = O\left(\frac{t^2}{\alpha}\log n\right)$.

Proof. Let $F = \{i \in [t] \mid \widehat{f}(\{i\}) \neq 0\}$. Given that $phdeg(f) \leq 1$, this set is non-empty. Consider the following protocol: Alice picks a subset $I \subseteq [n]$ of indices uniformly at random using shared randomness, where $|I|$ will be determined later, and sends the indices and corresponding bitvalues to Bob. Let $\{x_i\}_{i\in I}$ be the bitvalues sent, and let $j(i) = \lceil \sigma(i)/t \rceil$ and $k(i) \equiv \sigma(i) \mod t$ for all $i \in I$, where $\sigma \in S_n$ is Bob's permutation. The probability that none of the indices sent by Alice are matched to a non-zero Fourier coefficient according to Bob's permutation, within one of the $\alpha n/t$ blocks he has, is

$$\Pr_\sigma[k(i) \notin F,\ \forall i \in I] \leq \left(1 - \alpha\frac{|F|}{t}\right)^{|I|} \leq e^{-\alpha|I||F|/t}$$

which we can make almost arbitrarily small by choosing $|I|$ to be sufficiently large. (Note that the first inequality above would be an equality if we chose the elements of I with replacement, and choosing them without replacement cannot make $\Pr[k(i) \notin F,\ \forall i \in I]$ higher.) Hence with high probability $I \cap F \cap [\alpha n/t] \neq \emptyset$. Choose some $\ell \in I \cap F \cap [\alpha n/t]$. Bob computes $\mathrm{sgn}[\widehat{f}(\{k(\ell)\})] \cdot \sigma(x)_{k(\ell)}^{(j(\ell))} \cdot w_{j(\ell)}$: if it is $+1$, then he outputs that $B_f(x) = w$, and if it is -1, then he outputs that $B_f(x) = \overline{w}$.

To see why the protocol works, we calculate the probability that $\mathrm{sgn}[\widehat{f}(\{k(\ell)\})] \cdot \sigma(x)_{k(\ell)}^{(j(\ell))}$ is equal to $f(\sigma(x)^{(j(\ell))})$.

$$\begin{aligned}
&\Pr_x\left[\mathrm{sgn}[\widehat{f}(\{k(\ell)\})]\sigma(x)_{k(\ell)}^{(j(\ell))} = f(\sigma(x)^{(j(\ell))})\right] = \\
&\quad = \frac{1}{2} + \frac{1}{2^{t+1}} \sum_{x\in\{-1,1\}^t} \mathrm{sgn}[\widehat{f}(\{k(\ell)\})]\sigma(x)_{k(\ell)}^{(j(\ell))} f(\sigma(x)^{(j(\ell))}) \\
&\quad = \frac{1}{2} + \frac{1}{2}\,\mathrm{sgn}[\widehat{f}(\{k(\ell)\})] \cdot \widehat{f}(\{k(\ell)\}) \\
&\quad = \frac{1}{2} + \frac{1}{2}|\widehat{f}(\{k(\ell)\})|,
\end{aligned}$$

which is greater than $1/2$ and where we used in the first line that the distribution on Alice's inputs is uniform. Therefore, by a union bound, for sufficiently large $|I| = O\left(\frac{t}{\alpha}\log\frac{1}{|\widehat{f}(\{k(\ell)\})|}\right)$, the overall success probability of the protocol (i.e. $I \cap F \cap [\alpha n/t] \neq \emptyset$ and Bob's output equals f) is strictly greater than $1/2$. Since $|\widehat{f}(\{k(\ell)\})| \geq 2^{1-t}$ (as it is nonzero and is an average of 2^t ± 1's), this gives us the final overhead of $O(t^2/\alpha)$. ◀

