Electronic Colloquium on Computational Complexity, Report No. 130 (2021)

Matrix hypercontractivity, streaming algorithms and LDCs: the large alphabet case

Srinivasan Arunachalam
IBM Quantum.
IBM T.J. Watson Research Center
Yorktown Heights, USA
Srinivasan.Arunachalam@ibm.com

João F. Doriguello
Center for Quantum Technologies
National University of Singapore, Singapore
joaofd@nus.edu.sg

September 7, 2021

Abstract

Hypercontractive inequalities for real-valued functions over the Boolean cube play an important role in theoretical computer science. In this work, we prove a hypercontractive inequality for matrix-valued functions defined over large alphabets, generalizing the result of Ben-Aroya, Regev, de Wolf (FOCS'08) for the Boolean alphabet. To obtain our result we prove a generalization of the powerful 2-uniform convexity inequality for trace norms of Ball, Carlen, Lieb (Inventiones Mathematicae'94). We give two applications of this hypercontractive inequality.

Locally decodable codes. We present a lower bound for locally decodable codes (LDC) over large alphabets. An LDC $C : \mathbb{Z}_r^n \to \mathbb{Z}_r^N$ is an encoding of $x \in \mathbb{Z}_r^n$ into a codeword $C(x)$ in such a way that one can recover an arbitrary $x_i \in \mathbb{Z}_r$ (with probability at least $1/r + \varepsilon$) by making only a few queries to a corrupted codeword. The main question in LDCs is the trade-off between N and n. By using hypercontractivity, we give an exponential lower bound $N = 2^{\Omega(\varepsilon^4 n/r^4)}$ for 2-query (possibly non-linear) LDCs over $\mathbb{Z}_r$. Previously exponential lower bounds were known for $r = 2$ (Kerenidis and de Wolf (JCSS'04)) and for *linear* codes (Dvir and Shpilka (SICOMP'07)).

Streaming algorithms. We present upper and lower bounds for the communication complexity of the Hidden Hypermatching problem when defined over large alphabets, which generalizes the well-known Boolean Hidden Matching problem. We then consider streaming algorithms for approximating the value of Unique Games on a t-hyperedge hypergraph: in this direction a simple edge-counting argument gives an r-approximation with $O(\log n)$ space. On the other hand, we use our communication lower bound to show that every streaming algorithm in the adversarial model achieving a $(r - \varepsilon)$-approximation of this value requires $\Omega(n^{1-1/t})$ classical space or $\Omega(n^{1-2/t})$ quantum space. In this setting, these results simplify and generalize the seminal work of Kapralov, Khanna and Sudan (SODA'15) and Kapravol and Krachun (STOC'19) for the case $r = 2$.

1 Introduction

In this paper we prove new results in two areas of theoretical computer science that have received a lot of attention recently: *streaming algorithms* and *locally decodable codes*.

Streaming algorithms is a model of computation introduced by Alon, Matias and Szegedy [AMS99] (for which they won the Gödel Prize in 2005) in order to understand the space complexity of ap-

ISSN 1433-8092

proximation algorithms to solve problems. In the last decade, there have been several results in the direction of proving upper and lower bounds for streaming algorithms for combinatorial optimization problems [VY11, GKK12, KKS14, GVV17, KKSV17, KK19, GT19, CGV20, CGSV21a, CGSV21b, AD21, CKP+21]. The goal here is to obtain a $1/\gamma$ approximation (for some $\gamma \leq 1$) of the optimum value of the combinatorial optimization problem with as little space as possible. One favourite problem considered by many works is the well-known *Max-Cut*, or its generalization over large alphabets $\mathbb{Z}_r$, *Unique Games*. Here, giving a 2-approximation algorithm for Max-Cut on n vertices can be done in logarithmic space, while a sequence of works [KKS14, KKSV17, KK19] showed that getting a $(2-\varepsilon)$-approximation requires linear space, matching the upper bound by [AGM12]. A similar, but less optimized, scenario was observed for Unique Games, i.e., there is a threshold behaviour in complexity going from r to $(r-\varepsilon)$-approximation. Curiously, many of these lower bounds were proven via variants of a problem called *Boolean Hidden Matching* (BHM) and it is well known that BHM can be solved using logarithmic *quantum space*, so a natural question is, could quantum space help solving these combinatorial optimization problems? One corollary from [KKS14, SWY12] is that obtaining the *strong* $(1+\varepsilon)$-approximation factor for Max-Cut and Unique Games streaming algorithms is quantum-hard. However, understanding the space complexity of streaming in the widely-studied, weaker regime of $(2-\varepsilon)$-approximation (for Maxcut) or $(r-\varepsilon)$-approximation (for Unique games over $\mathbb{Z}_r$) algorithms, it is still unclear whether there could be any savings in the quantum regime.

Locally decodable codes (LDCs) are error correcting codes $C : \Sigma^n \to \Gamma^N$ (for alphabets Σ, Γ) that allow transmission of information over noisy channels. By querying a few locations of a noisy codeword $\tilde{C}(x)$, one needs to reconstruct an arbitrary coordinate of $x \in \Sigma^n$ with probability at least $1/|\Sigma| + \varepsilon$. The main goal in this field is to understand trade-offs between N and n. LDCs have found several applications in pseudorandom generators, hardness amplification, private information retrieval schemes, cryptography, complexity theory (refer to [Yek12, Gop18] for a detailed exposition). Despite their ubiquity, LDCs are not well understood, even with the simplest of case of 2-*query* LDC*s*. For the case when $\Sigma = \Gamma = \{0,1\}$, exponential lower bounds of $N = 2^{\Omega(n)}$ were established over two decades back [GKST02, KW04, DS07]. In contrast, a breakthrough result of Dvir and Gopi [DG16] in 2015 showed how to construct 2-query LDCs with *subexponential* length in the regime when $\Sigma = \{0,1\}$ and Γ is a finite field $\mathbb{F}_N$. Despite these results, our knowledge of such N and n trade-offs for 2-query LDCs is still lacking, specially for the not very well studied case when $\Sigma = \Gamma = \mathbb{Z}_r$.

Prior works that handled simpler versions of the questions above used one technical tool successfully: *hypercontractivity* for real-valued functions over the Boolean cube. Since we are concerned with proving quantum lower bounds for streaming algorithms and establishing lower bounds for LDCs when the input alphabet is over $\mathbb{Z}_r$, it leads us to the following main question: *Is there a version of hypercontractivity for matrix-valued functions over* $\mathbb{Z}_r$*?*

1.1 Our results

Summarizing our main contributions, we first prove a version of hypercontractivity for matrix-valued functions $f : \mathbb{Z}_r^n \to \mathbb{C}^{m\times m}$. The proof of this crucially relies on proving uniform convexity for trace norms of r matrices, which in turn generalizes the powerful 2-uniform convexity by Ball, Carlen and Lieb [BCL94]. Using this new hypercontractivity theorem, we prove our two applications.

First, we prove a quantum space lower bound for streaming algorithms. It is easy to see that obtaining a 2-approximation algorithm for Max-k-Cut on n vertices in the classical streaming model

can be done in $O(\log n)$ space, and we show that obtaining a 1.99-approximation algorithm in the adversarial model requires $\Omega(n^{1-2/t})$ quantum space or $\Omega(n^{1-1/t})$ classical space. As far as we are aware, this is the first quantum space lower bound for an optimization problem. Although our lower bounds apply to the adversarial model, while prior works of Kapralov, Khanna and Sudan [KKS14] and the mathematical tour-de-force result of Kapralov and Krachun [KK19] obtained an $\Omega(n)$ classical space lower bound for $(2-\varepsilon)$-approximation in the *random* model, our proofs are significantly simpler. We further generalize our results to the case of t-hyperedge hypergraphs with vertices taking values over $\mathbb{Z}_r$. These hypergraphs can naturally be viewed as instances of Unique Games wherein the constraints are over $\mathbb{Z}_r$. Here again, we prove that obtaining an r-approximation algorithm requires $O(\log n)$ classical space and obtaining a $(r-\varepsilon)$-approximation algorithm requires $\Omega(n^{1-1/t})$ classical space or $\Omega(n^{1-2/t})$ quantum space.

Second, we show an $N = 2^{\Omega(n/r^4)}$ lower bound for (even non-linear) LDCs over $\mathbb{Z}_r$. In particular, for all r smaller than $n^{1/4}$, we prove an exponential in n lower bound for LDCs over $\mathbb{Z}_r$. Previous main results in this direction were by Goldreich et al. [GKST02] for $r = 2$ and linear LDCs, Kerenidis and de Wolf [KW04] for $r = 2$ and *non-linear* LDCs, Wehner and de Wolf [WdeW05] for non-linear LDCs from $\{0,1\}^n \to \mathbb{Z}_r^N$ and finally by Dvir and Shpilka [DS07] for $r > 2$ but linear LDCs. Apart from the result of [DS07], we are not aware of any lower bounds for non-linear LDCs from $\mathbb{Z}_r^n \to \mathbb{Z}_r^N$, even though it is a very natural question with connections to other fundamental problems, such as polynomial identity testing [DS07], private information retrieval [KT00, GKST02], additive combinatorics [BDG16] and quantum complexity theory [Aar18], to cite a few. Furthermore, we are not aware of a formal reduction between LDCs with $\Sigma = \{0,1\}$ and $\Sigma = \mathbb{Z}_r$, specially with recovery probability $1/|\Sigma| + \varepsilon$. Moreover, some past works define LDCs over general Σ with success probability $\geq \Pr[\text{wrong output}] + \varepsilon$ [Gop18], $\geq 1/2 + \varepsilon$ [GKST02] or $\geq 1 - \varepsilon$ [Dvi11]. These alternative definitions are encompassed by ours by considering ε a constant large enough. In the remaining part of the introduction, we describe these contributions in more detail.

1.2 Matrix hypercontractive inequality (over large alphabets)

Fourier analysis on the Boolean cube. We first discuss the basics of Fourier analysis before stating our result. Let $f : \{0,1\}^n \to \mathbb{R}$ be a function, then the Fourier decomposition of f is

$$f(x) = \sum_{S \in \{0,1\}^n} \widehat{f}(S)(-1)^{S \cdot x},$$

where $S \cdot x = \sum_{i=1}^n S_i x_i$ (where the sum is over $\{0,1\}$) and the *Fourier coefficients* of f are defined as $\widehat{f}(S) = \mathbb{E}_x[f(x)(-1)^{S \cdot x}]$, the expectation taken over uniformly random $x \in \{0,1\}^n$. One of the technical tools in the area of theoretical computer science is the hypercontractivity theorem proven by Bonami and Beckner [Bon70, Bec75]. In order to understand the hypercontractivity theorem, we first need to define the noise operator: for a noise parameter $\rho \in [-1,1]$, let T_ρ be an operator on the space of functions $f : \{0,1\}^n \to \mathbb{R}$ defined as

$$(\mathrm{T}_\rho f)(x) = \mathop{\mathbb{E}}_{y \sim \mathcal{N}_\rho(x)}[f(y)],$$

where $y \sim \mathcal{N}_\rho(x)$ denotes that the random string $y \in \{0,1\}^n$ is drawn as $y_i = x_i$ with probability $\frac{1}{2} + \frac{1}{2}\rho$ and as $y_i = x_i \oplus 1$ with probability $\frac{1}{2} - \frac{1}{2}\rho$ for each $i \in [n]$ independently. One can show that the Fourier expansion of $\mathrm{T}_\rho f$ can be written as

$$(\mathrm{T}_\rho f)(x) = \sum_{S \in \{0,1\}^n} \rho^{|S|} \widehat{f}(S)(-1)^{S \cdot x}.$$

One way to intuitively view this expression is that "large-weight" Fourier coefficients are reduced by an exponential factor while "small-weight" Fourier coefficients remain approximately the same. Consequently, it is not hard to see that $\|\mathrm{T}_\rho f\|_p \leq \|f\|_p$ for every $p \geq 1$, where $\|f\|_p := \left(\mathbb{E}_x[|f(x)|^p]\right)^{1/p}$ is the standard normalized p-norm of the function f. The main hypercontractivity theorem states that the previous inequality holds true even if we increase the left-hand size by a larger norm (meaning that norms under the noise operator are not just contractive, but *hypercontractive*), i.e., for every $p \in [1,2]$ and $\rho \leq \sqrt{p-1}$, we have that $\|\mathrm{T}_\rho f\|_2 \leq \|f\|_p$,[1] which can alternatively be written as

$$\left(\sum_{S\in\{0,1\}^n} \rho^{2|S|}\widehat{f}(S)^2\right)^{1/2} \leq \left(\frac{1}{2^n}\sum_{x\in\{0,1\}^n}|f(x)|^p\right)^{1/p}. \tag{1}$$

This inequality has found several applications in approximation theory [KKMO07, DS05], expander graphs [HLW06], circuit complexity [LMN93], coding theory [CL93], quantum computing [GKK+07, Mon11] (for more applications we refer the reader to [deW08, O'D14, Mon12]). All these applications deal with understanding the effect of noise on real-valued functions on the Boolean cube.

Generalizations of hypercontractivity. There are two natural generalizations of hypercontractivity: (i) a hypercontractivity statement for arbitrary product probability spaces. In this direction, it is possible to prove a similar hypercontractive inequality: for every $p \in [1,2]$ and $f \in L^2(\Omega_1 \times \cdots \times \Omega_n, \pi_1 \otimes \cdots \otimes \pi_n)$, we have

$$\|\mathrm{T}_\rho f\|_2 \leq \|f\|_p \;\text{ for }\; \rho \leq \sqrt{p-1}\cdot\lambda^{1/p-1/2}, \tag{2}$$

where λ is the smallest probability in any of the finite probability spaces (Ω_i, π_i) (see [O'D14, Chapter 10]). As a corollary, one gets a hypercontractive inequality for $f : \mathbb{Z}_r^n \to \mathbb{R}$; (ii) a hypercontractivity statement for matrix-valued functions $f : \{0,1\}^n \to \mathbb{C}^{m\times m}$, where the Fourier coefficients $\widehat{f}(S) = \mathbb{E}_x[f(x)(-1)^{S\cdot x}]$ are now $m \times m$ complex matrices. This was considered by Ben-Aroya, Regev and de Wolf [BRdeW08], who proved a hypercontractivity statement by using the powerful inequality of Ball, Carlen and Lieb [BCL94].

However, is there a generalization of hypercontractivity in both directions, i.e., a matrix-valued hypercontractivity for functions over $\mathbb{Z}_r$? This is open as far as we are aware and is our first main technical result.

Result 1. *For any $f : \mathbb{Z}_r^n \to \mathbb{C}^{m\times m}$, $p \in [1,2]$ and $\rho \leq \sqrt{\frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}}$,*

$$\left(\sum_{S\in\mathbb{Z}_r^n} \rho^{2|S|}\|\widehat{f}(S)\|_p^2\right)^{1/2} \leq \left(\frac{1}{r^n}\sum_{x\in\mathbb{Z}_r^n}\|f(x)\|_p^p\right)^{1/p}, \tag{3}$$

where $\|M\|_p := \left(\sum_i \sigma_i(M)^p\right)^{1/p}$ is the Schatten p-norm defined from the singular values $\{\sigma_i(M)\}_i$ of the matrix M and $|S| := |\{i \in [n] : S_i \neq 0\}|$ is the Hamming weight of $S \in \mathbb{Z}_r^n$.

The above result can be seen as an analogue of Eq. (1) where the absolute values are replaced with Schatten norms. We now make a couple of remarks. First, when $m = 1$ our result compares

[1] The hypercontractivity theorem can be stated for arbitrary $1 \leq p \leq q$ and $\rho \leq \sqrt{(p-1)/(q-1)}$, here we state it for $q = 2$ since we will be concerned with this setting.

to the one in Eq. (2) for $f : \mathbb{Z}_r^n \to \mathbb{R}$, but with a slightly worse ρ parameter compared to the $(1/r)^{1/p-1/2}$ factor. Second, for $r = 2$ we recover the same inequality from [BRdeW08]. The proof of this result is rather mathematical and not-so-intuitive. To this end, as in the proof of hypercontractive inequalities [O'D14, BRdeW08], our result follows by induction on n. It so happens that the base case is the most non-trivial step in the proof. So for now, let us assume $n = 1$, i.e., our goal is to prove Eq. (3) for $n = 1$. We now consider two special *simple* cases of the inequality.

(i) $r = 2$ and $\mathbb{C}^{m\times m}$ is replaced with real numbers: in this case, this is the well-known two-point inequality by Gross [Gro75] used for understanding the Logarithmic Sobolev inequalities. A proof of this inequality can also be easily viewed from a geometric perspective. As far as we are aware, there is no generalized r-point inequality for $r > 2$.

(ii) $r = 2$ and $\mathbb{C}^{m\times m}$ are arbitrary matrices: in this case, we only need to deal with two matrices $f(0), f(1)$ and Eq. (3) is exactly a powerful inequality in functional analysis, called the 2-*uniform convexity* of trace norms,

$$\left(\frac{\|X+Y\|_p^p + \|X-Y\|_p^p}{2}\right)^{2/p} \geq \|X\|_p^2 + (p-1)\|Y\|_p^2.$$

This inequality was first proven for certain values of p by Tomczak-Jaegermann [TJ74] before being extended for all $p \in [1, 2]$ by Ball, Carlen and Lieb [BCL94] in 1994. Since then it has found several applications, e.g. an optimal hypercontractivity inequality for Fermi fields [CL93], regularized convex optimization [DSSST10] and metric embedding [LN04, Nao16]. 2-uniform convexity can also be used to prove a variety of other inequalities, for example, Khintchine inequality [TJ74, DGTJ84], Hoeffding and Bennett-style bounds [Pin94, HRMS20]. Moreover, the above result could be seen as a corollary of Hanner's inequality for matrices (originally proven for Lebesgue spaces L_p [Han56]), but, unfortunately, Hanner's inequality for Schatten trace ideals are only proven for $p \leq 4/3$ (see more in [BCL94]). As far as we are aware, a generalization of the above inequality when considering r matrices was unknown.

One contribution in our work is the following generalization of a result from Ball, Carlen and Lieb [BCL94] (note it also implies a generalization of the two-point inequality), which we believe may be of independent interest.

Result 2. *Let $r \in \mathbb{Z}$, $r \geq 2$. Let $\omega_r := e^{2i\pi/r}$, $A_0, \ldots, A_{r-1} \in \mathbb{C}^{n\times n}$ and $p \in [1, 2]$, then*

$$\left(\frac{1}{r}\sum_{k=0}^{r-1}\left\|\sum_{j=0}^{r-1}\omega_r^{jk}A_j\right\|_p^p\right)^{2/p} \geq \|A_0\|_p^2 + \frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}\sum_{k=1}^{r-1}\|A_k\|_p^2. \tag{4}$$

Now that we have established Result 2, the proof of Result 1 is a simple induction argument on n: for the base case $n = 1$, Result 1 is exactly Result 2, and proving the induction step requires an application of Minkowski inequality. Since this proof is very similar to the one in [BRdeW08], we omit the details here.

1.3 Application 1: Streaming algorithms

Approximation algorithms for combinatorial optimization problems have been a rich area of study in theoretical computer science. One of the most famous approximation algorithms is by Goemans and Williamson [GW95] who proved that one can obtain a 1/0.878-approximation algorithm in

polynomial time for Max-Cut using semi-definite programming and this is believed to be optimal assuming the Unique Games conjecture is true [KKMO07]. In the past few years there has been a sequence of works [KKS14, GVV17, GT19, KK19, CGV20, CGSV21a, CGSV21b] that tried to prove *unconditional* hardness of combinatorial optimization (e.g. the Max-Cut problem) in the well-known streaming model of computation by Alon, Matias and Szegedy [AMS99].

In the streaming model, the goal is to optimize the amount of *space* needed to solve a problem rather than time, and output a value which is at least a fraction $1/\gamma$ of the optimum value with high probability. Many recent works referenced above have shown interesting threshold theorems, for example, for the Max-Cut problem, getting a 2-approximation algorithm using $O(\log n)$ classical space is easy: one simply counts the number of edges in the graph (which requires only a counter of size $2\log n$) and outputs half this count. Moreover, one can obtain a graph sparsifier using $O(n/\varepsilon^2)$ space [AGM12] and, from it, a $(1+\varepsilon)$-approximation for the Max-Cut value. On the other hand, Kapralov et al. [KKS14] initiated the study of proving streaming lower bounds for Max-Cut in the random-edge model (where inputs arrive randomly, and not necessarily adversarially), and in this work they showed that one requires $\Omega(\sqrt{n})$ space for $(2-\varepsilon)$-approximations in an n-vertex graph, together with a classical lower bound $\Omega(n^{1-\varepsilon})$ for $(1+\varepsilon)$-approximations in the adversarial model (their proof, together with Result 4 below, immediately implies a similar quantum lower bound for $(1+\varepsilon)$-approximation). After a sequence of works, Kapralov and Krachun [KK19] finally obtained an $\Omega(n)$ space lower bound for $(2-\varepsilon)$-approximations even in the random-edge model.

A common technique to prove streaming lower bounds is via communication complexity. To see this, suppose a problem P has inputs (X,Y) and the goal is to find space-efficient streaming algorithms to compute $P(X,Y)$, when X,Y are presented in a stream (i.e., presented bit-by-bit). Then, one way to lower bound the *space complexity* is to prove lower bounds on the following problem: consider the one-way communication problem where Alice gets the input X, Bob gets the input Y, their goal is to compute $P(X,Y)$ and only Alice is allowed to communicate to Bob. Then one can show that any lower bound for randomized one-way communication implies an equivalent lower bound for streaming algorithms. This technique has been used by a sequence of papers to prove lower bounds on space complexity of Max-Cut [KKS14, KK19, GT19], matching [GKK12], Max-CSP [GVV17, CGV20, CGSV21a, CGSV21b] and counting cycles [VY11, AD21]. One problem that is used often in this direction is a variant of the Boolean Hidden Matching.

1.3.1 Hidden Matching and its variants

The Boolean Hidden Matching (BHM) problem was introduced by Bar-Yossef et al. [BYJK04] (which was in turn inspired by Kerenidis and de Wolf [KW04] for proving LDC lower bounds) in order to prove exponential separations between quantum and classical one-way communication complexities. Below we described the generalized Hidden Matching problem over larger alphabets and hypermatching. The r-ary Hidden Hypermatching (r-HH(α,t,n)) problem is a two-party communication problem between Alice and Bob: Alice is given $x\in\mathbb{Z}_r^n$ and Bob is given a string $w\in\mathbb{Z}_r^{\alpha n/t}$ and $\alpha n/t$-many disjoint t-tuples (for $\alpha\in(0,1]$), i.e., hyperedges of an α-partial hypermatching, which can also be viewed as an incident matrix $M\in\{0,1\}^{\alpha n/t\times n}$ (each row corresponding to a hyperedge). In the YES instance it is promised that $w=Mx$ (over $\mathbb{Z}_r$), while in the NO instance it is promised that w is uniformly random, and the goal is to decide which is the case using a message sent from Alice to Bob.

