An Exponential Lower Bound for Linear 3-Query Locally Correctable Codes

Pravesh K. Kothari
praveshk@cs.cmu.edu
Carnegie Mellon University

Peter Manohar
pmanohar@cs.cmu.edu
Carnegie Mellon University

November 1, 2023

Abstract

We prove that the blocklength n of a linear 3-query locally correctable code (LCC) $\mathcal{L} \colon \mathbb{F}^k \to \mathbb{F}^n$ with distance δ must be at least $n \geq 2^{\Omega\left(\left(\frac{\delta^2 k}{(|\mathbb{F}|-1)^2}\right)^{1/8}\right)}$. In particular, the blocklength of a linear 3-query LCC with constant distance over any small field grows *exponentially* with k. This improves on the best prior lower bound of $n \geq \tilde{\Omega}(k^3)$ [AGKM23], which holds even for the weaker setting of 3-query locally *decodable* codes (LDCs), and comes close to matching the best-known construction of 3-query LCCs based on binary Reed–Muller codes, which achieve $n \leq 2^{O(k^{1/2})}$. Because there is a 3-query LDC with a strictly subexponential blocklength [Yek08, Efr09], as a corollary we obtain the first strong separation between q-query LCCs and LDCs for any constant $q \geq 3$.

Our proof is based on a new upgrade of the method of spectral refutations via *Kikuchi matrices* developed in recent works [GKM22, HKM23, AGKM23] that reduces establishing (non-)existence of combinatorial objects to proving unsatisfiability of associated XOR instances. Our key conceptual idea is to apply this method with XOR instances obtained via *long-chain derivations* — a structured variant of low-width resolution for XOR formulas from proof complexity [Gri01, Sch08].

Keywords: Locally Correctable Codes, Locally Decodable Codes, Kikuchi Matrices

Contents

1 Introduction

A *locally correctable code* (LCC) is an error correcting code that admits, in addition, a *local* correction (a.k.a. *self correction*) algorithm that can recover any symbol of the original codeword by querying only a small number of randomly chosen symbols from the received corrupted codeword. More formally, we say that a code $\mathcal{L}: \{0,1\}^k \rightarrow \{0,1\}^n$ is q-locally correctable if for any codeword x, a corruption y of x, and input $u \in [n]$, the local correction algorithm reads at most q symbols (typically a small constant such as 2 or 3) of y and recovers the bit x_u with probability $1/2 + \varepsilon$ whenever $\Delta(x, y) := |\{v \in [n] : x_v \neq y_v\}| \leq \delta n$, where δ, the "distance" of the code, and ε, the decoding accuracy, are constants. The central question about LCCs is to determine the smallest possible blocklength n as a function of the message length k for a fixed number of queries q.

Local correction was first introduced for *program checking* [BK95], and early applications utilized that Reed–Muller codes are locally correctable via polynomial interpolation. Since then, LCCs have been a mainstay in complexity and algorithmic coding theory with a long array of applications. An abridged list (the surveys [Tre04, Yek12, Dvi12] provide details) of applications includes sublinear algorithms and property testing [RS96, BLR93], probabilistically checkable proofs [ALM+98, AS98], IP=PSPACE [LFKN90, Sha90], worst-case to average-case reductions [BFNW93], constructions of explicit rigid matrices [Dvi10], and t-private information retrieval protocols [IK99, BIW10]. The existence of LCCs turns out to have natural connections to incidence geometry [Dvi12], additive combinatorics [BDL13], and the theory of block designs [BIW10].

For any constant $q \in \mathbb{N}$, Reed–Muller codes (i.e., evaluations of $(q-1)$-degree polynomials) yield binary, linear[1] q-LCCs with a blocklength $n \leq 2^{O(k^{\frac{1}{q-1}})}$. Given their extensive applications and connections, finding LCCs of smaller blocklength has been a major project in theoretical computer science over the past three decades with some remarkable successes over the years. For example, *multiplicity codes* [KSY14] significantly beat the blocklength of Reed–Muller codes in the *super-constant* query regime. In the constant-query regime, *matching vector codes* [Efr09, Yek08] use a strictly sub-exponential (i.e., $n \leq \exp(\exp(O(\sqrt{\log k} \log \log k))))$ blocklength to obtain 3-query locally *decodable* codes — a relaxation of LCCs where the local correction property holds only for the k message bits. To sidestep the difficulty of finding more efficient LCCs, the work of [BGH+04] introduced *relaxed LCCs* that soften the local correction property and has seen exciting recent developments [GRR20, AS21, CGS20, KM23, CY23]. These successes notwithstanding, constructing better constant-query LCCs has remained a major open question (see, e.g., Chapter 8 in [Yek12]).

LCC lower bounds. The lack of progress on finding better constant-query LCCs has motivated a long investigated conjecture that Reed–Muller codes might be *optimal* constant query LCCs. The work of [KW04, GKST06] essentially confirmed this conjecture for the "base case" of $q = 2$ by proving that $n \geq 2^{\Omega(k)}$ for any two-query LCC, matching the construction of Hadamard codes, which are 2-LCCs with $n = 2^k$. For $q \geq 3$, however, only a polynomial lower bound is known. The works of [KW04, Woo07] prove that q-LCCs must have $n \geq \tilde{\Omega}(k^{1/(1-1/\lceil \frac{q}{2} \rceil)})$,[2] and for the specific case

---

[1]A code is *linear* over a field $\mathbb{F}$ if the encoding map $\mathcal{L}$ is an $\mathbb{F}$-linear map.

[2]These lower bounds all hold for non-linear codes over small (i.e., polylog(k)) size alphabets. A weaker polynomial lower bound [KT00, IS18] is known to hold for linear codes over all fields and for the specific case of $q = 3$, [Woo10]

of $q = 3$, a recent work [AGKM23] (which, like this work, is based on the *Kikuchi* matrix method) obtained a polynomial improvement on this bound, showing that $n \geq \tilde{\Omega}(k^3)$.