4 Conclusions

We proposed a very broad generalization of the famous Boolean Hidden (Hyper)Matching problem, which we called the f-Boolean Hidden Partition (f-BHP$_n^{\alpha,t}$) problem. Instead of using the Parity function to arrive at the final bit-string that Alice and Bob wish to explore, we use a generic Boolean function f. We partially characterize the communication complexity of the whole problem in terms of one property of f: its sign degree. We proved that if $sdeg(f) \leq 1$, then there exists an efficient bounded-error classical protocol that solves the f-BHP$_n^{\alpha,t}$ with $O(\log n)$ bits. Similarly to the classical case, we proved that if $sdeg(f) \leq 2$, then there exists an efficient bounded-error quantum protocol that solves the f-BHP$_n^{\alpha,t}$ with $O(\log n)$ qubits. We then pursued a classical-quantum communication gap by proving classical and quantum lower bounds for cases of the problem where $sdeg(f) \geq 2$. First we noted that the f-BHP$_n^{\alpha,t}$ problem is hard for almost all symmetric functions with $sdeg(f) \geq 2$ via a simple reduction from the Boolean Hidden Matching problem. And second we generalized previous communication complexity lower bounds based on Fourier analysis to prove that functions with $phdeg(f) = d \geq 2$ lead to a classical $\Omega(n^{1-1/d})$ communication cost and functions with $phdeg(f) = d \geq 3$ lead to a quantum $\Omega(n^{1-2/d})$ communication cost for the f-BHP$_n^{\alpha,t}$ problem.

It is known that $phdeg(f) \leq sdeg(f)$, but our lower bounds are probably not tight for *all* functions with sign degree ≥ 2. We proved that this is an inherent limitation of the chosen distribution for Alice and Bob's inputs during the proof, since under the uniform distribution it is possible to solve the problem with $O(\log n)$ bits of communication if $phdeg(f) \leq 1$. We then make the following conjectures.

▶ **Conjecture 13.** $R_\epsilon^1(f\text{-BHP}_n^{\alpha,t}) = \Omega(n^{1-1/d})$ *if* $sdeg(f) = d \geq 2$.

▶ **Conjecture 14.** $Q_\epsilon^1(f\text{-BHP}_n^{\alpha,t}) = \Omega(n^{1-2/d})$ *if* $sdeg(f) = d \geq 3$.

A proof of these results would require a non-uniform distribution on Alice and Bob's inputs.

We hope that these conjectures help motivate the development of necessary quantum lower bound techniques.

References

1 Scott Aaronson and Andris Ambainis. Forrelation: A problem that optimally separates quantum from classical computing. *SIAM Journal on Computing*, 47(3):982–1038, 2018. arXiv:1411.5729.

2 Scott Aaronson, Andris Ambainis, Janis Iraids, Martins Kokainis, and Juris Smotrovs. Polynomials, quantum query complexity, and Grothendieck's inequality. In *31st Conference on Computational Complexity*, 2016. arXiv:1511.08682.

3 James Aspnes, Richard Beigel, Merrick Furst, and Steven Rudich. The expressive power of voting polynomials. *Combinatorica*, 14(2):135–148, 1994.

4 Ziv Bar-Yossef, Thathachar S. Jayram, and Iordanis Kerenidis. Exponential separation of quantum and classical one-way communication complexity. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pages 128–137. ACM, 2004.

5 William Beckner. Inequalities in Fourier analysis. *Ann. of Math.*, 102:159–182, 1975.

6 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. arXiv:0705.3806.

7 Aline Bonami. Étude des coefficients Fourier des fonctions de $L^p(G)$. In *Annales de l'institut Fourier*, volume 20(2), pages 335–402, 1970.

8 Harry Buhrman, Richard Cleve, Serge Massar, and Ronald de Wolf. Nonlocality and communication complexity. *Reviews of modern physics*, 82(1):665, 2010. arXiv:0907.3584.

9 Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. *Physical Review Letters*, 87(16):167902, 2001. quant-ph/0102001.

10 Harry Buhrman, Nikolay Vereshchagin, and Ronald de Wolf. On computation and communication with small bias. In *Proc. 22nd Annual IEEE Conf. Computational Complexity*, pages 24–32, 2007.

11 João F. Doriguello and Ashley Montanaro. Exponential quantum communication reductions from generalizations of the boolean hidden matching problem. *arXiv preprint arXiv:2001.05553*, 2020.

12 João Fernando Doriguello and Ashley Montanaro. Quantum sketching protocols for Hamming distance and beyond. *Phys. Rev. A*, 99:062331, 2019. arXiv:1810.12808.

13 Devdatt P. Dubhashi and Alessandro Panconesi. *Concentration of measure for the analysis of randomized algorithms*. Cambridge University Press, 2009.

14 Dmitry Gavinsky. Classical interaction cannot replace a quantum message. In *Proc. 40th Annual ACM Symp. Theory of Computing*, pages 95–102, 2008. quant-ph/0703215.

15 Dmitry Gavinsky, Julia Kempe, Iordanis Kerenidis, Ran Raz, and Ronald de Wolf. Exponential separations for one-way quantum communication complexity, with applications to cryptography. *SIAM Journal on Computing*, 38(5):1695–1708, 2008. quant-ph/0611209.