There have been a few lines of work in understanding the problem of Hidden Hypermatching: (i) the seminal work of Bar-Yossef et al. [BYJK04] and Gavinsky et al. [GKK+07] showed that,

for $r = t = 2$, BHM can be solved using $O(\log n)$ qubits but requires $\Omega(\sqrt{n})$ classical bits of communication; (*ii*) Verbin and Yu [VY11] considered the problem where $r = 2$ and $t \geq 2$ (which in fact inspired many follow-up works on using hypermatching for classical streaming lower bounds) and showed a classical lowed bound of $\Omega(n^{1-1/t})$, which was subsequently generalized to a $\Omega(n^{1-2/t})$ quantum lower bound by Shi, Wu and Yu [SWY12]; (*iii*) Guruswami and Tao [GT19] studied the problem for when $t = 2$ and $r \geq 2$, proving a classical $\Omega(\sqrt{n})$ lower bound. A natural question is then, what is the quantum and classical communication complexities for $r, t \geq 2$? In this paper, we give both upper and lower bounds for the Hidden Hypermatching problem for every r and t.

Upper bounds on Hidden Hypermatching. For a given $t \geq 2$, the same classical communication protocol for $r = 2$ can be used for general $r > 2$. The idea is that Alice picks $O((n/\alpha)^{1-1/t})$ entries of x uniformly at random to send to Bob. By the Birthday Paradox, with high probability Bob will obtain all the values from one of his hyperedges i, and thus can compare $(Mx)_i$ with the corresponding w_i. If they are equal, he outputs YES, otherwise he outputs NO, which leads to an one-side error of $O(1/r)$. The total amount of communication is $O(\log(rn)(n/\alpha)^{1-1/t})$ bits.[2] The situation is more interesting in the quantum setting. For $t = 2$, we prove that Hidden Hypermatching can be solved using only a logarithmic amount of qubits for every $r = \text{poly}(n)$.

Result 3. *There is a protocol for r-HH$(\alpha, 2, n)$ with one-sided error $1/3$ using $O(\log(nr)/\alpha)$ qubits.*

The above bound uses a non-trivial procedure that allows to learn the sum of two numbers modulo r by using just one "query" and crucially uses the knowledge of the string w: given a suitable superposition of two numbers, one can obtain their sum with one-sided error by using one measurement. As far as we are aware, such a statement was not known prior to our work. However, the knowledge of w is vital, which means that the protocol does not work for more general settings where there is no promise on the inputs (e.g. a relational version of the r-ary Hidden Hypermatching problem where Bob must output one hyperedge i and its corresponding value $(Mx)_i$), and it also cannot be used as a building block for the general case $t, r > 2$. The current upper bound on the quantum communication complexity of the r-HH(α, t, n) problem with $t, r > 2$ thus matches the classical one. In view of the lower bounds stated below, we hence make the following conjecture.

Conjecture 1. *If $t, r > 2$, there is a protocol for r-HH(α, t, n) using $O(\log(rn)(n/\alpha)^{1-1/\lceil t/2 \rceil})$ qubits.*

Lower bounds on Hidden Hypermatching. The standard approach for proving a lower bound on the amount of communication required to solve the Hidden Hypermatching problem is via Fourier analysis. In the classical proofs of Gavinsky et al. [GKK+07], Verbin and Yu [VY11] and Guruswami and Tao [GT19], the total variation distance between the probability distributions arising from the YES and NO instances is bounded using the inequality of Kahn, Kalai and Linial [KKL89] (which can be seen as a corollary of the hypercontractivity inequality). On the other hand, Shi, Wu and Yu [SWY12] obtained a quantum lower bound by bounding the Schatten 1-norm between the possible density matrices received by Bob in both YES and NO instances via the matrix-valued hypercontractivity from Ben-Aroya, Regev and de Wolf [BRdeW08]. We follow a similar approach by using our *generalized matrix-valued hypercontractive* inequality from Result 1 in order to obtain the following lower bound (note that, for $r = 2$, our lower bound is exponential better in α compared to [SWY12]).

Result 4. *Every constant-bias protocol for the r-HH(α, t, n) problem with $t, r \geq 2$ requires at least $\Omega((n/t)^{1-2/t}/\alpha^{2/t})$ qubits of communication or $\Omega((n/t)^{1-1/t}/\alpha^{1/t})$ bits of communication.*

[2] One can further improve this complexity to $O(\log(n \log r) + (\log r) \cdot (n/\alpha)^{1-1/t})$ by Newman's theorem [New91].

1.3.2 Relation to streaming lower bounds

As mentioned at the start of this section, using one-way communication complexity lower bounds has been a common technique used by several recent works [VY11, KKS14, GVV17, GT19, CGV20] to prove streaming lower bounds. Using our classical and quantum communication lower bound we present two lower bounds for streaming problems.

There are a few natural generalizations to Max-Cut. One is Max-k-Cut, i.e., finding the maximum cut value on a hypergraph with k-sized hyperedges. Clearly, the lower bound of [KK19] holds true for Max-k-Cut, but could one prove better lower bound depending on k? Another is the Unique Games problem, a constraint satisfaction problem defined on a graph, where a linear constraint (a permutation) over $\mathbb{Z}_r$ is specified on each edge and the goal is to find a vertex assignment over $\mathbb{Z}_r$ that maximizes the number of satisfied constraints. When $r = 2$, Unique Games reduces to Max-Cut. Guruswami and Tao [GT19] studied the streaming complexity of the Unique Games problem and proved a lower bound of $\Omega(\sqrt{n})$ in the adversarial model by using a reduction to Hidden Matching over $\mathbb{Z}_r$ and the same bound was obtained in [CGSV21b] for a larger set of problems including Unique Games.

Here we join both directions, i.e., Max-k-Cut and the standard Unique Games problem, into a generalized version of Unique Games defined on a hypergraph and obtain streaming classical and quantum lower bounds in the adversarial model for any value $k, r \geq 2$.

Result 5. *Every streaming algorithm giving a $(r-\varepsilon)$-approximation for Unique Games on k-hyperedge n-vertex hypergraphs over $\mathbb{Z}_r$ uses $\Omega(n^{1-2/k})$ quantum space or $\Omega(n^{1-1/k})$ classical space.*

The above result clearly generalizes the work of Guruswami and Tao [GT19]. Compared to Kapralov and Krachun [KK19], on the one hand our results are for the weaker adversarial model, and they obtained a stronger linear lower bound, but on the other hand, their result does not immediately generalize for $\mathbb{Z}_r$ and is a purely classical proof (in fact we remark that our classical lower bound is significantly simpler than their work). As far as we are aware, these are the first quantum lower bounds for Unique Games and Max-k-Cut in the streaming model.

1.4 Application 2: Locally decodable codes

A locally decodable code (LDC) is an error correcting code that allows to retrieve a single bit of the original message (with high probability) by only examining a few bits in a corrupted codeword. More formally, a (q, ε, δ)-LDC was defined by Katz and Trevisan [KT00] as a function $C : \mathbb{Z}_r^n \to \mathbb{Z}_r^N$ that satisfies the following: for all $x \in \mathbb{Z}_r^n$, $i \in [n]$ and $y \in \mathbb{Z}_r^N$ that satisfies $d(C(x), y) \leq \delta$ (i.e., a δ-fraction of the elements of $C(x)$ are corrupted), there exists an algorithm $\mathcal{A}$ that makes q queries to y non-adaptively and outputs $\mathcal{A}^y(i) \in \mathbb{Z}_r$ such that $\Pr[\mathcal{A}^y(i) = x_i] \geq 1/r + \varepsilon$ (where the probability is over the randomness of $\mathcal{A}$). Over $\{0, 1\}$, LDCs have found several applications in private information retrieval [CGKS95], multiparty computation [IK04], data structures [CGdeW09] and average-case complexity theory [Tre04].

The natural question in constructing LDCs is the trade-off between N and n. A well-known 2-query LDC is the Hadamard encoding that maps $x \in \mathbb{Z}_r^n$ into the string $C(x) = (\langle x, y\rangle)_{y\in\{0,1\}^n}$: on input $i \in [n]$, a decoding algorithm queries $C(x)$ at a uniformly random y and $y + e_i$ and retrieves $C(x) = \langle x, y \oplus e_i\rangle - \langle x, y\rangle$, where $e_i = 0^{i-1}10^{n-i}$. Here the encoding length is $N = 2^n$, and an important question is, are there 2-query LDCs with $N \ll 2^n$? For the case $r = 2$, Goldreich et al. [GKST02] showed a lower bound $N = 2^{\Omega(n)}$ for *linear codes*, which was later improved by

Obata [Oba02]. Later, Kerenidis and de Wolf [KW04] proved an exponential lower bound for *non-linear codes* using a quantum argument![3]

This left open the setting where $r > 2$. Following these works, for 2-query *non-linear* LDCs $C : \{0,1\}^n \to \mathbb{Z}_r^N$ (note the inputs are over $\{0,1\}$ and not $\mathbb{Z}_r$), Wehner and de Wolf [WdeW05] proved the lower bound $N = 2^{\Omega(n/r^2)}$. On the other hand, Dvir and Shpilka [DS07] showed a lower bound of $N = 2^{\Omega(n)}$ for every 2-query *linear* LDC $C : \mathbb{Z}_r^n \to \mathbb{Z}_r^N$, even independent of the field size. To prove their result, they crucially observed that, given a linear LDC over $\mathbb{Z}_r$, one can construct a linear LDC over $\{0,1\}$ (with almost the same parameters) and then invoked the result of Goldreich et al. [GKST02]. This reduction, however, fails for non-linear codes and motivates if there are *non-linear* LDCs $C : \mathbb{Z}_r^n \to \mathbb{Z}_r^N$ with $N \ll 2^n$?

The main contribution here is a lower bound for *non-linear* LDCs over $\mathbb{Z}_r$ that scale as $2^{\Omega(n/r^4)}$, and which gives a super-polynomial lower bound for $r = o(n^{1/4})$. Our lower bound comes from using our hypercontractive inequality in Result 1. The idea is similar to the one from [BRdeW08], but more technical as a result of optimizing the dependence on r. A large $r^2N \times r^2N$ matrix with rank 1 is constructed from a given 2-query LDC. By considering its Fourier transform over $\mathbb{Z}_r$, there exist various entries of the form $\mathbb{E}_x\left[\omega_r^{k_1 C(x)_{j_1} + k_2 C(x)_{j_2} - x_i}\right]$, whose absolute values are bounded by a technical result generalizing a few different ideas from [KT00, KW04, BRdeW08]. It is possible then to lower bound the Schatten norm of the Fourier transformed matrix. On the other hand, its rank 1 implies a simple expression for the original matrix's Schatten norm. The hypercontractive inequality connects both quantities and leads to the following final result.

Result 6. *If $C : \mathbb{Z}_r^n \to \mathbb{Z}_r^N$ is a $(2, \delta, \varepsilon)$-*LDC*, then $N = 2^{\Omega(\delta^2\varepsilon^4 n/r^4)}$.*

We briefly mention that, if one requires the success probability to be larger than, for example, $1/2 + \varepsilon$ instead of $1/r + \varepsilon$, so that plurality vote can be used and the success probability amplified, then ε becomes a constant bounded away from $1/r$ (if $r > 2$) and our lower bound is no longer dependent on ε.

Further applications (Private information retrieval) Katz and Trevisan [KT00], and Goldreich et al. [GKST02] established a nice connection between LDCs and private information retrieval (PIR) protocols. We do not define these PIR schemes here and refer the reader to Section 6. Almost as a black-box, using Result 6, we get the following lower bound for PIR schemes over $\mathbb{Z}_r$.

Result 7. *A classical 2-server PIR scheme with query size t, answer size a and recovery probability $1/r + \varepsilon$, satisfies $t = \Omega\big(\delta^2\varepsilon^4 n/r^4 - a\big)$.*

After completion of this work. After completing this work, Chou et al. [CGS+21] put up an online preprint in which they improve our classical streaming lower bound to $\Omega(n)$ for a broad class of problems, including Unique Games. As far as we are aware, our quantum streaming lower bound is the first for hypergraphs over $\mathbb{Z}_r$. Additionally, after completion, Jop Briët (private communication) gave an alternate proof of $N = 2^{\Omega(n/r^2)}$ for 2-query LDCs over $\mathbb{Z}_r$ using the non-commutative Khintchine inequality.

1.5 Future work

Our work open up a few directions of research.

[3] For simplicity in exposition, we omit the dependence on δ, ε in these lower bounds.

1. *Proving* LDC *lower bounds.* The first natural open question is, can we prove a lower bound of $N = 2^{\Omega(n/r)}$ for LDCs over $\mathbb{Z}_r$, or, more ambitiously, prove that $N = 2^{\Omega(n)}$? As far as we are aware, there are no super-polynomial lower bounds for N even for $r = \omega(\sqrt{n})$. Similarly, can one also prove a lower bound of $N = 2^{\Omega(n \log r)}$ for *non-linear* locally-*correctable* codes over $\mathbb{Z}_r$ (thereby matching a similar lower bound for linear case [BDSS11]).

2. *Communication complexity of r-ary Hidden Hypermatching.* Our communication protocol behind Result 3 relies on the promise on the inputs, i.e., on the string $w \in \mathbb{Z}_r^{\alpha n/t}$ that either satisfies $w = Mx$ or is uniformly random. Is there a protocol with the same complexity which does not explicitly use w? More generally, what is the communication complexity of a relational version of the r-$\mathsf{HH}(\alpha, 2, n)$ problem in which Bob outputs a hyperedge and the corresponding entry of Mx? Moreover, is it possible to match the quantum lower bounds from Result 4?

3. *Better bounds on streaming algorithms.* What is the quantum space complexity of approximating Max-Cut or Unique Games? Is it possible to obtain some saving in space complexity, e.g. an upper bound of $O(n^{1-2/t})$ that matches our lower bound, or is the quantum space complexity $\Omega(n)$? The former would be interesting because advantage in quantum space complexity are only handful (for contrived problems) and the latter would require proving new quantum lower bounds for the communication problems introduced in [KKSV17, KK19, CGSV21a, CGSV21b, CGS+21].

4. *Generalized hypercontractivity.* Another open question is regarding our main Result 1, which shows a form of $(2, q)$-hypercontractivity, since the result works for all Schatten p-norms with $p \in [1, 2]$. Can we prove a general (q, p)-hypercontractive statement for matrices, firstly for matrix-valued functions over $\{0, 1\}$, and then further generalize that to functions over $\mathbb{Z}_r$? Proving this might also require a generalization of the powerful inequality of Ball, Carlen and Lieb [BCL94] in a different direction.

Acknowledgements. SA firstly thanks T.S. Jayram for introducing him to this problem (and several discussions thereafter) on proving quantum bounds for streaming algorithms while participating in the program "Quantum Wave in Computing" held at Simons Institute for the Theory for Computing. We thank Jop Briët and Ronald de Wolf for many clarifications and discussions regarding hypercontractivity and LDCs, and Mario Szegedy for discussions during the initial stages of this project. We are also very thankful to Keith Ball and Eric Carlen for the help in understanding their proof of uniform convexity for trace ideals. JFD was supported by the Singapore National Research Foundation, the Prime Minister's Office, Singapore and the Ministry of Education, Singapore under the Research Centres of Excellence programme under research grant R 710-000-012-135.

2 Preliminaries

Let $[n] := \{1, \ldots, n\}$. For $r \in \mathbb{Z}$, $r \geq 2$, we let $\mathbb{Z}_r := \{0, \ldots, r-1\}$ be the ring with addition and multiplication modulo r, and let $\omega_r := e^{2\pi i/r}$. Given $S \in \mathbb{Z}_r^n$, we write $|S| := |\{i \in [n] : S_i \neq 0\}|$ for its Hamming weight. Let $\mathrm{D}(\mathbb{C}^m)$ be the set of all quantum states over $\mathbb{C}^m$, i.e., the set of positive semi-definite matrices with trace 1. For a matrix $M \in \mathbb{C}^{m \times m}$, the (unnormalized) Schatten p-norm is defined as $\|M\|_p := (\operatorname{Tr}|M|^p)^{1/p} = \left(\sum_i \sigma_i(M)^p\right)^{1/p}$, where $\{\sigma_i(M)\}_i$ are the singular values of M, i.e., the eigenvalues of the positive semi-definite operator $|M| := \sqrt{M^\dagger M}$. We also define the normalized Schatten p-norm as $\|M\|_p := \left(\frac{1}{m}\operatorname{Tr}|M|^p\right)^{1/p} = \left(\frac{1}{m}\sum_i \sigma_i(M)^p\right)^{1/p}$. Throughout the paper we shall use the unnormalized Schatten norm, unless stated otherwise. Given a vector

$v \in \mathbb{C}^m$, its p-norm is $\|v\|_p := \left(\sum_{i=1}^m |v_i|^p\right)^{1/p}$. Given two probability distributions P and Q on the same finite set, their total variation distance is $\|P - Q\|_{\text{tvd}} := \sum_i |P(i) - Q(i)|$ (we might abuse notation and use random variables instead of their probability distributions in $\|\cdot\|_{\text{tvd}}$). For a probability $p = 1/r + \varepsilon$ with fixed $r \in \mathbb{Z}$, we refer to ε as its *advantage*, and to 2ε as its *bias*.

The Fourier transform of a matrix-valued function $f : \mathbb{Z}_r^n \to \mathbb{C}^{m\times m}$ is a function $\widehat{f} : \mathbb{Z}_r^n \to \mathbb{C}^{m\times m}$ defined by

$$\widehat{f}(S) = \frac{1}{r^n} \sum_{x\in\mathbb{Z}_r^n} f(x)\omega_r^{-S\cdot x},$$

where $S \cdot x = \sum_{i=1}^n S_i x_i$ is a sum over $\mathbb{Z}_r$. Here the Fourier coefficients $\widehat{f}(S)$ are $m \times m$ complex matrices and we can write $f : \mathbb{Z}_r^n \to \mathbb{C}^{m\times m}$ as

$$f(x) = \sum_{S\in\mathbb{Z}_r^n} \widehat{f}(S)\omega_r^{S\cdot x}.$$

We will need the Holevo-Helstrom theorem [Hel76] which characterizes the optimal success probability of distinguishing between two quantum states.

Fact 2 ([Wat18, Theorem 3.4])**.** *Let ρ_0, ρ_1 be two quantum states that appear with probability p and $1-p$, respectively. The optimal success probability of predicting which state it is by a POVM is*

$$\frac{1}{2} + \frac{1}{2}\|p\rho_0 - (1-p)\rho_1\|_1.$$

3 Hypercontractive Inequality

In this section we prove our main result, a hypercontractive inequality for matrix-valued functions over $\mathbb{Z}_r$, generalizing a result from [BRdeW08]. The proof is by induction on n and the base case $n = 1$ is proven in Section 3.1, which is a generalization of Ball, Carlen and Lieb [BCL94] when considering r matrices. After this, the induction is fairly straightforward and is described in Section 3.2.

3.1 Generalizing Ball, Carlen and Lieb

We first state the powerful inequality of Ball, Carlen and Lieb [BCL94, Theorem 1].

Theorem 3 (Optimal 2-uniform convexity)**.** *Let $A, B \in \mathbb{C}^{n\times n}$, and $p \in [1, 2]$. Then*

$$\left(\frac{\|A+B\|_p^p + \|A-B\|_p^p}{2}\right)^{2/p} \geq \|A\|_p^2 + (p-1)\|B\|_p^2.$$

As previously mentioned in the introduction, this inequality was first proven by Tomczak-Jaegermann [TJ74] for $p \leq 4/3$, before being generalized by Ball, Carlen and Lieb [BCL94] for all $p \in [1, 2]$ in 1994. Since then it has found several applications [CL93, DSSST10, LN04, Nao16]. The above result can be recast in a slightly different way.

Theorem 4. *Let $p \in [1, 2]$ and $Z, W \in \mathbb{C}^{n\times n}$ such that $\text{Tr}[|Z|^{p-1}ZW^\dagger] = \text{Tr}[|Z|^{p-1}WZ^\dagger] = 0$ (where $|Z|^{p-1} = (ZZ^\dagger)^{(p-1)/2}$). Then*

$$\|Z + W\|_p^2 \geq \|Z\|_p^2 + (p-1)\|W\|_p^2.$$

Theorem 4 is implicit in the proof of [BCL94, Theorem 1], and it is where most of the difficulty lies, while the reduction from Theorem 3 to Theorem 4 is done by defining

$$Z = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}, \qquad W = \begin{bmatrix} B & 0 \\ 0 & -B \end{bmatrix}.$$

Nonetheless, Theorem 4 holds more generally for any $Z, W \in \mathbb{C}^{n\times n}$ that satisfy $\mathrm{Tr}[|Z|^{p-1}ZW^\dagger] = \mathrm{Tr}[|Z|^{p-1}WZ^\dagger] = 0$. By using this result, we can prove the following generalization of Theorem 3.

Theorem 5 (A generalization of [BCL94])**.** *Let $r \in \mathbb{Z}$, $r \geq 2$. Let $\omega_r = e^{2i\pi/r}$, $A_0, \ldots, A_{r-1} \in \mathbb{C}^{n\times n}$ and $p \in [1, 2]$, then*

$$\left(\frac{1}{r}\sum_{j=0}^{r-1}\|A_j\|_p^p\right)^{2/p} \geq \left\|\frac{1}{r}\sum_{j=0}^{r-1}A_j\right\|_p^2 + \frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}\sum_{k=1}^{r-1}\left\|\frac{1}{r}\sum_{j=0}^{r-1}\omega_r^{-jk}A_j\right\|_p^2, \tag{5a}$$

$$\left(\frac{1}{r}\sum_{k=0}^{r-1}\left\|\sum_{j=0}^{r-1}\omega_r^{jk}A_j\right\|_p^p\right)^{2/p} \geq \|A_0\|_p^2 + \frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}\sum_{k=1}^{r-1}\|A_k\|_p^2. \tag{5b}$$

Notice that for $r = 2$ we recover Theorem 3, since $\frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)} = p - 1$.