Limitations of prior lower bound techniques. Beyond the weakness in the quantitative results, all the above lower bounds suffer from an important inherent limitation — they all hold even for the weaker setting of locally decodable codes (LDCs). As we mentioned above, there are sub-exponential length (and thus substantially beating Reed–Muller) 3-query binary, linear codes that are locally decodable [Yek08, Efr09]. Indeed, characterizing the limitations of prior proof techniques and finding methods that could separate LCCs and LDCs itself has been a major research goal. For example, Dvir, Gopi, Gu and Wigderson [DGGW19] formalize the limitations of prior lower bound techniques for LCCs by showing that the "random restriction" approach in [KT00] applies to a more general setting of "spanoids" where they are, in fact, tight. On the other hand, to show a strong separation between LCCs and LDCs, Barkol, Ishai and Weinreb [BIW10] build an approach for stronger LCC lower bounds via connections to the well-studied Hamada conjecture ([Ham73], see lecture notes [Ton11]) and its generalizations in the theory of block designs, while Dvir, Saraf and Wigderson [DSW14] develop new geometric techniques to prove a slightly superquadratic lower bound for an appropriate formulation of 3-LCCs over the reals.

To summarize: there is an exponential gap between best-known constructions and lower bounds for q-LCCs for $q \geq 3$. Further, the best known lower bound techniques for q-LCCs apply also to q-LDCs and thus provably cannot yield an exponential lower bound.

Our result. In this work, we prove an *exponential* lower bound for linear 3-query LCCs. We note that the best-known constructions of LCCs (and also LDCs) namely Reed–Muller codes and matching vector codes, are $\mathbb{F}_2$-linear.

Theorem 1. *Let $\mathcal{L} \colon \mathbb{F}^k \to \mathbb{F}^n$ be a linear $(3, \delta, \varepsilon)$-LCC. Then, $n \geq 2^{\Omega((\delta^2 k/(|\mathbb{F}|-1)^2)^{1/8})}$. In particular, if $\mathcal{L} \colon \mathbb{F}_2^k \to \mathbb{F}_2^n$ is a $(3, \delta, \varepsilon)$-LCC where δ is constant, then $n \geq 2^{\Omega(k^{1/8})}$.*

Theorem 1 improves on the prior best lower bound of $n \geq \tilde{\Omega}(k^3)$ [AGKM23] and comes close to matching the blocklength $n = \exp(O(\sqrt{k}))$ of 3-query LCCs based on Reed–Muller codes; in Section 9, we comment on potential strengthenings of our argument to come closer and even match (up to constants in the exponent) the $\exp(O(\sqrt{k}))$ bound.

Theorem 1 also yields the first strong *separation* between 3-LCCs and 3-LDCs. No such separation was known for q-LDCs and q-LCCs for any constant $q \geq 3$.[3] In particular, Theorem 1 implies that matching vector codes that yield linear 3-LDCs over $\mathbb{F}_2$ of sub-exponential blocklength, such as the codes in [Yek08, Efr09], *cannot* admit a local correction algorithm, answering a question of Yekhanin (see Chapter 8 in [Yek12]).

Our proof is based on the method of spectral refutation via Kikuchi matrices developed in prior works [GKM22, HKM23, AGKM23]. The key idea in this method is to associate the existence of a combinatorial object (e.g., a 3-LCC) to the satisfiability of a family of XOR formulas and find a *spectral* refutation (i.e., certificate of unsatisfiability) for a randomly chosen member of the family.

---

shows a lower bound of $\Omega(k^2)$ for linear 3-LDCs over all fields.

[3]The work of [BGT17] shows a separation between 2-LCCs and 2-LDCs over poly(n)-sized alphabets. For 2-LCCs on small alphabets, a strong separation cannot exist, e.g., on $\mathbb{F}_2$, the Hadamard code gives both an essentially optimal 2-LCC and 2-LDC.

Our key new conceptual idea is to apply an appropriate version of the Kikuchi matrix method to XOR formulas obtained by *long chain derivations* — a structured variant of low-width XOR resolution refutations in proof complexity [Gri01, Sch08] — to the naive XOR instances obtained from the query sets of a purported linear 3-LCC. These new XOR formulas allow us to utilize the additional structural in 3-LCCs and, in particular, significantly surpass the cubic lower bound [AGKM23] for 3-LDCs that also used the Kikuchi matrix method. We discuss the new challenges that arise in analyzing spectral refutations of XOR instances produced by such long chain derivations and our technical ideas for handling them in Sections 3 and 4.

1.1 Roadmap

The rest of the paper is organized as follows. First, in Section 2, we introduce some notation and recall basic facts about LCCs that we shall use in the proof. Then, in Section 3, we give a detailed overview of the proof. In Section 4, we give an essentially complete proof of a new lower bound of $n \geq \tilde{\Omega}(k^4)$ for binary linear 3-LCCs as a warmup. Following the warmup, in Sections 5 to 8 we prove Theorem 1 for binary 3-LCCs, i.e., when $\mathbb{F} = \mathbb{F}_2$; we handle the case of arbitrary finite fields in Appendix A. Finally, in Section 9 we conclude with some remarks on the proof of Theorem 1, possible strengthenings, and extensions.

2 Preliminaries

2.1 Basic notation

We let $[n]$ denote the set $\{1, \ldots, n\}$. For two subsets $S, T \subseteq [n]$, we let $S \oplus T$ denote the symmetric difference of S and T, i.e., $S \oplus T := \{i : (i \in S \land i \notin T) \lor (i \notin S \land i \in T)\}$. For a natural number $t \in \mathbb{N}$, we let $\binom{[n]}{t}$ be the collection of subsets of $[n]$ of size exactly t. Given variables $x_1, \ldots, x_n$ and a subset $C \subseteq [n]$, we let $x_C := \prod_{v \in C} x_v$.