16 Carl Wilhelm Helstrom. *Quantum detection and estimation theory*. Academic press, 1976.

17 Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on boolean functions. In *Proc. 29th Annual Symp. Foundations of Computer Science*, pages 68–80. IEEE, 1988.

18 Hartmut Klauck. Lower bounds for quantum communication complexity. In *Proceedings 42nd IEEE Symposium on Foundations of Computer Science*, pages 288–297. IEEE, 2001. quant-ph/0106160.

19 Ashley Montanaro. A new exponential separation between quantum and classical one-way communication complexity. *Quantum Inf. Comput.*, 11(7&8):574–591, 2011. arXiv:1007.3587.

20 Ilan Newman. Private vs. common random bits in communication complexity. *Information Processing Letters*, 39(2):67–71, 1991.

21 Ran Raz. Fourier analysis for probabilistic communication complexity. *Computational Complexity*, 5(3-4):205–221, 1995.

22 Ran Raz. Exponential separation of quantum and classical communication complexity. In *STOC*, volume 99, pages 358–367. Citeseer, 1999.

23 Oded Regev and Bo'az Klartag. Quantum one-way communication can be exponentially stronger than classical communication. In *Proc. 43rd Annual ACM Symp. Theory of Computing*, pages 31–40, 2011. arXiv:1009.3640.

24 Yaoyun Shi, Xiaodi Wu, and Wei Yu. Limits of quantum one-way communication by matrix hypercontractive inequality, 2012.

25 Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26:1484–1509, 1997. quant-ph/9508027.

26 Elad Verbin and Wei Yu. The streaming complexity of cycle counting, sorting by reversals, and other problems. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 11–25. Society for Industrial and Applied Mathematics, 2011.

27 Andrew Chi-Chih Yao. Probabilistic computations: Toward a unified measure of complexity. In *18th Annual Symposium on Foundations of Computer Science (sfcs 1977)*, pages 222–227. IEEE, 1977.

28 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. ACM, 1979.

29 Andrew Chi-Chih Yao. Quantum circuit complexity. In *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science*, pages 352–361. IEEE, 1993.

A Proof of Upper Bounds

In this and the following appendices, denote by $R^1_\epsilon(\mathcal{P})$ and $Q^1_\epsilon(\mathcal{P})$ the classical and quantum communication cost of the protocol $\mathcal{P}$ in bits and qubits, respectively, and denote by $R^1_\epsilon(f) = \min_{\mathcal{P}} R^1_\epsilon(\mathcal{P})$ and $Q^1_\epsilon(f) = \min_{\mathcal{P}} Q^1_\epsilon(\mathcal{P})$ the minimum classical and quantum communication cost, respectively, over all one-way protocols $\mathcal{P}$ without shared randomness that solve a communication problem f with failure probability $0 < \epsilon < 1/2$.

A.1 Classical Upper Bound

Consider the f-BHP$^{\alpha,t}_n$ problem for $f : \{-1,1\}^t \to \{-1,1\}$ with $sdeg(f) \leq 1$. Now let $p : \{-1,1\}^t \to [-1,1]$ be a normalized sign-representing polynomial for f. Hence we can write

$$p(x) = \alpha_0 + \sum_{i=1}^{t} \alpha_i x_i$$

with $(\alpha_i)_{i=0}^t \in \mathbb{R}$. Let $\beta = \min_x |p(x)|$ be the bias of p.

▶ **Theorem 5.** $R^1_\epsilon(f\text{-BHP}^{\alpha,t}_n) = O\left((\frac{t}{\alpha\beta})^2 \log\frac{1}{\epsilon} \log n\right)$ *if* $sdeg(f) \leq 1$.

Proof. Consider the following protocol: Alice picks $m = O\big((\frac{t}{\alpha\beta})^2 \log\frac{1}{\epsilon}\big)$ bits from x uniformly at random (with replacement) and sends them to Bob, together with their indices. Let I and $\{x_i\}_{i\in I}$ be the indices and bitvalues sent, respectively. Let $j(i) = \lceil \sigma(i)/t \rceil$ and $k(i) \equiv \sigma(i) \mod t$ for all $i \in I$, where $\sigma \in S_n$ is Bob's permutation. Define the random variable $X(i) = (\alpha_{k(i)} x_i + \alpha_0/t) w_{j(i)}$ if $\sigma(i) \in [\alpha n/t]$ and $X(i) = 0$ if $\sigma(i) \notin [\alpha n/t]$, where α_0 and α_k are the zeroth order and x_k's coefficients, respectively, from the sign-representing polynomial p, and define $X = \sum_{i\in I} X(i)$. Bob then computes $\mathrm{sgn}(X)$. If the sign is 1, then he outputs $B_f(x) = w$, and if the sign is -1, then he outputs $B_f(x) = \overline{w}$.