Proof. In order to prove this theorem, first note that both inequalities are equivalent: just define $A'_k = \frac{1}{r}\sum_{j=0}^{r-1}\omega_r^{-jk}A_j \iff A_k = \sum_{j=0}^{r-1}\omega_r^{jk}A'_j$. Therefore we shall focus on Eq. (5b). In order to prove it, let us first define the $rn \times rn$ matrices

$$Z_j := \mathrm{diag}(\{\omega_r^{jk}A_j\}_{k=0}^{r-1}) = \begin{bmatrix} A_j & 0 & 0 & \ldots & 0 \\ 0 & \omega_r^j A_j & 0 & \ldots & 0 \\ 0 & 0 & \omega_r^{2j}A_j & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \omega_r^{(r-1)j}A_j \end{bmatrix} \tag{6}$$

for $j \in \{0, \ldots, r-1\}$. Now, since the trace is additive for block matrices, we have

$$\mathrm{Tr}\left|\sum_{j=0}^{r-1}Z_j\right|^p = \sum_{k=0}^{r-1}\mathrm{Tr}\left|\sum_{j=0}^{r-1}\omega_r^{jk}A_j\right|^p. \tag{7}$$

Moreover, observe that

$$\|Z_j\|_p^2 = \left(\sum_{k=0}^{r-1}\mathrm{Tr}\,|\omega_r^{jk}A_j|^p\right)^{2/p} = (r\,\mathrm{Tr}\,|A_j|^p)^{2/p} = r^{2/p}\|A_j\|_p^2.$$

Therefore we can rewrite Eq. (5b) as

$$\left\|\sum_{j=0}^{r-1}Z_j\right\|_p^2 \geq \|Z_0\|_p^2 + \frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}\sum_{j=1}^{r-1}\|Z_j\|_p^2.$$

The above can be proven by repeated applications of Theorem 4 as follows: consider a permutation of $[r-1]$ given by $(k_1, \dots, k_{r-1})$. Since $\mathrm{Tr}[|Z_j|^{p-1} Z_j Z_k^\dagger] = \mathrm{Tr}[|Z_j|^{p-1} Z_k Z_j^\dagger] = 0$ for any $j \neq k$, then (define $k_0 := 0$)

$$\mathrm{Tr}\left[|Z_{k_j}|^{p-1} Z_{k_j} \left(\sum_{l=j+1}^{r-1} Z_{k_l}\right)^\dagger\right] = \mathrm{Tr}\left[|Z_{k_j}|^{p-1} \left(\sum_{l=j+1}^{r-1} Z_{k_l}\right) Z_{k_j}^\dagger\right] = 0$$

for every $j \in \{0, 1, \dots, r-2\}$, meaning that Theorem 4 can be applied, which implies

$$\begin{aligned}\left\|\sum_{j=0}^{r-1} Z_j\right\|_p^2 &\geq \|Z_0\|_p^2 + (p-1)\left\|\sum_{j=1}^{r-1} Z_j\right\|_p^2 \\ &\geq \|Z_0\|_p^2 + (p-1)\|Z_{k_1}\|_p^2 + (p-1)^2\left\|\sum_{j=2}^{r-1} Z_{k_j}\right\|_p^2 \geq \|Z_0\|_p^2 + \sum_{j=1}^{r-1}(p-1)^j \|Z_{k_j}\|_p^2.\end{aligned}$$

Averaging the above inequality over all the $(r-1)!$ permutations of the set $[r-1]$, we obtain

$$\begin{aligned}\left\|\sum_{j=0}^{r-1} Z_j\right\|_p^2 &\geq \|Z_0\|_p^2 + \frac{1}{(r-1)!}\sum_{j=1}^{r-1}\|Z_j\|_p^2 \sum_{k=1}^{r-1}(r-2)!(p-1)^k \\ &= \|Z_0\|_p^2 + \frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}\sum_{j=1}^{r-1}\|Z_j\|_p^2,\end{aligned}$$

proving our theorem statement. □

Remark 1. *It is not hard to see that* $\frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)} \geq \frac{p-1}{r-1}$ *and* $\lim_{p\to 2} \frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)} = 1$.

Observe that $t \mapsto t^{p/2}$ is concave for $p \in [1,2]$, hence Theorem 5 implies the seemingly weaker

$$\frac{1}{r}\sum_{k=0}^{r-1}\left\|\sum_{j=0}^{r-1}\omega_r^{jk} A_j\right\|_p^2 \geq \|A_0\|_p^2 + \frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}\sum_{k=1}^{r-1}\|A_k\|_p^2 \tag{8}$$

for $p \in [1,2]$. Nonetheless, the above inequality also implies Theorem 5 (this fact was already pointed out for $r = 2$ by [BCL94]). Indeed, consider again the $rn \times rn$ matrices Z_j from Eq. (6). Then, similar to Eq. (7) (which only considered the $\ell = 0$ case below), for any $\ell \in \mathbb{Z}_r$ we have

$$\mathrm{Tr}\left|\sum_{j=0}^{r-1}\omega_r^{j\ell} Z_j\right|^p = \sum_{k=0}^{r-1}\mathrm{Tr}\left|\sum_{j=0}^{r-1}\omega_r^{jk} A_j\right|^p \implies \left\|\sum_{j=0}^{r-1}\omega_r^{j\ell} Z_j\right\|_p^2 = \left(\sum_{k=0}^{r-1}\left\|\sum_{j=0}^{r-1}\omega_r^{jk} A_j\right\|_p^p\right)^{2/p}.$$

Since $\|Z_j\|_p^2 = r^{2/p}\|A_j\|_p^2$ for $j \in \mathbb{Z}_r$, Eq. (8) implies (define $\zeta := \frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}$ for simplicity)

$$\|A_0\|_p^2 + \zeta\sum_{k=1}^{r-1}\|A_k\|_p^2 = \frac{\|Z_0\|_p^2}{r^{2/p}} + \zeta\sum_{k=1}^{r-1}\frac{\|Z_k\|_p^2}{r^{2/p}} \leq \frac{r^{-2/p}}{r}\sum_{\ell=0}^{r-1}\left\|\sum_{j=0}^{r-1}\omega_r^{j\ell} Z_j\right\|_p^2 = \left(\frac{1}{r}\sum_{k=0}^{r-1}\left\|\sum_{j=0}^{r-1}\omega_r^{jk} A_j\right\|_p^p\right)^{2/p},$$

which is exactly Theorem 5.

3.2 Proving $(2,p)$-hypercontractive inequality over $\mathbb{Z}_r$

Having proven the base case of our main theorem statement, we are now ready to prove our hypercontractivity theorem for matrix-valued functions over $\mathbb{Z}_r$.

Theorem 6. *Let $p \in [1,2]$. For every $f : \mathbb{Z}_r^n \to \mathbb{C}^{m\times m}$ and*

$$\rho \le \sqrt{\frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}},$$

we have

$$\left(\sum_{S\in\mathbb{Z}_r^n} \rho^{2|S|}\|\widehat{f}(S)\|_p^2\right)^{1/2} \le \left(\frac{1}{r^n}\sum_{x\in\mathbb{Z}_r^n}\|f(x)\|_p^p\right)^{1/p},$$

where $|S| := |\{i \in [n] : S_i \neq 0\}|$.

Proof. For ease of notation, define $\zeta := \frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}$. It suffices to prove the inequality for $\rho = \sqrt{\zeta}$. Our proof closely follows the one in [BRdeW08] and is by induction on n. For $n = 1$, the desired statement is

$$\sum_{S\in\mathbb{Z}_r} \zeta^{|S|}\|\widehat{f}(S)\|_p^2 \le \left(\frac{1}{r}\sum_{x\in\mathbb{Z}_r}\|f(x)\|_p^p\right)^{2/p}. \tag{9}$$

Consider the matrices $A_0,\dots,A_{r-1}$ such that $f(k) = \sum_{j=0}^{r-1}\omega_r^{jk}A_j$ for all $k \in \mathbb{Z}_r$, so that Eq. (9) can be written as

$$\|A_0\|_p^2 + \zeta\sum_{k=1}^{r-1}\|A_k\|_p^2 \le \left(\frac{1}{r}\sum_{k=0}^{r-1}\left\|\sum_{j=0}^{r-1}\omega_r^{jk}A_j\right\|_p^p\right)^{2/p},$$

using the fact that $\widehat{f}(j) = \frac{1}{r}\sum_{k=0}^{r-1} f(k)\omega_r^{-jk} = A_j$, which is precisely Theorem 5.

We now assume the inequality holds for n and prove it for $n+1$. Let $f : \mathbb{Z}_r^{n+1} \to \mathbb{C}^{m\times m}$ and $g_i = f|_{x_{n+1}=i}$ for $i \in \{0,\dots,r-1\}$ be the function obtained by fixing the last bit of $f(\cdot)$ to i. By the induction hypothesis we have that, for every $i \in \{0,\dots,r-1\}$ and $p \in [1,2]$,

$$\sum_{S\in\mathbb{Z}_r^n} \zeta^{|S|}\|\widehat{g_i}(S)\|_p^2 \le \left(\frac{1}{r^n}\sum_{x\in\mathbb{Z}_r^n}\|g_i(x)\|_p^p\right)^{2/p}.$$

We now take the ℓ_p average of each of these r inequalities to obtain

$$\left(\frac{1}{r}\sum_{i=0}^{r-1}\left(\sum_{S\in\mathbb{Z}_r^n}\zeta^{|S|}\|\widehat{g_i}(S)\|_p^2\right)^{p/2}\right)^{2/p} \le \left(\frac{1}{r}\sum_{i=0}^{r-1}\frac{1}{r^n}\sum_{x\in\mathbb{Z}_r^n}\|g_i(x)\|_p^p\right)^{2/p} = \left(\frac{1}{r^{n+1}}\sum_{x\in\mathbb{Z}_r^{n+1}}\|f(x)\|_p^p\right)^{2/p}. \tag{10}$$

The right-hand side of the inequality above is exactly the right-hand side of the conjectured hypercontractive inequality. Below, we show how to lower bound the left-hand side of the inequality above by the desired left-hand side of the conjectured statement. To do so, we will need the following Minkowski's inequality.

Lemma 7 (Minkowski's inequality, [HLP52, Theorem 26]). *For any $r_1 \times r_2$ matrix whose rows are given by $u_1, \dots, u_{r_1}$ and whose columns are given by $v_1, \dots, v_{r_2}$, and any $1 \le q_1 \le q_2 \le \infty$,*

$$\|(\|v_1\|_{q_2}, \dots, \|v_{r_2}\|_{q_2})\|_{q_1} \le \|(\|u_1\|_{q_1}, \dots, \|u_{r_2}\|_{q_1})\|_{q_2} .$$

Now, consider the $r^n \times r$ matrix whose entries are given by $c_{S,i} = r^{n/2} \left\| \zeta^{|S|/2} \widehat{g_i}(S) \right\|_p$ for every $i \in \{0, \dots, r-1\}$ and $S \in \mathbb{Z}_r^n$. Then the left-hand side of Eq. (10) can be written as

$$\begin{aligned}
\left(\frac{1}{r} \sum_{i=0}^{r-1} \left(\sum_{S \in \mathbb{Z}_r^n} \zeta^{|S|} \|\widehat{g_i}(S)\|_p^2 \right)^{p/2} \right)^{1/p} &= \left(\frac{1}{r} \sum_{i=0}^{r-1} \left(\frac{1}{r^n} \sum_{S \in \mathbb{Z}_r^n} c_{S,i}^2 \right)^{p/2} \right)^{1/p} \\
&\ge \left(\frac{1}{r^n} \sum_{S \in \mathbb{Z}_r^n} \left(\frac{1}{r} \sum_{i=0}^{r-1} c_{S,i}^p \right)^{2/p} \right)^{1/2} \\
&= \left(\sum_{S \in \mathbb{Z}_r^n} \zeta^{|S|} \left(\frac{1}{r} \sum_{i=0}^{r-1} \|\widehat{g_i}(S)\|_p^p \right)^{2/p} \right)^{1/2} , \qquad (11)
\end{aligned}$$

where the first inequality follows from Lemma 7 with $q_1 = p$ and $q_2 = 2$.

Now, for a fixed $S \in \mathbb{Z}_r^n$, we use the base case $n = 1$, i.e., Eq. (9), on the functions $h(i) = \widehat{g_i}(S)$ in order to get

$$\left(\frac{1}{r} \sum_{i=0}^{r-1} \|\widehat{g_i}(S)\|_p^p \right)^{2/p} \ge \sum_{i=0}^{r-1} \zeta^{|i|} \left\| \frac{1}{r} \sum_{j=0}^{r-1} h(j) \omega_r^{-ij} \right\|_p^2 = \sum_{i=0}^{r-1} \zeta^{|i|} \left\| \frac{1}{r} \sum_{j=0}^{r-1} \widehat{g_j}(S) \omega_r^{-ij} \right\|_p^2 .$$

Plugging this back into Eq. (11), we have

$$\begin{aligned}
\left(\sum_{S \in \mathbb{Z}_r^n} \zeta^{|S|} \left(\frac{1}{r} \sum_{i=0}^{r-1} \|\widehat{g_i}(S)\|_p^p \right)^{2/p} \right)^{1/2} &\ge \left(\sum_{S \in \mathbb{Z}_r^n} \sum_{i=0}^{r-1} \zeta^{|S|+|i|} \left\| \frac{1}{r} \sum_{j=0}^{r-1} \widehat{g_j}(S) \omega_r^{-ij} \right\|_p^2 \right)^{1/2} \\
&= \left(\sum_{S \in \mathbb{Z}_r^{n+1}} \zeta^{|S|} \|\widehat{f}(S)\|_p^2 \right)^{1/2} ,
\end{aligned}$$

where we used the fact that $g_j = f|_{x_{n+1}=j}$, so, for every $i \in \mathbb{Z}_r$ and $S \in \mathbb{Z}_r^n$, we have that $\widehat{f}(S,i) = \frac{1}{r} \sum_{j=0}^{r-1} \widehat{g_j}(S) \omega_r^{-ij}$. The lower bound we obtained above is exactly the left-hand side of the conjectured hypercontractive inequality, which proves the theorem statement. □

Remark 2 (Comparison with hypercontractivity for real numbers). *For real functions $f : \mathbb{Z}_r^n \to \mathbb{R}$, it is known that [LO00, Wol07] (see also [O'D14, Theorem 10.18])*

$$\left(\sum_{S \in \mathbb{Z}_r^n} \rho^{2|S|} |\widehat{f}(S)|^2 \right)^{1/2} \le \left(\frac{1}{r^n} \sum_{x \in \mathbb{Z}_r^n} |f(x)|^p \right)^{1/p} ,$$

where $\rho \le \sqrt{\frac{(r-1)^{1-1/p} - (r-1)^{-(1-1/p)}}{(r-1)^{1/p} - (r-1)^{-1/p}}}$. Moreover, this bound on ρ is perfectly sharp, meaning that our bound $\rho \le \sqrt{\frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}}$ in Theorem 6 can possibly be improved.

4 Hidden Hypermatching Problem

The Boolean Hidden Matching (BHM) problem is a canonical problem in one-way communication complexity. Here, Alice is given a string $x \in \{0,1\}^n$, while Bob is given a string $w \in \{0,1\}^{\alpha n/2}$ and a sequence of $\alpha n/2$ disjoint pairs $(i_1, j_1), \ldots, (i_{\alpha n/2}, j_{\alpha n/2}) \in [n]^2$ (called α-partial matching), where $\alpha \in (0,1]$. Let $z \in \{0,1\}^{\alpha n/2}$ be the string defined as $z_\ell = x_{i_\ell} \oplus x_{j_\ell}$ for $\ell \in [\alpha n/2]$. It is promised that $z \oplus w = b^{\alpha n/2}$ for some $b \in \{0,1\}$. By sending a single message from Alice to Bob, their task is to output b, i.e., to decide whether $z \oplus w$ equals the all 0 string or the all 1 string.

The BHM problem was proposed by Bar-Yossef *et al.* [BYJK04], where they showed a simple quantum protocol using only $O(\log n)$ qubits of communication. Later, Gavinsky *et al.* [GKK+07], by using Fourier techniques, specially the inequality of Kahn, Kalai and Linial [KKL89], proved that any classical protocol needs to communicate $\Omega(\sqrt{n})$ bits in order to solve the problem. Since then, many generalizations of the BHM problem were proposed. Verbin and Yu [VY11] extended the $\alpha n/2$ disjoint pairs received by Bob to $\alpha n/t$ disjoint t-tuples $(M_{1,1}, \ldots, M_{1,t}), \ldots, (M_{\alpha n/t,1}, \ldots, M_{\alpha n/t,t})$ (called "hypermatching"). The main task now is to compute the parity $z_\ell = \bigoplus_{k=1}^{t} x_{M_{\ell,k}}$ of a "hyperedge". Verbin and Yu named the resulting problem Boolean Hidden Hypermatching (2-HH(α, t, n)),[4] proved a lower bound $\Omega(n^{1-1/t})$ on any classical communication protocol and used this to bound the amount of space required in streaming algorithms. A quantum lower bound $\Omega(n^{1-2/t})$ on the 2-HH(α, t, n) problem was later proven by Shi, Wu and Yu [SWY12].

Subsequently, Kapralov, Khanna and Sudan [KKS14] proposed the Boolean Hidden Partition, where Bob does not receive a matching anymore, but the edges of any graph G. It is promised that either $Mx = w$, where M is the edge incidence matrix of G, or w is taken uniformly at random independently on x, and Alice and Bob's task is to decide which is the correct case. In another line, Guruswami and Tao [GT19] introduced the r-ary Hidden Matching (r-HH$(\alpha, 2, n)$) problem, where now x and w are over $\mathbb{Z}_r$ instead of $\{0,1\}$, Bob receives a matching M (and not a general graph), and either $Mx = w$ or w is drawn uniformly at random. Finally, Doriguello and Montanaro [DM20] expanded the 2-HH(α, t, n) problem to computing a fixed Boolean function on the hyperedges of Bob's hypermatching instead of the Parity function. Here we shall consider the standard Hidden Hypermatching problem over a larger alphabet.

In the following, an α-partial t-hypermatching $M \in \mathcal{M}^{\alpha}_{t,n}$ on n vertices is defined as a sequence of $\alpha n/t$ disjoint hyperedges $(M_{1,1}, \ldots, M_{1,t}), \ldots, (M_{\alpha n/t,1}, \ldots, M_{\alpha n/t,t}) \in [n]^t$ with t vertices each, where $\mathcal{M}^{\alpha}_{t,n}$ is the set of all such hypermatchings. If $\alpha = 1$, we shall write $\mathcal{M}_{t,n}$.

Definition 8. *Let $n, t \in \mathbb{N}$ be such that $t|n$ and $\alpha \in (0,1]$. In the r-ary Hidden Hypermatching (r-HH(α, t, n)) problem, Alice gets $x \in \mathbb{Z}_r^n$, Bob gets an α-partial t-hypermatching $M \in \mathcal{M}^{\alpha}_{t,n}$ and a string $w \in \mathbb{Z}_r^{\alpha n/t}$. The hyperedges of M are $(M_{1,1}, \ldots, M_{1,t}), \ldots, (M_{\alpha n/t,1}, \ldots, M_{\alpha n/t,t})$. Let $M \in \{0,1\}^{\alpha n/t \times n}$ also be the incident matrix of Bob's hypermatching. Consider the distributions:*

1. YES *distribution $\mathcal{D}^{\mathsf{YES}}$, let $w = Mx$ (where the matrix product Mx is over $\mathbb{Z}_r$);*
2. NO *distribution $\mathcal{D}^{\mathsf{NO}}$, w is uniformly random in $\mathbb{Z}_r^{\alpha n/t}$.*

In the r-ary Hidden Hypermatching problem, Alice sends a message to Bob who needs to decide with high probability if w is drawn from $\mathcal{D}^{\mathsf{YES}}$ or $\mathcal{D}^{\mathsf{NO}}$.

[4] We use the notation r-HH(α, t, n) for simplicity in exposition throughout.

4.1 Quantum protocol for r-ary Hidden Hypermatching

For $t = 2$, we obtain a efficient quantum communication protocol to solve the r-ary Hidden Hypermatching problem.

Theorem 9. *Given $\varepsilon > 0$, there is a protocol for the r-$\mathsf{HH}(\alpha, 2, n)$ problem with one-sided error ε and $O(\frac{1}{\alpha}\log(nr)\log(1/\varepsilon))$ qubits of communication from Alice to Bob.*

Proof. Let $M \in \mathcal{M}^{\alpha}_{2,n}$ be Bob's matching with edges $(M_{1,1}, M_{1,2}), \ldots, (M_{\alpha n/2,1}, M_{\alpha n/2,2})$. Alice sends the following state to Bob,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|x_i, i\rangle,$$

who measures it with the POVM $\{E_1, \ldots, E_{\alpha n/2}, \mathbb{I} - \sum_{i=1}^{\alpha n/2} E_i\}$, where

$$E_i := |M_{i,1}\rangle\langle M_{i,1}| + |M_{i,2}\rangle\langle M_{i,2}|$$

for $i \in \{1, \ldots, \alpha n/2\}$. With probability $1 - \alpha$ the POVM outputs the final outcome, and with probability α he will obtain a measurement outcome E_i with $i \in [\alpha n/2]$ and get the state

$$|\psi\rangle := \frac{1}{\sqrt{2}}(|x_{M_{i,1}}, M_{i,1}\rangle + |x_{M_{i,2}}, M_{i,2}\rangle).$$

By repeating the procedure $O(1/\alpha)$ times, Bob obtains an outcome $i \in [\alpha n/2]$ with high probability.

For the ease of notation, we can write $M_{i,1} = 0$ and $M_{i,2} = 1$ (note that Bob knows the values of both $M_{i,1}, M_{i,2}$ explicitly). Bob now attaches a $\lceil \log_2 r \rceil$-qubit register in the state $|0\rangle$ to $|\psi\rangle$ and applies a Fourier transform Q_r over $\mathbb{Z}_r$ to it to obtain

$$|0\rangle|\psi\rangle \to \frac{1}{\sqrt{r}}\sum_{k=0}^{r-1}|k\rangle|\psi\rangle.$$

From now on we shall consider a parameter $\ell \in \mathbb{Z}_r$ to be determined later. Let X be the usual Pauli operator and let S_ℓ and P be the shift and phase operators over $\mathbb{Z}_r$ defined as $S_\ell|k\rangle = |\ell - k\rangle$ and $P|k\rangle = \omega_r^k|k\rangle$ for $k \in \mathbb{Z}_r$. Let $C_\ell := PS_\ell P \otimes X$. Bob applies the controlled unitary U_ℓ defined as $U_\ell|k\rangle|\psi\rangle = |k\rangle C_\ell^k|\psi\rangle$ on his state, followed by an inverse Fourier transform $Q_r^\dagger$ on his first register to get

$$\frac{1}{\sqrt{r}}\sum_{k=0}^{r-1}U_\ell|k\rangle|\psi\rangle = \frac{1}{\sqrt{r}}\sum_{k=0}^{r-1}|k\rangle C_\ell^k|\psi\rangle \xrightarrow{Q_r^\dagger\otimes\mathbb{I}} \frac{1}{r}\sum_{j=0}^{r-1}\sum_{k=0}^{r-1}\omega_r^{-jk}|j\rangle C_\ell^k|\psi\rangle.$$

Let us calculate $C_\ell|\psi\rangle$ and $C_\ell^2|\psi\rangle$. We have

$$\begin{aligned}
C_\ell|\psi\rangle &= \frac{1}{\sqrt{2}}(PS_\ell P \otimes X)(|x_0, 0\rangle + |x_1, 1\rangle) \\
&= \frac{1}{\sqrt{2}}(PS_\ell \otimes \mathbb{I})(\omega_r^{x_0}|x_0, 1\rangle + \omega_r^{x_1}|x_1, 0\rangle) \\
&= \frac{1}{\sqrt{2}}(P \otimes \mathbb{I})(\omega_r^{x_0}|\ell - x_0, 1\rangle + \omega_r^{x_1}|\ell - x_1, 0\rangle) \\
&= \frac{\omega_r^\ell}{\sqrt{2}}(|\ell - x_1, 0\rangle + |\ell - x_0, 1\rangle) \qquad (12)
\end{aligned}$$

and

$$\begin{aligned} C_\ell^2|\psi\rangle &= \frac{\omega_r^\ell}{\sqrt{2}}(PS_\ell P \otimes X)(|\ell - x_1, 0\rangle + |\ell - x_0, 1\rangle) \\ &= \frac{\omega_r^\ell}{\sqrt{2}}(PS_\ell \otimes \mathbb{I})(\omega_r^{\ell - x_1}|\ell - x_1, 1\rangle + \omega_r^{\ell - x_0}|\ell - x_0, 0\rangle) \\ &= \frac{\omega_r^\ell}{\sqrt{2}}(P \otimes \mathbb{I})(\omega_r^{\ell - x_1}|x_1, 1\rangle + \omega_r^{\ell - x_0}|x_0, 0\rangle) \\ &= \omega_r^{2\ell}|\psi\rangle. \end{aligned}$$

We can see from the above that $C_\ell^{2k}|\psi\rangle = \omega_r^{2k\ell}|\psi\rangle$. By defining $\Delta_\ell := \ell - (x_0 + x_1)$ and $\delta_k = 1$ if k is odd and 0 otherwise, Bob's final state is

$$\frac{1}{r}\sum_{j=0}^{r-1}\sum_{k=0}^{r-1}\omega_r^{k(\ell-j)}|j\rangle\frac{1}{\sqrt{2}}(|x_0 + \Delta_\ell\delta_k, 0\rangle + |x_1 + \Delta_\ell\delta_k, 1\rangle). \tag{13}$$

Now observe that, if $\ell = x_0 + x_1$, then $C_\ell|\psi\rangle = \omega_r^\ell|\psi\rangle$ in Eq. (12). This means that Bob's state in Eq. (13) becomes $|x_0 + x_1\rangle|\psi\rangle$, and if he measures his first register, he obtains $x_0 + x_1 \pmod r$ with certainty.