For a rectangular matrix $A \in \mathbb{R}^{m \times n}$, we let $\|A\|_2 =:= \max_{x \in \mathbb{R}^m, y \in \mathbb{R}^n : \|x\|_2 = \|y\|_2 = 1} x^\top A y$ denote the spectral norm of A, and $\|A\|_{\infty \to 1} := \max_{x \in \{-1,1\}^m, y \in \{-1,1\}^n} x^\top A y$. We note that $\|A\|_{\infty \to 1} \leq \sqrt{nm} \|A\|_2$.

2.2 XOR formulas

An XOR instance ψ on n variables $x_1, x_2, \ldots, x_n$ taking values in $\{-1, 1\}$ is a collection of constraints of the form $\{x_C = b_C\}$ where $C \in \mathcal{H}$ where $\mathcal{H} \subseteq 2^{[n]}$ is the *constraint hypergraph*. The *arity* of a constraint $\{x_C = b_C\}$ equals $|C|$. The arity of ψ is the maximum arity of any constraint in it. The XOR formula associated with ψ is the expression $\psi(x) = \sum_{C \in \mathcal{H}} b_C x_C$ seen as a polynomial over $\{-1, 1\}^n$. Notice that $\psi(x) = |\mathcal{H}|$ if x satisfies all the constraints of ψ and in general evaluates to (number of constraints satisfied by x) - (number of constraints violated by x). The *value* $\mathrm{val}(\psi)$ of a XOR instance ψ (or, of the associated formula $\psi(x)$) is the maximum of $\psi(x)$ as x ranges over $\{-1, 1\}^n$. More generally, for a function $f(x)$, we shall let $\mathrm{val}(f) := \max_{x \in \{-1,1\}^n} f(x)$.

2.3 Locally correctable codes

We refer the reader to the survey [Yek12] for background.

Definition 2.1 (Locally correctable code). A map $\mathcal{L}\colon \mathbb{F}^k \to \mathbb{F}^n$ is a (q, δ, ε)-locally correctable code if there exists a randomized decoding algorithm $\mathrm{Dec}(\cdot)$ that takes input an oracle access to some $y \in \mathbb{F}^n$ and a $u \in [n]$, (1) makes at most q queries to the string y, and (2) for all $b \in \mathbb{F}^k$, $u \in [n]$, and all $y \in \mathbb{F}^n$ such that $\Delta(y, \mathcal{L}(b)) \leq \delta n$, $\Pr[\mathrm{Dec}^y(u) = \mathcal{L}(b)_u] \geq \frac{1}{2} + \varepsilon$. Here, $\Delta(x, y)$ denotes the Hamming distance between x and y, i.e., the number of indices $v \in [n]$ where $x_v \neq y_v$. We will use q-LCCs to denote (q, δ, ϵ)-LCCs where δ, ϵ are some fixed small constants.

$\mathcal{L}$ is *linear* if the map $\mathcal{L}$ is a linear map. We note that for linear codes, $k = \dim(\mathcal{V})$, where $\mathcal{V}$ is the image of $\mathbb{F}^k$ under the map $\mathcal{L}$. Without the loss of generality, a linear $\mathcal{L}$ is systematic, i.e., $\mathcal{L}(b)_i = b_i$ for $i \in [k]$. By a slight abuse of notation, we will also use $\mathcal{L}$ to denote the set of all codewords, i.e., elements in the range of the map $\mathcal{L}$.

For the Boolean case, i.e., when $\mathbb{F} = \mathbb{F}_2$, it shall be more convenient to think of the map $\mathcal{L}$ as a function from $\{-1, 1\}^k$ to $\{-1, 1\}^n$, defined via the mapping $0 \leftrightarrow 1$ and $1 \leftrightarrow -1$.

We next discuss a combinatorial characterization of locally correctable codes. To begin with, we recall basic notions about hypergraphs.

Definition 2.2. A q-uniform hypergraph $\mathcal{H}$ on vertex set $[n]$ is a collection of subsets $C \subseteq [n]$ of size q called hyperedges. We say that $\mathcal{H}$ is a *matching* if all the hyperedges in $\mathcal{H}$ are disjoint. For a subset $Q \subseteq [n]$, we define the degree of Q in $\mathcal{H}$, denoted $\deg_{\mathcal{H}}(Q)$, to be $|\{C \in \mathcal{H} : Q \subseteq C\}|$.

LCCs admit a standard combinatorial characterization (formalized in the definition below).

Definition 2.3 (Linear LCC in normal form). A linear code $\mathcal{L}\colon \mathbb{F}^k \to \mathbb{F}^n$ is (q, δ)-normally correctable if for each $u \in [n]$, there is a q-uniform hypergraph matching $\mathcal{H}_u$ with at least δn hyperedges such that for every $C = \{v_1, \ldots, v_q\} \in \mathcal{H}_u$, there are coefficients $\alpha_1, \ldots, \alpha_q \in \mathbb{F} \setminus \{0\}$ such that, for any $b \in \mathbb{F}^k$, $x = \mathcal{L}(b)$ satisfies $x_u = \alpha_1 x_{v_1} + \cdots + \alpha_q x_{v_q}$.

Fact 2.4 (Reduction to LCC normal form, Theorem 8.1 in [Dvi16]). *Let* $\mathcal{L}\colon \mathbb{F}^k \to \mathbb{F}^n$ *be a linear code that is* (q, δ, ε)-*locally correctable. Then, there is a linear code* $\mathcal{L}'\colon \mathbb{F}^k \to \mathbb{F}^{2n}$ *that is* (q, δ')-*normally correctable, with* $\delta' \geq \delta/2q$.