To see why the protocol works, we calculate the expectation value of random variable X.

$$\begin{aligned}
\mathbb{E}[X] &= m \cdot \mathbb{E}_i[X(i)] \\
&= \alpha m \cdot \mathbb{E}_i[(\alpha_{k(i)} x_i + \alpha_0/t) w_{j(i)}] \\
&= \alpha m \cdot \mathbb{E}_j\big[\mathbb{E}_k[\alpha_k \sigma(x)^{(j)}_k + \alpha_0/t] w_j\big] \\
&= \alpha m \cdot \mathbb{E}_j\left[\frac{p(\sigma(x)^{(j)})}{t} w_j\right] \\
&= \alpha m \frac{t}{n} \sum_{j=1}^{n/t} \frac{p(\sigma(x)^{(j)})}{t} w_j \\
&= \frac{\alpha m}{n}\left[\sum_{j:w_j=1} p(\sigma(x)^{(j)}) - \sum_{j:w_j=-1} p(\sigma(x)^{(j)})\right].
\end{aligned}$$

If $f(\sigma(x)^{(j)}) = w_j$, then $w_j = 1 \implies p(\sigma(x)^{(j)}) \geq \beta > 0$ and $w_j = -1 \implies p(\sigma(x)^{(j)}) \leq -\beta < 0$. Therefore

$$\mathbb{E}[X] \geq \frac{\alpha m}{n}\left[\sum_{j:w_j=0} \beta - \sum_{j:w_j=1} -\beta\right] = \alpha m \frac{\beta}{t}.$$

If, on the other hand, $f(\sigma(x)^{(j)}) = -w_j$, then $w_j = 1 \implies p(\sigma(x)^{(j)}) \leq -\beta < 0$ and $w_j = -1 \implies p(\sigma(x)^{(j)}) \geq \beta > 0$. Therefore

$$\mathbb{E}[X] \leq \frac{\alpha m}{n}\left[\sum_{j:w_j=0} -\beta - \sum_{j:w_j=1} \beta\right] = -\alpha m \frac{\beta}{t}.$$

By using a Chernoff bound [13] of the type $\Pr[X > \mathbb{E}[X] + u], \Pr[X < \mathbb{E}[X] - u] \leq e^{-2u^2/m}$ with $u > 0$ and setting $u = \pm\mathbb{E}[X] > 0$, we can make

$$\Pr[X > 0 \mid B_f(x) = \overline{w}],\ \Pr[X < 0 \mid B_f(x) = w] \leq \epsilon$$

by taking $m = O\big((\frac{t}{\alpha\beta})^2 \log \frac{1}{\epsilon}\big)$. Therefore Alice and Bob can decide if $B_f(x) = w$ or $B_f(x) = \overline{w}$ with error probability ϵ and $O\big((\frac{t}{\alpha\beta})^2 \log \frac{1}{\epsilon} \log n\big)$ bits of communication. ◀

A.2 Quantum Upper Bound

Consider the f-BHP$_n^{\alpha,t}$ problem for $f : \{-1,1\}^t \to \{-1,1\}$ with $sdeg(f) = 2$. Let $p : \{-1,1\}^t \to [-1,1]$ be a normalized sign-representing polynomial for f. Let $\beta = \min_x |p(x)|$ be the bias of p.

We say that a polynomial q of degree k is block-multilinear if its variables $x_1, \ldots, x_N$ can be partitioned into k blocks $R_1, \ldots, R_k$, such that every monomial of q contains exactly one variable from each block. As a special case, a block-multilinear polynomial q of degree 2 can be written as

$$q(x_1, \ldots, x_n, y_1, \ldots, y_m) = \sum_{\substack{i\in[n]\\ j\in[m]}} a_{ij} x_i y_j$$

with variables in the first block labeled as $x_1, \ldots, x_n$ and the variables in the second block labeled as $y_1, \ldots, y_m$. Defining the matrix $A = (a_{ij})_{i\in[n],j\in[m]}$, then

$$q(x, y) = x^T A y$$

for all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. We say that q is *bounded* if $|q(x, y)| \leq 1$ for all $x \in \{-1,1\}^n, y \in \{-1,1\}^m$. This translates to

$$\max_{\substack{x\in\{-1,1\}^n\\ y\in\{-1,1\}^m}} \left|\sum_{\substack{i\in[n]\\ j\in[m]}} a_{ij} x_i y_j\right| \leq 1,$$

i.e., $\|A\|_{\infty\to 1} \leq 1$.

