ECCC

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

SRINIVASAN ARUNACHALAM, IBM Quantum, IBM T.J. Watson Research Center, USA

JOAO F. DORIGUELLO, HUN-REN Alfréd Rényi Institute of Mathematics, Hungary and Centre for Quantum Technologies, National University of Singapore, Singapore

We prove a hypercontractive inequality for matrix-valued functions defined over large alphabets. In order to do so, we prove a generalization of the powerful 2-uniform convexity inequality for trace norms of Ball, Carlen, Lieb (Inventiones Mathematicae'94). Using our hypercontractive inequality, we present upper and lower bounds for the communication complexity of the Hidden Hypermatching problem defined over large alphabets. We then consider streaming algorithms for approximating the value of Unique Games on a hypergraph with t-size hyperedges. By using our communication lower bound, we show that every streaming algorithm in the adversarial model achieving an $(r - \varepsilon)$-approximation of this value requires $\Omega(n^{1-2/t})$ quantum space, where r is the alphabet size. We next present a lower bound for locally decodable codes (LDC) $\mathbb{Z}_r^n \to \mathbb{Z}_r^N$ over large alphabets with recoverability probability at least $1/r + \varepsilon$. Using hypercontractivity, we give an exponential lower bound $N = 2^{\Omega(\varepsilon^4 n/r^4)}$ for 2-query (possibly non-linear) LDCs over $\mathbb{Z}_r$ and using the non-commutative Khintchine inequality we prove an improved lower bound of $N = 2^{\Omega(\varepsilon^2 n/r^2)}$.

CCS Concepts: • **Theory of computation** → **Streaming models**; **Quantum communication complexity**; **Communication complexity**; **Error-correcting codes**.

Additional Key Words and Phrases: Streaming Algorithms, Locally Decodable Codes, Quantum Communication Complexity, Hypercontractivity, Max-Cut, Unique Games

21

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 [3] (for which they won the Gödel Prize in 2005) in order to understand the space complexity of approximation 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 [4, 16, 20, 21, 31, 36, 46–48, 68]. 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 problem*, or its generalization over large alphabets $\mathbb{Z}_r$, *Unique Games*. There is a 2-approximation algorithm for Max-Cut on n vertices using logarithmic space: 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 [2] and, from it, a $(1 + \varepsilon)$-approximation for the Max-Cut value. On the other hand, a sequence of works [19, 46–48] showed that getting a $(2 - \varepsilon)$-approximation requires linear space. A similar, but less optimized, phenomenon was observed for Unique Games, i.e., there is a threshold behaviour in complexity going from r to $(r - \varepsilon)$-approximation, where r is the alphabet size. 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 solve these combinatorial optimization problems? One corollary from [46, 64] 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 Max-Cut) or $(r - \varepsilon)$-approximation (for Unique Games over $\mathbb{Z}_r$), 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 $\widetilde{C}(x)$, 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, and complexity theory (we refer to [34, 73] for detailed expositions). Despite their ubiquity, LDCs are not well understood, even with the simplest case of 2-*query* LDCs. For the case when $\Sigma = \Gamma = \{0, 1\}$, exponential lower bounds of $N = 2^{\Omega(n)}$ were established over two decades back [28, 32, 50]. In contrast, a breakthrough result of Dvir and Gopi [27] 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 of size $2^{n^{o(1)}}$. 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 $\Sigma = \Gamma = \mathbb{Z}_r$.[1] What is the relation between n and N in this case?

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 LDCs when the input alphabet is $\mathbb{Z}_r$, 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 a generalization of the powerful 2-uniform convexity by Ball, Carlen, and Lieb [5]. Using this new hypercontractivity theorem, we prove a quantum space lower bound for streaming algorithms. It is easy to see that obtaining a 2-approximation algorithm for Max-t-Cut on n vertices and t-sized hyperedges 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. Compared to the prior works of Kapralov, Khanna, and Sudan [46] and of Kapralov and

---

[1]We shall think of an r-size alphabet set as the ring $\mathbb{Z}_r$ since its structure will be extensively used in our analyses.

Krachun [48], who obtained an $\Omega(\sqrt{n})$ and $\Omega(n)$ classical space lower bounds in the *random* and *adversarial* models, respectively, for $(2 - \varepsilon)$-approximation algorithms, our lower bounds apply to the adversarial model and are derived through significantly simpler proofs. We further generalize our results to the case of hypergraphs with t-size hyperedges and 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 while a $(r - \varepsilon)$-approximation algorithm requires $\Omega(n^{1-2/t})$ quantum space (hiding the dependence on r and ε).

We next show an $N = 2^{\Omega(n/r^2)}$ lower bound for (even non-linear) LDCs over $\mathbb{Z}_r$. In particular, for all r smaller than $\sqrt{n}$, 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. [32] for $r = 2$ and linear LDCs, Kerenidis and de Wolf [50] for $r = 2$ and *non-linear* LDCs, Wehner and de Wolf [70] for non-linear LDCs from $\{0, 1\}^n \to \mathbb{Z}_r^N$ and finally by Dvir and Shpilka [28] for $r > 2$ but linear LDCs. Apart from the result of [28], 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 private information retrieval [32, 49], additive combinatorics [13], and quantum complexity theory [1], to cite a few. Moreover, the setting of equal alphabets $\Sigma = \Gamma = \mathbb{Z}_r$ is particularly relevant and natural in the context of polynomial identity testing and $\Sigma\Pi\Sigma$ circuits, which has been previously considered by Dvir and Shpilka [28] for linear LDCs. Furthermore, we are not aware of a formal reduction between LDCs with $\Sigma = \{0, 1\}$ and $\Sigma = \mathbb{Z}_r$, specially with recovery probability $1/|\Sigma| + \varepsilon$. Finally, some past works define LDCs over general Σ with success probability $\geq \Pr[\text{wrong output}] + \varepsilon$ [34], $\geq 1/2 + \varepsilon$ [32] or $\geq 1 - \varepsilon$ [26]. 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 $S \cdot x = \sum_{i=1}^n S_i x_i \in \{0, 1\}$ and the *Fourier coefficients* $\widehat{f}(S)$ of f are $\widehat{f}(S) = \mathbb{E}_x[f(x)(-1)^{S \cdot x}]$, 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 [8, 12]. 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 the operator on the space of functions $f : \{0, 1\}^n \to \mathbb{R}$ defined as

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

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 (see, e.g., [60, Proposition 2.47])

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

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. Another property of the noise operator is that it is a contraction, i.e., $\|\mathrm{T}_\rho f\|_p \leq \|f\|_p$ for every

$p \geq 1$, where $\|f\|_p \triangleq \left(\mathbb{E}_x[|f(x)|^p]\right)^{1/p}$ is the standard normalized p-norm of the function f. The main hypercontractivity theorem states that this 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 $\|T_\rho f\|_2 \leq \|f\|_p$,[2] 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 [23, 51], expander graphs [41, Section 11.5], circuit complexity [54], coding theory [14], quantum computing [30, 55] (for more applications we refer the reader to [56, 60, 71]). 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

$$\|T_\rho f\|_2 \leq \|f\|_p \quad \text{for} \quad \rho \leq \sqrt{p-1} \cdot \lambda^{1/p-1/2}, \tag{2}$$

where λ is the smallest probability in any of the finite probability spaces (Ω_i, π_i) (see [60, 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 [10], who proved a hypercontractivity statement by using the powerful inequality of Ball, Carlen, and Lieb [5].

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 (THEOREM 8). *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 $\|M\|_p \triangleq \left(\sum_i \sigma_i(M)^p\right)^{1/p}$ is the Schatten p-norm defined from the singular values $\{\sigma_i(M)\}_i$ of the matrix M and $|S| \triangleq |\{i \in [n] : S_i \neq 0\}|$ is the Hamming weight of $S \in \mathbb{Z}_r^n$.

The above result can be seen as an analogue of Equation (1) where the absolute values are replaced with Schatten norms. We now make a couple of remarks. First, when $m = 1$ our result compares to the one in Equation (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 [10]. To this end, as in the proof of hypercontractive inequalities [10, 60], 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 Equation (3) for $n = 1$. We now consider two special *simple* cases of the inequality.

---

[2]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.

(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 [35] 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 Equation (3) is exactly a powerful inequality in functional analysis, called the 2-*uniform convexity* of trace norms,[3]

$$\left(\frac{\|X + Y\|_p^p + \|X - Y\|_p^p}{2} \right)^{2/p} \geq \|X\|_p^2 + (p-1)\|Y\|_p^2. \tag{4}$$

This inequality was first proven for certain values of p by Tomczak-Jaegermann [65] before being extended for all $p \in [1, 2]$ by Ball, Carlen, and Lieb [5] in 1994. Since then it has found several applications, e.g., an optimal hypercontractivity inequality for Fermi fields [14], regularized convex optimization [25], and metric embedding [53, 57]. 2-uniform convexity can also be used to prove a variety of other inequalities, e.g., Khintchine's inequality [22, 65] and Hoeffding and Bennett-style bounds [42, 62]. Moreover, the above result could be seen as a corollary of Hanner's inequality for matrices [5]

$$\|X + Y\|_p^p + \|X - Y\|_p^p \geq |\|X\|_p + \|Y\|_p|^p + |\|X\|_p - \|Y\|_p|^p$$

(originally proven for Lebesgue spaces L_p [38]), but, unfortunately, Hanner's inequality for Schatten trace ideals is only proven for $1 \leq p \leq 4/3$ and $p \geq 4$ [5] (see [15] for recent progress on Hanner's inequality for matrices). As far as we are aware, a generalization of Equation (4) involving r matrices was unknown.

One contribution in our work is the following generalization of a result from Ball, Carlen, and Lieb [5] (note it also implies a generalization of the two-point inequality), which we believe may be of independent interest.

RESULT 2 (THEOREM 6). *Let* $r \in \mathbb{Z}$, $r \geq 2$, $\omega_r \triangleq e^{2i\pi/r}$, $A_0, \ldots, A_{r-1} \in \mathbb{C}^{n \times n}$, *and* $p \in [1, 2]$, *then*

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

Once Result 2 is established, 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's inequality. Since this proof is similar to the one in [10], we omit the details here.

1.3 Application: communication complexity and streaming algorithms

1.3.1 Communication complexity: Hidden Matching and its variants. The Boolean Hidden Matching (BHM) problem was introduced by Bar-Yossef et al. [6] (which was in turn inspired by Kerenidis and de Wolf [50] for proving LDC lower bounds) in order to prove exponential separations between quantum and classical one-way communication complexities. We describe below the generalized Hidden Matching problem over larger alphabets and hypermatching. Here a hypermatching is a set of hyperedges without common vertices. 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 from $[n]$ (where $\alpha \in (0, 1]$). The disjoint t-tuples from $[n]$ can be viewed as hyperedges of size t, which, in turn, can also be viewed as an

---

[3]The 2 in 2-uniform convexity comes from a lower bound on the modulus of convexity of the normed space L_p with $p \in (1, 2]$ (see [5] for more details) and not from our parameter r in Result 1 being equal to 2.

incidence 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 on understanding the problem of Hidden Hypermatching: (*i*) the seminal work of Bar-Yossef et al. [6] and Gavinsky et al. [30] 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 [68] 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 [64]; (*iii*) Guruswami and Tao [36] studied the problem for when $t = 2$ and $r \geq 2$, proving a classical $\Omega(\sqrt{n})$ lower bound; (*iv*) apart from these, there have been more exotic generalizations of the Hidden Hypermatching problem, e.g., Kapralov, Khanna, and Sudan [46] proposed the Boolean Hidden Partition, where Bob does not receive a matching anymore (a graph made up of pair-wise disjoint edges), but the edges of any graph G, and Doriguello and Montanaro [24] 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 and give both upper and lower bounds for its quantum and classical communication complexities for every $t, r \geq 2$.

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.[4] 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 (THEOREM 12). *There is a one-way 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 one 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 and the known upper bound $O(n^{1-1/\lceil t/2 \rceil} \log n)$ for $r = 2$ [50] (note that $\lceil t/2 \rceil$ is the number of quantum queries required to learn the parity of t bits [7, 29]), we 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. [30], Verbin and Yu [68], and Guruswami and

---

[4]One can further improve this complexity to $O(\log(n \log r) + (\log r) \cdot (n/\alpha)^{1-1/t})$ by Newman's theorem [58].