On the other hand, if $\ell \neq x_0 + x_1$, then the probability of measuring the first register and obtaining the outcome $m \in \mathbb{Z}_r$ is

$$\begin{aligned} \Pr[m] &= \frac{1}{2r^2}\sum_{k_1,k_2=0}^{r-1}\omega_r^{(\ell-m)(k_1-k_2)}(\langle x_0 + \Delta_\ell\delta_{k_2}|x_0 + \Delta_\ell\delta_{k_1}\rangle + \langle x_1 + \Delta_\ell\delta_{k_2}|x_1 + \Delta_\ell\delta_{k_1}\rangle) \\ &= \frac{1}{r^2}\sum_{k_1,k_2\ \text{even}}^{r-1}\omega_r^{(\ell-m)(k_1-k_2)} + \frac{1}{r^2}\sum_{k_1,k_2\ \text{odd}}^{r-1}\omega_r^{(\ell-m)(k_1-k_2)} \\ &= \left|\frac{1}{r}\sum_{k\ \text{even}}^{r-1}\omega_r^{k(\ell-m)}\right|^2 + \left|\frac{1}{r}\sum_{k\ \text{odd}}^{r-1}\omega_r^{k(\ell-m)}\right|^2. \end{aligned}$$

It is not hard to see that the above probability is maximum for when $m = \ell$, in which case

$$\Pr[m = \ell] = \frac{1}{r^2}\left\lfloor\frac{r+1}{2}\right\rfloor^2 + \frac{1}{r^2}\left\lfloor\frac{r}{2}\right\rfloor^2 = \begin{cases} \frac{1}{2} & r \text{ even}, \\ \frac{1}{2} + \frac{1}{2r^2} & r \text{ odd}. \end{cases}$$

Given the considerations above, Bob uses the following strategy: he picks ℓ as the corresponding entry w_i from $w \in \mathbb{Z}_r^{\alpha n/2}$ given the measured hyperedge $(M_{i,1}, M_{i,2})$. If the outcome m from measuring his final state in Eq. (13) equals w_i, then he outputs YES, otherwise he outputs NO. Indeed, in the YES instance, $w_i = x_{M_{i,1}} + x_{M_{i,2}}$ and so m equals w_i with probability 1, while in the NO instance, m equals w_i with probability at most $\frac{1}{2} + \frac{1}{2r^2}$. Thus the communication protocol has one-sided error at most $\frac{1}{2} + \frac{1}{2r^2}$, i.e., $\Pr[\text{error}|\text{YES}] = 0$ and $\Pr[\text{error}|\text{NO}] \leq \frac{1}{2} + \frac{1}{2r^2}$. By repeating the whole protocol $O(\log(1/\varepsilon))$ more times, the one-sided error probability can be decreased to ε: if in any of the repetitions the final measurement outcome is different from w_i, then Bob knows that NO is the correct answer. □

4.2 Quantum lower bound on r-ary Hidden Hypermatching

In this section we shall turn our attention to proving quantum and classical lower bounds on the amount of communication required by the r-$\mathsf{HH}(\alpha, t, n)$ problem, but first we need the following lemma.

Lemma 10. *Let $f : \mathbb{Z}_r^n \to \mathrm{D}(\mathbb{C}^{2^m})$ be any mapping from an n-bit alphabet to m-qubit density matrices. Then for any $\delta \in [0, 1/(r-1)]$, we have*

$$\sum_{S \in \mathbb{Z}_r^n} \delta^{|S|} \|\widehat{f}(S)\|_1^2 \leq 2^{2(r-1)\delta m}.$$

Proof. Let $p := 1 + (r-1)\delta$. First note that, given the eigenvalues $\sigma_1, \ldots, \sigma_{2^m}$ from $f(x)$, which are non-negative reals that sum to 1, we have

$$\|f(x)\|_p^p = \sum_{i=1}^{2^m} \sigma_i^p \leq \sum_{i=1}^{2^m} \sigma_i = 1.$$

Using Theorem 6 and Remark 1, we now get

$$\sum_{S \in \mathbb{Z}_r^n} \left(\frac{p-1}{r-1}\right)^{|S|} \|\widehat{f}(S)\|_p^2 \leq \left(\frac{1}{r^n} \sum_{x \in \mathbb{Z}_r^n} \|f(x)\|_p^p\right)^{2/p} \leq \left(\frac{1}{r^n} \cdot r^n\right)^{2/p} = 1.$$

On the other hand, the normalized Schatten norm $2^{-m/p}\|\widehat{f}(S)\|_p$ is non-decreasing with p, since $p \leq q \implies \left(\frac{1}{2^m}\sum_{i=1}^{2^m} \sigma_i^p\right)^{1/p} \leq \left(\frac{1}{2^m}\sum_{i=1}^{2^m} \sigma_i^q\right)^{1/q}$ by Hölder's inequality, hence

$$\sum_{S \in \mathbb{Z}_r^n} \left(\frac{p-1}{r-1}\right)^{|S|} 2^{-2m/p} \|\widehat{f}(S)\|_p^2 \geq \sum_{S \in \mathbb{Z}_r^n} \left(\frac{p-1}{r-1}\right)^{|S|} 2^{-2m} \|\widehat{f}(S)\|_1^2.$$

Rearranging the inequalities leads to

$$\sum_{S \in \mathbb{Z}_r^n} \left(\frac{p-1}{r-1}\right)^{|S|} \|\widehat{f}(S)\|_1^2 \leq 2^{2m(1-1/p)} \leq 2^{2m(p-1)}.$$ □

We are now ready to state and prove our main quantum communication complexity lower bound for the r-ary Hidden Hypermatching problem.

Theorem 11. *Any quantum protocol that achieves advantage $\varepsilon > 0$ for the r-$\mathsf{HH}(\alpha, t, n)$ problem with $t \geq 3$ and $\alpha \leq \min(1/2, (r-1)^{-1/2})$ requires at least $m = \Omega(r^{-(1+1/t)}(\varepsilon^2/\alpha)^{2/t}(n/t)^{1-2/t})$ qubits of communication from Alice to Bob.*

Notice that for $r = 2$ our lower bound reads $\Omega(\alpha^{-2/t}(n/t)^{1-2/t})$, which has a better dependence on α compared to the lower bound $\Omega(\log(1/\alpha)(n/t)^{1-2/t})$ from [SWY12]. Also, see Remark 3 at the end of the section for an improvement on the requirement $\alpha \leq \min(1/2, (r-1)^{-1/2})$.

Proof. Consider an m-qubit communication protocol. An arbitrary m-qubit protocol can be viewed as Alice sending an encoding of her input $x \in \mathbb{Z}_r^n$ into a quantum state so that Bob can distinguish if his w was drawn from $\mathcal{D}^{\mathsf{YES}}$ or $\mathcal{D}^{\mathsf{NO}}$. Let $\rho : \mathbb{Z}_r^n \to \mathrm{D}(\mathbb{C}^{2^m})$ be Alice's encoding function. For

our 'hard' distribution, Alice and Bob receive $x \in \mathbb{Z}_r^n$ and $M \in \mathcal{M}_{t,n}^{\alpha}$, respectively, uniformly at random, while Bob's input $w \in \mathbb{Z}_r^{\alpha n/t}$ is drawn from the distribution $\mathcal{D} := \frac{1}{2}\mathcal{D}^{\mathsf{YES}} + \frac{1}{2}\mathcal{D}^{\mathsf{NO}}$, i.e., with probability $1/2$ is comes from $\mathcal{D}^{\mathsf{YES}}$, and with probability $1/2$ it comes from $\mathcal{D}^{\mathsf{NO}}$. Let $p_x := r^{-n}$, $p_M := |\mathcal{M}_{t,n}^{\alpha}|^{-1}$ and $p_w := r^{-\alpha n/t}$, then our hard distribution $\mathcal{P}$ is

$$\Pr[x, \mathsf{YES}, M, w] = \frac{1}{2} p_x \cdot p_M \cdot [Mx = w], \quad \Pr[x, \mathsf{NO}, M, w] = \frac{1}{2} p_x \cdot p_M \cdot p_w. \tag{14}$$

Conditioning on Bob's input (M, w), from his perspective, Alice sends the message ρ_x with probability $\Pr[x|M, w]$. Therefore, conditioned on an instance of the problem (YES or NO), Bob receives one of the following two quantum states $\rho_{\mathsf{YES}}^{M,w}$ and $\rho_{\mathsf{NO}}^{M,w}$, each appearing with probability $\Pr[\mathsf{YES}|M, w]$ and $\Pr[\mathsf{NO}|M, w]$, respectively,

$$\begin{aligned} \rho_{\mathsf{YES}}^{M,w} &= \sum_{x \in \mathbb{Z}_r^n} \Pr[x|\mathsf{YES}, M, w] \cdot \rho_x = \frac{1}{\Pr[\mathsf{YES}, M, w]} \sum_{x \in \mathbb{Z}_r^n} \Pr[x, \mathsf{YES}, M, w] \cdot \rho_x, \\ \rho_{\mathsf{NO}}^{M,w} &= \sum_{x \in \mathbb{Z}_r^n} \Pr[x|\mathsf{NO}, M, w] \cdot \rho_x = \frac{1}{\Pr[\mathsf{NO}, M, w]} \sum_{x \in \mathbb{Z}_r^n} \Pr[x, \mathsf{NO}, M, w] \cdot \rho_x. \end{aligned} \tag{15}$$

Bob's best strategy to determine the distribution of w conditioning on his input (M, w) is no more than the chance to distinguish between these two quantum states $\rho_{\mathsf{YES}}^{M,w}$ and $\rho_{\mathsf{NO}}^{M,w}$.

Now let ε_{bias} be the bias of the protocol that distinguishes between $\rho_{\mathsf{YES}}^{M,w}$ and $\rho_{\mathsf{NO}}^{M,w}$. According to Lemma 2, the bias ε_{bias} of any quantum protocol for a fixed M and w can be upper bounded as

$$\varepsilon_{bias} \leq \left\| \Pr[\mathsf{YES}|M, w] \cdot \rho_{\mathsf{YES}}^{M,w} - \Pr[\mathsf{NO}|M, w] \cdot \rho_{\mathsf{NO}}^{M,w} \right\|_1.$$

We prove in Theorem 12 below that, if $m \leq \frac{\gamma}{r^{1+1/t}} (\frac{\varepsilon^2}{\alpha})^{2/t} (n/t)^{1-2/t}$ for a universal constant γ, then the average bias over M and w is at most ε^2, i.e.,

$$\mathop{\mathbb{E}}_{(M,w) \sim \mathcal{P}_{M,w}} [\varepsilon_{bias}] \leq \varepsilon^2,$$

where $\mathcal{P}_{M,w}$ is the marginal distribution of $\mathcal{P}$. Therefore, by Markov's inequality, for at least a $(1-\varepsilon)$-fraction of M and w, the bias in distinguishing between $\rho_{\mathsf{YES}}^{M,w}$ and $\rho_{\mathsf{NO}}^{M,w}$ is ε small. Therefore, Bob's advantage over randomly guessing the right distribution will be at most ε (for the event that M and w are such that the distance between $\rho_{\mathsf{YES}}^{M,w}$ and $\rho_{\mathsf{NO}}^{M,w}$ is more than ε) plus $\varepsilon/2$ (for the advantage over random guessing when $\varepsilon_{bias} \leq \varepsilon$), and so $m = \Omega(r^{-(1+1/t)} (\varepsilon^2/\alpha)^{2/t} (n/t)^{1-2/t})$. □

Theorem 12. *For $x \in \mathbb{Z}_r^n$, $M \in \mathcal{M}_{t,n}^{\alpha}$, $w \in \mathbb{Z}_r^{\alpha n/t}$ and $b \in \{\mathsf{YES}, \mathsf{NO}\}$, consider the probability distribution $\mathcal{P}$ defined in Eq. (14). Given an encoding function $\rho : \mathbb{Z}_r^n \to \mathrm{D}(\mathbb{C}^{2^m})$, consider the quantum states $\rho_{\mathsf{YES}}^{M,w}$ and $\rho_{\mathsf{NO}}^{M,w}$ from Eq. (15). If $\alpha \leq \min(1/2, (r-1)^{-1/2})$, there is a universal constant $\gamma > 0$ (independent of n, t, r and α), such that, for all $\varepsilon > 0$, if $m \leq \frac{\gamma}{r^{1+1/t}} (\frac{\varepsilon^2}{\alpha})^{2/t} (n/t)^{1-2/t}$, then*

$$\mathop{\mathbb{E}}_{(M,w) \sim \mathcal{P}_{M,w}} \left[\left\| \Pr[\mathsf{YES}|M, w] \cdot \rho_{\mathsf{YES}}^{M,w} - \Pr[\mathsf{NO}|M, w] \cdot \rho_{\mathsf{NO}}^{M,w} \right\|_1 \right] \leq \varepsilon^2.$$

Proof. For the ease of notation, we shall write

$$\varepsilon_{bias} := \mathop{\mathbb{E}}_{(M,w) \sim \mathcal{P}_{M,w}} \left[\left\| \Pr[\mathsf{YES}|M, w] \cdot \rho_{\mathsf{YES}}^{M,w} - \Pr[\mathsf{NO}|M, w] \cdot \rho_{\mathsf{NO}}^{M,w} \right\|_1 \right].$$

Therefore, we have that

$$\begin{aligned}
\varepsilon_{bias} &= \sum_{M\in\mathcal{M}^{\alpha}_{t,n}} \sum_{w\in\mathbb{Z}_r^{\alpha n/t}} \Pr[M,w]\cdot \left\| \Pr[\mathsf{YES}|M,w]\cdot\rho^{M,w}_{\mathsf{YES}} - \Pr[\mathsf{NO}|M,w]\cdot\rho^{M,w}_{\mathsf{NO}}\right\|_1 \\
&= \sum_{M\in\mathcal{M}^{\alpha}_{t,n}} \sum_{w\in\mathbb{Z}_r^{\alpha n/t}} \left\| \sum_{x\in\mathbb{Z}_r^n} \Big(\Pr[x,\mathsf{YES},M,w]\cdot\rho_x - \Pr[x,\mathsf{NO},M,w]\cdot\rho_x\Big)\right\|_1 \\
&= \sum_{M\in\mathcal{M}^{\alpha}_{t,n}} \sum_{w\in\mathbb{Z}_r^{\alpha n/t}} \left\| \sum_{x\in\mathbb{Z}_r^n} \frac{1}{2} p_x\cdot p_M\left(\big[Mx=w\big]-p_w\right)\rho_x\right\|_1 \qquad \text{(By Eqs. (14), (15))} \\
&= \sum_{M\in\mathcal{M}^{\alpha}_{t,n}} \sum_{w\in\mathbb{Z}_r^{\alpha n/t}} \left\| \sum_{x\in\mathbb{Z}_r^n} \frac{1}{2} p_x\cdot p_M\left(\big[Mx=w\big]-p_w\right)\cdot\sum_{S\in\mathbb{Z}_r^n}\widehat{\rho}(S)\omega_r^{S\cdot x}\right\|_1 \\
&\qquad\qquad \text{(Fourier decomposition of } \rho) \\
&= \sum_{M\in\mathcal{M}^{\alpha}_{t,n}} \sum_{w\in\mathbb{Z}_r^{\alpha n/t}} \left\| \sum_{S\in\mathbb{Z}_r^n} u(M,w,S)\widehat{\rho}(S)\right\|_1 \\
&\le \sum_{S\in\mathbb{Z}_r^n}\sum_{M\in\mathcal{M}^{\alpha}_{t,n}}\sum_{w\in\mathbb{Z}_r^{\alpha n/t}} |u(M,w,S)|\cdot\|\widehat{\rho}(S)\|_1,
\end{aligned}$$

where we defined

$$u(M,w,S) := \frac{1}{2}\sum_{x\in\mathbb{Z}_r^n} p_x\cdot p_M\cdot\omega_r^{S\cdot x}\left(\big[Mx=w\big]-p_w\right). \tag{16}$$

Next, we upper bound the quantity $u(M,w,S)$ using the lemma below. In the following lemma, given an α-partial hypermatching $M\in\mathcal{M}^{\alpha}_{t,n}$, we can, without lost of generality, complete M with $(1-\alpha)n/t$ remaining hyperedges and turn it into a perfect hypermatching, i.e., we can assume that $M\in\mathcal{M}_{t,n}$. Moreover, we shall write $S|_{M_i} = S_{M_{i,1}}S_{M_{i,2}}\ldots S_{M_{i,t}}\in\mathbb{Z}_r^t$ to denote the string S restricted to the hyperedge $M_i=(M_{i,1},\ldots,M_{i,t})$, where $S_{M_{i,j}}$ is the $M_{i,j}$-th entry of S. The same applies to $x\in\mathbb{Z}_r^n$.

Lemma 13. *Let $M\in\mathcal{M}_{t,n}$, $w\in\mathbb{Z}_r^{\alpha n/t}$ and $S\in\mathbb{Z}_r^n$. Define the set*

$$\begin{aligned}
\Delta(M) = \{S\in\mathbb{Z}_r^n\setminus\{0^n\} \mid\ & S_{M_{i,1}} = S_{M_{i,2}} = \cdots = S_{M_{i,t}} \text{ for every } i\in[\alpha n/t] \\
& \text{and } S|_{M_i} = 0^t \text{ for every } i > \alpha n/t\}.
\end{aligned}$$

Given $u(M,w,S)$ as defined in Eq. (16), we have $u(M,w,S) = \frac{1}{2}\cdot r^{-\alpha n/t}\cdot p_M$ if $S\in\Delta(M)$ and 0 if $S\notin\Delta(M)$.

Proof. Recall the definition of u:

$$u(M,w,S) = \frac{1}{2}\sum_{x\in\mathbb{Z}_r^n} p_x\cdot p_M\cdot\omega_r^{S\cdot x}\left(\big[Mx=w\big]-p_w\right).$$

In order to understand this expression, we start with the following:

$$\sum_{x\in\mathbb{Z}_r^n}\omega_r^{S\cdot x}\big[Mx=w\big] = \sum_{x\in\mathbb{Z}_r^n}\omega_r^{S\cdot x}\prod_{i=1}^{\alpha n/t}\big[(Mx)_i=w_i\big]$$

$$= \sum_{x\in\mathbb{Z}_r^n} \omega_r^{S\cdot x} \prod_{i=1}^{\alpha n/t} \left[\sum_{j=1}^{t} x_{M_{i,j}} \equiv w_i \pmod r\right]$$
$$= \sum_{x\in\mathbb{Z}_r^n} \omega_r^{\sum_{i=1}^{n/t}\sum_{j=1}^{t} S_{M_{i,j}} x_{M_{i,j}}} \prod_{i=1}^{\alpha n/t} \left[\sum_{j=1}^{t} x_{M_{i,j}} \equiv w_i \pmod r\right]$$
$$= \sum_{x\in\mathbb{Z}_r^n} \omega_r^{\sum_{i=1}^{n/t}\sum_{j=1}^{t} S_{M_{i,j}} x_{(i-1)t+j}} \prod_{i=1}^{\alpha n/t} \left[\sum_{j=1}^{t} x_{(i-1)t+j} \equiv w_i \pmod r\right],$$

where we reordered $x \in \mathbb{Z}_r^n$ in the last step. Therefore

$$\sum_{x\in\mathbb{Z}_r^n} \omega_r^{S\cdot x}\big[Mx = w\big] = \left(\prod_{i=1}^{\alpha n/t} \sum_{x\in\mathbb{Z}_r^t} \omega_r^{\sum_{j=1}^{t} S_{M_{i,j}} x_j} \left[\sum_{j=1}^{t} x_j \equiv w_i \pmod r\right]\right)\left(\prod_{i>\alpha n/t}^{n/t} \sum_{x\in\mathbb{Z}_r^t} \omega_r^{\sum_{j=1}^{t} S_{M_{i,j}} x_j}\right)$$
$$= r^{n(1-\alpha)} \prod_{i=1}^{\alpha n/t} \sum_{x\in\mathbb{Z}_r^t} \omega_r^{S|_{M_i}\cdot x} \left[\sum_{j=1}^{t} x_j \equiv w_i \pmod r\right],$$

where $S|_{M_i} = 0^t$ for all $i > \alpha n/t$, otherwise the expression above is 0. Now we use that

$$\sum_{j=1}^{t} x_j \equiv w_i \pmod r \implies x_t \equiv w_i - \sum_{j=1}^{t-1} x_j \pmod r,$$

and so

$$S|_{M_i}\cdot x = \sum_{j=1}^{t} S_{M_{i,j}} x_i = \sum_{j=1}^{t-1} S_{M_{i,j}} x_j + S_{M_{i,t}}\left(w_i - \sum_{j=1}^{t-1} x_j\right) = S_{M_{i,t}} w_i + \sum_{j=1}^{t-1} (S_{M_{i,j}} - S_{M_{i,t}}) x_j$$

modulo r. This leads to

$$\sum_{x\in\mathbb{Z}_r^n} \omega_r^{S\cdot x}\big[Mx = w\big] = r^{n(1-\alpha)} \prod_{i=1}^{\alpha n/t} \omega_r^{S_{M_{i,t}} w_i} \sum_{x\in\mathbb{Z}_r^{t-1}} \omega_r^{\sum_{j=1}^{t-1}(S_{M_{i,j}} - S_{M_{i,t}})x_j} = \frac{r^n}{r^{\alpha n/t}} \prod_{i=1}^{\alpha n/t} \omega_r^{S_{M_{i,t}} w_i} \tag{17}$$

if, for all $i \in [\alpha n/t]$, $S_{M_{i,j}}$ is constant for all $j \in [t]$, i.e., if $S_{M_{i,1}} = S_{M_{i,2}} = \cdots = S_{M_{i,t}}$ for any $i \in [\alpha n/t]$. Otherwise the above expression is 0. Thus, if $S_{M_{i,1}} = S_{M_{i,2}} = \cdots = S_{M_{i,t}}$ for any $i \in [\alpha n/t]$ and $S|_{M_i} = 0^t$ for $i > \alpha n/t$, then we can use Eq. (17) to get (remember that $p_x := r^{-n}$ and $p_w := r^{-\alpha n/t}$)

$$|u(M,w,S)| = \frac{1}{2}\left|\sum_{x\in\mathbb{Z}_r^n} p_x p_M \omega_r^{S\cdot x}\left(\big[Mx = w\big] - p_w\right)\right| = \frac{1}{2}\frac{1}{r^{\alpha n/t}} p_M \left|\prod_{i=1}^{\alpha n/t} \omega_r^{S_{M_{i,t}} w_i} - [S = 0^n]\right|$$
$$= \begin{cases} 0 & \text{if } S = 0^n, \\ \frac{1}{2} r^{-\alpha n/t} p_M & \text{if } S \neq 0^n. \end{cases}$$

Hence, we have

$$|u(M,w,S)| = \begin{cases} 0 & \text{if } S = 0^n, \\ \frac{1}{2} r^{-\alpha n/t} p_M & \text{if } S_{M_{i,1}} = S_{M_{1,2}} = \cdots = S_{M_{i,t}}\ \forall i \in [\alpha n/t] \text{ and } S|_{M_i} = 0^t\ \forall i > \alpha n/t, \\ 0 & \text{otherwise,} \end{cases}$$

proving the lemma statement □

We now proceed to upper bound ε_{bias} using the expression for $|u(M,w,S)|$ from Lemma 13. For $S \in \mathbb{Z}_r^n$, let $|S| := |\{i \in [n] : S_i \neq 0\}|$. Notice that, if $S \in \Delta(M)$, then $|S| = kt$ for some $k \in [\alpha n/t]$. Hence, we have that

$$\varepsilon_{bias} \leq \frac{1}{2} \sum_{S \in \mathbb{Z}_r^n} \sum_{\substack{M \in \mathcal{M}_{t,n}^\alpha \\ S \in \Delta(M)}} p_M \sum_{w \in \mathbb{Z}_r^{\alpha n/t}} \frac{1}{r^{\alpha n/t}} \|\widehat{\rho}(S)\|_1 = \frac{1}{2} \sum_{k=1}^{\alpha n/t} \sum_{\substack{S \in \mathbb{Z}_r^n \\ |S|=kt}} \sum_{\substack{M \in \mathcal{M}_{t,n}^\alpha \\ S \in \Delta(M)}} p_M \|\widehat{\rho}(S)\|_1$$
$$= \frac{1}{2} \sum_{k=1}^{\alpha n/t} \sum_{\substack{S \in \mathbb{Z}_r^n \\ |S|=kt}} \Pr_{M \sim \mathcal{M}_{t,n}^\alpha} [S \in \Delta(M)] \cdot \|\widehat{\rho}(S)\|_1,$$

using that

$$\sum_{\substack{M \in \mathcal{M}_{t,n}^\alpha \\ S \in \Delta(M)}} p_M = \Pr_{M \sim \mathcal{M}_{t,n}^\alpha} [S \in \Delta(M)].$$

We now upper bound this probability using the following lemma.