We note that there is slight difference in Fact 2.4 compared to Theorem 8.1 in [Dvi16]. In Fact 2.4, we require that the matchings are q-uniform and all coefficients α are nonzero, and we obtain $\delta' \geq \delta/2q$. On the other hand, [Dvi16] allows for hyperedges of size $\leq q$, i.e., some coefficients α may be zero, and obtains $\delta' \geq \delta/q$. We remark that Fact 2.4 immediately follows from [Dvi16] by "padding" the code with n 0's. This loses an additional factor of 2 in δ, but allows one to make all hyperedges have size exactly q by querying the padded 0 entries.

Finally, we recall the lower bound for linear 2-LDCs from [GKST06].

Fact 2.5 (Lemma 3.3, Claim 4.4 in [GKST06]). *Let* $\mathcal{L}\colon \mathbb{F}^k \to \mathbb{F}^n$ *be a linear map, and let* $G_1, \ldots, G_k$ *be matchings on n vertices such that for every* $b \in \mathbb{F}^k$ *and every* $i \in [k]$ *and every* $(u, v) \in G_i$, *it holds that* $x_u - x_v = b_i$, *where* $x = \mathcal{L}(b)$. *Suppose that* $\frac{1}{k} \sum_{i=1}^{k} |G_i| \geq \delta n$. *Then,* $\delta k \leq 2 \log_2 n$.

When $\mathbb{F} \neq \mathbb{F}_2$, the above lower bound, as stated, only applies to the setting where the decoder is a linear function with the added restriction that each non-zero coefficient of the linear combination is in $\{-1, 1\}$. It is not hard to remove this restriction on coefficients, but, in our setting, we obtain a better dependence on $|\mathbb{F}|$ in Theorem 1 by applying this more specialized lemma.

2.4 Concentration inequalities

We will need the following standard concentration inequalities.

Fact 2.6 (Chernoff Bound). *Let $x_1, \ldots, x_n$ be i.i.d. Bernoulli random variables with mean p, and let $\mu = pn$. Then, for any $\delta \geq 0$,*

$$\Pr[\sum_{i=1}^{n} x_i \geq (1 + \delta)\mu] \leq \exp(-\delta^2 \mu / (2 + \delta)) \ .$$

Fact 2.7 (Scalar Bernstein inequality). *Let $x_1, \ldots, x_n$ be independent mean 0 random variables satisfying $|x_i| \leq M$ almost surely for every i. Let $\sigma^2 \geq \sum_{i=1}^{n} \mathbb{E}[x_i^2]$, for every $i \in [n]$. Then, for all $t \geq 0$, it holds that*

$$\Pr[\sum_{i=1}^{n} x_i \geq t] \leq \exp\left(-\frac{\frac{1}{2}t^2}{\sigma^2 + \frac{1}{3}Mt}\right) \ .$$

We will use the following non-commutative Khintchine inequality [LP91].

Fact 2.8 (Rectangular Matrix Khintchine inequality, Theorem 4.1.1 of [Tro15]). *Let $X_1, \ldots, X_k$ be fixed $d_1 \times d_2$ matrices and $b_1, \ldots, b_k$ be i.i.d. from $\{-1, 1\}$. Let $\sigma^2 \geq \max(\|\sum_{i=1}^{k} X_i X_i^\top\|\|_2, \|\sum_{i=1}^{k} X_i^\top X_i\|\|_2)$. Then*

$$\mathbb{E}\left[\left\| \sum_{i=1}^{k} b_i X_i \right\|_2 \right] \leq \sqrt{2\sigma^2 \log(d_1 + d_2)} \ .$$

Tail Bounds for r-partite non-negative polynomials. We give an elementary proof of a concentration inequality for r-partite polynomials with non-negative coefficients. Such inequalities are the subject of the celebrated work of Kim and Vu [KV00] (with tightenings due to Schudy and Sviridenko [SS12]). For r-partite polynomials, our inequality below saves a crucial $2^{O(r)}$ factor in the estimate of the typical value when compared to a blackbox application of the above results (without which, we can only obtain a quasi-polynomial lower bound for 3-LCCs).

Lemma 2.9 (Tail Bounds from bounded expected derivatives). *Let $x = \{x_j^{(i)}\}_{\substack{1 \leq i \leq r \\ 1 \leq j \leq n}}$ be nr independent p-biased Bernoulli random variables. Let $P(x) = P(x^{(1)}, x^{(2)}, \ldots, x^{(r)})$ be a r-partite multilinear polynomial of degree $\leq r$ with nonnegative coefficients. That is, each monomial with a non-zero coefficient in P has degree at most 1 in each $x^{(i)}$ for $1 \leq i \leq r$. For $Z \in ([n] \cup \{\star\})^r$, let $\mu_Z(P)$ be the expected partial derivative of P with respect to the variables $\{x_{Z_h}^{(h)} \mid 1 \leq h \leq r, \ Z_h \neq \star\}$. Suppose that there exists a $\mu, \gamma > 0$ such that for every Z, $\mu_Z(P) \leq \mu \cdot \gamma^{|Z|}$, where $|Z|$ denotes the number of non $\star$ coordinates in Z.*
    Then, for every $\beta > 0$,

$$\Pr_y[P(y) \geq (1 + \beta)^r \mu] \leq r(n + 1)^r \alpha \ ,$$

where $\alpha = \exp\left(-\frac{\frac{1}{2}\beta^2}{2\gamma + \frac{1}{3}\gamma\beta}\right)$.

Proof. We will consider the random process that samples $x \in \{0,1\}^{nr}$ by fixing $x^{(i)}$ to a random draw from their distribution one at a time. At each step, we obtain a new polynomial of smaller degree obtained by fixing one additional set of variables to a fixed value in P. We understand how the parameters μ_Z of the polynomials so generated evolve via the Bernstein inequality Fact 2.7.