Tao [36], 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 [44] (which can be seen as a corollary of the hypercontractivity inequality). On the other hand, Shi, Wu, and Yu [64] 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 [10]. We follow a similar approach by using our *generalized matrix-valued hypercontractive* inequality from Result 1 in order to obtain the following lower bound (for $r = 2$ our lower bound is exponentially better in α than [64]).

RESULT 4 (THEOREM 14, THEOREM 39). *Every constant-bias protocol for r-HH(α, t, n) with $t, r \geq 2$ requires $\Omega((n/t)^{1-2/t}/\alpha^{2/t})$ qubits of communication or $\Omega((n/t)^{1-1/t}/\alpha^{1/t})$ bits of communication.*

The phase transition from $t = 2$ to $t > 2$ in the quantum communication complexity of r-HH(α, t, n) is reminiscent of the well-known fact that parity of n bits can be computed exactly using $\lceil n/2 \rceil$ quantum queries [7, 29]: a function requiring more than one query within the same hyperedge will lead to a large communication complexity since, by design of the Hidden Matching problem, it is hard to extract information from the same hyperedge more than once.

1.3.2 Streaming lower bounds. One-way communication complexity lower bounds has been used by several recent works to prove streaming lower bounds [21, 36, 37, 46, 68]. 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 [36, 46, 48], matching [31], Max-CSP [20, 21, 37] and counting cycles [4, 68]. One problem that is used often in this direction is the aforementioned Boolean Hidden Matching and related variants. Using our classical and quantum communication lower bounds (Result 4) we present two lower bounds for streaming problems on generalizations to Max-Cut.

There are a few natural generalizations to Max-Cut. One is Max-t-Cut, i.e., finding the maximum cut value on a hypergraph with t-sized hyperedges, to which the $\Omega(n)$ classical lower bound from [48] already applies. 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 [36] 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 from Hidden Matching over $\mathbb{Z}_r$. This bound was improved to $\Omega(n)$ by Chou et al. [19] for a larger set of problems including Unique Games.

Here we join both directions, i.e., Max-t-Cut and the standard Unique Games problem, into a generalized version of Unique Games defined on a hypergraph and obtain a streaming *quantum* lower bound in the adversarial model for any value $t, r \geq 2$.[5]

RESULT 5 (THEOREM 22). *Every streaming algorithm giving a $(r - \varepsilon)$-approximation for Unique Games on t-hyperedge n-vertex hypergraphs over $\mathbb{Z}_r$ uses $\Omega(n^{1-2/t})$ quantum space.*

---

[5]We note that Kapralov et al. [46] proved a classical lower bound $\Omega(n^{1-\varepsilon})$ for $(1 + \varepsilon)$-approximations of Max-Cut in the adversarial model and their proof, together with Result 4, readily implies a similar quantum lower bound.

The above result follows from generalizing the proof of Guruswami and Tao [36]. As far as we are aware, these are the first quantum lower bounds for Unique Games in the streaming model.

1.4 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 [49] 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 [18], multiparty computation [43], data structures [17], and average-case complexity theory [66].

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 $y \oplus e_i$ 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. [32] showed a lower bound $N = 2^{\Omega(n)}$ for *linear codes*, which was later improved by Obata [59]. Later, Kerenidis and de Wolf [50] proved an exponential lower bound for *non-linear codes* using a quantum argument![6]

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 [70] proved the lower bound $N = 2^{\Omega(n/r^2)}$. On the other hand, Dvir and Shpilka [28] 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. [32]. This reduction, however, fails for non-linear codes and motivates the question of whether 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 $2^{\Omega(n/r^2)}$, and which gives a super-polynomial lower bound for $r = o(n^{1/2})$. Our lower bound comes from using the non-commutative Khintchine's inequality [67].[7]

RESULT 6 (THEOREM 31). *If* $C : \mathbb{Z}_r^n \to \mathbb{Z}_r^N$ *is a* $(2, \delta, \varepsilon)$-LDC*, then* $N = 2^{\Omega(\delta^2 \varepsilon^2 n/r^2)}$.

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. Moreover, we stress that the success probability is bounded by $1/r + \varepsilon$, meaning that the lower bound of Wehner and de Wolf [70] is not directly applicable by restricting the messages to binary strings. Finally, we remark that in an earlier version of the paper, we proved a similar lower bound using our matrix-valued hypercontractivity via the proof technique in [10, Theorem 11]. This leads, however, to a weaker bound, $N = 2^{\Omega(\delta^2 \varepsilon^4 n/r^4)}$. After our preprint appeared online, we were made aware of the non-commutative Khintchine's inequality which leads us to the improved lower bound of $2^{\Omega(\delta^2 \varepsilon^2 n/r^2)}$.

---

[6]For simplicity in exposition, we omit the dependence on δ, ε in these lower bounds.
[7]We thank Jop Brië for helping us prove this theorem.

Further applications (Private information retrieval). Katz and Trevisan [49] and Goldreich et al. [32] 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. Using Result 6 and other auxiliary results, we get the following lower bound for PIR schemes over $\mathbb{Z}_r$ when $r \geq 2$ is prime.

RESULT 7 (THEOREM 36). *Let $r \geq 2$ be prime. A classical 2-server PIR scheme with alphabet $\mathbb{Z}_r$, query size t, answer size a, and recovery probability $1/r + \varepsilon$ satisfies $t = \Omega\big(\varepsilon^2 n/r^{4a+6} - a\big)$.*

Previous lower bounds on 2-server PIR protocols with alphabet size $r = 2$ include, among others, $t = \Omega(n\varepsilon^2/2^{5a})$ [50], $t = \Omega(n\varepsilon^2/2^{2a})$ [70], and (for a restricted model of 2-server PIR) $t + a = \Omega(n^{1/3})$ [63]. Other bounds in various settings include [9, 33, 74]. We note that Result 7 does not contradict Dvir and Gopi's [27] 2-server PIR scheme with query size $t = n^{O(\sqrt{\log \log n/\log n})} = n^{o(1)}$ since it uses answer size $a = O(t)$, in which case our bound becomes trivial.

After completion of this work. Kallaugher and Parekh [45] recently put out an online preprint in which they improve our quantum streaming lower bound for Max-t-Cut to $\Omega(n)$. As far as we know, our quantum streaming lower bound for $\mathbb{Z}_r$ when $r > 2$ was not known before.

1.5 Future work

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 [11])?

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-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 Unique Games? Given the recent work of Kallaugher and Parekh [45], who proved a $\Omega(n)$ streaming lower bound for Max-t-Cut, we conjecture that the same bound applies to Unique Games. A proof might require obtaining new quantum lower bounds for the communication problems introduced in [19, 20, 47, 48].

4. Generalized hypercontractivity. Another open question is regarding our main Result 1, which shows a form of $(2, p)$-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 different generalization of the inequality of Ball, Carlen, and Lieb [5].

2 PRELIMINARIES

Let $[n] \triangleq \{1, \ldots, n\}$. For $r \in \mathbb{Z}$, $r \geq 2$, we let $\mathbb{Z}_r \triangleq \{0, \ldots, r-1\}$ be the ring with addition and multiplication modulo r, and let $\omega_r \triangleq e^{2\pi i/r}$. Given $S \in \mathbb{Z}_r^n$, we write $|S| \triangleq |\{i \in [n] : S_i \neq 0\}|$ for its Hamming weight. Let $D(\mathbb{C}^m)$ 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 $M \in \mathbb{C}^{m \times m}$, its (unnormalized) Schatten p-norm is defined as $\|M\|_p \triangleq (\text{Tr} |M|^p)^{1/p} = \big(\sum_{i=1}^m \sigma_i(M)^p\big)^{1/p}$, where $\{\sigma_i(M)\}_i$ are the singular values of M, i.e., the eigenvalues of the positive semi-definite operator $|M| \triangleq \sqrt{M^\dagger M}$. Moreover,

let $\|M\| \triangleq \max_{i \in [m]} \sigma_i(M)$ and $\|M\|_F \triangleq \sqrt{\sum_{i,j=1}^m |M_{ij}|^2}$ be its spectral and Frobenious norms, respectively (notice that $\|M\| = \|M\|_\infty$). Given a vector $v \in \mathbb{C}^m$, its p-norm is $\|v\|_p \triangleq \left(\sum_{i=1}^m |v_i|^p \right)^{1/p}$. Given two probability distributions P and Q on the same finite set, their total variation distance is $\|P - Q\|_{\text{tvd}} \triangleq \sum_i |P(i) - Q(i)|$ (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 $\varepsilon r/(r-1)$ 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 shall need the Holevo-Helstrom theorem [40] which characterizes the optimal success probability of distinguishing between two quantum states.

FACT 2 ([69, 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*

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

We will need the following result derived from the non-commutative Khintchine's inequality.

LEMMA 3. *Given $A_1, \ldots, A_n \in \mathbb{C}^{N \times N}$, then*

$$\mathop{\mathbb{E}}_{k \sim \mathbb{Z}_r^n} \left\| \sum_{i=1}^n \omega_r^{k_i} A_i \right\| \leq 2\sqrt{2\log(2N)} \sqrt{\sum_{i=1}^n \|A_i\|^2}.$$

PROOF. $\displaystyle \mathop{\mathbb{E}}_{k \sim \mathbb{Z}_r^n} \left\| \sum_{i=1}^n \omega_r^{k_i} A_i \right\| = \mathop{\mathbb{E}}_{k \sim \mathbb{Z}_r^n} \left\| \sum_{i=1}^n \left(\omega_r^{k_i} A_i - \mathop{\mathbb{E}}_{k' \sim \mathbb{Z}_r^n} \left[\omega_r^{k_i'} A_i \right] \right) \right\|$

$$\leq \mathop{\mathbb{E}}_{k,k' \sim \mathbb{Z}_r^n} \left\| \sum_{i=1}^n (\omega_r^{k_i} - \omega_r^{k_i'}) A_i \right\|$$

$$= \mathop{\mathbb{E}}_{k,k' \sim \mathbb{Z}_r^n} \mathop{\mathbb{E}}_{\varepsilon \sim \{\pm 1\}^n} \left\| \sum_{i=1}^n \varepsilon_i (\omega_r^{k_i} - \omega_r^{k_i'}) A_i \right\|$$

$$\leq 2 \mathop{\mathbb{E}}_{k \sim \mathbb{Z}_r^n} \mathop{\mathbb{E}}_{\varepsilon \sim \{\pm 1\}^n} \left\| \sum_{i=1}^n \varepsilon_i \omega_r^{k_i} A_i \right\|$$

$$\leq 2\sqrt{2\log(2N)} \sqrt{\sum_{i=1}^n \|A_i\|^2},$$

using the non-commutative Khintchine's inequality [67, Theorem 4.1.1] in the last step. □

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 [10]. 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 [5] 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 [5, Theorem 1].

Theorem 4 (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 [65] for $p \leq 4/3$, before being generalized by Ball, Carlen, and Lieb [5] for all $p \in [1, 2]$ in 1994. Since then it has found several applications [14, 25, 53, 57]. The above result can be recast in a slightly different way.

Theorem 5. *Let $p \in [1, 2]$ and $Z, W \in \mathbb{C}^{n \times n}$ such that $\operatorname{Tr}[|Z|^{p-2} Z W^\dagger] = \operatorname{Tr}[|Z|^{p-2} W Z^\dagger] = 0$ (where $|Z|^{p-2} = (ZZ^\dagger)^{p/2-1}$). Then*

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

Theorem 5 is implicit in the proof of [5, Theorem 1], and it is where most of the difficulty lies, while the reduction from Theorem 4 to Theorem 5 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 5 holds more generally for any $Z, W \in \mathbb{C}^{n \times n}$ that satisfy $\operatorname{Tr}[|Z|^{p-2} Z W^\dagger] = \operatorname{Tr}[|Z|^{p-2} W Z^\dagger] = 0$. By using this result, we can prove the following generalization of Theorem 4.