Lemma 14. *Let $t \in \mathbb{Z}$. Let $S \in \mathbb{Z}_r^n$ with $k_j := \frac{1}{t} \cdot |\{i \in [n] : S_i = j\}| \in \mathbb{Z}$ for $j \in \{1, \ldots, r-1\}$. Let $k := \sum_{j=1}^{r-1} k_j$. For any $M \in \mathcal{M}_{t,n}^\alpha$, let $\Delta(M)$ be the set from Lemma 13. Then*

$$\Pr_{M \sim \mathcal{M}_{t,n}^\alpha} [S \in \Delta(M)] = \frac{\binom{\alpha n/t}{k}}{\binom{n}{kt}} \frac{k!}{(kt)!} \prod_{j=1}^{r-1} \frac{(k_j t)!}{k_j!}.$$

Proof. We can assume without lost of generality that $S = 1^{k_1 t} 2^{k_2 t} \ldots (r-1)^{k_{r-1} t} 0^{n-kt}$. First note that the total number $|\mathcal{M}_{t,n}^\alpha|$ of α-partial hypermatchings is $n!/\big((t!)^{\alpha n/t}(\alpha n/t)!(n-\alpha n)!\big)$. This can be seen as follows: pick a permutation of n, view the first $\alpha n/t$ tuples of length t as $\alpha n/t$ hyperedges, and ignore the ordering within each hyperedge, the ordering of the $\alpha n/t$ hyperedges and the ordering of the last $n - \alpha n$ vertices. Now, given our particular S, notice that $S \in \Delta(M)$ if, for $j \in [r-1]$, M has exactly k_j hyperedges in

$$\left\{ 1 + t\sum_{i=1}^{j-1} k_i,\ 2 + t\sum_{i=1}^{j-1} k_i,\ 3 + t\sum_{i=1}^{j-1} k_i, \ldots, (k_j - 1) + t\sum_{i=1}^{j-1} k_i,\ t\sum_{i=1}^{j-1} k_i,\ t\sum_{i=1}^{j} k_i \right\},$$

i.e., k_1 hyperedges in $\{1, \ldots, k_1 t\}$, k_2 hyperedges in $\{k_1 t + 1, \ldots, (k_2 + k_1)t\}$, etc., and also $\alpha n/t - k$ hyperedges in $[n] \setminus [kt]$. The number of ways to pick k_j hyperedges in $\left\{1 + t\sum_{i=1}^{j-1} k_i, \ldots, t\sum_{i=1}^{j} k_i\right\}$ is $(k_j t)!/((t!)^{k_j} k_j!)$. The number of ways to pick the remaining $\alpha n/t - k$ hyperedges in $[n] \setminus [kt]$ is $(n-kt)!/((t!)^{\alpha n/t - k}(\alpha n/t - k)!(n-\alpha n)!)$. Hence $\Pr_{M \sim \mathcal{M}_{t,n}^\alpha}[S \in \Delta(M)]$ equals

$$\frac{\frac{(n-kt)!}{(t!)^{\alpha n/t-k}(\alpha n/t-k)!(n-\alpha n)!}}{\frac{n!}{(t!)^{\alpha n/t}(\alpha n/t)!(n-\alpha n)!}} \prod_{j=1}^{r-1} \frac{(k_j t)!}{(t!)^{k_j} k_j!} = \frac{(n-kt)!(\alpha n/t)!}{n!(\alpha n/t - k)!} \prod_{j=1}^{r-1} \frac{(k_j t)!}{k_j!} = \frac{\binom{\alpha n/t}{k}}{\binom{n}{kt}} \frac{k!}{(kt)!} \prod_{j=1}^{r-1} \frac{(k_j t)!}{k_j!}. \qquad \square$$

By using Lemma 14 and the notation $|S|_i := |\{j \in [n] : S_j = i\}|$, we continue upper bounding ε_{bias} as follows

$$\varepsilon_{bias} \leq \frac{1}{2} \sum_{k=1}^{\alpha n/t} \frac{\binom{\alpha n/t}{k}}{\binom{n}{kt}} \sum_{\substack{k_1, \ldots, k_{r-1} \geq 0 \\ \sum_{j=1}^{r-1} k_j = k}} \sum_{\substack{S \in \mathbb{Z}_r^n \\ |S|_i = k_i t,\ i \in [r-1]}} \frac{k!}{(kt)!} \left(\prod_{j=1}^{r-1} \frac{(k_j t)!}{k_j!} \right) \|\widehat{\rho}(S)\|_1$$

$$\leq \frac{1}{2}\sum_{k=1}^{\alpha n/t}\frac{\binom{\alpha n/t}{k}}{\binom{n}{kt}}\sqrt{\sum_{\substack{k_1,\dots,k_{r-1}\geq 0\\ \sum_{j=1}^{r-1}k_j=k}}\sum_{\substack{S\in\mathbb{Z}_r^n\\ |S|_i=k_it,\ i\in[r-1]}}\frac{k!^2}{(kt)!^2}\prod_{j=1}^{r-1}\frac{(k_jt)!^2}{k_j!^2}}\sqrt{\sum_{\substack{k_1,\dots,k_{r-1}\geq 0\\ \sum_{j=1}^{r-1}k_j=k}}\sum_{\substack{S\in\mathbb{Z}_r^n\\ |S|_i=k_it,\ i\in[r-1]}}\|\widehat{\rho}(S)\|_1^2} \tag{18}$$

$$\leq \frac{1}{2}\sum_{k=1}^{\alpha n/t}\frac{\binom{\alpha n/t}{k}}{\binom{n}{kt}}\sqrt{\sum_{\substack{k_1,\dots,k_{r-1}\geq 0\\ \sum_{j=1}^{r-1}k_j=k}}\sum_{\substack{S\in\mathbb{Z}_r^n\\ |S|_i=k_it,\ i\in[r-1]}}\frac{k!^2}{(kt)!^2}\prod_{j=1}^{r-1}\frac{(k_jt)!^2}{k_j!^2}}\sqrt{\sum_{\substack{S\in\mathbb{Z}_r^n\\ |S|=kt}}\|\widehat{\rho}(S)\|_1^2}$$

$$= \frac{1}{2}\sum_{k=1}^{\alpha n/t}\frac{\binom{\alpha n/t}{k}}{\sqrt{\binom{n}{kt}}}\sqrt{\sum_{\substack{k_1,\dots,k_{r-1}\geq 0\\ \sum_{j=1}^{r-1}k_j=k}}\frac{k!^2}{(kt)!}\prod_{j=1}^{r-1}\frac{(k_jt)!}{k_j!^2}}\sqrt{\sum_{\substack{S\in\mathbb{Z}_r^n\\ |S|=kt}}\|\widehat{\rho}(S)\|_1^2}, \tag{19}$$

where Eqs. (18) and (19) used Cauchy-Schwarz inequality and $\sum_{\substack{S\in\mathbb{Z}_r^n\\ |S|_i=k_it}} 1 = \binom{n}{kt}(kt)!\prod_{j=1}^{r-1}\frac{1}{(k_jt)!}$, respectively. We now use the multinomial theorem in

$$\sum_{\substack{k_1,\dots,k_{r-1}\geq 0\\ \sum_{j=1}^{r-1}k_j=k}}\frac{k!^2}{(kt)!}\prod_{j=1}^{r-1}\frac{(k_jt)!}{k_j!^2} = \sum_{\substack{k_1,\dots,k_{r-1}\geq 0\\ \sum_{j=1}^{r-1}k_j=k}}\frac{\binom{k}{k_1,\dots,k_{r-1}}^2}{\binom{kt}{k_1t,\dots,k_{r-1}t}} \leq \sum_{\substack{k_1,\dots,k_{r-1}\geq 0\\ \sum_{j=1}^{r-1}k_j=k}}\binom{k}{k_1,\dots,k_{r-1}} = (r-1)^k, \tag{20}$$

which leads to

$$\varepsilon_{bias} \leq \frac{1}{2}\sum_{k=1}^{\alpha n/t}\alpha^k\frac{\binom{n/t}{k}}{\sqrt{\binom{n}{kt}}}(r-1)^{k/2}\sqrt{\sum_{\substack{S\in\mathbb{Z}_r^n\\ |S|=kt}}\|\widehat{\rho}(S)\|_1^2},$$

where we also used that $\binom{\alpha n/t}{k} \leq \alpha^k\binom{n/t}{k}$ for $\alpha\in[0,1]$. In order to compute the above sum, we shall split it into two parts: one in the range $1\leq k<4rm$, and the other in the range $4rm\leq k\leq \alpha n/t$.

Sum I ($1\leq k<4rm$): in order to upper bound each term, pick $\delta = k/(4rm)$ in Lemma 10, so

$$\sum_{\substack{S\in\mathbb{Z}_r^n\\ |S|=kt}}\|\widehat{\rho}(S)\|_1^2 \leq \frac{1}{\delta^{kt}}\sum_{S\in\mathbb{Z}_r^n}\delta^{|S|}\|\widehat{f}(S)\|_1^2 \leq \frac{1}{\delta^{kt}}2^{2r\delta m} = \left(\frac{2^{1/(2t)}4rm}{k}\right)^{kt}.$$

Therefore, and by using that $m\leq \frac{\gamma}{r^{1+1/t}}(\frac{\varepsilon^2}{\alpha})^{2/t}(n/t)^{1-2/t}$ and $\binom{q}{s}^2\binom{\ell q}{\ell s}^{-1} \leq (\frac{s}{q})^{(\ell-2)s}$ (see [SWY12, Appendix A.5]) for $q=n/t, s=k, \ell=t$, we have

$$\begin{aligned}\frac{1}{2}\sum_{k=1}^{4rm-1}\alpha^k\frac{\binom{n/t}{k}}{\sqrt{\binom{n}{kt}}}(r-1)^{k/2}\sqrt{\sum_{\substack{S\in\mathbb{Z}_r^n\\ |S|=kt}}\|\widehat{\rho}(S)\|_1^2} &\leq \frac{1}{2}\sum_{k=1}^{4rm-1}\alpha^k(r-1)^{k/2}\left(\frac{kt}{n}\right)^{(1-2/t)kt/2}\left(\frac{2^{1/(2t)}4rm}{k}\right)^{kt/2}\\ &\leq \frac{1}{2}\sum_{k=1}^{4rm-1}\alpha^k(r-1)^{k/2}\left(\frac{2^{1/(2t)}4\gamma\varepsilon^{4/t}}{\alpha^{2/t}r^{1/t}k^{2/t}}\right)^{kt/2}\\ &\leq \frac{1}{2}\sum_{k=1}^{4rm-1}\left(\frac{2^{1/4}(4\gamma)^{t/2}\varepsilon^2}{k}\right)^k \leq \frac{\varepsilon^2}{2}\end{aligned}$$

for sufficiently small γ.

Sum II ($4rm \le k \le \alpha n/t$): first we note that the function $g(k) := \alpha^k (r-1)^{k/2} \binom{n/t}{k} / \sqrt{\binom{n}{kt}}$ is non-increasing in the interval $1 \le k \le \alpha n/t \le n/(2t)$. That is because $\alpha\sqrt{r-1} \le 1$, and so

$$\frac{g(k-1)}{g(k)} \ge \frac{\binom{n/t}{k-1}}{\sqrt{\binom{n}{kt-t}}} \frac{\sqrt{\binom{n}{kt}}}{\binom{n/t}{k}} = \sqrt{\frac{kt}{n-kt+t} \prod_{j=1}^{t-1} \frac{n-kt+j}{kt-j}} \ge \sqrt{\frac{kt}{n-kt+t} \prod_{j=1}^{t-1} \frac{n-kt+j+1}{kt-j+1}}$$
$$= \sqrt{\prod_{j=1}^{t-2} \frac{n-kt+j+1}{kt-j}} \ge 1,$$

where we used that $\frac{a}{b} \ge \frac{a+s}{b+s}$ for all $a, b, s > 0$ with $a \ge b$. Hence, and with the aid once more of Lemma 10 with $\delta = 1$ and the inequality $\binom{q}{s}^2 \binom{\ell q}{\ell s}^{-1} \le \left(\frac{s}{q}\right)^{(\ell-2)s}$ (for $q = n/t, s = 2m, \ell = t$) in order to bound $g(4rm)$,

$$\begin{aligned}
\frac{1}{2} \sum_{k=4rm}^{\alpha n/t} \alpha^k \frac{\binom{n/t}{k}}{\sqrt{\binom{n}{kt}}} (r-1)^{k/2} \sqrt{\sum_{\substack{S \in \mathbb{Z}_r^n \\ |S|=kt}} \|\widehat{\rho}(S)\|_1^2} &\le \frac{1}{2} g(4rm) \sum_{k=4rm}^{\alpha n/t} \sqrt{\sum_{\substack{S \in \mathbb{Z}_r^n \\ |S|=kt}} \|\widehat{\rho}(S)\|_1^2} \\
&\le \frac{1}{2} g(4rm) \sqrt{\frac{\alpha n}{t}} \sqrt{\sum_{S \in \mathbb{Z}_r^n} \|\widehat{\rho}(S)\|_1^2} \qquad (21) \\
&\le \frac{1}{2} \left(\alpha\sqrt{r-1}\right)^{4rm} \left(\frac{4rm}{n/t}\right)^{2(t-2)rm} \sqrt{\frac{\alpha n}{t}} 2^{(r-1)m} \\
&\le \frac{1}{2} \left(2^{1/4}\alpha\sqrt{r-1}\right)^{4rm} \left(\frac{(4\gamma)^{t/2}\varepsilon^2}{\alpha\sqrt{r}(n/t)}\right)^{4(1-2/t)rm} \sqrt{\frac{\alpha n}{t}} \\
&\le \frac{\varepsilon^2}{2},
\end{aligned}$$

where Eq. (21) comes from Cauchy-Schwarz, and in the last step we used that $m \ge 1 \implies 4(1-2/t)m \ge 1$ (so n is in the denominator and $\varepsilon^{4(1-2/t)m} \le \varepsilon$) and picked γ sufficiently small.

Finally, merging both results, we get that, if $m \le \frac{\gamma}{r^{1+1/t}} (\frac{\varepsilon^2}{\alpha})^{2/t} (n/t)^{1-2/t}$, then $\varepsilon_{bias} \le \varepsilon^2$. □

A very similar classical communication lower bound for the r-$\mathsf{HH}(\alpha, t, n)$ problem can be proven.

Theorem 15. *Any one-way classical protocol that achieves advantage $\varepsilon > 0$ for the r-$\mathsf{HH}(\alpha, t, n)$ problem with $t \ge 2$ and $\alpha \le 1/2$ requires at least $\Omega(r^{-1}(\varepsilon^4/\alpha)^{1/t}(n/t)^{1-1/t})$ bits of communication.*

The proof is very similar to that of past works [GKK+07, VY11, GT19] and we include it in Appendix A for completeness. We now conclude this section with a remark that improves the r dependence of the α parameter.

Remark 3. *The dependence of α on r can be improved. For example, we can improve the bound in Eq. (20) by observing that $\binom{k}{k_1,\ldots,k_{r-1}}^2 \binom{kt}{k_1 t,\ldots,k_{r-1} t}^{-1} \le 1$, which can be seen from the identity*

$\binom{k}{k_1,\ldots,k_{r-1}} = \binom{k_1}{k_1}\binom{k_1+k_2}{k_2}\cdots\binom{k_1+k_2+\cdots+k_{r-1}}{k_{r-1}}$ *and the inequality* $\binom{q}{s}^2\binom{\ell q}{\ell s}^{-1} \le (\frac{s}{q})^{(\ell-2)s} \le 1$. *Hence*

$$\sum_{\substack{k_1,\ldots,k_{r-1}\ge 0\\ \sum_{j=1}^{r-1}k_j=k}} \frac{k!^2}{(kt)!}\prod_{j=1}^{r-1}\frac{(k_jt)!}{k_j!^2} = \sum_{\substack{k_1,\ldots,k_{r-1}\ge 0\\ \sum_{j=1}^{r-1}k_j=k}} \frac{\binom{k}{k_1,\ldots,k_{r-1}}^2}{\binom{kt}{k_1t,\ldots,k_{r-1}t}} \le \sum_{\substack{k_1,\ldots,k_{r-1}\ge 0\\ \sum_{j=1}^{r-1}k_j=k}} 1 = \binom{k+r-2}{k},$$

which is better than $(r-1)^k$. *By bounding*

$$\binom{k+r-2}{k} \le e^k\left(1+\frac{r-2}{k}\right)^k,$$

the new function $g(k) := \alpha^k\sqrt{\binom{k+r-2}{k}}\binom{n/t}{k}/\sqrt{\binom{n}{kt}}$ *is still non-increasing in the interval* $4rm \le k \le \alpha n/t \le n/(2t)$ *if now*

$$\alpha \le e^{-1/2}\min_{4rm\le k\le \alpha n/t}\sqrt{\frac{k}{k+r-2}} = e^{-1/2}\sqrt{\frac{4rm}{4rm+r-2}}.$$

For $m \gg 1$, α *is essentially independent of* r, *and hence* $\alpha \le \min(1/2, e^{-1/2}) = 1/2$.

4.3 Quantum streaming lower bound for Unique Games on hypergraphs

The Unique Games problem is a generalization of the classical Max-Cut and can in fact be viewed as constraint satisfaction problems on a graph but over a larger alphabet. Consider a graph on n vertices $x_1,\ldots,x_n$ and edges in E. The constraint on an arbitrary edge $(i,j)\in E$ is specified by a permutation $\pi_{i,j}:\mathbb{Z}_r\to\mathbb{Z}_r$ and the goal is to find an assignment of $x_1,\ldots,x_n\in\mathbb{Z}_r$ that maximizes

$$\sum_{(i,j)\in E}[\pi_{i,j}(x_i)=x_j].$$

In this section, we consider a generalization of Unique Games to hypergraphs.

Definition 16 (Unique Games instance on hypergraphs)**.** *A hypergraph* $H=(V,E)$ *is defined on a vertex set* V *of size* n *with* t*-sized hyperedges* E *(i.e.,* t*-sized subsets of* V*). Given a linear constraint on a hyperedge* $e\in E$, *i.e., a linear function* $\pi_e:\mathbb{Z}_r^t\to\{0,1\}$, *the goal is to compute*

$$\max_{x\in\mathbb{Z}_r^n}\sum_{e\in E}\pi_e(x_e),$$

where x_e *corresponds to the set of vertex-assignment in the hyperedge* $e\in E$.

Definition 17. *Let* $H=(V,E)$ *be a hypergraph and let* OPT *be the optimal value of the Unique Games on* H. *A randomized algorithm gives a* γ*-approximation to a Unique Games instance with failure probability* $\delta\in[0,1/2)$ *if, on any input hypergraph* H, *it outputs a value in the interval* $[\mathrm{OPT}/\gamma,\mathrm{OPT}]$ *with probability at least* $1-\delta$.

A uniformly random assignment of $x\in\mathbb{Z}_r^n$ to the vertex set V will satisfy a $1/r$-fraction of the hyperedges, since each linear constraint $\pi_e(x_e)$ is satisfied with probability $1/r$. This gives a trivial r-approximation algorithm for the problem above. Below we show that any better than trivial approximation requires space that scales as n^β for constant $\beta>0$.

Theorem 18. *Let $r, t \geq 2$ be integers. Every quantum streaming algorithm giving a $(r-\varepsilon)$-approximation for Unique Games on hypergraphs (as in Definition 16) with at most t-sized hyperedges with alphabet size r and success probability at least $2/3$ over its internal randomness, needs $\Omega((n/t)^{1-2/t})$ space (which hides dependence on r, ε).*

The proof of this theorem combines techniques used by Guruswami and Tao [GT19] and Kapralov, Khanna and Sudan [KKS14]. Akin to these works, based on the Hidden Matching problem, we will construct instances of the hypergraph for which a Unique Games instance is hard to solve space-efficiently in the streaming model.