Formally, fix a $0 \le t$. Let $(Z_{t+1}, \ldots, Z_r) \in [n]^{r-t}$ be a tuple of length $r - t$. We define the quantity $\mu_{y^{(1)},\ldots,y^{(t)},Z_{t+1},\ldots,Z_r}$ to be the quantity $\mu_{Z_{t+1},\ldots,Z_r}(P_t)$ where $P_t = P(y^{(1)},\ldots,y^{(t)},x^{(t+1)},\ldots,x^{(r)})$. Here, we use the notation $y^{(i)}$ to denote sampled values for $x^{(i)}$. Note that P_t has $r - t$ "free" groups of variables $x^{(t+1)},\ldots,x^{(r)}$.

Let $t \in \{0,\ldots,r\}$. We will show that with probability at least $1 - t(n+1)^r \alpha$ over the draw of $y^{(1)},\ldots,y^{(t)}$, it holds that for every $Z_{t+1},\ldots,Z_r$ with $|Z_h| \in \{0,1\}$ for all $h = t+1,\ldots r$, we have $\mu_{y^{(1)},\ldots,y^{(t)},Z_{t+1},\ldots,Z_r} \le (1+\beta)^t \mu \cdot \gamma^{\sum_{h=t+1}^r |Z_h|}$.

We prove this by induction. The base case of $t = 0$ forms the hypothesis of the lemma. We now prove the inductive step. Let $t \ge 1$, and suppose that with probability at least $1 - (t-1)(n+1)^r \alpha$ over the draw of $y^{(1)},\ldots,y^{(t-1)}$, it holds that for tuple of length $r - t + 1$ $Z_t,\ldots,Z_r$, we have $\mu_{y^{(1)},\ldots,y^{(t-1)},Z_t,\ldots,Z_r} \le (1+\beta)^{t-1} \mu \cdot \gamma^{\sum_{h=t}^r |Z_h|}$.

Fix $Z_{t+1},\ldots,Z_r$ with $|Z_h| \in \{0,1\}$ for all $h = t+1,\ldots r$. We now show that with probability at least $1 - \alpha$ over the draw of $y^{(t)}$, it holds that $\mu_{y^{(1)},\ldots,y^{(t-1)},y^{(t)},Z_{t+1},\ldots,Z_r} \le (1+\beta)\mu'$, where $\mu' = (1+\beta)^{t-1} \mu \cdot \gamma^{\sum_{h=t+1}^r |Z_h|}$. The lemma then follows by union bound over the (crudely) at most $(n+1)^r$ choices for $Z_{t+1},\ldots,Z_r$.

For an assignment $y^{(t)}$, we have that $\mu_{y^{(1)},\ldots,y^{(t-1)},y^{(t)},Z_{t+1},\ldots,Z_r} = f(y^{(t)})$, where $f(x^{(t)}) := \sum_{u=1}^n c_u x^{(t)}$ is a linear polynomial with nonnegative coefficients $c_u := \mu_{S_1,\ldots,S_{t-1},\{u\},Z_{t+1},\ldots,Z_r}$. We note that the mean is $\mathbb{E}_{y^{(t)}}[f(x^{(t)})] = \mu_{S_1,\ldots,S_{t-1},\emptyset,Z_{t+1},\ldots,Z_r} \le \mu'$, by the induction hypothesis. We also have that $c_u = \mu_{S_1,\ldots,S_{t-1},\{u\},Z_{t+1},\ldots,Z_r} \le (1+\beta)^{t-1} \mu \cdot \gamma^{1+\sum_{h=t+1}^r |Z_h|} = \gamma\mu'$, again by the induction hypothesis.

We now bound the polynomial by using the Bernstein Inequality. Let $y'^{(t)}$ be the centered version of $y^{(t)}$, i.e., $y_u'^{(t)} = y_u^{(t)} - p$, so that $y_u'^{(t)} = 1 - p$ with probability p, and $-p$ with probability $1 - p$. Then, $\mathbb{E}[(y_u'^{(t)} c_u)^2] = c_u^2((1-p)^2 p + p^2(1-p)) \le 2pc_u^2$. Further, we observe that $|y_u'^{(t)} c_u| \le c_u \le \gamma\mu' =: M$ always holds. We also note that

$$\sigma^2 := \sum_{u=1}^n \mathbb{E}[(y_u'^{(t)} c_u)^2] \le \sum_{u=1}^n 2pc_u^2 \le 2(\max_u c_u)(p \sum_{u=1}^n c_u) \le 2M \cdot \mu_{S_1,\ldots,S_{t-1},\emptyset,Z_{t+1},\ldots,Z_r}$$

$$\le 2(\gamma\mu')\mu' = 2\gamma\mu'^2 \ .$$

Thus, by the Bernstein Inequality, we have

$$\Pr[f(y'^{(t)}) \ge \lambda] \le \exp\left(-\frac{\frac{1}{2}\lambda^2}{\sigma^2 + \frac{1}{3}M\lambda}\right) \ ,$$

and therefore

$$\Pr[f(y'^{(t)}) \ge \beta\mu'] \le \exp\left(-\frac{\frac{1}{2}\beta^2}{2\gamma + \frac{1}{3}\gamma\beta}\right) = \alpha \ .$$

6

Note that since f is linear, $f(y'^{(t)}) = f(y^{(t)}) - \mu_{S_1,\dots,S_{t-1},\emptyset,Z_{t+1},\dots,Z_r}$, and so it follows that $\Pr[f(y^{(t)}) \geq (1+\beta)\mu'] \leq \alpha$, which finishes the proof. $\qquad\square$

3   Proof overview

In this section, we will focus on the case of $\mathbb{F} = \mathbb{F}_2$ to give a high-level overview of the main ideas in the proof of Theorem 1. Without loss of generality, we can assume that $\mathcal{L}$ is a systematic linear map $\mathcal{L}: \{-1,1\}^k \to \{-1,1\}^n$, so that the first k bits in any codeword are the message bits themselves, i.e., for any $b \in \{-1,1\}^k$, $x = \mathcal{L}(b)$ satisfies $x_i = b_i$ for all $i \in [k]$. In this section and the next, we will use the notation $\gtrsim$ and $\lesssim$ to suppress a multiplicative $\mathrm{polylog}(n)$ factor.