Theorem 6 (A generalization of [5]). *Let $r \in \mathbb{Z}, r \geq 2$. Let $\omega_r \triangleq e^{2i\pi/r}, A_0, \ldots, A_{r-1} \in \mathbb{C}^{n \times n}$, and $p \in [1, 2]$, then*

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

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

Notice that for $r = 2$ we recover Theorem 4, 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 Equation (5b). In order

to prove it, let us first define the $rn \times rn$ matrices

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

for $j \in \{0, \dots, r-1\}$. 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 Equation (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 5 as follows: consider a permutation of $[r-1]$ given by $(k_1, \dots, k_{r-1})$. Since

$$\mathrm{Tr}[|Z_j|^{p-2} Z_j Z_k^\dagger] = \mathrm{Tr} \begin{bmatrix} |A_j|^{p-2} A_j A_k^\dagger & 0 & \dots & 0 \\ 0 & \omega_r^{j-k} |A_j|^{p-2} A_j A_k^\dagger & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \omega_r^{(r-1)(j-k)} |A_j|^{p-2} A_j A_k^\dagger \end{bmatrix}$$

$$= \mathrm{Tr}[|A_j|^{p-2} A_j A_k^\dagger] \sum_{\ell=0}^{r-1} \omega_r^{\ell(j-k)} = 0$$

for any $j \neq k$ (and similarly $\mathrm{Tr}[|Z_j|^{p-2} Z_k Z_j^\dagger] = 0$), then (define $k_0 \triangleq 0$)

$$\mathrm{Tr} \left[|Z_{k_j}|^{p-2} Z_{k_j} \left(\sum_{l=j+1}^{r-1} Z_{k_l} \right)^\dagger \right] = \mathrm{Tr} \left[|Z_{k_j}|^{p-2} \left(\sum_{l=j+1}^{r-1} Z_{k_l} \right) Z_{k_j}^\dagger \right] = 0$$

for every $j \in \{0, 1, \dots, r-2\}$, meaning that Theorem 5 can be applied, which implies

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

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

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

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

$$= \|Z_0\|_p^2 + \frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)} \sum_{j=1}^{r-1} \|Z_j\|_p^2,$$

proving our theorem statement. □

REMARK 7. *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 6 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 6 (this fact was already pointed out for $r=2$ by [5]). Indeed, consider again the $rn \times rn$ matrices Z_j from Equation (6). Then, similar to Equation (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 $j \in \mathbb{Z}_r$, Equation (8) implies (define $\zeta \triangleq \frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}$ 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 6.

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

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

we have

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

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

PROOF. For ease of notation, define $\zeta \triangleq \frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}$. It suffices to prove the inequality for $\rho = \sqrt{\zeta}$. Our proof closely follows the one in [10] and is by induction on n. For $n = 1$, the desired

statement is

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

$$\Updownarrow$$

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

where we used that $f(k) = \sum_{j=0}^{r-1} \omega_r^{jk} \widehat{f}(j)$. But this is precisely Theorem 6 with $A_k = \widehat{f}(k)$, $k \in \mathbb{Z}_r$.

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 $i \in \{0, \ldots, r - 1\}$ be the function obtained by fixing the last coordinate of $f(\cdot)$ to i. By the induction hypothesis we have that, for every $i \in \{0, \ldots, r - 1\}$ and $p \in [1, 2]$,

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

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

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

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 above inequality by the desired left-hand side of the conjectured statement. To do so, we will need the following Minkowski's inequality.

LEMMA 9 (MINKOWSKI'S INEQUALITY, [39, THEOREM 26]). *For any $r_1 \times r_2$ matrix whose rows are given by $u_1, \ldots, u_{r_1}$ and whose columns are given by $v_1, \ldots, v_{r_2}$, and any $1 \leq q_1 \leq q_2 \leq \infty$,*

$$\left\| (\|v_1\|_{q_2}, \ldots, \|v_{r_2}\|_{q_2}) \right\|_{q_1} \geq \left\| (\|u_1\|_{q_1}, \ldots, \|u_{r_1}\|_{q_1}) \right\|_{q_2}.$$

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

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

$$\geq \left(\frac{1}{r^n} \sum_{S \in \mathbb{Z}_r^n} \left(\frac{1}{r} \sum_{i=0}^{r-1} c_{S,i}^p \right)^{2/p} \right)^{1/2}$$

(Lemma 9 with $q_1 = p$ and $q_2 = 2$)

$$= \left(\sum_{S \in \mathbb{Z}_r^n} \zeta^{|S|} \left(\frac{1}{r} \sum_{i=0}^{r-1} \|\widehat{g_i}(S)\|_p^p \right)^{2/p} \right)^{1/2}. \tag{11}$$

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

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

Plugging this back into Equation (11), we have

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

$$= \left(\sum_{S \in \mathbb{Z}_r^{n+1}} \zeta^{|S|} \|\widehat{f}(S)\|_p^2 \right)^{1/2} ,$$

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. $\qquad\square$

REMARK 10 (COMPARISON WITH HYPERCONTRACTIVITY FOR REAL NUMBERS). *For real functions $f : \mathbb{Z}_r^n \to \mathbb{R}$, it is known that [52, 72] (see also [60, Theorem 10.18])*

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

where $\rho \leq \sqrt{\frac{(r-1)^{1-1/p} - (r-1)^{-(1-1/p)}}{(r-1)^{1/p} - (r-1)^{-1/p}}}$. Moreover, this bound on ρ is perfectly sharp, meaning that our bound $\rho \leq \sqrt{\frac{(p-1)(1-(p-1)^{r-1})}{(r-1)(2-p)}}$ in Theorem 8 can possibly be improved. We note that, for $r = 2$, both bounds become $\rho \leq \sqrt{p-1}$.

4 HIDDEN HYPERMATCHING PROBLEM

In the following, an α-partial t-hypermatching $M \in \mathcal{M}_{t,n}^\alpha$ 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 $\mathcal{M}_{t,n}^\alpha$ is the set of all such hypermatchings. If $\alpha = 1$, we shall write $\mathcal{M}_{t,n}$.

DEFINITION 11. *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 $M \in \mathcal{M}_{t,n}^\alpha$ 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 incidence matrix of Bob's hypermatching. Consider the distributions:*

(1) *YES distribution $\mathcal{D}^{\text{YES}}$, let $w = Mx$ (where the matrix product Mx is over $\mathbb{Z}_r$);*

(2) *NO distribution $\mathcal{D}^{\text{NO}}$, 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 $\mathcal{D}^{\text{YES}}$ or $\mathcal{D}^{\text{NO}}$.

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 12. *Given $\varepsilon > 0$, there is a protocol for the r-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 $M \in \mathcal{M}_{2,n}^{\alpha}$ be Bob's matching with disjoint edges $(M_{1,1}, M_{1,2}), \ldots, (M_{\alpha n/2,1}, M_{\alpha n/2,2}) \in [n]^2$. Alice first sends the following state in $\mathbb{C}^r \otimes \mathbb{C}^n$ to Bob,

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

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

$$E_i \triangleq |M_{i,1}\rangle\langle M_{i,1}| + |M_{i,2}\rangle\langle M_{i,2}|, \qquad \forall i \in [\alpha n/2]$$

(note that $|M_{i,j}\rangle \in \mathbb{C}^n$ for $i \in [\alpha n/2], j \in [2]$). 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

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

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 quantum Fourier transform Q_r over $\mathbb{Z}_r$ to it to obtain

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

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 $C_\ell \triangleq PS_\ell P \otimes X$. 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 quantum 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

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

and

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

$$= \frac{\omega_r^\ell}{\sqrt{2}} (PS_\ell \otimes \mathbb{I})(\omega_r^{\ell-x_1} |\ell - x_1, 1\rangle + \omega_r^{\ell-x_0} |\ell - x_0, 0\rangle)$$

$$= \frac{\omega_r^\ell}{\sqrt{2}} (P \otimes \mathbb{I})(\omega_r^{\ell-x_1} |x_1, 1\rangle + \omega_r^{\ell-x_0} |x_0, 0\rangle)$$

$$= \omega_r^{2\ell} |\psi\rangle.$$

We can see from the above that $C_\ell^{2k} |\psi\rangle = \omega_r^{2k\ell} |\psi\rangle$. By defining $\Delta_\ell \triangleq \ell - (x_0 + x_1)$ 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 Equation (12). This means that Bob's state in Equation (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

$$\Pr[m] = \frac{1}{2r^2} \sum_{k_1, k_2=0}^{r-1} \omega_r^{(\ell-m)(k_1-k_2)} (\langle x_0 + \Delta_\ell \delta_{k_2} | x_0 + \Delta_\ell \delta_{k_1}\rangle + \langle x_1 + \Delta_\ell \delta_{k_2} | x_1 + \Delta_\ell \delta_{k_1}\rangle)$$

$$= \frac{1}{r^2} \sum_{k_1, k_2 \text{ even}}^{r-1} \omega_r^{(\ell-m)(k_1-k_2)} + \frac{1}{r^2} \sum_{k_1, k_2 \text{ odd}}^{r-1} \omega_r^{(\ell-m)(k_1-k_2)}$$

$$= \left| \frac{1}{r} \sum_{k \text{ even}}^{r-1} \omega_r^{k(\ell-m)} \right|^2 + \left| \frac{1}{r} \sum_{k \text{ odd}}^{r-1} \omega_r^{k(\ell-m)} \right|^2.$$

It is not hard to see that the above probability is maximal 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 Equation (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-HH(α, t, n) problem, but first we need the following lemma.

LEMMA 13. *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 $p \triangleq 1 + (r-1)\delta$. 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 8 and Remark 7, 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, note $p \leq q \implies \left(\frac{1}{2^m} \sum_{i=1}^{2^m} \sigma_i^p\right)^{1/p} \leq \left(\frac{1}{2^m} \sum_{i=1}^{2^m} \sigma_i^q\right)^{1/q}$ 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

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

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

THEOREM 14. *Any quantum protocol that achieves advantage $\varepsilon > 0$ for the r-$\mathrm{HH}(\alpha, t, n)$ problem with $t \geq 3$ and $\alpha \leq \min(1/2, (r-1)^{-1/2})$ requires $\Omega(r^{-(1+1/t)}(\varepsilon^2/\alpha)^{2/t}(n/t)^{1-2/t})$ 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 [64]. Also, see Remark 19 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 $\mathcal{D}^{\mathrm{YES}}$ or $\mathcal{D}^{\mathrm{NO}}$. 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 $\mathcal{D} \triangleq \frac{1}{2}\mathcal{D}^{\mathrm{YES}} + \frac{1}{2}\mathcal{D}^{\mathrm{NO}}$, i.e., with probability $1/2$ is comes from $\mathcal{D}^{\mathrm{YES}}$, and with probability $1/2$ it comes from $\mathcal{D}^{\mathrm{NO}}$. Let $p_x \triangleq r^{-n}$, $p_M \triangleq |\mathcal{M}_{t,n}^\alpha|^{-1}$ and $p_w \triangleq r^{-\alpha n/t}$, then our hard distribution $\mathcal{P}$ is

$$\Pr[x, \mathrm{YES}, M, w] = \frac{1}{2} p_x \cdot p_M \cdot [Mx = w], \qquad \Pr[x, \mathrm{NO}, M, w] = \frac{1}{2} p_x \cdot p_M \cdot p_w. \qquad (14)$$

Conditioning on Bob's input (M, w), from his perspective, Alice sends the message $\rho(x)$ 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 $\rho_{\text{YES}}^{M,w}$ and $\rho_{\text{NO}}^{M,w}$, each appearing with probability $\Pr[\text{YES}|M,w]$ and $\Pr[\text{NO}|M,w]$, respectively,

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

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 $\rho_{\text{YES}}^{M,w}$ and $\rho_{\text{NO}}^{M,w}$.