Input distributions. To this end, we construct two distributions $\mathcal{Y}$ and $\mathcal{N}$ such that $\mathcal{Y}$ is supported on satisfiable Unique Games instances and $\mathcal{N}$ is supported on instances for which at most an $O(1/r)$-fraction of the constraints is satisfied. We now define these instances in a multi-stage way (using k stages). First, sample k independent α-partial t-hypermatchings on n vertices and then construct a hypergraph G by putting together all the hyperedges from these k stages. Note that G still has n vertices, while the number of hyperedges is $k \cdot \alpha n/t$ (since each stage has $\alpha n/t$ many hyperedges and we allow multiple hyperedges should they be sampled). Now we specify the constraints π_e in Definition 16 for the $\mathcal{Y}, \mathcal{N}$ distributions:

- $\mathcal{Y}$ distribution: sample $z \in \mathbb{Z}_r^n$ and for each $e \in E$, let $\pi_e(x_e) = \left[\sum_{i \in e} x_i = \sum_{i \in e} z_i\right]$ (where by $i \in e$ we mean all the vertices in the hyperedge e).
- $\mathcal{N}$ distribution: for each $e \in E$, pick a uniform $q \in \mathbb{Z}_r$ and let $\pi_e(x_e) = \left[\sum_{i \in e} x_i = q\right]$.

It is clear that, in the $\mathcal{Y}$ distribution, the optimal solution is when all the $x_1, \ldots, x_n$ are just set to $z_1, \ldots, z_n$. Below we show that for the $\mathcal{N}$ distribution, the value of the optimal solution is at most $(1+\varepsilon)/r$ with high probability.

Lemma 19. *Let $\varepsilon \in (0,1)$. If $k = O(r(\log r)t/(\alpha\varepsilon^2))$, then for the Unique Games instance sampled from $\mathcal{N}$ distribution above, the optimal fraction of satisfiable constraints (i.e., number of hyperedges $e \in E$ for which $\pi_e(\cdot)$ evaluates to 1) over all possible vertex labelling is at most $(1+\varepsilon)/r$ with high probability.*

Proof. The proof of this lemma is similar to the proof in [GT19, Lemma 4.1]. Fix an assignment $x \in \mathbb{Z}_r^n$. Let X_e^ℓ be the random variable that indicates that the hyperedge $e \in E$ appears in ℓ-th stage and is satisfied by x. Let $S = \sum_{\ell,e} X_e^\ell$. The expectation of S is $k\alpha n/t \cdot 1/r$, since the total number of hyperedges is $\alpha n/t$ for each of the k stages and the probability that a uniform x satisfies a t-hyperedge (i.e., probability that $\sum_{e \in E} x_e = q$ for some fixed q) is $1/r$. Using the same analysis in [GT19], we can show that the variables X_e^ℓ are negatively correlated. Indeed, first note that hyperedges from different stages are independent. Now suppose we know that the random variables $X_{e_1}^\ell, \ldots, X_{e_s}^\ell$ have value 1, and we also know a hyperedge $e \in E$. If $e \cap e_u \neq \emptyset$ for some $u \in [s]$, then $X_e^\ell = 0$, since the hyperedges of a given stage form a matching. Otherwise, the conditional expectation of X_e^ℓ (conditioned on $e \cap e_u = \emptyset$ for all $u \in [s]$) is $\frac{\alpha n/t - s}{r}\binom{n-ts}{t}^{-1}$, which is less than its unconditional expectation of $\frac{\alpha n/t}{r}\binom{n}{t}^{-1}$. Therefore, in all cases one has $\mathbb{E}[X_e^\ell | X_{e_1}^\ell = \cdots = X_{e_s}^\ell = 1] \leq \mathbb{E}[X_e^\ell]$, which means negative correlation.

Hence, using a Chernoff bound for negative-correlated variables leads to

$$\Pr[S \geq (1+\varepsilon)(k\alpha n/t)/r] \leq \exp(-\varepsilon^2 k\alpha n/(3rt)) = \exp(-O(n \log r)),$$

where the inequality used the choice of k. Applying a union bound over the set of $x \in \mathbb{Z}_r^n$ concludes the proof of the lemma. □

Reduction to Hypermatching. The reduction to r-ary Hidden Hypermatching is similar to the analysis used by Guruswami and Tao [GT19], but now it is from quantum streaming algorithms to one-way quantum communication complexity. The main lemma that we need is the following.

Lemma 20. *Let $\varepsilon > 0$. If there is a streaming algorithm using at most c qubits of space that distinguishes between the $\mathcal{Y}$ and $\mathcal{N}$ distributions on Unique Games instances (with k stages) with bias $1/3$, then there is a c-qubit protocol that distinguish between the YES and NO distributions of r-$\mathsf{HH}(\alpha, t, n)$ with bias $\Omega(1/k)$.*

In order to prove this lemma we need a few definitions and facts. First, towards proving the lemma above, let us assume there is a c-qubit streaming $\mathcal{A}$ for Lemma 20. During the execution of the streaming protocol on instances from the $\mathcal{Y}$ and $\mathcal{N}$ distributions, let the memory content after receiving the ith stage constraints be given by the c-qubit quantum states $|\phi_i^{\mathcal{Y}}\rangle$ and $|\phi_i^{\mathcal{N}}\rangle$, respectively.[5] Assume that $|\phi_0^{\mathcal{Y}}\rangle = |\phi_0^{\mathcal{N}}\rangle = 0$. Using the notion of informative index from [KKS14, Definition 6.2], we say an index $j \in \{0, \ldots, k-1\}$ is δ-informative if

$$\left\| |\phi_{j+1}^{\mathcal{Y}}\rangle - |\phi_{j+1}^{\mathcal{N}}\rangle \right\|_1 \geq \left\| |\phi_j^{\mathcal{Y}}\rangle - |\phi_j^{\mathcal{N}}\rangle \right\|_1 + \delta.$$

With this definition it is not hard to see the following fact, which follows from a simple triangle inequality.

Fact 21. *Suppose there exists a streaming protocol for distinguishing the $\mathcal{Y}, \mathcal{N}$ distributions with advantage $\geq 1/3$, then there exists a $\Omega(1/k)$-informative index.*

Suppose j^* is an $\Omega(1/k)$-informative index for the streaming protocol $\mathcal{A}$. Using this we devise a communication protocol for r-$\mathsf{HH}(\alpha, t, n)$ with bias $\Omega(1/k)$ as follows: suppose Alice has a string $x \in \mathbb{Z}_r^n$ and Bob has $w \in \mathbb{Z}_r^{\alpha n/t}$ and a hypermatching $M \in \mathcal{M}_{t,n}^{\alpha}$.

1. Alice samples j^* many α-partial t-hypermatchings and runs the streaming algorithm $\mathcal{A}$ on Unique Games constraints for the first j^* stages that follow the $\mathcal{Y}$ distribution with $z = x$. She then sends the memory contents after these j^* stages to Bob.

2. Bob assigns the constraints $\sum_{i \in e} x_i = w_e$, where $e \in M$, according to his inputs w, M. He then continues running $\mathcal{A}$ on these constraints as the (j^*+1)th stage.

   Let $|s\rangle$ be the quantum state that Bob gets after running $\mathcal{A}$.

3. Let $|\phi^{\mathsf{YES}}\rangle$ and $|\phi^{\mathsf{NO}}\rangle$ be the resulting quantum states under the two cases, depending on w's distribution (these can be computed by Bob since $\mathcal{A}$ is known). Bob can distinguish between $|\phi^{\mathsf{YES}}\rangle$ and $|\phi^{\mathsf{NO}}\rangle$ with bias $\frac{1}{2}\left\| |\phi^{\mathsf{YES}}\rangle - |\phi^{\mathsf{NO}}\rangle \right\|_1$ by measuring the state $|s\rangle$ with a suitable POVM, according to Lemma 2.

We are now ready to prove Lemma 20.

[5] Without loss of generality, we assume they are pure states– this only affects the cost of the protocol by a constant factor (since one can always purify mixed quantum states by doubling the dimension).

Proof of Lemma 20. We argue that the above protocol achieves a $\Omega(1/k)$ bias in distinguishing between the YES and NO distributions from r-HH(α, t, n). To this end, let U be the unitary that maps the quantum state after stage j^* and constraints of stage j^*+1 (which is classical) to the quantum state after j^*+1. Thus we have $|\phi^{\mathsf{YES}}\rangle = |\phi^{\mathcal{Y}}_{j^*+1}\rangle = U|\phi^{\mathcal{Y}}_{j^*}, C^{\mathcal{Y}}\rangle$ and $|\phi^{\mathsf{NO}}\rangle = U|\phi^{\mathcal{Y}}_{j^*}, C^{\mathcal{N}}\rangle$, where $C^{\mathcal{Y}}$ and $C^{\mathcal{N}}$ are the constraints corresponding to the YES and NO distributions, respectively, and, similarly, we have $|\phi^{\mathcal{N}}_{j^*+1}\rangle = U|\phi^{\mathcal{N}}_{j^*}, C^{\mathcal{N}}\rangle$. Then, we have

$$\begin{aligned}\left\||\phi^{\mathsf{YES}}\rangle - |\phi^{\mathsf{NO}}\rangle\right\|_1 &\geq \left\||\phi^{\mathcal{Y}}_{j^*+1}\rangle - |\phi^{\mathcal{N}}_{j^*+1}\rangle\right\|_1 - \left\||\phi^{\mathsf{NO}}\rangle - |\phi^{\mathcal{N}}_{j^*+1}\rangle\right\|_1 \\ &\geq \left\||\phi^{\mathcal{Y}}_{j^*+1}\rangle - |\phi^{\mathcal{N}}_{j^*+1}\rangle\right\|_1 - \left\||\phi^{\mathcal{Y}}_{j^*}\rangle - |\phi^{\mathcal{N}}_{j^*}\rangle\right\|_1 = \Omega(1/k),\end{aligned}$$

where the second inequality used that $\left\||\phi^{\mathsf{NO}}\rangle - |\phi^{\mathcal{N}}_{j^*+1}\rangle\right\|_1 = \left\|U|\phi^{\mathcal{Y}}_{j^*}, C^{\mathcal{N}}\rangle - U|\phi^{\mathcal{N}}_{j^*}, C^{\mathcal{N}}\rangle\right\|_1 \leq \||\phi^{\mathcal{Y}}_{j^*}\rangle - |\phi^{\mathcal{N}}_{j^*}\rangle\|_1$ (since unitaries preserve norms) and the third inequality is because j^* is an informative index. Hence in Step (3) of the procedure above, the bias of Bob in obtaining the right outcome is $\Omega(1/k)$. □

Proof of Theorem 18. Finally, by picking $k = O(r(\log r)t/(\alpha\varepsilon^2))$ in order to invoke Lemma 19 and using our lower bound in Theorem 11 with $\alpha = O(1)$, we get our desired lower bound of

$$\Omega(r^{-(1+1/t)}(k^2\alpha)^{-2/t}(n/t)^{1-2/t}) = \Omega((n/t)^{1-2/t}).$$ □

It is possible to prove a classical version of Theorem 18.

Theorem 22. *Let $r, t \geq 2$ be integers. Every classical streaming algorithm giving an $(r-\varepsilon)$-approximation for Unique Games on hypergraphs (as in Definition 16) with at most t-sized hyperedges with alphabet size r and success probability at least $2/3$ over its internal randomness, needs $\Omega((n/t)^{1-1/t})$ space (which hides dependence on r, ε).*

Proof. Since the proof is very similar to the one of Theorem 18, we shall just point out the few required modifications. The main idea is still to reduce a streaming algorithm for Unique Games to a communication protocol for r-HH(α, t, n). The distributions $\mathcal{Y}$ and $\mathcal{N}$ on the Unique Games inputs are the same. Let $S^{\mathcal{Y}}_i$ and $S^{\mathcal{N}}_i$ be the memory after receiving the ith stage constraints. The notion of information index is similarly defined for $S^{\mathcal{Y}}_i$ and $S^{\mathcal{N}}_i$, i.e., an index $j \in \{0, \ldots, k-1\}$ is δ-informative if

$$\left\|S^{\mathcal{Y}}_{j+1} - S^{\mathcal{N}}_{j+1}\right\|_{\mathrm{tvd}} \geq \left\|S^{\mathcal{Y}}_j - S^{\mathcal{Y}}_j\right\|_{\mathrm{tvd}} + \delta.$$

The communication protocol for r-HH(α, t, n) is basically the same as in the quantum case, using an $\Omega(1/k)$-informative index j^*. At the end of Step (2), Bob gets the memory s. Let S^{YES} and S^{NO} be the resulting memory distributions under the two cases depending on w's distribution. Bob outputs 1 if $\Pr[S^{\mathsf{YES}} = s] \geq \Pr[S^{\mathsf{NO}} = s]$, and 0 otherwise. The bias of distinguishing between S^{YES} and S^{NO} is $\frac{1}{2}\|S^{\mathsf{YES}} - S^{\mathsf{NO}}\|_{\mathrm{tvd}}$, which can be shown to be at least $\Omega(1/k)$, similarly to the quantum case. Indeed, let f be the function that maps the memory after stage j^* and constraints $C^{\mathcal{Y}}$ or $C^{\mathcal{N}}$ of stage (j^*+1) to the memory after stage (j^*+1). Then $S^{\mathsf{YES}} = S^{\mathcal{Y}}_{j^*+1} = f(S^{\mathcal{Y}}_{j^*}, C^{\mathcal{Y}})$ and $S^{\mathsf{NO}} = f(S^{\mathcal{Y}}_{j^*}, C^{\mathcal{N}})$. By using Lemma 23 below, we can show that

$$\begin{aligned}\left\|S^{\mathsf{YES}} - S^{\mathsf{NO}}\right\|_{\mathrm{tvd}} &\geq \left\|S^{\mathcal{Y}}_{j^*+1} - S^{\mathcal{N}}_{j^*+1}\right\|_{\mathrm{tvd}} - \left\|S^{\mathsf{NO}} - S^{\mathcal{N}}_{j^*+1}\right\|_{\mathrm{tvd}} \\ &= \left\|S^{\mathcal{Y}}_{j^*+1} - S^{\mathcal{N}}_{j^*+1}\right\|_{\mathrm{tvd}} - \left\|f(S^{\mathcal{Y}}_{j^*}, C^{\mathcal{Y}}) - f(S^{\mathcal{N}}_{j^*}, C^{\mathcal{N}})\right\|_{\mathrm{tvd}} \\ &\geq \left\|S^{\mathcal{Y}}_{j^*+1} - S^{\mathcal{N}}_{j^*+1}\right\|_{\mathrm{tvd}} - \left\|S^{\mathcal{Y}}_{j^*} - S^{\mathcal{N}}_{j^*}\right\|_{\mathrm{tvd}} \geq \Omega(1/k).\end{aligned}$$

Lemma 23 ([KKS14, Claim 6.5]). *Let X, Y be two random variables and W be independent of (X, Y). Then, for any function f,*

$$\|f(X,W) - f(Y,W)\|_{\text{tvd}} \leq \|X - Y\|_{\text{tvd}}.$$

Therefore Bob can distinguish between S^{YES} and S^{NO} with bias at least $\Omega(1/k)$, meaning that there is a c-bit protocol that distinguish between the YES and NO distributions of r-$\mathsf{HH}(\alpha, t, n)$ with bias $\Omega(1/k)$. By picking $k = O(r(\log r)t/(\alpha\varepsilon^2))$ in order to invoke Lemma 19 and using the classical lower bound on r-$\mathsf{HH}(\alpha, t, n)$, we get our desired lower bound of

$$\Omega(r^{-1}(k^4\alpha)^{-1/t}(n/t)^{1-1/t}) = \Omega((n/t)^{1-1/t}). \qquad \square$$

5 Locally Decodable Codes

In this section we prove our lower bound on locally decodable codes over $\mathbb{Z}_r$. Before that, let us first formally define an LDC.

Definition 24 (Locally decodable code). *A (q, δ, ε)-locally decodable code over $\mathbb{Z}_r$ is a function $C : \mathbb{Z}_r^n \to \mathbb{Z}_r^N$ that satisfies the following: for every $x \in \mathbb{Z}_r^n$ and $i \in [n]$, there exists a (randomized) algorithm $\mathcal{A}$ that, on any input $y \in \mathbb{Z}_r^N$ that satisfies $d(y, C(x)) \leq \delta N$, makes q queries to y non-adaptively and outputs a number $\mathcal{A}^y(i) \in \mathbb{Z}_r$ that satisfies $\Pr[\mathcal{A}^y(i) = x_i] \geq 1/r + \varepsilon$ (where the probability is only taken over the randomness of $\mathcal{A}$).*

As is often the case when proving LDC lower bounds, we use the useful fact proven by Katz and Trevisan [KT00] that, without loss of generality, one can assume that an LDC is smooth, i.e., the queries made by $\mathcal{A}$ have "reasonable" probability over all indices, and that $\mathcal{A}$ makes queries to a codeword (and not a corrupted codeword). We first formally define a smooth code below.

Definition 25 (Smooth code). *We say $C : \mathbb{Z}_r^n \to \mathbb{Z}_r^N$ is a (q, c, ε)-smooth code if there exists a decoding algorithm $\mathcal{A}$ that satisfies the following: for every $x \in \mathbb{Z}_r^n$ and $i \in [n]$, $\mathcal{A}$ makes at most q non-adaptive queries to $C(x)$ and outputs $\mathcal{A}^{C(x)}(i) \in \mathbb{Z}_r$ such that $\Pr[\mathcal{A}^{C(x)}(i) = x_i] \geq 1/r + \varepsilon$ (where the probability is only taken over the randomness of $\mathcal{A}$). Moreover, for every $x \in \mathbb{Z}_r^n$, $i \in [n]$ and $j \in [N]$, on input i, the probability that $\mathcal{A}$ queries the index j in $C(x) \in \mathbb{Z}_r^N$ is at most c/N.*

Crucially note that smooth codes only require a decoder to recover x_i when given access to an actual codeword, unlike the standard definition of LDC where a decoder is given a noisy codeword. With this definition in hand, we state a theorem of Katz and Trevisan.

Theorem 26 ([KT00]). *A (q, δ, ε)-LDC $C : \mathbb{Z}_r^n \to \mathbb{Z}_r^N$ is a $(q, q/\delta, \varepsilon)$-smooth code.*

We remark that a converse to this theorem holds: a (q, c, ε)-smooth code is a $(q, \delta, \varepsilon - c\delta)$-$\mathsf{LDC}$, since the probability that the decoder queries one of δN corrupted positions is at most $(c/N)(\delta N) = c\delta$.

5.1 Smooth codes over large alphabets

Katz and Trevisan [KT00] observed that a (q, c, ε)-smooth code over $\{0, 1\}$ is a $(q, q, \varepsilon^2/2c)$-smooth code that is good on *average*, i.e., that there is a decoder $\mathcal{A}$ such that, for all $i \in [n]$,

$$\frac{1}{2^n} \sum_{x \in \{0,1\}^n} \Pr[\mathcal{A}^{C(x)}(i) = x_i] \geq \frac{1}{2} + \frac{\varepsilon^2}{2c}.$$

This comes from the observation that a q-decoder can partition the set $[N]$ into q-tuples, pick one of such tuples uniformly at random and continue as the original decoder by querying the elements of the picked tuple at the cost of a slightly worse success probability. These ideas are formally explained in the result below, where we already generalize them to large alphabets $\mathbb{Z}_r$ (the overall presentation is inspired in [BRdeW08, Theorem 15]).

Theorem 27. *Suppose $C : \mathbb{Z}_r^n \to \mathbb{Z}_r^N$ is a (q, c, ε)-smooth code. Then for every $i \in [n]$, there exists a set M_i consisting of at least $\varepsilon N/(2cq)$ disjoint sets of at most q elements of $[N]$ each such that, for every $Q \in M_i$, there exists a function $f_Q : \mathbb{Z}_r^{|Q|} \to \mathbb{Z}_r$ with the property*

$$\sum_{k=1}^{r-1} \mathop{\mathbb{E}}_{x\sim\mathbb{Z}_r^n} \left[\omega_r^{k(f_Q(C(x)_Q)-x_i)}\right] \geq \frac{r}{2}\varepsilon.$$

Here $C(x)_Q$ is the restriction of $C(x)$ to the bits in Q.

Proof. Fix some $i \in [n]$. In order to decode x_i, we can assume, without lost of generality, that the decoder $\mathcal{A}$ picks some set $Q \subseteq [N]$ (of at most q indices) with probability $p(Q)$, queries those bits, and then outputs a random variable (not yet a function) $f_Q(C(x)_Q) \in \mathbb{Z}_r$ that depends on the query-outputs. Call such a Q "good" if

$$\frac{1}{r} + \frac{\varepsilon}{2} \leq \Pr_{x\sim\mathbb{Z}_r^n}[f_Q(C(x)_Q) = x_i] = \mathop{\mathbb{E}}_{\substack{x\sim\mathbb{Z}_r^n \\ k\sim\mathbb{Z}_r}} \left[\omega_r^{k(f_Q(C(x)_Q)-x_i)}\right] \iff \frac{r}{2}\varepsilon \leq \sum_{k=1}^{r-1} \mathop{\mathbb{E}}_{x\sim\mathbb{Z}_r^n} \left[\omega_r^{k(f_Q(C(x)_Q)-x_i)}\right].$$

Now construct the hypergraph $H_i = (V, E_i)$ with $V = [N]$ and edge-set E_i consisting of all good sets Q. The probability that the decoder queries any $Q \in E_i$ is $p(E_i) := \sum_{Q\in E_i} p(Q)$. If it queries some $Q \in E_i$, then

$$\Pr_{x\sim\mathbb{Z}_r^n}[f_Q(C(x)_Q) = x_i] \leq 1 \iff \sum_{k=1}^{r-1} \mathop{\mathbb{E}}_{x\sim\mathbb{Z}_r^n} \left[\omega_r^{k(f_Q(C(x)_Q)-x_i)}\right] \leq r-1,$$

and if it queries some $Q \notin E_i$, then $\sum_{k=1}^{r-1} \mathbb{E}_{x\sim\mathbb{Z}_r^n} \left[\omega_r^{k(f_Q(C(x)_Q)-x_i)}\right] < \frac{r}{2}\varepsilon$. Given the smooth code property of outputting x_i with probability at least $\frac{1}{r} + \varepsilon$ for every x, we have

$$r\varepsilon \leq \sum_{k=1}^{r-1} \mathop{\mathbb{E}}_{x,Q} \left[\omega_r^{k(f_Q(C(x)_Q)-x_i)}\right] < p(E_i)(r-1) + (1-p(E_i))\frac{r}{2}\varepsilon = \frac{r}{2}\varepsilon + p(E_i)\left(r - 1 - \frac{r}{2}\varepsilon\right),$$

hence

$$p(E_i) > \frac{\varepsilon}{2 - 2/r - \varepsilon} \geq \frac{\varepsilon}{2}.$$

Since C is also smooth, for every $j \in [N]$ we have

$$\sum_{Q\in E_i : j\in Q} p(Q) \leq \sum_{Q : j\in Q} p(Q) = \Pr[\mathcal{A} \text{ queries } j] \leq \frac{c}{N}.$$

Let M_i be a matching in H_i of maximal size. We want to show that $|M_i| \geq \varepsilon N/(2cq)$. To do so, define $T := \bigcup_{Q\in M_i} Q$. Observe that the set T has at most $q|M_i|$ elements, and intersects each $Q \in E_i$ (otherwise M_i would not be maximal). The size of M_i can be lower bounded as follows:

$$\frac{\varepsilon}{2} < p(E_i) = \sum_{Q : Q\in E_i} p(Q) \overset{(a)}{\leq} \sum_{j\in T} \sum_{Q\in E_i : j\in Q} p(Q) \leq \frac{c|T|}{N} \leq \frac{cq|M_i|}{N},$$

where (a) holds because each $Q \in E_i$ is counted exactly once on the left and at least once on the right (since T intersects each $Q \in E_i$). Hence $|M_i| \geq \varepsilon N/(2cq)$. Finally, the random variables $f_Q(C(x)_Q)$ can be fixed in $\mathbb{Z}_r$ without reducing the probability $\Pr_{x\sim\mathbb{Z}_r^n}[f_Q(C(x)_Q) = x_i]$. □

As previously mentioned, the above result tells us that a decoder can focus its queries on one of the q-tuples Q. We can go one step further and show that the decoder can restrict itself to computing a linear function of the queried bits while still maintaining a good correlation with the target bit x_i at the cost of decreasing the average success probability.