In order to prove the quantum upper bound, we will need the following results. In what comes, define $\widetilde{x} = (1, x_1, \ldots, x_t)$.

▶ **Lemma 15** ([2]). *Given a $m \times m$ complex matrix M, there exists a unitary U (on a possibly larger space with basis $|1\rangle, \ldots, |k\rangle$ for some $k \geq m$) such that, for any unit vector $|y\rangle = \sum_{i=1}^m \alpha_i |i\rangle$, $U|y\rangle = \frac{M|y\rangle}{\|M\|} + |\phi\rangle$, where $|\phi\rangle$ consists of basis states $|i\rangle$, $i > m$ only.*

▶ **Theorem 16** ([2]). *Let $p : \{-1,1\}^t \to [-1,1]$ be a sign-representing polynomial for f with $sdeg(f) = 2$. Then there is a block-multilinear polynomial $\widetilde{p} : \mathbb{R}^{2(t+1)} \to \mathbb{R}$ such that $\widetilde{p}(\widetilde{x}, \widetilde{x}) = p(x)$ for any $x \in \{-1,1\}^t$, and $|\widetilde{p}(y)| \leq 3$ for any $y \in \{-1,1\}^{2(t+1)}$.*

Let $\widetilde{p} : \mathbb{R}^{2(t+1)} \to \mathbb{R}$ be the block-multilinear polynomial of degree 2 obtained from the sign-representing polynomial p of f according to Theorem 16. It can be written as

$$\widetilde{p}(x,y) = \sum_{i,j\in[t+1]} a_{ij}x_iy_j = x^TAy, \tag{3}$$

where $A = (a_{ij})_{i,j\in[t+1]}$.

With these in hands, we present our upper bound.

▶ **Theorem 6.** $Q^1_\epsilon(f\text{-BHP}^{\alpha,t}_n) = O\left((\frac{t}{\alpha\beta})^2 \log\frac{1}{\epsilon}\log n\right)$ *if* $sdeg(f) \leq 2$.

Proof. Consider the following protocol: Alice sends to Bob $m = O\big((\frac{t}{\alpha\beta})^2\log\frac{1}{\epsilon}\big)$ copies of the quantum state of $O(\log n)$ qubits

$$|\psi_A\rangle = \frac{1}{\sqrt{n+n/t}}\left(\sum_{i=1}^{n} x_i|i\rangle + \sum_{i=1}^{n/t}|n+i\rangle\right).$$

Bob measures each of them by using the POVM

$$\left\{|n+j\rangle\langle n+j| + \sum_{i=(j-1)t+1}^{jt}|\sigma^{-1}(i)\rangle\langle\sigma^{-1}(i)|\right\}_{j\in[n/t]},$$

where $\sigma \in S_n$ is his permutation, and attaches a qubit in the state $|+\rangle$ to each of the final states. Let $I \subseteq [n/t]$ be the sequence of indices from his measurements. Then his final state is

$$|\psi_B\rangle = \bigotimes_{j\in I}|+\rangle|\psi^{(j)}\rangle,$$

where

$$|\psi^{(j)}\rangle = \frac{1}{\sqrt{t+1}}\left(|n+j\rangle + \sum_{i=(j-1)t+1}^{jt} x_{\sigma^{-1}(i)}|\sigma^{-1}(i)\rangle\right).$$

Let A be the $(t+1)\times(t+1)$ matrix from the representation of $\widetilde{p}$ according to Eq. 3. Lemma 15 guarantees the existence of a unitary U_j such that $U_j|\psi^{(j)}\rangle = \frac{A|\psi^{(j)}\rangle}{\|A\|} + |\phi^{(j)}\rangle$, with $\langle\phi^{(j)}|\psi^{(j)}\rangle = 0$. Bob then applies a controlled U_j gate onto each $|+\rangle_j|\psi^{(j)}\rangle$ to obtain

$$\bigotimes_{j\in I} CU_j|\psi_B\rangle = \bigotimes_{j\in I}\left(\frac{1}{\sqrt{2}}|0\rangle|\psi^{(j)}\rangle + \frac{1}{\sqrt{2}}|1\rangle U_j|\psi^{(j)}\rangle\right)$$

and then performs a Hadamard gate on the first qubit of each of the subsystems I and measures them. Let $m_j \in \{0,1\}$ be the result of the measurement for block $j \in I$. Define the random variable $X(j) = -(-1)^{m_j}w_j$ if $j \in [\alpha n/t]$ and $X(j) = 0$ if $j \notin [\alpha n/t]$, and define $X = \sum_{j\in I} X(j)$. Bob then computes $\mathrm{sgn}(X)$: if $\mathrm{sgn}(X) > 0$, he outputs that $B_f(x) = w$, and if $\mathrm{sgn}(X) < 0$, he outputs that $B_f(x) = \overline{w}$.