The Kikuchi matrix method.   Our proof uses the Kikuchi matrix method developed in prior works [GKM22, HKM23, AGKM23] for finding extremal trade-offs for combinatorial structures in hypergraphs. This method works in two steps: (1) formulate a hypergraph possessing some relevant structure as a family of satisfiable XOR formulas, and, (2) construct a spectral refutation (i.e., a certificate of unsatisfiability) of a randomly chosen member of this family. The spectral refutations in the second step rely on appropriate *Kikuchi* matrices — a term that we loosely use to describe induced subgraphs of an appropriately chosen Cayley graph associated with the hypergraph. The success of the spectral refutation naturally relies on the structure of the XOR instances. The power of the method comes from the ease (at least in hindsight, given [GKM22, HKM23, AGKM23]) in identifying the relevant combinatorial structure that is sufficient for the success of the spectral refutations. This method has been used to prove Feige's conjecture [Fei08] on the hypergraph Moore bound (extremal girth vs. density trade-off) [GKM22, HKM23], improved lower bounds for 3-LDCs [AGKM23], and generalizations of Szemeredi's theorems for arithmetic progressions with restricted common differences [BC23] (which closely follows the argument in [AGKM23]).

   Our proof can be seen as an upgrade on a recent work [AGKM23] that showed a lower bound of $n \geq \tilde{\Omega}(k^3)$ on the block length n of a code of dimension k and constant distance.[4] The key conceptual idea that helps us move beyond the cubic to an exponential lower bound (a bound that provably cannot hold for 3-LDCs [Efr09, Yek08]) is a new family of XOR instances that crucially exploits the additional structure in LCCs. Our new family of XOR instances is produced by performing a certain structured variant of *low-width* resolution (well-studied in proof complexity [Gri01, Sch08]) on the "basic" family. We call this process *long chain derivations*.

   In the following, we will first recall the conceptual crux of the lower bound for q-LDCs in [AGKM23] and then use it to motivate our approach for 3-LCCs.

3.1   The naive XOR instance and LDC lower bounds

Let's first consider the case of 3-LDCs and start by recalling the combinatorial characterization (formalized as the *normal form* in Definition 2.3). A code $\mathcal{L}: \{-1,1\}^k \to \{-1,1\}^n$ is a (q,δ)-LDC if for every $1 \leq i \leq k$, there exists a q-uniform hypergraph matching H_i over $[n]$ of size δn such that for every $b \in \{-1,1\}^k$ and codeword $x = \mathcal{L}(b)$, for every $i \in [k]$ and every $C \in H_i$, it holds

---

[4]Their result extends to non-linear codes but we omit this distinction here.

that $x_C = b_i$. The combinatorial characterization above can be easily seen to be equivalent to the satisfiability of a family of q-XOR instances.

Observation 3.1 (LDCs and a Family of XOR Instances). Let $H_1, H_2, \ldots, H_k$ be q-uniform hypergraph matchings on $[n]$ of size δn. For every $b \in \{-1, 1\}^k$, define the following q-XOR instance Φ_b in n variables $x_1, x_2, \ldots, x_n$.

$$\forall i \in [k], \ \forall C \in H_i \,, x_C = b_i \,. \tag{1}$$

Then, there exists a (normal form) linear LDC $\mathcal{L} : \{-1, 1\}^k \to \{-1, 1\}^n$ described by the collection of q-uniform matchings $H_1, H_2, \ldots, H_k$ on $[n]$ if and only if Φ_b is satisfiable for every $b \in \{-1, 1\}^k$.

If $\mathcal{L}$ is a (q, δ)-LDC described by matchings $H_1, H_2, \ldots, H_k$, then $x = \mathcal{L}(b)$ satisfies all the constraints in Φ_b. Conversely, if Φ_b is satisfiable for every b, then one can easily construct a linear map $\mathcal{L}$ (easily seen to be a linear (q, δ)-LDC) where $\mathcal{L}(b)$ is some satisfying assignment to Φ_b.

The main idea of [AGKM23] is to show that for any collection of δn-size q-matchings $H_1, H_2, \ldots, H_k$, if k is large enough as a function of n, then for a randomly chosen b, Φ_b is unsatisfiable with high probability. This implies an upper bound on k. Now, when b is random, Φ_b is XOR formula generated via $k \ll n$ bits, i.e., much smaller than the number of variables. Thus, a naive union bound argument cannot establish unsatisfiability of Φ_b. The work of [AGKM23] establishes unsatisfiability of Φ_b for a random b via a *spectral refutation* using *Kikuchi* matrices.

Spectral refutations for Φ_b. Let us now recall how the spectral refutation in [AGKM23] works. Their main result is for the case of $q = 3$ (where they obtained improvements on prior works). However, for our purpose of illustrating the conceptual idea, we will focus on the simpler setting of even q and sketch their proof that $k \leq \tilde{O}(n^{1-2/q})$ for q-LDCs.

First, we observe that for the XOR instance Φ_b, there is an associated "instance polynomial" $\Phi_b(x) := \sum_{i=1}^{k} \sum_{C \in H_i} b_i x_C$. We note that $\Phi_b(x)$ is the number of constraints satisfied by x minus the number of constraints violated, and thus Φ_b is unsatisfiable if and only if $\mathrm{val}(\Phi_b) := \max_{x \in \{-1,1\}^n} \Phi_b(x)$ is less than $\sum_{i=1}^{k} |H_i| = k \cdot \delta n$. Thus, to show that Φ_b is unsatisfiable, we will bound $\mathrm{val}(\Phi_b)$.