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

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

We prove in Theorem 15 below that, if $m \leq \frac{\gamma}{r^{1+1/t}} \left(\frac{\varepsilon^2}{\alpha} \right)^{2/t} (n/t)^{1-2/t}$ for a universal constant γ, then the average bias over M and w is at most ε^2, i.e.,

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

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 $\rho_{\text{YES}}^{M,w}$ and $\rho_{\text{NO}}^{M,w}$ 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 $\rho_{\text{YES}}^{M,w}$ and $\rho_{\text{NO}}^{M,w}$ is more than ε) plus $\varepsilon/2$ (for the advantage over random guessing when $\varepsilon_{\text{bias}} \leq \varepsilon$), and so $m = \Omega(r^{-(1+1/t)} (\varepsilon^2/\alpha)^{2/t} (n/t)^{1-2/t})$. □

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

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

PROOF. For the ease of notation, we shall write

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

Therefore, we have that

$$
\begin{aligned}
\varepsilon_{\text{bias}} &= \sum_{M \in \mathcal{M}_{t,n}^\alpha} \sum_{w \in \mathbb{Z}_r^{\alpha n/t}} \Pr[M, w] \cdot \left\| \Pr[\text{YES}|M,w] \cdot \rho_{\text{YES}}^{M,w} - \Pr[\text{NO}|M,w] \cdot \rho_{\text{NO}}^{M,w} \right\|_1 \\
&= \sum_{M \in \mathcal{M}_{t,n}^\alpha} \sum_{w \in \mathbb{Z}_r^{\alpha n/t}} \left\| \sum_{x \in \mathbb{Z}_r^n} \left(\Pr[x, \text{YES}, M, w] \cdot \rho(x) - \Pr[x, \text{NO}, M, w] \cdot \rho(x) \right) \right\|_1 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(Equation (15) and conditional probability)} \\
&= \sum_{M \in \mathcal{M}_{t,n}^\alpha} \sum_{w \in \mathbb{Z}_r^{\alpha n/t}} \left\| \sum_{x \in \mathbb{Z}_r^n} \frac{1}{2} p_x \cdot p_M \left([Mx = w] - p_w \right) \rho(x) \right\|_1 \qquad\qquad \text{(Equation (14))}
\end{aligned}
$$

$$= \sum_{M \in \mathcal{M}_{t,n}^{\alpha}} \sum_{w \in \mathbb{Z}_r^{\alpha n/t}} \left\| \sum_{x \in \mathbb{Z}_r^n} \frac{1}{2} p_x \cdot p_M \left([Mx = w] - p_w \right) \cdot \sum_{S \in \mathbb{Z}_r^n} \widehat{\rho}(S) \omega_r^{S \cdot x} \right\|_1$$

(Fourier decomposition of ρ)

$$= \sum_{M \in \mathcal{M}_{t,n}^{\alpha}} \sum_{w \in \mathbb{Z}_r^{\alpha n/t}} \left\| \sum_{S \in \mathbb{Z}_r^n} u(M, w, S) \widehat{\rho}(S) \right\|_1$$

$$\leq \sum_{S \in \mathbb{Z}_r^n} \sum_{M \in \mathcal{M}_{t,n}^{\alpha}} \sum_{w \in \mathbb{Z}_r^{\alpha n/t}} |u(M, w, S)| \cdot \|\widehat{\rho}(S)\|_1,$$

where we defined

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

Next, we upper bound the quantity $u(M, w, S)$ using the lemma below. In the following lemma, let $I(M) = \bigcup_{i \in [\alpha n/t], j \in [t]} \{M_{i,j}\}$ be the set of indices of the α-partial matching $M \in \mathcal{M}_{n,t}^{\alpha}$. 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 16. *Let $M \in \mathcal{M}_{t,n}^{\alpha}$ and $I(M) = \bigcup_{i \in [\alpha n/t], j \in [t]} \{M_{i,j}\}$. Define the set*

$$\Delta(M) = \{S \in \mathbb{Z}_r^n \setminus \{0^n\} \mid S_{M_{i,1}} = S_{M_{i,2}} = \cdots = S_{M_{i,t}} \; \forall i \in [\alpha n/t] \text{ and } S_j = 0 \; \forall j \notin I(M)\}.$$

Given $u(M, w, S)$ as defined in Equation (16) for $w \in \mathbb{Z}_r^{\alpha n/t}$ and $S \in \mathbb{Z}_r^n$, 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:

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

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

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

$$= \sum_{x \in \mathbb{Z}_r^n} \omega_r^{S \cdot x} \prod_{i=1}^{\alpha n/t} \left[\sum_{j=1}^t x_{M_{i,j}} = w_i \right]$$

$$= \sum_{x \in \mathbb{Z}_r^n} \left(\prod_{j \notin I(M)} \omega_r^{S_j x_j} \right) \left(\prod_{i=1}^{\alpha n/t} \omega_r^{S|_{M_i} \cdot x|_{M_i}} \left[\sum_{j=1}^t x_{M_{i,j}} = w_i \right] \right)$$

$$= \left(\prod_{j \notin I(M)} \sum_{x \in \mathbb{Z}_r} \omega_r^{S_j x} \right) \left(\prod_{i=1}^{\alpha n/t} \sum_{x \in \mathbb{Z}_r^t} \omega_r^{S|_{M_i} \cdot x} \left[\sum_{j=1}^t x_j = w_i \right] \right) \quad \text{(Relabelling } x \in \mathbb{Z}_r^n\text{)}$$

$$= r^{n(1-\alpha)} [S_j = 0 \; \forall j \notin I(M)] \prod_{i=1}^{\alpha n/t} \sum_{x \in \mathbb{Z}_r^t} \omega_r^{S|_{M_i} \cdot x} \left[\sum_{j=1}^t x_j = w_i \right].$$

Now we use that $\sum_{j=1}^t x_j = w_i \implies x_t = w_i - \sum_{j=1}^{t-1} x_j$ modulo r, so that

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

modulo r. This leads to (assuming that $S_j = 0$ for all $j \notin I(M)$)

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

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 all $i \in [\alpha n/t]$. Otherwise the above expression is 0. Thus, if $S_{M_{i,1}} = S_{M_{i,2}} = \cdots = S_{M_{i,t}}$ for all $i \in [\alpha n/t]$ and $S_j = 0$ for $j \notin I(M)$, then we can use Equation (17) to get (remember that $p_x \triangleq r^{-n}$ and $p_w \triangleq r^{-\alpha n/t}$)

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

$$= \begin{cases} 0 & \text{if } S = 0^n, \\ \frac{1}{2} r^{-\alpha n/t} p_M & \text{if } S \neq 0^n. \end{cases}$$

Hence, we have

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

proving the lemma statement □

We now proceed to upper bound $\varepsilon_{\text{bias}}$ using the expression for $|u(M, w, S)|$ from Lemma 16. For $S \in \mathbb{Z}_r^n$, let $|S| \triangleq |\{i \in [n] : S_i \neq 0\}|$. Notice that, if $S \in \Delta(M)$, then $|S| = kt$ for some $k \in [\alpha n/t]$. Hence, we have that

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

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

using that

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

We now upper bound this probability using the following lemma.

LEMMA 17. Let $t \in \mathbb{Z}$. Let $S \in \mathbb{Z}_r^n$ with $k_j \triangleq \frac{1}{t} \cdot |\{i \in [n] : S_i = j\}| \in \mathbb{Z}$ for $j \in \{1, \ldots, r-1\}$. Let $k \triangleq \sum_{j=1}^{r-1} k_j$. For any $M \in \mathcal{M}_{t,n}^\alpha$, let $\Delta(M)$ be the set from Lemma 16. 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 loss 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 $n! / ((t!)^{\alpha n/t} (\alpha n/t)! (n - \alpha n)!)$. 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

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

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

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

Using Lemma 17 and the notation $|S|_i \triangleq |\{j \in [n] : S_j = i\}|$, we continue bounding $\varepsilon_{\text{bias}}$ as

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

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

(Cauchy-Schwarz)

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

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

where the last equality uses $\sum_{S \in \mathbb{Z}_r^n : |S|_i = k_i t} 1 = \binom{n}{kt}(kt)! \prod_{j=1}^{r-1} \frac{1}{(k_j t)!}$. Since

$$\sum_{\substack{k_1,\ldots,k_{r-1} \geq 0 \\ \sum_{j=1}^{r-1} k_j = k}} \frac{k!^2}{(kt)!} \prod_{j=1}^{r-1} \frac{(k_j t)!}{k_j!^2} = \sum_{\substack{k_1,\ldots,k_{r-1} \geq 0 \\ \sum_{j=1}^{r-1} k_j = k}} \frac{\binom{k}{k_1,\ldots,k_{r-1}}^2}{\binom{kt}{k_1 t,\ldots,k_{r-1} t}} \leq \sum_{\substack{k_1,\ldots,k_{r-1} \geq 0 \\ \sum_{j=1}^{r-1} k_j = k}} \binom{k}{k_1,\ldots,k_{r-1}} = (r-1)^k, \quad (18)$$

where the last equality follows from the the multinomial theorem, then

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

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 ($1 \le k < 4rm$): in order to upper bound each term, pick $\delta = k/(4rm)$ in Lemma 13, so

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

By using that $m \le \frac{\gamma}{r^{1+1/t}} (\frac{\varepsilon^2}{\alpha})^{2/t} (n/t)^{1-2/t}$ and $\binom{q}{s}^2 \binom{\ell q}{\ell s}^{-1} \le (\frac{s}{q})^{(\ell-2)s}$ (see [64, Appendix A.5]) for $q = n/t, s = k, \ell = t$, we thus have

$$\sum_{k=1}^{4rm-1} \alpha^k \frac{\binom{n/t}{k}}{\sqrt{\binom{n}{kt}}} (r-1)^{k/2} \sqrt{\sum_{S \in \mathbb{Z}_r^n : |S| = kt} \|\widehat{\rho}(S)\|_1^2} \le \sum_{k=1}^{4rm-1} \alpha^k (r-1)^{k/2} \left(\frac{kt}{n}\right)^{(1-2/t)kt/2} \left(\frac{2^{1/(2t)} 4rm}{k}\right)^{kt/2}$$

$$\le \sum_{k=1}^{4rm-1} \alpha^k (r-1)^{k/2} \left(\frac{2^{1/(2t)} 4\gamma \varepsilon^{4/t}}{\alpha^{2/t} r^{1/t} k^{2/t}}\right)^{kt/2}$$

$$\le \sum_{k=1}^{4rm-1} \left(\frac{2^{1/4} (4\gamma)^{t/2} \varepsilon^2}{k}\right)^k \le \varepsilon^2$$

for sufficiently small γ.

Sum II ($4rm \le k \le \alpha n/t$): first we note that the function $g(k) \triangleq \alpha^k (r-1)^{k/2} \binom{n/t}{k} / \sqrt{\binom{n}{kt}}$ 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 13 with $\delta = 1$ and the inequality $\binom{q}{s}^2 \binom{\ell q}{\ell s}^{-1} \le (\frac{s}{q})^{(\ell-2)s}$ (for $q = n/t, s = 2m, \ell = t$) in order to bound $g(4rm)$,

$$\sum_{k=4rm}^{\alpha n/t} \alpha^k \frac{\binom{n/t}{k}}{\sqrt{\binom{n}{kt}}} (r-1)^{k/2} \sqrt{\sum_{S \in \mathbb{Z}_r^n : |S| = kt} \|\widehat{\rho}(S)\|_1^2} \le g(4rm) \sum_{k=4rm}^{\alpha n/t} \sqrt{\sum_{S \in \mathbb{Z}_r^n : |S| = kt} \|\widehat{\rho}(S)\|_1^2}$$

$$\le g(4rm) \sqrt{\frac{\alpha n}{t}} \sqrt{\sum_{S \in \mathbb{Z}_r^n} \|\widehat{\rho}(S)\|_1^2} \quad \text{(Cauchy-Schwarz)}$$

$$\le \left(\alpha\sqrt{r-1}\right)^{4rm} \left(\frac{4rm}{n/t}\right)^{2(t-2)rm} \sqrt{\frac{\alpha n}{t}} 2^{(r-1)m}$$

$$\le \left(2^{1/4} \alpha\sqrt{r-1}\right)^{4rm} \left(\frac{(4\gamma)^{t/2} \varepsilon^2}{\alpha\sqrt{r}(n/t)}\right)^{4(1-2/t)rm} \sqrt{\frac{\alpha n}{t}}$$

$$\le \varepsilon^2,$$

where 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 $\varepsilon_{\text{bias}} \le \varepsilon^2$. $\qquad \square$

A very similar classical communication lower bound for the r-HH(α, t, n) problem can be proven.