To see why the protocol works, first note that the probability of measuring 1 is

$$\begin{aligned}\Pr[1] &= \frac{1}{2}\left(1 + \langle\psi^{(j)}|U|\psi^{(j)}\rangle\right) = \frac{1}{2}\left(1 + \frac{\langle\psi^{(j)}|A|\psi^{(j)}\rangle}{\|A\|}\right)\\ &= \frac{1}{2}\left(1 + \frac{\widetilde{p(\sigma(x)^{(j)}, \sigma(x)^{(j)})}}{\|A\|(t+1)}\right) = \frac{1}{2}\left(1 + \frac{p(\sigma(x)^{(j)})}{\|A\|(t+1)}\right).\end{aligned}$$

The remainder of the argument is similar to the classical upper bound proof. Recalling that $m = |I|$, the expectation value of X is

$$\begin{aligned}\mathbb{E}[X] &= m \cdot \mathbb{E}_j[X(j)] \\ &= \alpha m \cdot \mathbb{E}_j[-(-1)^{m_j} w_j] \\ &= \alpha m \frac{t}{n} \sum_{j=1}^{n/t} \left(\Pr[m_j = 1] - \Pr[m_j = 0]\right) w_j \\ &= \alpha m \frac{t}{n} \left[\sum_{j:w_j=1} \frac{p(\sigma(x)^{(j)})}{\|A\|(t+1)} - \sum_{j:w_j=-1} \frac{p(\sigma(x)^{(j)})}{\|A\|(t+1)} \right].\end{aligned}$$

If $f(\sigma(x)^{(j)}) = w_j$, then $w_j = 1 \implies p(\sigma(x)^{(j)}) \geq \beta > 0$ and $w_j = -1 \implies p(\sigma(x)^{(j)}) \leq -\beta < 0$. Therefore

$$\mathbb{E}[X] \geq \alpha m \frac{t}{n} \frac{1}{\|A\|(t+1)} \left[\sum_{j:w_j=1} \beta - \sum_{j:w_j=-1} -\beta \right] = \frac{\alpha m \beta}{\|A\|(t+1)}.$$

If, on the other hand, $f(\sigma(x)^{(j)}) = -w_j$, then $w_j = 1 \implies p(\sigma(x)^{(j)}) \leq -\beta < 0$ and $w_j = -1 \implies p(\sigma(x)^{(j)}) \geq \beta > 0$. Therefore

$$\mathbb{E}[X] \leq \alpha m \frac{t}{n} \frac{1}{\|A\|(t+1)} \left[\sum_{j:w_j=1} -\beta - \sum_{j:w_j=-1} \beta \right] = -\frac{\alpha m \beta}{\|A\|(t+1)}.$$

By using a Chernoff bound [13] of the type $\Pr[X > \mathbb{E}[X] + u], \Pr[X < \mathbb{E}[X] - u] \leq e^{-2u^2/m}$ with $u > 0$ and setting $u = \pm\mathbb{E}[X] > 0$, we can make

$$\Pr[X > 0 \mid B_f(x) = \overline{w}],\ \Pr[X < 0 \mid B_f(x) = w] \leq \epsilon$$

by taking $m = O\big((\frac{t}{\alpha\beta})^2 \log \frac{1}{\epsilon}\big)$, where we use that $\|A\| \leq \|A\|_{\infty\to 1} \leq 3$ according to Theorem 16 (note that $\frac{\|Ax\|_2}{\|x\|_2} \leq \frac{\|Ax\|_1}{\|x\|_\infty}$, and taking maximums over all x on both sides gives $\|A\| \leq \|A\|_{\infty\to 1}$). Therefore Alice and Bob can decide if $B_f(x) = w$ or $B_f(x) = \overline{w}$ with error probability ϵ and $O\big((\frac{t}{\alpha\beta})^2 \log \frac{1}{\epsilon} \log n\big)$ qubits of communication. ◀