Theorem 28. *Suppose $C : \mathbb{Z}_r^n \to \mathbb{Z}_r^N$ is a (q, c, ε)-smooth code. Then for every $i \in [n]$, there exists a set M_i consisting of at least $\varepsilon N/(2cq)$ disjoint sets of at most q elements of $[N]$ each such that, for every $Q \in M_i$,*

$$\sum_{k=1}^{r-1} \sum_{S\in\mathbb{Z}_r^{|Q|}} \left| \mathop{\mathbb{E}}_{x\sim\mathbb{Z}_r^n} \left[\omega_r^{S\cdot C(x)_Q - kx_i}\right]\right| \geq \frac{\varepsilon r}{2}.$$

Here $C(x)_Q$ is the restriction of $C(x)$ to the bits in Q.

Proof. Fix $i \in [n]$ and take the set M_i produced by Theorem 27. For every $Q \in M_i$ we have

$$\sum_{k=1}^{r-1} \mathop{\mathbb{E}}_{x\sim\mathbb{Z}_r^n} \left[\omega_r^{k(f_Q(C(x)_Q)-x_i)}\right] \geq \frac{\varepsilon r}{2}.$$

For $k \in \{1, \ldots, r-1\}$, define the function $h_{Q,k} : \mathbb{Z}_r^{|Q|} \to \mathbb{C}$ by $h_{Q,k}(x) = \omega_r^{kf_Q(x)}$. Consider its Fourier transform $\widehat{h}_{Q,k} : \mathbb{Z}_r^{|Q|} \to \mathbb{C}$. Hence we can write

$$h_{Q,k}(x) = \sum_{S\in\mathbb{Z}_r^{|Q|}} \widehat{h}_{Q,k}(S)\omega_r^{S\cdot x}.$$

Finally, using that $|\widehat{h}_{Q,k}(S)| \in [0, 1]$ for all $S \in \mathbb{Z}_r^{|Q|}$, we can upper bound $\varepsilon r/2$ by

$$\sum_{k=1}^{r-1} \mathop{\mathbb{E}}_{x} \left[\omega_r^{k(f_Q(C(x)_Q)-x_i)}\right] = \sum_{S\in\mathbb{Z}_r^{|Q|}} \sum_{k=1}^{r-1} \widehat{h}_{Q,k}(S) \mathop{\mathbb{E}}_{x} \left[\omega_r^{S\cdot C(x)_Q - kx_i}\right] \leq \sum_{S\in\mathbb{Z}_r^{|Q|}} \sum_{k=1}^{r-1} \left|\mathop{\mathbb{E}}_{x} \left[\omega_r^{S\cdot C(x)_Q - kx_i}\right]\right|. \quad \square$$

5.2 An exponential lower bound for LDCs

In this section, we use our results from matrix-valued hypercontractivity to obtain our lower bound for LDCs over $\mathbb{Z}_r$.

Theorem 29. *If $C : \mathbb{Z}_r^n \to \mathbb{Z}_r^N$ is a $(2, \delta, \varepsilon)$-LDC, then $N = 2^{\Omega(\delta^2\varepsilon^4 n/r^4)}$.*

Proof. In this proof we shall use the normalized Schatten norm. Fix $x \in \mathbb{Z}_r^n$. Define the vector $v_x \in \mathbb{C}^{r^2N}$

$$v_x = \left(1, \ldots, 1, \omega_r^{C(x)_1}, \ldots, \omega_r^{C(x)_N}, \omega_r^{2C(x)_1}, \ldots, \omega_r^{2C(x)_N}, \ldots, \omega_r^{(r-1)C(x)_1}, \ldots, \omega_r^{(r-1)C(x)_N}\right),$$

where each sequence $\omega_r^{jC(x)_1}, \ldots, \omega_r^{jC(x)_N}$ is repeated r times consecutively. Let $R := r^2N$ and define the $R \times R$ symmetric matrix $f(x) := v_x^{\mathrm{T}} \cdot v_x$ whose $(N(rj_1 + m_1) + \ell_1, N(rj_2 + m_2) + \ell_2)$-entry

is $\omega_r^{j_1 C(x)_{\ell_1} + j_2 C(x)_{\ell_2}}$, where $j_1, j_2, m_1, m_2 \in \mathbb{Z}_r$ and $\ell_1, \ell_2 \in [N]$ (note that there are r repeated entries $\omega_r^{j_1 C(x)_{\ell_1} + j_2 C(x)_{\ell_2}}$ in each row and column). Since $f(x)$ has rank 1 and its R^2 entries have absolute value 1, its only non-zero singular value is R. Hence $\|f(x)\|_p^p = R^{p-1}$ for every $x \in \mathbb{Z}_r^n$.

Fix $i \in [n]$. For every $k \in \{1, \ldots, r-1\}$ consider the $R \times R$ matrices $\widehat{f}(0^{i-1}k0^{n-i})$ that are the Fourier transform of f at the strings in $\mathbb{Z}_r^n$ which are zero in all but the ith coordinate:

$$\widehat{f}(0^{i-1}k0^{n-i}) = \frac{1}{r^n} \sum_{x \in \mathbb{Z}_r^n} f(x) \omega_r^{-kx_i}.$$

We shall lower bound $\sum_{k=1}^{r-1} \left\| \widehat{f}(0^{i-1}k0^{n-i}) \right\|_p^p$.

By Theorem 28, there is a set M_i consisting of at least $\delta\varepsilon N/8$ disjoint sets of indices in $[N]$, each with cardinality at most 2,[6] such that $\sum_{k=1}^{r-1} \sum_{S \in \mathbb{Z}_r^{|Q|}} \left| \mathbb{E}_{x \sim \mathbb{Z}_r^n} \left[\omega_r^{S \cdot C(x)_Q - kx_i} \right] \right| \in [\varepsilon r/2, r^{|Q|}(r-1)]$. Given $S = (S_1, S_2)$, consider $Q = (Q_1, Q_2) \in M_i$[7] and the following 2×2 submatrix in $f(x)$

$$\begin{pmatrix} \omega_r^{2S_1 C(x)_{Q_1}} & \omega_r^{S_1 C(x)_{Q_1} + S_2 C(x)_{Q_2}} \\ \omega_r^{S_1 C(x)_{Q_1} + S_2 C(x)_{Q_2}} & \omega_r^{2S_2 C(x)_{Q_2}} \end{pmatrix}.$$

Observe that this submatrix clearly exists in $f(x)$, and comes from the rows and columns $N(rS_1 + m_1) + Q_1$ and $N(rS_2 + m_2) + Q_2$ for any $m_1, m_2 \in \mathbb{Z}_r$. In particular, we can take $m_1 = S_2$ and $m_2 = S_1$, so that such submatrix does not have overlapping rows or columns with any other submatrix similarly defined from different S' or Q'. Hence the corresponding 2×2 submatrix of $\widehat{f}(0^{i-1}k0^{n-i})$ is

$$\begin{pmatrix} \alpha & \mathbb{E}_{x \sim \mathbb{Z}_r^n} \left[\omega_r^{S \cdot C(x)_Q - kx_i} \right] \\ \mathbb{E}_{x \sim \mathbb{Z}_r^n} \left[\omega_r^{S \cdot C(x)_Q - kx_i} \right] & \beta \end{pmatrix},$$

for some $\alpha, \beta \in \mathbb{C}$ (in this proof we will not be concerned with the value of α, β). Let P be the $R \times R$ permutation matrix that, for every $Q = (Q_1, Q_2)$ and $S = (S_1, S_2)$, swaps rows $N(rS_1 + S_2) + Q_1$ and $N(rS_2 + S_1) + Q_2$. We define the matrices $F_i(k) := P\widehat{f}(0^{i-1}k0^{n-i})$ for $k \in \{1, \ldots, r-1\}$. Because we previously chose $m_1 = S_2$ and $m_2 = S_1$, for each of the at least $\delta\varepsilon N/8$ sets $Q \in M_i$, $F_i(k)$ has diagonal entries $\mathbb{E}_{x \sim \mathbb{Z}_r^n} \left[\omega_r^{S \cdot C(x)_Q - kx_i} \right]$ for all $S \in \mathbb{Z}_r^{|Q|}$ (each entry is repeated twice). In other words, $F_i(k)$ has at least $\delta\varepsilon N r^2/4$ entries $\mathbb{E}_{x \sim \mathbb{Z}_r^n} \left[\omega_r^{S \cdot C(x)_Q - kx_i} \right]$ for $Q \in M_i$ and $S \in \mathbb{Z}_r^{|Q|}$.

The Schatten norm $\|\cdot\|_p$ is *unitarily invariant*: $\|UAV\|_p = \|A\|_p$ for every matrix A and unitaries U, V. We shall use the following lemma. Its proof is left to the end of the section.

Lemma 30 ([Bha13, Eq. (IV.52)])**.** *Let $\|\cdot\|$ be a unitarily-invariant norm on $\mathbb{C}^{d \times d}$. If $A \in \mathbb{C}^{d \times d}$ and $\mathrm{diag}(A)$ is the matrix obtained from A by setting its off-diagonal entries to 0, then $\|\mathrm{diag}(A)\| \leq \|A\|$.*

Using this lemma, we obtain

$$\sum_{k=1}^{r-1} \left\| \widehat{f}(0^{i-1}k0^{n-i}) \right\|_p^p = \sum_{k=1}^{r-1} \|F_i(k)\|_p^p \geq \sum_{k=1}^{r-1} \|\mathrm{diag}(F_i(k))\|_p^p \geq \frac{2}{R} \sum_{k=1}^{r-1} \sum_{Q \in M_i} \sum_{S \in \mathbb{Z}_r^{|Q|}} \left| \mathop{\mathbb{E}}_{x \sim \mathbb{Z}_r^n} \left[\omega_r^{S_Q \cdot C(x)_Q - kx_i} \right] \right|^p,$$

[6] Here we used Theorem 26 in order to invoke Theorem 28 with $c = q/\delta$.

[7] If Q is a singleton, take $Q = (Q_1, Q_1)$ and $S = (S_1, 0)$.

but, by Hölder's inequality,

$$\sum_{k=1}^{r-1}\sum_{S\in\mathbb{Z}_r^{|Q|}}\left|\mathop{\mathbb{E}}_{x\sim\mathbb{Z}_r^n}\left[\omega_r^{S_Q\cdot C(x)_Q-kx_i}\right]\right|^p \geq \frac{1}{r^{3(p-1)}}\left(\sum_{k=1}^{r-1}\sum_{S\in\mathbb{Z}_r^{|Q|}}\left|\mathop{\mathbb{E}}_{x\sim\mathbb{Z}_r^n}\left[\omega_r^{S_Q\cdot C(x)_Q-kx_i}\right]\right|\right)^p \geq \frac{1}{r^{3(p-1)}}\left(\frac{\varepsilon r}{2}\right)^p,$$

hence

$$\sum_{k=1}^{r-1}\left\|\widehat{f}(0^{i-1}k0^{n-i})\right\|_p^p \geq \frac{2}{R}\frac{\delta\varepsilon N}{8}\frac{1}{r^{3(p-1)}}\left(\frac{\varepsilon r}{2}\right)^p = \frac{\delta\varepsilon}{4r^{2p-1}}\left(\frac{\varepsilon}{2}\right)^p,$$

which implies

$$\sum_{k=1}^{r-1}\left\|\widehat{f}(0^{i-1}k0^{n-i})\right\|_p^2 \geq \frac{1}{r^{2/p-1}}\left(\sum_{k=1}^{r-1}\left\|\widehat{f}(0^{i-1}k0^{n-i})\right\|_p^p\right)^{2/p} \geq \frac{1}{r^3}\left(\frac{\delta\varepsilon}{4}\right)^{2/p}\left(\frac{\varepsilon}{2}\right)^2,$$

where we used Hölder's inequality again. Now, using the hypercontractive inequality, we have for any $p\in[1,2]$ that

$$n(p-1)\frac{1}{r^4}\left(\frac{\delta\varepsilon}{4}\right)^{2/p}\left(\frac{\varepsilon}{2}\right)^2 \leq \sum_{i=1}^{n}\sum_{k=1}^{r-1}\frac{p-1}{r-1}\left\|\widehat{f}(0^{i-1}k0^{n-i})\right\|_p^2 \leq \left(\frac{1}{r^n}\sum_{x\in\mathbb{Z}_r^n}\|f(x)\|_p^p\right)^{2/p} = R^{2(p-1)/p}.$$

Choosing $p=1+1/\log R$ gives us

$$\frac{n}{\log R}\frac{1}{r^4}\left(\frac{\delta\varepsilon}{4}\right)^2\left(\frac{\varepsilon}{2}\right)^2 \leq R^{2/(1+\log R)} = 4^{\log R/(1+\log R)} \implies R \geq \frac{2^{\delta^2\varepsilon^4 n/(2^6 r^4)}}{2^{4^{\log R/(1+\log R)}}} = 2^{\Omega(\delta^2\varepsilon^4 n/r^4)}.$$

Since $R=r^2N$, we have the desired lower bound by adjusting the constant in the $\Omega(\cdot)$ in the exponent. □

Proof of Lemma 30. The proof sets the off-diagonal entries of A to 0 recursively without increasing its norm. Start with the off-diagonal entries in the dth row and column. Define D_d be the diagonal matrix by $D_{d,d}=-1$ and $D_{i,i}=1$ for $i<d$. Note that D_dAD_d is the same as A, except that the off-diagonal entries of the dth row and column are multiplied by -1. Hence $A_{d-1}:=(A+D_dAD_d)/2$ is the matrix obtained from A by setting those entries to 0 (this does not affect the diagonal). Since D_d is unitary, by the triangle inequality

$$\|A_{d-1}\| = \|(A+D_dAD_d)/2\| \leq \frac{1}{2}(\|A\|+\|D_dAD_d\|) = \|A\|.$$

Continuing in this manner for $i=1,\ldots,d-1$, we can set the off-diagonal entries in the $(d-i)$th row and column of A_{d-i} to 0 by using the diagonal matrix D_{d-i} which has a -1 only on its $(d-i)$th position and without increasing its norm. □

6 2-server private information retrieval

As mentioned in the introduction, the connection between LDCs and PIR is well known since the results of [KT00, GKST02]. In general, upper bounds on LDCs are derived via PIR schemes, which in turn means that our LDC lower bounds translate to PIR lower bounds, which we illustrate below. We first define the notion of private information retrieval.

Definition 31. *A one-round, $(1-\delta)$-secure, k-server private information retrieval* (PIR) *scheme with recovery probability $1/r+\varepsilon$, query size t and answer size a, consists of a randomized user and k deterministic algorithms $S_1,\ldots,S_k$ (the servers) that satisfy the following:*

1. *On input $i\in[n]$, the user produces k queries $q_1,\ldots,q_k\in\mathbb{Z}_r^t$ and sends them to the k servers respectively. The servers reply back with a string $a_j=S_j(x,q_j)\in\mathbb{Z}_r^a$, and based on $a_1,\ldots,a_k$ and i, the user outputs $b\in\mathbb{Z}_r$.*
2. *For every $x\in\mathbb{Z}_r^n$ and $i\in[n]$, the output b of the user satisfies $\Pr[b=x_i]\geq 1/r+\varepsilon$.*
3. *For every $x\in\mathbb{Z}_r^n$ and $j\in[k]$, the distributions over q_j (over the user's randomness) are δ-close for different $i\in[n]$.*

We crucially remark that for the lower bounds that we present below, the function S_j could be an arbitrary (not necessarily linear) function over $x_1,\ldots,x_n\in\mathbb{Z}_r$.

Our PIR lower bound follows almost immediately from the following immediate consequence of Goldreich et al. [GKST02, Lemma 5.1]. In the following we shall assume $\delta=0$.

Lemma 32 ([GKST02])**.** *If there is a classical 2-server* PIR *scheme with query size t, answer size a and recovery probability $1/r+\varepsilon$, then there is a $(2,3,\varepsilon)$-smooth code $C:\mathbb{Z}_r^n\to(\mathbb{Z}_r^a)^m$ with $m\leq 6r^t$.*

We remark that Goldreich et al. [GKST02] state the lemma above only for $r=2$, but the exact same analysis carries over to the large alphabet case. We now get the following main theorem.

Theorem 33. *A classical 2-server* PIR *scheme with query size t, answer size a and recovery probability $1/r+\varepsilon$ satisfies $t\geq\Omega\big((\delta^2\varepsilon^4 n/r^4-a)/\log r\big)$.*

Proof. By using Lemma 32, there is a $(2,3,\varepsilon)$-smooth code $C:\mathbb{Z}_r^n\to(\mathbb{Z}_r^a)^m$ with $m\leq 6r^t$. In order to apply Theorem 29, we form a new code C' by transforming each old string $C(x)_j\in\mathbb{F}_r^a$ using the Hadamard code into $C'(x)_j\in\{0,1\}^{2^a}\subseteq\mathbb{Z}_r^{2^a}$. The total length of C' is $m'=m2^a$. By using Theorem 29 on C' (note that the theorem can be applied directly to a smooth code), this gives us

$$m'a\geq 2^{\Omega(\delta^2\varepsilon^4 n/r^4)},$$

and since $m=O(r^t)$, we get the desired lower bound in the theorem statement. □

References

[Aar18] Scott Aaronson. PDQP/qpoly= All. *arXiv preprint arXiv:1805.08577*, 2018. 3

[AD21] Sepehr Assadi and Aditi Dudeja. A simple semi-streaming algorithm for global minimum cuts. In *Symposium on Simplicity in Algorithms (SOSA)*, pages 172–180. SIAM, 2021. 2, 6

[AGM12] Kook Jin Ahn, Sudipto Guha, and Andrew McGregor. Graph sketches: sparsification, spanners, and subgraphs. In *Proceedings of the 31st ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, pages 5–14, 2012. 2, 6

[AMS99] Noga Alon, Yossi Matias, and Mario Szegedy. The space complexity of approximating the frequency moments. *Journal of Computer and system sciences*, 58(1):137–147, 1999. 1, 6

[BCL94] Keith Ball, Eric A. Carlen, and Elliott H. Lieb. Sharp uniform convexity and smoothness inequalities for trace norms. *Inventiones mathematicae*, 115(1):463–482, 1994. 2, 4, 5, 10, 11, 12, 13

[BDG16] Jop Briët, Zeev Dvir, and Sivakanth Gopi. Outlaw distributions and locally decodable codes. *arXiv preprint arXiv:1609.06355*, 2016. 3

[BDSS11] Arnab Bhattacharyya, Zeev Dvir, Amir Shpilka, and Shubhangi Saraf. Tight lower bounds for 2-query LCCs over finite fields. In *Proceedings of the 52nd IEEE Annual Symposium on Foundations of Computer Science*, pages 638–647. IEEE, 2011. 10

[Bec75] William Beckner. Inequalities in Fourier analysis. *Annals of Mathematics*, pages 159–182, 1975. 3

[Bha13] Rajendra Bhatia. *Matrix analysis*, volume 169. Springer Science & Business Media, 2013. 33

[Bon70] Aline Bonami. Étude des coefficients de Fourier des fonctions de $L^p(G)$. In *Annales de l'institut Fourier*, volume 20, pages 335–402, 1970. 3

[BRdeW08] Avraham Ben-Aroya, Oded Regev, and Ronald de Wolf. A hypercontractive inequality for matrix-valued functions with applications to quantum computing and LDCs. In *Proceedings of the 49th IEEE Annual Symposium on Foundations of Computer Science*, pages 477–486. IEEE, 2008. 4, 5, 7, 9, 11, 14, 31

[BYJK04] Ziv Bar-Yossef, Thathachar S. Jayram, and Iordanis Kerenidis. Exponential separation of quantum and classical one-way communication complexity. In *Proceedings of the 36th Annual ACM Symposium on Theory of computing*, pages 128–137, 2004. 6, 16

[CGKS95] Benny Chor, Oded Goldreich, Eyal Kushilevitz, and Madhu Sudan. Private information retrieval. In *Proceedings of the 36th IEEE Annual Foundations of Computer Science*, pages 41–50. IEEE, 1995. 8

[CGS+21] Chi-Ning Chou, Alexander Golovnev, Madhu Sudan, Ameya Velingker, and Santhoshini Velusamy. Linear space streaming lower bounds for approximating CSPs. *arXiv preprint arXiv:2106.13078*, 2021. 9, 10

[CGSV21a] Chi-Ning Chou, Alexander Golovnev, Madhu Sudan, and Santhoshini Velusamy. Approximability of all Boolean CSPs in the dynamic streaming setting. *arXiv preprint arXiv:2102.12351*, 2021. 2, 6, 10

[CGSV21b] Chi-Ning Chou, Alexander Golovnev, Madhu Sudan, and Santhoshini Velusamy. Approximability of all finite CSPs in the dynamic streaming setting. *arXiv preprint arXiv:2105.01161*, 2021. 2, 6, 8, 10

[CGV20] Chi-Ning Chou, Alexander Golovnev, and Santhoshini Velusamy. Optimal streaming approximations for all Boolean Max-2CSPs and Max-kSAT. In *Proceedings of the 61st IEEE Annual Symposium on Foundations of Computer Science*, pages 330–341. IEEE, 2020. 2, 6, 8

[CGdeW09] Victor Chen, Elena Grigorescu, and Ronald de Wolf. Efficient and error-correcting data structures for membership and polynomial evaluation. *arXiv preprint arXiv:0909.3696*, 2009. 8

[CKP+21] Lijie Chen, Gillat Kol, Dmitry Paramonov, Raghuvansh R Saxena, Zhao Song, and Huacheng Yu. Almost optimal super-constant-pass streaming lower bounds for reachability. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 570–583, 2021. 2

[CL93] Eric A. Carlen and Elliott H. Lieb. Optimal hypercontractivity for Fermi fields and related non-commutative integration inequalities. *Communications in Mathematical Physics*, 155(1):27–46, 1993. 4, 5, 11