THEOREM 18. *Any one-way classical protocol that achieves advantage $\varepsilon > 0$ for the r-HH(α, t, n) problem with $t \geq 2$ and $\alpha \leq 1/2$ requires $\Omega(r^{-1}(\varepsilon^4/\alpha)^{1/t}(n/t)^{1-1/t})$ bits of communication.*

The proof is very similar to that of past works [30, 36, 68] 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 19. *The dependence of α on r can be improved. For example, we can improve the bound in Equation (18) by observing that $\binom{k}{k_1,\ldots,k_{r-1}}^2 \binom{kt}{k_1 t,\ldots,k_{r-1} t}^{-1} \leq 1$, 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 $\binom{q}{s}^2\binom{\ell q}{\ell s}^{-1} \leq (\frac{s}{q})^{(\ell-2)s} \leq 1$. Hence*

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

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

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

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

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

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

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 20 (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 $\pi_e : \mathbb{Z}_r^t \to \mathbb{Z}_r$, the goal is to compute*

$$\max_{x \in \mathbb{Z}_r^n} \sum_{e \in E} [\pi_e(x_e) = 0],$$

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

DEFINITION 21. *Let $H = (V, E)$ be a hypergraph and let OPT be the optimal value of the Unique Games on H. For $\gamma \geq 1$, 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 $[OPT/\gamma, OPT]$ 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 22. *Let $r, t \geq 2$ be integers. Every quantum streaming algorithm giving an $(r - \varepsilon)$-approximation for Unique Games on hypergraphs (as in Definition 20) 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 [36] and Kapralov, Khanna, and Sudan [46]. 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 20 for the $\mathcal{Y}, \mathcal{N}$ distributions:

- $\mathcal{Y}$ distribution: sample $z \in \mathbb{Z}_r^n$ and for each $e \in E$, let $\pi_e(x_e) = \sum_{i \in e}(x_i - z_i)$ (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 $\pi_e(x_e) = q - \sum_{i \in e} x_i$.

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 23. *Let $\varepsilon \in (0, 1)$. If $k = O(rt \log(r)/(\alpha \varepsilon^2))$, then for the Unique Games instance sampled from the distribution $\mathcal{N}$ above, the optimal fraction of satisfiable constraints (i.e., number of hyperedges $e \in E$ for which $\pi_e(\cdot)$ evaluates to 0) 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 [36, 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 the 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 as in [36], 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, a Chernoff bound for negative-correlated variables (see, e.g., [61, Theorem 3.2]) yields

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

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. $\qquad\square$

Reduction to Hypermatching. The reduction to r-ary Hidden Hypermatching is similar to the analysis used by Guruswami and Tao [36], but now it is from quantum streaming algorithms to one-way quantum communication complexity. The main lemma that we need is the following.

LEMMA 24. *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-HH(α, 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 24. 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.[8] Assume that $|\phi_0^{\mathcal{Y}}\rangle = |\phi_0^{\mathcal{N}}\rangle = 0$. Using the notion of informative index from [46, 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 25. *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-HH(α, 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 $|\phi^{\text{YES}}\rangle$ and $|\phi^{\text{NO}}\rangle$ 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 $|\phi^{\text{YES}}\rangle$ and $|\phi^{\text{NO}}\rangle$ with bias $\frac{1}{2}\left\| |\phi^{\text{YES}}\rangle - |\phi^{\text{NO}}\rangle \right\|_1$ by measuring the state $|s\rangle$ with a suitable POVM, according to Fact 2.

We are now ready to prove Lemma 24.

PROOF OF Lemma 24. 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 $|\phi^{\text{YES}}\rangle = |\phi_{j^*+1}^{\mathcal{Y}}\rangle = U|\phi_{j^*}^{\mathcal{Y}}, C^{\mathcal{Y}}\rangle$ and $|\phi^{\text{NO}}\rangle = U|\phi_{j^*}^{\mathcal{Y}}, C^{\mathcal{N}}\rangle$,

---

[8]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).

where $C^{\mathcal{Y}}$ and $C^{\mathcal{N}}$ are the constraints corresponding to the YES and NO distributions, respectively, and, similarly, we have $|\phi_{j^*+1}^{\mathcal{N}}\rangle = U|\phi_{j^*}^{\mathcal{N}}, C^{\mathcal{N}}\rangle$. Then, we have

$$\left\| |\phi^{\text{YES}}\rangle - |\phi^{\text{NO}}\rangle \right\|_1 \geq \left\| |\phi_{j^*+1}^{\mathcal{Y}}\rangle - |\phi_{j^*+1}^{\mathcal{N}}\rangle \right\|_1 - \left\| |\phi^{\text{NO}}\rangle - |\phi_{j^*+1}^{\mathcal{N}}\rangle \right\|_1$$

$$\geq \left\| |\phi_{j^*+1}^{\mathcal{Y}}\rangle - |\phi_{j^*+1}^{\mathcal{N}}\rangle \right\|_1 - \left\| |\phi_{j^*}^{\mathcal{Y}}\rangle - |\phi_{j^*}^{\mathcal{N}}\rangle \right\|_1 = \Omega(1/k),$$

where the second inequality used that $\left\| |\phi^{\text{NO}}\rangle - |\phi_{j^*+1}^{\mathcal{N}}\rangle \right\|_1 = \left\| U|\phi_{j^*}^{\mathcal{Y}}, C^{\mathcal{N}}\rangle - U|\phi_{j^*}^{\mathcal{N}}, C^{\mathcal{N}}\rangle \right\|_1 \leq \left\| |\phi_{j^*}^{\mathcal{Y}}\rangle - |\phi_{j^*}^{\mathcal{N}}\rangle \right\|_1$ (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 22. Finally, by picking $k = O(rt \log(r)/(\alpha \varepsilon^2))$ in order to invoke Lemma 23 and using our lower bound in Theorem 14 with $\alpha = O(1)$, we get our desired lower bound of

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

REMARK 26. *It is possible to prove a classical version of* Theorem 22 *by using the classical lower bound on r-HH(α, t, n), which leads to $\Omega((n/t)^{1-1/t})$ classical space (hiding dependence on r, ε). Such bound, though, is already subsumed by $\Omega(n)$ from Chou et al. [19].*

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 27 (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 [49] 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 28 (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 29 ([49]). *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 $(q, \delta, \varepsilon - c\delta)$-LDC, since the probability that the decoder queries one of δN corrupted positions is at most $(c/N)(\delta N) = c\delta$.

We now present our lower bound for LDCs over $\mathbb{Z}_r$ using the non-commutative Khintchine's inequality. We thank Jop Briët for the proof.

REMARK 30. *It is possible to prove a weaker lower bound $N = 2^{\Omega(\delta^2 \varepsilon^4 n/r^4)}$ using our matrix-valued hypercontractivity via the proof technique in [10, Theorem 11].*

THEOREM 31. *If $C : \mathbb{Z}_r^n \to \mathbb{Z}_r^N$ is a $(2, \delta, \varepsilon)$-LDC, then $N = 2^{\Omega(\delta^2 \varepsilon^2 n/r^2)}$.*

PROOF. We know that C is also a $(2, 2/\delta, \varepsilon)$-smooth code. Fix some $i \in [n]$. In order to decode x_i, we can assume, without loss of generality, that the decoder $\mathcal{A}$ picks some set $\{u, v\}$, where $u, v \in [N]$, with probability $p(u, v)$, queries those bits, and then outputs $f_i^{u,v}(C(x)_u, C(x)_v) \in \mathbb{Z}_r$ that depends on the query-outputs. Given the smooth code property of outputting x_i with probability at least $\frac{1}{r} + \varepsilon$ for every x, we have

$$\frac{1}{r} + \varepsilon \leq \Pr_{u,v \sim [N]}[f_i^{u,v}(C(x)_u, C(x)_v) = x_i] = \frac{1}{r} \sum_{k=0}^{r-1} \sum_{u,v \in [N]} p(u, v) \omega_r^{k(f_i^{u,v}(C(x)_u, C(x)_v) - x_i)} \iff$$

$$r\varepsilon \leq \sum_{k=1}^{r-1} \sum_{u,v \in [N]} p(u, v) \omega_r^{k(f_i^{u,v}(C(x)_u, C(x)_v) - x_i)},$$

where the if and only if comes from separating the term $k = 0$, which equals $\sum_{u,v \in [N]} p(u, v) = 1$.

For $k \in [r - 1]$, define the function $h_{i;k}^{u,v} : \mathbb{Z}_r^2 \to \mathbb{C}$ by $h_{i;k}^{u,v}(x) = \omega_r^{k f_i^{u,v}(x)}$. Consider its Fourier transform, $\widehat{h_{i;k}^{u,v}} : \mathbb{Z}_r^2 \to \mathbb{C}$, defined by

$$\widehat{h_{i;k}^{u,v}}(S) = \frac{1}{r^2} \sum_{x \in \mathbb{Z}_r^2} h_{i;k}^{u,v}(x) \omega_r^{-S \cdot x}.$$

Hence we can write

$$\omega_r^{k f_i^{u,v}(C(x)_u, C(x)_v)} = \sum_{a,b \in \mathbb{Z}_r} \widehat{h_{i;k}^{u,v}}(a, b) \omega_r^{aC(x)_u + bC(x)_v}.$$

Let $F_k^i \in \mathbb{C}^{rN \times rN}$ be the matrix defined as $(F_k^i)_{(u,a),(v,b)} = p(u, v) \widehat{h_{i;k}^{u,v}}(a, b)$ and let $C(x) \in \mathbb{C}^{rN}$ be the vector defined as $C(x)_{(u,a)} = \omega_r^{aC(x)_u}$. Then

$$\sum_{u,v \in [N]} p(u, v) \omega_r^{k f_i^{u,v}(C(x)_u, C(x)_v)} = C(x)^\top F_k^i C(x),$$

which means that

$$\sum_{k=1}^{r-1} \mathbb{E}_{x \sim \mathbb{Z}_r^n} \left[\omega_r^{-kx_i} C(x)^\top F_k^i C(x) \right] = \sum_{k=1}^{r-1} \sum_{u,v \in [N]} p(u, v) \mathbb{E}_{x \sim \mathbb{Z}_r^n} \left[\omega_r^{k(f_i^{u,v}(C(x)_u, C(x)_v) - x_i)} \right] \geq r\varepsilon.$$

By summing over all $i \in [n]$, we finally get to

$$\sum_{k=1}^{r-1} \mathbb{E}_{x \sim \mathbb{Z}_r^n} \left[C(x)^\top \left(\sum_{i=1}^n \omega_r^{-kx_i} F_k^i \right) C(x) \right] \geq r\varepsilon n.$$

The left-hand side can be upper bounded as follows:

$$\sum_{k=1}^{r-1} \mathbb{E}_{x \sim \mathbb{Z}_r^n} \left[C(x)^\top \left(\sum_{i=1}^n \omega_r^{-kx_i} F_k^i \right) C(x) \right] \leq \sum_{k=1}^{r-1} \mathbb{E}_{x \sim \mathbb{Z}_r^n} \left[\left\| \sum_{i=1}^n \omega_r^{-kx_i} F_k^i \right\| \|C(x)\|_2^2 \right]$$

$$= rN \sum_{k=1}^{r-1} \mathbb{E}_{x \sim \mathbb{Z}_r^n} \left[\left\| \sum_{i=1}^n \omega_r^{-kx_i} F_k^i \right\| \right]$$

$$\leq 2rN \sqrt{2 \log(2rN)} \sum_{k=1}^{r-1} \sqrt{\sum_{i=1}^n \|F_k^i\|^2},$$

where we used Lemma 3 in the last step. Consider the submatrix $[p(u,v)\widehat{h_{i;k}^{u,v}}(a,b)]_{a,b}$ of F_k^i for fixed $u, v \in [N]$. Since

$$\|[p(u,v)\widehat{h_{i;k}^{u,v}}(a,b)]_{a,b}\|^2 \leq \|[p(u,v)\widehat{h_{i;k}^{u,v}}(a,b)]_{a,b}\|_F^2 = \sum_{a,b \in \mathbb{Z}_r} p(u,v)^2 |\widehat{h_{i;k}^{u,v}}(a,b)|^2 = p(u,v)^2 \leq \frac{4}{\delta^2 N^2},$$

where we used Parseval's identity and $p(u,v) \leq 2/\delta N$ for all $u, v \in [N]$ by the definition of smooth code, then $\|F_k^i\|^2 \leq \frac{4}{\delta^2 N^2}$. This finally leads to

$$2rN \sqrt{2 \log(2rN)}(r-1) \frac{2\sqrt{n}}{\delta N} \geq r\varepsilon n \implies \log(2rN) \geq \frac{\delta^2 \varepsilon^2}{8r^2} n \implies N \geq \frac{1}{2r} 2^{\delta^2 \varepsilon^2 n/8r^2}. \qquad \square$$

6   2-SERVER PRIVATE INFORMATION RETRIEVAL

As mentioned in the introduction, the connection between LDCs and PIR is well known since the results of [32, 49]. 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 32.   *A one-round, $(1 - \delta)$-secure, k-server private information retrieval (PIR) scheme with alphabet $\mathbb{Z}_r$, 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 from a result of Goldreich et al. [32] and from a generalization of [50, Lemma 2]. We remark that Goldreich et al. state the lemma below only for $r = 2$, but the exact same analysis carries over to $r > 2$. In the following we shall assume $\delta = 0$.