To do this, we define a matrix whose quadratic form is equal to $\Phi_b(x)$.

Definition 3.2 (Kikuchi matrix and graphs). Let $C \in \binom{[n]}{q}$, let ℓ be a parameter, and let $N := \binom{n}{\ell}$. Let $A_C \in \{0, 1\}^{N \times N}$ be the matrix indexed by sets $S \in \binom{[n]}{\ell}$ where $A_C(S, T) = 1$ if $S \oplus T = C$, and 0 otherwise. Let $A_i := \sum_{C \in H_i} A_C$, and let $A := \sum_{i=1}^{k} b_i A_i$. We naturally interpret (and by abuse of notation, also call) A_C, A_i and A as adjacency matrices of "Kikuchi graphs" on the vertex set $\binom{[n]}{\ell}$.

Observe that A_C is a matching on vertex set $\binom{[n]}{\ell}$ of size $D = \binom{n-q}{q/2}\binom{q}{q/2}$. For any $x \in \{-1, 1\}^n$, let $x^{\circ \ell}$ denote the ℓ-wise monomial vector indexed by $S \in \binom{[n]}{\ell}$ with corresponding entry equal to x_S. Then, $x^{\circ \ell \top} A_C x^{\circ \ell} = D x_C$. Consequently, $x^{\circ \ell \top} A x^{\circ \ell} = D \Phi_b(x)$. Thus, if $x \in \{-1, 1\}^n$ satisfies Φ_b, then we have the following inequality that upper bounds k in terms of $\|A\|_2$:

$$k \delta n = \Phi_b(x) \leq \frac{1}{D} \left\| x^{\circ \ell} \right\|_2^2 \|A\|_2 = \frac{\binom{n}{\ell}}{D} \|A\|_2 \leq O((n/\ell)^{q/2}) \|A\|_2 \,. \tag{2}$$

8

We now choose $b \in \{-1, 1\}^k$ uniformly at random and consider $A = \sum_i b_i A_i$, which is a matrix Rademacher series of the A_i's. By the matrix Khintchine inequality, $\|A\|_2 \leq O(\sqrt{\log N}) \|\sum_i A_i^2\|_2^{1/2}$ with high probability.

A combinatorial proxy for $\|A\|_2$. Let Δ_i be the maximum degree of any node in the Kikuchi graph A_i, and let $\Delta = \max_{1 \leq i \leq k} \Delta_i$. Then, we can naively bound $\|\sum_i A_i^2\|_2 \leq \sum_i \|A_i\|_2^2 \leq k\Delta^2$. Thus, the maximum degree of the A_i's naturally controls the spectral norm of A as $\|A\|_2 \leq \Delta \cdot O(\sqrt{k\ell \log n})$.

Let us now investigate bounds on Δ. Since for each $C \in H_i$, A_C contributes D edges to A_i, the average degree of A_i is clearly $\delta n D / N \sim n(\ell/n)^{q/2}$. Thus, $\Delta \geq O(1) \max\{1, n(\ell/n)^{q/2}\}$. If Δ happens to be equal to this minimum possible value, then plugging it in Eq. (2) yields:

$$k\delta n \leq O(1) \left(\frac{n}{\ell}\right)^{q/2} \sqrt{k\ell \log n} \cdot \max\{1, n(\ell/n)^{q/2}\},$$

which implies that $k \leq O(\ell \log n) \cdot \max\{n^{q-2}/\ell^q, 1\}$. This is minimized at $\ell = n^{1-2/q}$ to give the lower bound of $k \leq \tilde{O}(n^{1-2/q})$, i.e., $n \geq \tilde{\Omega}(k^{q/(q-2)})$.

Handling irregularities: row pruning via polynomial concentration. We will now (for the first time in the argument) use that the H_i's are matchings to argue that while the A_i's are certainly not approximately regular (i.e., max degree Δ_i at most a polylog(n) factor larger than the average-degree), there is only a small fraction of nodes in any A_i that have a large degree. Of course, a small fraction of rows can still cause $\|A\|_2$ to be too large. In order to circumvent this issue, we observe that the argument in Eq. (2) works even if we were to replace $N \|A\|_2$ (maximum over arbitrary quadratic forms) by $\|A\|_{\infty \rightarrow 1}$ (maximum over quadratic forms on ±1-coordinate vectors). The latter quantity is insensitive to dropping a small fraction of rows since ±1-coordinate vectors when restricted to a small number of rows must have correspondingly small ℓ_2-norm.

To prove that only a small fraction of nodes can have a large degree in any A_i, we view the degree of any node S as a polynomial in the corresponding indicator variables $z \in \{0, 1\}^n$ with $\sum_i z_i = \ell$ and use tail inequalities for low-degree polynomials (that generalize concentration of Lipschitz functions) of Kim and Vu and extensions [KV00, SS12] to bound the chance that it takes a value polylog(n) times the average. This relies on establishing strong bounds on the expected partial derivatives of the degree polynomial by using that the H_i's are matchings.

The key heuristic: high density for Kikuchi graphs at low levels. Let's summarize the crucial steps of the above argument as follows: (1) q-LDCs naturally yields XOR instances of arity q, (2) to obtain our lower bound, we need that the Kikuchi matrices A_i corresponding to a matching H_i are approximately regular (after dropping a negligible fraction of rows), and (3) the argument can only yield a bound of the form $k \lesssim \ell$ where ℓ is the smallest level of the Kikuchi graphs A_i with an average degree $\gg 1$. More precisely, if there are m_i constraints of arity q in H_i, then the threshold ℓ is the smallest integer satisfying $m_i(\ell/n)^{q/2} \gg 1$ for all $i \in [k]$. Note that this threshold ℓ increases as q increases.