[DG16] Zeev Dvir and Sivakanth Gopi. 2-server PIR with subpolynomial communication. *Journal of the ACM (JACM)*, 63(4):1–15, 2016. 2

[DGTJ84] William J Davis, DJH Garling, and Nicole Tomczak-Jaegermann. The complex convexity of quasi-normed linear spaces. *Journal of functional analysis*, 55(1):110–150, 1984. 5

[DM20] João F. Doriguello and Ashley Montanaro. Exponential quantum communication reductions from generalizations of the Boolean Hidden Matching problem. In *15th Conference on the Theory of Quantum Computation, Communication and Cryptography*, volume 158 of *LIPIcs*, pages 1:1–1:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. 16, 40

[DS05] Irit Dinur and Samuel Safra. On the hardness of approximating minimum vertex cover. *Annals of mathematics*, pages 439–485, 2005. 4

[DS07] Zeev Dvir and Amir Shpilka. Locally decodable codes with two queries and polynomial identity testing for depth 3 circuits. *SIAM Journal on Computing*, 36(5):1404–1434, 2007. 2, 3, 9

[DSSST10] John C. Duchi, Shai Shalev-Shwartz, Yoram Singer, and Ambuj Tewari. Composite objective mirror descent. In *COLT*, pages 14–26. Citeseer, 2010. 5, 11

[Dvi11] Zeev Dvir. On matrix rigidity and locally self-correctable codes. *Computational Complexity*, 20(2):367–388, 2011. 3

[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 39th Annual ACM Symposium on Theory of computing*, pages 516–525, 2007. 4, 6, 7, 16, 25, 40

[GKK12] Ashish Goel, Michael Kapralov, and Sanjeev Khanna. On the communication and streaming complexity of maximum bipartite matching. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 468–485. SIAM, 2012. 2, 6

[GKST02] Oded Goldreich, Howard Karloff, Leonard J. Schulman, and Luca Trevisan. Lower bounds for linear locally decodable codes and private information retrieval. In *Proceedings of the 17th IEEE Annual Conference on Computational Complexity*, pages 175–183. IEEE, 2002. 2, 3, 8, 9, 34, 35

[Gop18] Sivakanth Gopi. Locality in coding theory. 2018. 2, 3

[Gro75] Leonard Gross. Logarithmic Sobolev inequalities. *American Journal of Mathematics*, 97(4):1061–1083, 1975. 5

[GT19] Venkatesan Guruswami and Runzhou Tao. Streaming hardness of unique games. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM*, volume 145 of *LIPIcs*, pages 5:1–5:12. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. 2, 6, 7, 8, 16, 25, 27, 28, 40

[GVV17] Venkatesan Guruswami, Ameya Velingker, and Santhoshini Velusamy. Streaming complexity of approximating max 2CSP and max acyclic subgraph. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017. 2, 6, 8

[GW95] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995. 5

[Han56] Olof Hanner. On the uniform convexity of L_p and l_p. *Arkiv för Matematik*, 3(3):239–244, 1956. 5

[Hel76] Carl Wilhelm Helstrom. *Quantum detection and estimation theory*. Academic press, 1976. 11

[HLP52] G. H. Hardy, J. E. Littlewood, and G. Pólya. Inequalities. *Cambridge University Press*, 10:169, 1952. 15

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006. 4

[HRMS20] Steven R. Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform Chernoff bounds via nonnegative supermartingales. *Probability Surveys*, 17:257–317, 2020. 5

[IK04] Yuval Ishai and Eyal Kushilevitz. On the hardness of information-theoretic multiparty computation. In *International Conference on the Theory and Applications of Cryptographic Techniques*, pages 439–455. Springer, 2004. 8

[KK19] Michael Kapralov and Dmitry Krachun. An optimal space lower bound for approximating MAX-CUT. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 277–288, 2019. 2, 3, 6, 8, 10

[KKL89] Jeff Kahn, Gil Kalai, and Nathan Linial. *The influence of variables on Boolean functions*. Citeseer, 1989. 7, 16, 40

[KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, 2007. 4, 6

[KKS14] Michael Kapralov, Sanjeev Khanna, and Madhu Sudan. Streaming lower bounds for approximating MAX-CUT. In *Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete algorithms*, pages 1263–1282. SIAM, 2014. 2, 3, 6, 8, 16, 27, 28, 30

[KKSV17] Michael Kapralov, Sanjeev Khanna, Madhu Sudan, and Ameya Velingker. $(1+\Omega(1))$-approximation to MAX-CUT requires linear space. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1703–1722. SIAM, 2017. 2, 10

[KT00] Jonathan Katz and Luca Trevisan. On the efficiency of local decoding procedures for error-correcting codes. In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, pages 80–86, 2000. 3, 8, 9, 30, 34

[KW04] Iordanis Kerenidis and Ronald de Wolf. Exponential lower bound for 2-query locally decodable codes via a quantum argument. *Journal of Computer Systems and Science*, 69(3):395–420, 2004. 2, 3, 6, 9

[LMN93] Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, Fourier transform, and learnability. *Journal of the ACM (JACM)*, 40(3):607–620, 1993. 4

[LN04] James R. Lee and Assaf Naor. Embedding the diamond graph in L_p and dimension reduction in L_1. *Geometric & Functional Analysis GAFA*, 14(4):745–747, 2004. 5, 11

[LO00] Rafał Latała and Krzysztof Oleszkiewicz. Between Sobolev and Poincaré. In *Geometric aspects of functional analysis*, pages 147–168. Springer, 2000. 15

[Mon11] Ashley Montanaro. A new exponential separation between quantum and classical one-way communication complexity. *Quantum Inf. Comput.*, 11(7&8):574–591, 2011. 4

[Mon12] Ashley Montanaro. Some applications of hypercontractive inequalities in quantum information theory. *Journal of Mathematical Physics*, 53(12):122206, 2012. 4

[Nao16] Assaf Naor. A spectral gap precludes low-dimensional embeddings. *arXiv preprint arXiv:1611.08861*, 2016. 5, 11

[New91] Ilan Newman. Private vs. common random bits in communication complexity. *Information processing letters*, 39(2):67–71, 1991. 7

[Oba02] Kenji Obata. Optimal lower bounds for 2-query locally decodable linear codes. In *International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 39–50. Springer, 2002. 9

[O'D14] Ryan O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, 2014. 4, 5, 15

[Pin94] Iosif Pinelis. Optimum bounds for the distributions of martingales in Banach spaces. *The Annals of Probability*, pages 1679–1706, 1994. 5

[SWY12] Yaoyun Shi, Xiaodi Wu, and Wei Yu. Limits of quantum one-way communication by matrix hypercontractive inequality, 2012. 2, 7, 16, 19, 24, 42

[TJ74] Nicole Tomczak-Jaegermann. The moduli of smoothness and convexity and the Rademacher averages of the trace classes S_p $(1 \leq p < \infty)$. *Studia Mathematica*, 50(2):163–182, 1974. 5, 11

[Tre04] Luca Trevisan. Some applications of coding theory in computational complexity. *arXiv preprint cs/0409044*, 2004. 8

[VY11] Elad Verbin and Wei Yu. The streaming complexity of cycle counting, sorting by reversals, and other problems. In *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 11–25. SIAM, 2011. 2, 6, 7, 8, 16, 25, 40

[Wat18] John Watrous. *The theory of quantum information*. Cambridge University Press, 2018. 11

[Wol07] Paweł Wolff. Hypercontractivity of simple random variables. *Studia Mathematica*, 3(180):219–236, 2007. 15

[deW08] Ronald de Wolf. A brief introduction to Fourier analysis on the Boolean cube. *Theory of Computing*, pages 1–20, 2008. 4

[WdeW05] Stephanie Wehner and Ronald de Wolf. Improved lower bounds for locally decodable codes and private information retrieval. In *International Colloquium on Automata, Languages, and Programming*, pages 1424–1436. Springer, 2005. 3, 9

[Yek12] Sergey Yekhanin. Locally decodable codes. *Foundations and Trends® in Theoretical Computer Science*, 6(3):139–255, 2012. 2

A Classical Hidden Hypermatching lower bound

The general idea behind the proof of Theorem 15 is already well established and is a simple generalization of results from [GKK+07, VY11, GT19, DM20]. We shall need the following well-known fact and a generalization of the KKL inequality [KKL89].

Fact 34. *Given just one sample, the best success probability in distinguishing between two probability distributions p and q is $\frac{1}{2} + \frac{1}{4}\|p - q\|_{\mathrm{tvd}}$.*

Lemma 35 (Generalized KKL inequality)**.** *Let $A \subseteq \mathbb{Z}_r^n$ and let $f : \mathbb{Z}_r^n \to \{0,1\}$ be its indicator function ($f(x) = 1$ iff $x \in A$). Then, for every $\delta \in [0, 1/r]$,*

$$\sum_{S \in \mathbb{Z}_r^n} \delta^{|S|} |\widehat{f}(S)|^2 \leq \left(\frac{|A|}{r^n}\right)^{2/(1+r\delta)}.$$

Proof. Apply the hypercontractive inequality to real-valued functions with $p = 1 + r\delta$. □

Theorem 36. *Any classical protocol that achieves advantage $\varepsilon > 0$ for the r-$\mathsf{HH}(\alpha, t, n)$ problem with $t \geq 2$ and $\alpha \leq 1/2$ requires at least $c = \Omega(r^{-1}(\varepsilon^4/\alpha)^{1/t}(n/t)^{1-1/t})$ bits of communication from Alice to Bob.*

Proof. By the minimax principle, it suffices to analyse *deterministic* protocols under some 'hard' input distribution. For our input distribution, Alice and Bob receive $x \in \mathbb{Z}_r^n$ and $M \in \mathcal{M}_{t,n}^{\alpha}$, respectively, uniformly at random, while Bob's input $w \in \mathbb{Z}_r^{\alpha n/t}$ is drawn from the distribution $\mathcal{D} := \frac{1}{2}\mathcal{D}^{\mathsf{YES}} + \frac{1}{2}\mathcal{D}^{\mathsf{NO}}$, i.e., with probability $1/2$ is comes from $\mathcal{D}^{\mathsf{YES}}$, and with probability $1/2$ it comes from $\mathcal{D}^{\mathsf{NO}}$.

Fix a small constant $\varepsilon > 0$ and let $c = \gamma r^{-1}(\varepsilon^4/\alpha)^{1/t}(n/t)^{1-1/t}$ for some universal constant γ. Consider any classical deterministic protocol that communicates at most $C := c - \log(1/\varepsilon)$ bits. Such protocol partitions the set of all r^n x's into 2^C sets. These sets have size $r^n/2^C$ on average,

and by a counting argument, with probability $1-\varepsilon$, the set A corresponding to Alice's message has size at least $\varepsilon r^n/2^C = r^n/2^c$. Given Alice's message, Bob knows that the random variable X corresponding to her input was drawn uniformly at random from A, and he also knows his input M. Therefore his knowledge of the random variable MX is described by the distribution

$$p_M(z) := \Pr[MX = z|M, A] = \frac{|\{x \in A | Mx = z\}|}{|A|}.$$

Given one sample of $w \in \mathbb{Z}_r^{\alpha n/t}$, Bob must decide whether it came from $\mathcal{D}^{\mathsf{YES}}$ (the distribution MX) or from $\mathcal{D}^{\mathsf{NO}}$ (the uniform distribution U). According to Fact 34, the advantage of any classical protocol in distinguishing between p_M and U is upper bounded by $\frac{1}{4}\|p_M - U\|_{\mathrm{tvd}}$. We prove in Theorem 37 below that, if $\alpha \leq 1/2$ and $c \leq \frac{\gamma}{r}(\frac{\varepsilon^4}{\alpha})^{1/t}(n/t)^{1-1/t}$, then the average advantage over all hypermatchings M is at most $\varepsilon^2/4$, i.e.,

$$\mathop{\mathbb{E}}_{M\sim\mathcal{M}_{t,n}^{\alpha}} [\|p_M - U\|_{\mathrm{tvd}}] \leq \varepsilon^2.$$

Therefore, by Markov's inequality, for at least a $(1-\varepsilon)$-fraction of M, the advantage in distinguishing between p_M and U is $\varepsilon/4$ small. Hence, Bob's total advantage over randomly guessing the right distribution will be at most ε (for the event that A is too small) plus ε (for the event that M is such that the distance between MX and U is more than ε) plus $\varepsilon/4$ (for the advantage over random guessing when $\|p_M - U\|_{\mathrm{tvd}} \leq \varepsilon$), and so $c = \Omega(r^{-1}(\varepsilon^4/\alpha)^{1/t}(n/t)^{1-1/t})$. □

Theorem 37. *Let $x \in \mathbb{Z}_r^n$ be uniformly distributed over a set $A \subseteq \mathbb{Z}_r^n$ of size $|A| \geq r^n/2^c$ for some $c \geq 1$. If $\alpha \leq 1/2$, there is a universal constant $\gamma > 0$ (independent of n, t, r and α), such that, for all $\varepsilon > 0$, if $c \leq \frac{\gamma}{r}(\frac{\varepsilon^4}{\alpha})^{1/t}(n/t)^{1-1/t}$, then*

$$\mathop{\mathbb{E}}_{M\sim\mathcal{M}_{t,n}^{\alpha}} [\|p_M - U\|_{\mathrm{tvd}}] \leq \varepsilon^2.$$

Proof. Let $f : \mathbb{Z}_r^n \to \{0,1\}$ be the characteristic function of A, i.e., $f(x) = 1$ iff $x \in A$. We shall bound the Fourier coefficients of p_M, which are related to the Fourier coefficients of f as follows:

$$\begin{aligned}
\widehat{p}_M(V) = \frac{1}{r^{\alpha n/t}} \sum_{z\in\mathbb{Z}_r^n} p_M(z)\omega_r^{-V\cdot z} &= \frac{1}{|A| r^{\alpha n/t}} \sum_{z\in\mathbb{Z}_r^n} |\{x \in A | Mx = z\}| \cdot \omega_r^{-V\cdot z} \\
&= \frac{1}{|A| r^{\alpha n/t}} \sum_{k=0}^{r-1} |\{x \in A | (Mx)\cdot V = k\}| \cdot \omega_r^{-k} \\
&= \frac{1}{|A| r^{\alpha n/t}} \sum_{k=0}^{r-1} |\{x \in A | x\cdot (M^{\mathrm{T}}V) = k\}| \cdot \omega_r^{-k} \\
&= \frac{1}{|A| r^{\alpha n/t}} \sum_{x\in A} \omega_r^{-x\cdot(M^{\mathrm{T}}V)} \\
&= \frac{r^n}{|A| r^{\alpha n/t}} \widehat{f}(M^{\mathrm{T}}V).
\end{aligned}$$

We now start bounding the expected *squared* total variation distance,

$$\mathop{\mathbb{E}}_{M\sim\mathcal{M}_{t,n}^{\alpha}} [\|p_M - U\|_{\mathrm{tvd}}^2] \leq r^{2\alpha n/t} \mathop{\mathbb{E}}_{M\sim\mathcal{M}_{t,n}^{\alpha}} [\|p_M - U\|_2^2]$$

$$= r^{2\alpha n/t} \mathop{\mathbb{E}}_{M\sim\mathcal{M}^{\alpha}_{t,n}} \left[\sum_{V\in\mathbb{Z}_r^{\alpha n/t}\setminus\{0^{\alpha n/t}\}} |\widehat{p}_M(V)|^2 \right]$$
$$= \frac{r^{2n}}{|A|^2} \mathop{\mathbb{E}}_{M\sim\mathcal{M}^{\alpha}_{t,n}} \left[\sum_{V\in\mathbb{Z}_r^{\alpha n/t}\setminus\{0^{\alpha n/t}\}} |\widehat{f}(M^{\mathrm{T}}V)|^2 \right],$$

where we used Cauchy-Schwarz followed by Parseval's identity. Note that there is at most one $V \in \mathbb{Z}_r^{\alpha n/t}$ such that $S = M^{\mathrm{T}}V$ for a given $S \in \mathbb{Z}_r^n$ (and that the only V that makes $M^{\mathrm{T}}V = 0^n$ is $V = 0^{\alpha n/t}$). This allows us to transform the expectation over hypermatchings into a probability,

$$\mathop{\mathbb{E}}_{M\sim\mathcal{M}^{\alpha}_{t,n}} [\|p_M - U\|_{\mathrm{tvd}}^2] \leq \frac{r^{2n}}{|A|^2} \mathop{\mathbb{E}}_{M\sim\mathcal{M}^{\alpha}_{t,n}} \left[\sum_{S\in\mathbb{Z}_r^n\setminus\{0^n\}} |\{V \in \mathbb{Z}_r^{\alpha n/t} | M^{\mathrm{T}}V = S\}| \cdot |\widehat{f}(S)|^2 \right]$$
$$= \frac{r^{2n}}{|A|^2} \sum_{S\in\mathbb{Z}_r^n\setminus\{0^n\}} \Pr_{M\sim\mathcal{M}^{\alpha}_{t,n}} [\exists V \in \mathbb{Z}_r^{\alpha n/t} : M^{\mathrm{T}}V = S] \cdot |\widehat{f}(S)|^2.$$

Now observe that $\Pr_{M\sim\mathcal{M}^{\alpha}_{t,n}}[\exists V \in \mathbb{Z}_r^{\alpha n/t} : M^{\mathrm{T}}V = S]$ is exactly the probability from Lemma 14, i.e., given $S \in \mathbb{Z}_r^n$ with $k_j := \frac{1}{t} \cdot |\{i \in [n] : S_i = j\}| \in \mathbb{Z}$ for $j \in [r-1]$ (the number of entries from S equal to $j \neq 0$ must be a multiple of t), and defining $k := \sum_{j=1}^{r-1} k_j$, then

$$\Pr_{M\sim\mathcal{M}^{\alpha}_{t,n}} [\exists V \in \mathbb{Z}_r^{\alpha n/t} : M^{\mathrm{T}}V = S] = \frac{\binom{\alpha n/t}{k}}{\binom{n}{kt}} \frac{k!}{(kt)!} \prod_{j=1}^{r-1} \frac{(k_j t)!}{k_j!} \leq \frac{\binom{\alpha n/t}{k}}{\binom{n}{kt}},$$

and so

$$\mathop{\mathbb{E}}_{M\sim\mathcal{M}^{\alpha}_{t,n}} [\|p_M - U\|_{\mathrm{tvd}}^2] \leq \frac{r^{2n}}{|A|^2} \sum_{k=1}^{\alpha n/t} \frac{\binom{\alpha n/t}{k}}{\binom{n}{kt}} \sum_{\substack{S\in\mathbb{Z}_r^n \\ |S|=kt}} |\widehat{f}(S)|^2.$$

Similarly to the quantum proof, we shall split the above sum into two parts: one in the range $1 \leq k < 2rc$, and the other in the range $2rc \leq k \leq \alpha n/t$.

Sum I ($1 \leq k < 2rc$): in order to upper bound each term, pick $\delta = k/(2rc)$ in Lemma 35, thus

$$\frac{r^{2n}}{|A|^2} \sum_{\substack{S\in\mathbb{Z}_r^n \\ |S|=kt}} |\widehat{f}(S)|^2 \leq \frac{r^{2n}}{|A|^2} \frac{1}{\delta^{kt}} \sum_{S\in\mathbb{Z}_r^n} \delta^{|S|} |\widehat{f}(S)|^2 \leq \frac{1}{\delta^{kt}} \left(\frac{r^n}{|A|}\right)^{2r\delta/(1+r\delta)} \leq \frac{1}{\delta^{kt}} \left(\frac{r^n}{|A|}\right)^{2r\delta} \leq \left(\frac{2^{1/t} 2rc}{k}\right)^{kt}.$$

By using that $c \leq \frac{\gamma}{r}(\frac{\varepsilon^4}{\alpha})^{1/t}(n/t)^{1-1/t}$ and $\binom{q}{s}\binom{\ell q}{\ell s}^{-1} \leq (\frac{s}{q})^{(\ell-1)s}$ (see [SWY12, Appendix A.5]) for $q = n/t, s = k, \ell = t$, we therefore have

$$\frac{r^{2n}}{|A|^2} \sum_{k=1}^{2rc-1} \alpha^k \frac{\binom{n/t}{k}}{\binom{n}{kt}} \sum_{\substack{S\in\mathbb{Z}_r^n \\ |S|=kt}} |\widehat{f}(S)|^2 \leq \sum_{k=1}^{2rc-1} \alpha^k \left(\frac{kt}{n}\right)^{(1-1/t)kt} \left(\frac{2^{1/t} 2rc}{k}\right)^{kt} \leq \sum_{k=1}^{2rc-1} \left(\frac{2^{1/t} 2\gamma\varepsilon^{4/t}}{k^{1/t}}\right)^{kt} \leq \frac{\varepsilon^4}{2},$$

where we used that $\binom{\alpha n/t}{k} \leq \alpha^k \binom{n/t}{k}$ for $\alpha \in [0,1]$ at the beginning and picked γ sufficiently small.

Sum II ($2rc \le k \le \alpha n/t$): first note that the function $g(k) := \binom{\alpha n/t}{k}/\binom{n}{kt}$ is decreasing in the interval $1 \le k \le \alpha n/t$ (since $\alpha \le 1/2$). Hence, by using Parseval's identity $\sum_{S\in\mathbb{Z}_r^n} |\widehat{f}(S)|^2 = |A|/r^n$ and the inequality $\binom{q}{s}\binom{\ell q}{\ell s}^{-1} \le (\frac{s}{q})^{(\ell-1)s}$ (for $q = n/t, s = 2m, \ell = t$) in order to bound $g(2rc)$,

$$\frac{r^{2n}}{|A|^2}\sum_{k=2rc}^{\alpha n/t}\frac{\binom{\alpha n/t}{k}}{\binom{n}{kt}}\sum_{\substack{S\in\mathbb{Z}_r^n\\|S|=kt}}|\widehat{f}(S)|^2 \le 2^c g(2rc) \le 2^c\alpha^{2rc}\left(\frac{2rc}{n/t}\right)^{2(t-1)rc} = 2^c\alpha^{2rc/t}\left(\frac{2\gamma\varepsilon^{4/t}}{(n/t)^{1/t}}\right)^{2(t-1)rc} \le \frac{\varepsilon^4}{2},$$

where in the last step we used that $c \ge 1 \implies 2(1-1/t)c \ge 1$ (so $\varepsilon^{2(1-1/t)c} \le \varepsilon$) and picked γ sufficiently small.

Summing both results, if $c \le \frac{\gamma}{r}(\frac{\varepsilon^4}{\alpha})^{1/t}(n/t)^{1-1/t}$, then $\mathbb{E}_{M\sim\mathcal{M}_{t,n}^{\alpha}}[\|p_M - U\|_{\mathrm{tvd}}^2] \le \varepsilon^4$. By Jensen's inequality, we finally get $\mathbb{E}_{M\sim\mathcal{M}_{t,n}^{\alpha}}[\|p_M - U\|_{\mathrm{tvd}}] \le \sqrt{\mathbb{E}_{M\sim\mathcal{M}_{t,n}^{\alpha}}[\|p_M - U\|_{\mathrm{tvd}}^2]} \le \varepsilon^2$. □

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il