LEMMA 33 ([32, LEMMA 5.1]).   *If there is a classical 2-server PIR scheme with alphabet $\mathbb{Z}_r$, 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 $m \leq 6r^t$.*

LEMMA 34.   *Let $r \geq 2$ be prime. Let $C : \mathbb{Z}_r^n \to (\mathbb{Z}_r^a)^m$ be a $(2, c, \varepsilon)$-smooth code. Then there is a $(2, cr^a, 2\varepsilon/r^{a+2})$-smooth code $C' : \mathbb{Z}_r^n \to \mathbb{Z}_r^{mr^a}$ that is good on average, i.e., there is a decoder $\mathcal{A}$ such*

that, for all $i \in [n]$,

$$\mathop{\mathbb{E}}_{x \sim \mathbb{Z}_r^a}\left[\Pr[\mathcal{A}^{C'(x)}(i) = x_i]\right] \geq \frac{1}{r} + \frac{2\varepsilon}{r^{a+2}}.$$

PROOF. We form a new code C' by transforming each old string $C(x)_j \in \mathbb{Z}_r^a$ using the Hadamard code into $C'(x)_j \in \mathbb{Z}_r^{r^a}$, $C'(x) = (\langle C(x), y \rangle)_{y \in \mathbb{Z}_r^a}$. The total length of C' is mr^a. The new decoder uses the same randomness as the old one. Let $f : (\mathbb{Z}_r^a)^2 \to \mathbb{Z}_r$ be the output function of the old decoder. Fix the queries $j, k \in [m]$. We now describe the new decoding procedure.

First, if for j, k the function f is such that $\Pr_{x \sim \mathbb{Z}_r^n}[f(C(x)_j, C(x)_k) = x_i] \leq \frac{1}{r}$, then the new decoder outputs a random value in $\mathbb{Z}_r$, in which case it is as good as the old one for an average x. Consider now the case $\Pr_{x \sim \mathbb{Z}_r^n}[f(C(x)_j, C(x)_k) = x_i] = \frac{1}{r} + \eta$ for some $\eta > 0$. Then

$$\Pr_{x \sim \mathbb{Z}_r^n}[f(C(x)_j, C(x)_k) = x_i] = \frac{1}{r} + \eta \iff \sum_{\ell=1}^{r-1} \mathop{\mathbb{E}}_{x \sim \mathbb{Z}_r^n}\left[\omega_r^{\ell(f(C(x)_j, C(x)_k) - x_i)}\right] = r\eta.$$

For $\ell \in [r-1]$, define the function $h_\ell : (\mathbb{Z}_r^a)^2 \to \mathbb{C}$ by $h_\ell(a, b) = \omega_r^{\ell f(a,b)}$. Consider its Fourier transform $\widehat{h}_\ell : (\mathbb{Z}_r^a)^2 \to \mathbb{C}$. Hence we can write

$$h_\ell(x) = \sum_{S, T \in \mathbb{Z}_r^a} \widehat{h}_\ell(S, T) \omega_r^{S \cdot a + T \cdot b}$$

and then

$$r\eta = \sum_{\ell=1}^{r-1} \sum_{S, T \in \mathbb{Z}_r^a} \widehat{h}_\ell(S, T) \mathop{\mathbb{E}}_{x \sim \mathbb{Z}_r^n}\left[\omega_r^{S \cdot C(x)_j + T \cdot C(x)_k - \ell x_i}\right]$$

$$\leq \sqrt{\sum_{\ell=1}^{r-1} \sum_{S, T \in \mathbb{Z}_r^a} |\widehat{h}_\ell(S, T)|^2} \sqrt{\sum_{\ell=1}^{r-1} \sum_{S, T \in \mathbb{Z}_r^a} \left|\mathop{\mathbb{E}}_{x \sim \mathbb{Z}_r^n}\left[\omega_r^{S \cdot C(x)_j + T \cdot C(x)_k - \ell x_i}\right]\right|^2}$$

$$= \sqrt{r-1} \sqrt{\sum_{\ell=1}^{r-1} \sum_{S, T \in \mathbb{Z}_r^a} \left|\mathop{\mathbb{E}}_{x \sim \mathbb{Z}_r^n}\left[\omega_r^{S \cdot C(x)_j + T \cdot C(x)_k - \ell x_i}\right]\right|^2},$$

where we used Parseval's identity $\sum_{S, T \in \mathbb{Z}_r^a} |\widehat{h}_\ell(S, T)|^2 = 1$ for all $\ell \in [r-1]$. It then follows that there are $S_0, T_0 \in \mathbb{Z}_r^a$ and $\ell_0 \in [r-1]$ such that

$$\left|\mathop{\mathbb{E}}_{x \sim \mathbb{Z}_r^n}\left[\omega_r^{S_0 \cdot C(x)_j + T_0 \cdot C(x)_k - \ell_0 x_i}\right]\right| \geq \frac{\eta}{r^a}.$$

We now use the following useful lemma whose proof is left to the end of the section.

LEMMA 35. *Let $f : \mathbb{Z}_r^n \to \mathbb{Z}_r$ and $\alpha \in [0, 1]$. Suppose that $\left|\mathbb{E}_{x \sim \mathbb{Z}_r^n}\left[\omega_r^{f(x)}\right]\right| \geq \alpha$. Then*

$$\max_{y \in \mathbb{Z}_r} \Pr_{x \sim \mathbb{Z}_r^n}[f(x) = y] \geq \frac{1}{r} + \frac{2\alpha}{r^2}.$$

According to the above lemma, there is $y \in \mathbb{Z}_r$ such that

$$\Pr_{x \sim \mathbb{Z}_r^n}[S_0 \cdot C(x)_j + T_0 \cdot C(x)_k - y = \ell_0 x_i] \geq \frac{1}{r} + \frac{2\eta}{r^{a+2}}.$$

Since both numbers $S_0 \cdot C(x)_j$ and $T_0 \cdot C(x)_k$ are in the code C', it is possible to recover x_i (the inverse of ℓ_0 is well defined since r is prime) with just two queries and average probability at least $1/r + 2\eta/r^{a+2}$. Averaging over the classical randomness, i.e., j, k and f, gives the lemma. Finally, the c-smoothness of C translates into cr^a-smoothness of C'. □

We now get the following main theorem.

THEOREM 36. *Let $r \geq 2$ be prime. A classical 2-server* PIR *scheme with alphabet $\mathbb{Z}_r$, query size t, answer size a, and recovery probability $1/r + \varepsilon$ satisfies $t \geq \Omega\big((\varepsilon^2 n/r^{4a+6} - a)/\log r\big)$.*

PROOF. By Lemma 33, there is a $(2, 3, \varepsilon)$-smooth code $C : \mathbb{Z}_r^n \to (\mathbb{Z}_r^a)^m$ with $m \leq 6r^t$. Lemma 34 allows us to transform C into a $(2, 3r^a, 2\varepsilon/r^{a+2})$-smooth code $C' : \mathbb{Z}_r^n \to \mathbb{Z}_r^{mr^a}$. Using Theorem 31 on C' (note the theorem can be applied directly to a smooth code that is good on average) yields

$$mr^a \geq 2^{\Omega(\varepsilon^2 n/r^{4a+6})},$$

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

PROOF OF Lemma 35. First notice that

$$\left| \underset{x \sim \mathbb{Z}_r^n}{\mathbb{E}} \big[\omega_r^{f(x)} \big] \right| = \left| \sum_{k \in \mathbb{Z}_r} \underset{x \sim \mathbb{Z}_r^n}{\Pr}[f(x) = k] \omega_r^k \right|,$$

hence, for any $\beta \in \mathbb{R}$,

$$\sum_{k \in \mathbb{Z}_r} \left| \underset{x \sim \mathbb{Z}_r^n}{\Pr}[f(x) = k] - \beta \right| = \sum_{k \in \mathbb{Z}_r} \left| \underset{x \sim \mathbb{Z}_r^n}{\Pr}[f(x) = k] \omega_r^k - \beta \omega_r^k \right|$$

$$\geq \left| \sum_{k \in \mathbb{Z}_r} \left(\underset{x \sim \mathbb{Z}_r^n}{\Pr}[f(x) = k] \omega_r^k - \beta \omega_r^k \right) \right| = \left| \underset{x \sim \mathbb{Z}_r^n}{\mathbb{E}} \big[\omega_r^{f(x)} \big] \right| \geq \alpha.$$

Therefore, there exists $y \in \mathbb{Z}_r$ such that

$$\left| \underset{x \sim \mathbb{Z}_r^n}{\Pr}[f(x) = y] - \beta \right| \geq \frac{\alpha}{r}.$$

Take $\beta = \frac{1}{r} - \alpha \frac{r-2}{r^2}$. Then either $\Pr_{x \sim \mathbb{Z}_r^n}[f(x) = y] \geq \beta + \frac{\alpha}{r} = \frac{1}{r} + \frac{2\alpha}{r^2}$, and so the lemma follows, or $\Pr_{x \sim \mathbb{Z}_r^n}[f(x) = y] \leq \beta - \frac{\alpha}{r} = \frac{1}{r} - 2\alpha \frac{r-1}{r^2}$. In the second case,

$$\sum_{k \in \mathbb{Z}_r} \underset{x \sim \mathbb{Z}_r^n}{\Pr}[f(x) = k] = 1 \implies \sum_{k \neq y} \underset{x \sim \mathbb{Z}_r^n}{\Pr}[f(x) = k] \geq 1 - \frac{1}{r} + 2\alpha \frac{r-1}{r^2},$$

and so there exists $z \in \mathbb{Z}_r$ such that

$$\underset{x \sim \mathbb{Z}_r^n}{\Pr}[f(x) = z] \geq \frac{1}{r-1} - \frac{1}{r(r-1)} + \frac{2\alpha}{r^2} = \frac{1}{r} + \frac{2\alpha}{r^2}. \qquad \Box$$

ACKNOWLEDGMENTS

REFERENCES

[1] Scott Aaronson. 2018. PDQP/qpoly=All. *arXiv preprint arXiv:1805.08577* (2018). https://doi.org/10.48550/arXiv.1805.08577

[2] Kook Jin Ahn, Sudipto Guha, and Andrew McGregor. 2012. Graph sketches: sparsification, spanners, and subgraphs. In *Proceedings of the 31st ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems* (Scottsdale, Arizona, USA) *(PODS '12)*. Association for Computing Machinery, New York, NY, USA, 5–14. https://doi.org/10.1145/2213556.2213560

[3] Noga Alon, Yossi Matias, and Mario Szegedy. 1999. The Space Complexity of Approximating the Frequency Moments. *J. Comput. System Sci.* 58, 1 (1999), 137–147. https://doi.org/10.1006/jcss.1997.1545

[4] Sepehr Assadi and Aditi Dudeja. 2021. A Simple Semi-Streaming Algorithm for Global Minimum Cuts. In *SIAM Symposium on Simplicity in Algorithms*. https://api.semanticscholar.org/CorpusID:232264809

[5] Keith Ball, Eric A. Carlen, and Elliott H. Lieb. 1994. Sharp uniform convexity and smoothness inequalities for trace norms. *Inventiones mathematicae* 115, 1 (01 Dec 1994), 463–482. https://doi.org/10.1007/BF01231769

[6] Ziv Bar-Yossef, T. S. Jayram, and Iordanis Kerenidis. 2004. Exponential separation of quantum and classical one-way communication complexity. In *Proceedings of the Thirty-Sixth Annual ACM Symposium on Theory of Computing* (Chicago, IL, USA) *(STOC '04)*. Association for Computing Machinery, New York, NY, USA, 128–137. https://doi.org/10.1145/1007352.1007379

[7] Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf. 2001. Quantum lower bounds by polynomials. *J. ACM* 48, 4 (jul 2001), 778–797. https://doi.org/10.1145/502090.502097

[8] William Beckner. 1975. Inequalities in Fourier Analysis. *Annals of Mathematics* 102, 1 (1975), 159–182. http://www.jstor.org/stable/1970980

[9] Richard Beigel, Lance Fortnow, and William Gasarch. 2006. A tight lower bound for restricted PIR protocols. *computational complexity* 15, 1 (01 May 2006), 82–91. https://doi.org/10.1007/s00037-006-0208-3

[10] Avraham Ben-Aroya, Oded Regev, and Ronald de Wolf. 2008. A Hypercontractive Inequality for Matrix-Valued Functions with Applications to Quantum Computing and LDCs. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*. 477–486. https://doi.org/10.1109/FOCS.2008.45

[11] Arnab Bhattacharyya, Zeev Dvir, Amir Shpilka, and Shubhangi Saraf. 2011. Tight Lower Bounds for 2-query LCCs over Finite Fields. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*. 638–647. https://doi.org/10.1109/FOCS.2011.28

[12] Aline Bonami. 1970. Étude des coefficients de Fourier des fonctions de $L^p(G)$. *Annales de l'Institut Fourier* 20, 2 (1970), 335–402. https://doi.org/10.5802/aif.357

[13] Jop Briët, Zeev Dvir, and Sivakanth Gopi. 2017. Outlaw Distributions and Locally Decodable Codes. In *8th Innovations in Theoretical Computer Science Conference (ITCS 2017) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 67)*, Christos H. Papadimitriou (Ed.). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 20:1–20:19. https://doi.org/10.4230/LIPIcs.ITCS.2017.20

[14] Eric A. Carlen and Elliott H. Lieb. 1993. Optimal hypercontractivity for Fermi fields and related non-commutative integration inequalities. *Communications in Mathematical Physics* 155, 1 (01 Jul 1993), 27–46. https://doi.org/10.1007/BF02100048

[15] Victoria M. Chayes. 2020. Matrix Rearrangement Inequalities Revisited. *arXiv preprint arXiv:2009.04032* (2020). https://doi.org/10.48550/arXiv.2009.04032

[16] Lijie Chen, Gillat Kol, Dmitry Paramonov, Raghuvansh R. Saxena, Zhao Song, and Huacheng Yu. 2021. Almost optimal super-constant-pass streaming lower bounds for reachability. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing* (Virtual, Italy) *(STOC 2021)*. Association for Computing Machinery, New York, NY, USA, 570–583. https://doi.org/10.1145/3406325.3451038

[17] Victor Chen, Elena Grigorescu, and Ronald de Wolf. 2010. Efficient and Error-Correcting Data Structures for Membership and Polynomial Evaluation. In *27th International Symposium on Theoretical Aspects of Computer Science - STACS 2010 (Proceedings of the 27th Annual Symposium on the Theoretical Aspects of Computer Science)*, Jean-Yves Marion and Thomas Schwentick (Eds.). Inria Nancy Grand Est & Loria, Nancy, France, 203–214. https://inria.hal.science/inria-00455571

[18] Benny Chor, Eyal Kushilevitz, Oded Goldreich, and Madhu Sudan. 1998. Private information retrieval. *J. ACM* 45, 6 (nov 1998), 965–981. https://doi.org/10.1145/293347.293350

[19] Chi-Ning Chou, Alexander Golovnev, Madhu Sudan, Ameya Velingker, and Santhoshini Velusamy. 2022. Linear space streaming lower bounds for approximating CSPs. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing* (Rome, Italy) *(STOC 2022)*. Association for Computing Machinery, New York, NY, USA, 275–288. https://doi.org/10.1145/3519935.3519983

[20] Chi-Ning Chou, Alexander Golovnev, Madhu Sudan, and Santhoshini Velusamy. 2022. Approximability of all finite CSPs with linear sketches. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*. 1197–1208. https://doi.org/10.1109/FOCS52979.2021.00117

[21] Chi-Ning Chou, Alexander Golovnev, and Santhoshini Velusamy. 2020. Optimal Streaming Approximations for all Boolean Max-2CSPs and Max-ksat. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*. 330–341. https://doi.org/10.1109/FOCS46700.2020.00039

[22] William J Davis, D.J.H Garling, and Nicole Tomczak-Jaegermann. 1984. The complex convexity of quasi-normed linear spaces. *Journal of Functional Analysis* 55, 1 (1984), 110–150. https://doi.org/10.1016/0022-1236(84)90021-1

[23] Irit Dinur and Samuel Safra. 2005. On the Hardness of Approximating Minimum Vertex Cover. *Annals of Mathematics* 162, 1 (2005), 439–485. http://www.jstor.org/stable/3597377

[24] João F. Doriguello and Ashley Montanaro. 2020. Exponential Quantum Communication Reductions from Generalizations of the Boolean Hidden Matching Problem. In *15th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2020) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 158)*, Steven T. Flammia (Ed.). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 1:1–1:16. https://doi.org/10.4230/LIPIcs.TQC.2020.1

[25] John C. Duchi, Shai Shalev-Shwartz, Yoram Singer, and Ambuj Tewari. 2010. Composite objective mirror descent. In *COLT 2010 - The 23rd Conference on Learning Theory*. 14–26. https://www.learningtheory.org/colt2010/papers/057Duchi.pdf

[26] Zeev Dvir. 2011. On Matrix Rigidity and Locally Self-correctable Codes. *computational complexity* 20, 2 (01 Jun 2011), 367–388. https://doi.org/10.1007/s00037-011-0009-1

[27] Zeev Dvir and Sivakanth Gopi. 2016. 2-Server PIR with Subpolynomial Communication. *J. ACM* 63, 4, Article 39 (sep 2016), 15 pages. https://doi.org/10.1145/2968443

[28] Zeev Dvir and Amir Shpilka. 2005. Locally decodable codes with 2 queries and polynomial identity testing for depth 3 circuits. In *Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing* (Baltimore, MD, USA) *(STOC '05)*. Association for Computing Machinery, New York, NY, USA, 592–601. https://doi.org/10.1145/1060590.1060678

[29] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Michael Sipser. 1998. Limit on the Speed of Quantum Computation in Determining Parity. *Phys. Rev. Lett.* 81 (Dec 1998), 5442–5444. Issue 24. https://doi.org/10.1103/PhysRevLett.81.5442

[30] Dmitry Gavinsky, Julia Kempe, Iordanis Kerenidis, Ran Raz, and Ronald de Wolf. 2007. Exponential separations for one-way quantum communication complexity, with applications to cryptography. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing* (San Diego, California, USA) *(STOC '07)*. Association for Computing Machinery, New York, NY, USA, 516–525. https://doi.org/10.1145/1250790.1250866

[31] Ashish Goel, Michael Kapralov, and Sanjeev Khanna. 2012. On the communication and streaming complexity of maximum bipartite matching. In *Proceedings of the 2012 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 468–485. https://doi.org/10.1137/1.9781611973099.41

[32] O. Goldreich, H. Karloff, L.J. Schulman, and L. Trevisan. 2002. Lower bounds for linear locally decodable codes and private information retrieval. In *Proceedings 17th IEEE Annual Conference on Computational Complexity*. 175–183. https://doi.org/10.1109/CCC.2002.1004353

[33] Oded Goldreich, Howard Karloff, Leonard J. Schulman, and Luca Trevisan. 2006. Lower bounds for linear locally decodable codes and private information retrieval. *computational complexity* 15, 3 (01 Oct 2006), 263–296. https://doi.org/10.1007/s00037-006-0216-3

[34] Sivakanth Gopi. 2018. *Locality in coding theory*. Ph.D. Dissertation. Princeton University. https://www.proquest.com/dissertations-theses/locality-coding-theory/docview/2128009409/se-2?accountid=14496

[35] Leonard Gross. 1975. Logarithmic Sobolev Inequalities. *American Journal of Mathematics* 97, 4 (1975), 1061–1083. http://www.jstor.org/stable/2373688

[36] Venkatesan Guruswami and Runzhou Tao. 2019. Streaming Hardness of Unique Games. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 145)*, Dimitris Achlioptas and László A. Végh (Eds.). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 5:1–5:12. https://doi.org/10.4230/LIPIcs.APPROX-RANDOM.2019.5

[37] Venkatesan Guruswami, Ameya Velingker, and Santhoshini Velusamy. 2017. Streaming Complexity of Approximating Max 2CSP and Max Acyclic Subgraph. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2017) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 81)*, Klaus Jansen, José D. P. Rolim, David P. Williamson, and Santosh S. Vempala (Eds.). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 8:1–8:19. https://doi.org/10.4230/LIPIcs.APPROX-RANDOM.2017.8

[38] Olof Hanner. 1956. On the uniform convexity of L_p and l_p. *Arkiv för Matematik* 3, 3 (01 Feb 1956), 239–244. https://doi.org/10.1007/BF02589410

[39] G. H. Hardy, J. E. Littlewood, and G. Pólya. 1952. Inequalities. *Cambridge University Press* 10 (1952), 169.

[40] Carl W. Helstrom. 1969. Quantum detection and estimation theory. *Journal of Statistical Physics* 1, 2 (01 Jun 1969), 231–252. https://doi.org/10.1007/BF01007479

[41] Shlomo Hoory, Nathan Linial, and Avi Wigderson. 2006. Expander graphs and their applications. *Bull. Amer. Math. Soc.* 43, 4 (2006), 439–561. https://ams.org/journals/bull/2006-43-04/S0273-0979-06-01126-8/

[42] Steven R. Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. 2020. Time-uniform Chernoff bounds via nonnegative supermartingales. *Probability Surveys* 17 (2020), 257 – 317. https://doi.org/10.1214/18-PS321

[43] Yuval Ishai and Eyal Kushilevitz. 2004. On the Hardness of Information-Theoretic Multiparty Computation. In *Advances in Cryptology - EUROCRYPT 2004*, Christian Cachin and Jan L. Camenisch (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 439–455. https://doi.org/10.1007/978-3-540-24676-3_26

[44] J. Kahn, G. Kalai, and N. Linial. 1988. The influence of variables on Boolean functions. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science (SFCS '88)*. IEEE Computer Society, USA, 68–80. https://doi.org/10.1109/SFCS.1988.21923

[45] John Kallaugher and Ojas Parekh. 2022. The Quantum and Classical Streaming Complexity of Quantum and Classical Max-Cut. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*. 498–506. https://doi.org/10.1109/FOCS54457.2022.00054

[46] Michael Kapralov, Sanjeev Khanna, and Madhu Sudan. 2015. Streaming Lower Bounds for Approximating MAX-CUT. In *Proceedings of the 2015 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1263–1282. https://doi.org/10.1137/1.9781611973730.84

[47] Michael Kapralov, Sanjeev Khanna, Madhu Sudan, and Ameya Velingker. 2017. $(1 + \Omega(1))$-approximation to MAX-CUT requires linear space. In *Proceedings of the 2017 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1703–1722. https://doi.org/10.1137/1.9781611974782.112

[48] Michael Kapralov and Dmitry Krachun. 2019. An optimal space lower bound for approximating MAX-CUT. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* (Phoenix, AZ, USA) *(STOC 2019)*. Association for Computing Machinery, New York, NY, USA, 277–288. https://doi.org/10.1145/3313276.3316364

[49] Jonathan Katz and Luca Trevisan. 2000. On the efficiency of local decoding procedures for error-correcting codes. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing* (Portland, Oregon, USA) *(STOC '00)*. Association for Computing Machinery, New York, NY, USA, 80–86. https://doi.org/10.1145/335305.335315

[50] Iordanis Kerenidis and Ronald de Wolf. 2003. Exponential lower bound for 2-query locally decodable codes via a quantum argument. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing* (San Diego, CA, USA) *(STOC '03)*. Association for Computing Machinery, New York, NY, USA, 106–115. https://doi.org/10.1145/780542.780560

[51] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. 2007. Optimal Inapproximability Results for MAX-CUT and Other 2-Variable CSPs? *SIAM J. Comput.* 37, 1 (2007), 319–357. https://doi.org/10.1137/S0097539705447372

[52] R. Latała and K. Oleszkiewicz. 2000. Between Sobolev and Poincaré. In *Geometric Aspects of Functional Analysis: Israel Seminar 1996–2000*, Vitali D. Milman and Gideon Schechtman (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 147–168. https://doi.org/10.1007/BFb0107213

[53] James R. Lee and Assaf Naor. 2004. Embedding the diamond graph in L_p and dimension reduction in L_1. *Geometric & Functional Analysis GAFA* 14, 4 (01 Aug 2004), 745–747. https://doi.org/10.1007/s00039-004-0473-8

[54] Nathan Linial, Yishay Mansour, and Noam Nisan. 1993. Constant depth circuits, Fourier transform, and learnability. *J. ACM* 40, 3 (jul 1993), 607–620. https://doi.org/10.1145/174130.174138

[55] Ashley Montanaro. 2011. A new exponential separation between quantum and classical one-way communication complexity. *Quantum Info. Comput.* 11, 7–8 (jul 2011), 574–591. https://dl.acm.org/doi/abs/10.5555/2230916.2230919

[56] Ashley Montanaro. 2012. Some applications of hypercontractive inequalities in quantum information theory. *J. Math. Phys.* 53, 12 (12 2012), 122206. https://doi.org/10.1063/1.4769269

[57] Assaf Naor. 2017. A Spectral Gap Precludes Low-Dimensional Embeddings. In *33rd International Symposium on Computational Geometry (SoCG 2017) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 77)*, Boris Aronov and Matthew J. Katz (Eds.). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 50:1–50:16. https://doi.org/10.4230/LIPIcs.SoCG.2017.50

[58] Ilan Newman. 1991. Private vs. common random bits in communication complexity. *Inform. Process. Lett.* 39, 2 (1991), 67–71. https://doi.org/10.1016/0020-0190(91)90157-D

[59] Kenji Obata. 2002. Optimal Lower Bounds for 2-Query Locally Decodable Linear Codes. In *Randomization and Approximation Techniques in Computer Science*, José D. P. Rolim and Salil Vadhan (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 39–50. https://doi.org/10.1007/3-540-45726-7_4

[60] Ryan O'Donnell. 2014. *Analysis of Boolean functions*. Cambridge University Press. https://doi.org/10.1017/CBO9781139814782

[61] Alessandro Panconesi and Aravind Srinivasan. 1997. Randomized Distributed Edge Coloring via an Extension of the Chernoff–Hoeffding Bounds. *SIAM J. Comput.* 26, 2 (1997), 350–368. https://doi.org/10.1137/S0097539793250767

[62] Iosif Pinelis. 1994. Optimum Bounds for the Distributions of Martingales in Banach Spaces. *The Annals of Probability* 22, 4 (1994), 1679–1706. http://www.jstor.org/stable/2244912

[63] Alexander A. Razborov and Sergey Yekhanin. 2006. An $\Omega(n^{1/3})$ Lower Bound for Bilinear Group Based Private Information Retrieval. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*. 739–748. https://doi.org/10.1109/FOCS.2006.10

[64] Yaoyun Shi, Xiaodi Wu, and Wei Yu. 2012. Limits of quantum one-way communication by matrix Hypercontractive Inequality. https://cs.umd.edu/~xwu/papers/GHM_v4.pdf

[65] Nicole Tomczak-Jaegermann. 1974. The moduli of smoothness and convexity and the Rademacher averages of the trace classes S_p $(1 \le p < \infty)$. *Studia Mathematica* 50, 2 (1974), 163–182. https://eudml.org/doc/urn:eudml:doc:217886

[66] Luca Trevisan. 2004. Some applications of coding theory in computational complexity. *arXiv preprint cs/0409044* (2004). https://doi.org/10.48550/arXiv.cs/0409044

[67] Joel A. Tropp. 2015. An Introduction to Matrix Concentration Inequalities. *Foundations and Trends® in Machine Learning* 8, 1-2 (2015), 1–230. https://doi.org/10.1561/2200000048

[68] Elad Verbin and Wei Yu. 2011. The Streaming Complexity of Cycle Counting, Sorting By Reversals, and Other Problems. In *Proceedings of the 2011 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 11–25. https://doi.org/10.1137/1.9781611973082.2

[69] John Watrous. 2018. *The theory of quantum information*. Cambridge University Press. https://doi.org/10.1017/9781316848142

[70] Stephanie Wehner and Ronald de Wolf. 2005. Improved Lower Bounds for Locally Decodable Codes and Private Information Retrieval. In *Automata, Languages and Programming*, Luís Caires, Giuseppe F. Italiano, Luís Monteiro, Catuscia Palamidessi, and Moti Yung (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 1424–1436. https://doi.org/10.1007/11523468_115

[71] Ronald de Wolf. 2008. A brief introduction to Fourier analysis on the Boolean cube. *Theory of Computing* (2008), 1–20. https://theoryofcomputing.org/articles/gs001/

[72] Paweł Wolff. 2007. Hypercontractivity of simple random variables. *Studia Mathematica* 3, 180 (2007), 219–236. https://eudml.org/doc/284523

[73] Sergey Yekhanin. 2012. Locally Decodable Codes. *Foundations and Trends® in Theoretical Computer Science* 6, 3 (2012), 139–255. https://doi.org/10.1561/0400000030

[74] Kevin Yeo. 2023. Lower Bounds for (Batch) PIR with Private Preprocessing. In *Advances in Cryptology – EUROCRYPT 2023*, Carmit Hazay and Martijn Stam (Eds.). Springer Nature Switzerland, Cham, 518–550. https://doi.org/10.1007/978-3-031-30545-0_18

A CLASSICAL HIDDEN HYPERMATCHING LOWER BOUND

The general idea behind the proof of Theorem 18 is already well established and is a simple generalization of results from [24, 30, 36, 68]. We shall need the following well-known fact and a generalization of the KKL inequality [44].

FACT 37. *Given just one sample, the best success probability in distinguishing between two probability distributions p and q is $\frac{1}{2} + \frac{1}{4}\|p - q\|_{\text{tvd}}$.*

LEMMA 38 (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 $\delta \in [0, 1/r]$,*

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

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

THEOREM 39. *Any classical protocol that achieves advantage $\varepsilon > 0$ for the r-$\text{HH}(\alpha, t, n)$ problem with $t \ge 2$ and $\alpha \le 1/2$ needs $\Omega(r^{-1}(\varepsilon^4/\alpha)^{1/t}(n/t)^{1-1/t})$ 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 $\mathcal{D} \triangleq \frac{1}{2}\mathcal{D}^{\text{YES}} + \frac{1}{2}\mathcal{D}^{\text{NO}}$, i.e., with probability $1/2$ is comes from $\mathcal{D}^{\text{YES}}$, and with probability $1/2$ it comes from $\mathcal{D}^{\text{NO}}$.

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 $C \triangleq c - \log(1/\varepsilon)$ 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

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

Given one sample of $w \in \mathbb{Z}_r^{\alpha n/t}$, Bob must decide whether it came from $\mathcal{D}^{\text{YES}}$ (the distribution MX) or from $\mathcal{D}^{\text{NO}}$ (the uniform distribution U). According to Fact 37, the advantage of any classical protocol in distinguishing between p_M and U is upper bounded by $\frac{1}{4}\|p_M - U\|_{\text{tvd}}$. We prove in Theorem 40 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.,

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

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 $\|p_M - U\|_{\text{tvd}} \leq \varepsilon$), and so $c = \Omega(r^{-1}(\varepsilon^4/\alpha)^{1/t}(n/t)^{1-1/t})$. □

THEOREM 40. *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*

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

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:

$$\widehat{p}_M(V) = \frac{1}{r^{\alpha n/t}} \sum_{z \in \mathbb{Z}_r^n} p_M(z)\omega_r^{-V \cdot z} = \frac{1}{|A|r^{\alpha n/t}} \sum_{z \in \mathbb{Z}_r^n} |\{x \in A : Mx = z\}| \cdot \omega_r^{-V \cdot z}$$

$$= \frac{1}{|A|r^{\alpha n/t}} \sum_{k=0}^{r-1} |\{x \in A : (Mx) \cdot V = k\}| \cdot \omega_r^{-k}$$

$$= \frac{1}{|A|r^{\alpha n/t}} \sum_{k=0}^{r-1} |\{x \in A : x \cdot (M^\top V) = k\}| \cdot \omega_r^{-k}$$

$$= \frac{1}{|A|r^{\alpha n/t}} \sum_{x \in A} \omega_r^{-x \cdot (M^\top V)}$$

$$= \frac{r^n}{|A|r^{\alpha n/t}} \widehat{f}(M^\top V).$$

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

$$\mathbb{E}_{M \sim \mathcal{M}_{t,n}^\alpha}[\|p_M - U\|_{\text{tvd}}^2] \leq r^{2\alpha n/t} \mathbb{E}_{M \sim \mathcal{M}_{t,n}^\alpha}[\|p_M - U\|_2^2] \qquad \text{(Cauchy-Schwarz)}$$

$$= r^{2\alpha n/t} \mathop{\mathbb{E}}_{M \sim \mathcal{M}_{t,n}^{\alpha}} \left[\sum_{V \in \mathbb{Z}_r^{\alpha n/t} \setminus \{0^{\alpha n/t}\}} |\widehat{p}_M(V)|^2 \right] \qquad \text{(Parseval's identity)}$$

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

Note that there is at most one $V \in \mathbb{Z}_r^{\alpha n/t}$ such that $S = M^\top V$ for a given $S \in \mathbb{Z}_r^n$ (and that the only V that makes $M^\top V = 0^n$ is $V = 0^{\alpha n/t}$). This allows us to transform the expectation over hypermatchings into a probability,

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

$$= \frac{r^{2n}}{|A|^2} \sum_{S \in \mathbb{Z}_r^n \setminus \{0^n\}} \mathop{\Pr}_{M \sim \mathcal{M}_{t,n}^{\alpha}} [\exists V \in \mathbb{Z}_r^{\alpha n/t} : M^\top V = S] \cdot |\widehat{f}(S)|^2.$$

Now observe that $\Pr_{M \sim \mathcal{M}_{t,n}^{\alpha}}[\exists V \in \mathbb{Z}_r^{\alpha n/t} : M^\top V = S]$ is exactly the probability from Lemma 17, i.e., given $S \in \mathbb{Z}_r^n$ with $k_j \triangleq \frac{1}{t} \cdot |\{i \in [n] : S_i = j\}| \in \mathbb{Z}$ for $j \in [r-1]$ (the number of entries from S equal to $j \neq 0$ must be a multiple of t), and defining $k \triangleq \sum_{j=1}^{r-1} k_j$, then

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

and so

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

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 38, 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 [64, 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 ($2rc \leq k \leq \alpha n/t$): first note that the function $g(k) \triangleq \binom{\alpha n/t}{k} / \binom{n}{kt}$ is decreasing in the interval $1 \leq k \leq \alpha n/t$ (since $\alpha \leq 1/2$). Hence, by using Parseval's identity $\sum_{S \in \mathbb{Z}_r^n} |\widehat{f}(S)|^2 = |A|/r^n$

and the inequality $\binom{q}{s}\binom{\ell q}{\ell s}^{-1} \leq (\frac{s}{q})^{(\ell-1)s}$ (for $q = n/t, s = 2m, \ell = t$) in order to bound $g(2rc)$,

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

where the last step used that $2(1 - 1/t)c \geq 1 \implies \varepsilon^{2(1-1/t)c} \leq \varepsilon$ and picked γ sufficiently small.

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