We assert that even though the argument in [AGKM23] for the case when $q = 3$ requires more work (in both the design of the Kikuchi matrix itself and its analysis), the heuristic above continues to hold. Let us also note that ensuring approximate regularity is usually the trickiest aspect of the proof. In particular, while the heuristic above makes sense for all odd q (and not just $q = 3$), and

9

the work of [AGKM23] fails to obtain an improved lower bound for odd $q > 3$ because they were unable to find an appropriate "decomposition" that ensures approximate regularity of the resulting Kikuchi matrices.

Thus, in order to obtain an exponential lower bound, as in Theorem 1, via the schema above, we must construct Kikuchi graphs that have constant density (i.e., average degree) at much a lower level ℓ. Specifically, we will need to be able to take $\ell = \text{polylog}(n)$.[5]

3.2 Long chain derivations: improved spectral refutations by increasing density

Given the key heuristic above, we now show how to build XOR instances from 3-LCCs that yield constant density Kikuchi matrices at level $\ell = \text{polylog}(n)$. Our instances will balance two opposing concerns. On the one hand, they will be of large arity (in fact, $O(\log n)$ arity) which, given the discussion above, hurts the density at lower levels. Nonetheless, we will show that the number of higher arity constraints that we produce grows fast enough to compensate for this and gives us an overall increase in density at lower ℓ. We note (with the hope of pointing the reader to the trickiest part of the proof that motivates all our setup) that the analysis of "row pruning" i.e., arguing approximate regularity after removing a negligible fraction of rows, will get significantly more involved and motivates all our design choices. This includes the specific type of Kikuchi matrices that we will choose and a new decomposition for the constraints that, while a bit unnatural at the outset, helps guarantee approximate regularity. Let us see these ideas in more detail next.

Like 3-LDCs, 3-LCCs can, without loss of generality, be assumed to be $(3, \delta)$-normal. Thus, for any 3-LCC $\mathcal{L}: \{-1, 1\}^k \to \{-1, 1\}^n$, there are 3-uniform hypergraph matchings $H_1, \ldots, H_n$ on $[n]$, each of size δn, such that for every $b \in \{-1, 1\}^k$, $u \in [n]$, and $C \in H_u$, the encoding $x = \mathcal{L}(b)$ satisfies $x_C = x_u$. Note that the key difference between LCCs and LDCs is that here we have a "local correcting" hypergraph H_u for each $u \in [n]$, instead of only a hypergraph for each $u \in [k]$ in the case of LDCs.

The naive XOR instances. Similar to Observation 3.1, the combinatorial characterization yields that the XOR instance with constraints $x_C = x_u$ for every $C \in H_u$ and $u \in [n]$ (where on the right hand side, we set $x_u = b_u$ whenever $u \in [k]$) is satisfiable for every $b \in \{-1, 1\}^k$. If we focus only on the constraints corresponding to H_u for $u \in [k]$ (i.e., the "systematic" bits in the codeword), then we recover the same XOR instance as in the case of 3-LDCs and our method from above yields $k \leq \tilde{O}(n^{1/3})$ [AGKM23]. To improve on this significantly lossy formulation, we must make use of the additional constraints H_u for $u \notin [k]$. More specifically, if we were to only use the hypergraphs H_u for $u \in [k]$, then any lower bound we could prove would hold for LDCs as well, and in particular one could not hope to prove Theorem 1, which is false for LDCs.

Long chain derivations. We now show how to use the additional constraints in order to build a

---

[5]We note that while our lower bounds appear to get weaker as ℓ grows, generic convergence results about the Kikuchi matrices imply that taking $\ell \sim n$ and bounding Φ_b in terms of $\|A\|_2$ yields the *optimal* bound on k, whatever it may be! The reason the current argument (which is likely suboptimal) does not extend beyond $\ell = n^{1-2/q}$ is the potentially superfluous $\sqrt{\log N}$ multiplicative loss in the matrix Khintchine inequality. Investigating when this $\sqrt{\log N}$ factor (which is tight in the worst-case) can be removed is the topic of an ongoing research effort in random matrix theory [BBH23] and is naturally related to other problems such as resolving the matrix Spencer conjecture [Zou12, Mek14].
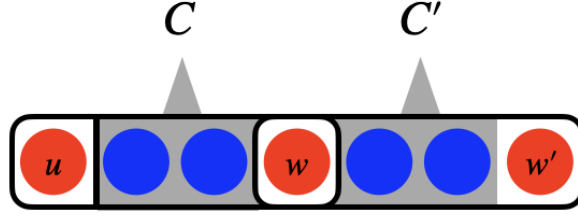
Figure 1: A 2-chain with head u. Note that $C \cup \{w\} \in H_u$ and $C' \cup \{w'\} \in H_w$, and that $x = \mathcal{L}(b)$ satisfies $x_C x_w = x_u$ and $x_{C'} x_{w'} = x_w$, and therefore $x_C x_{C'} x_{w'} = x_u$.

---

higher arity XOR instance that is (1) approximately regular (after an appropriate decomposition), and (2) results in high-density Kikuchi graphs at polylog(n) levels. We will construct higher arity XOR instances that use the additional constraints above using a structured variant of low-width XOR resolution [Gri01, Sch08] that we call *long chain derivations*.

Let us start by forming extra constraints via 2-chains. Observe that for any $u \in [n]$ and $C \in H_u$, we have that for any $b \in \{-1, 1\}^k$, $x = \mathcal{L}(b) \in \{-1, 1\}^n$ satisfies the equation $x_u x_C = 1$. Now, let us choose $w \in C$ and $C' \in H_w$. We also have that $x_w x_{C'} = 1$. As $x_C = x_{C \setminus \{w\}} x_w$, it follows that the "derivation" $x_u x_{C \setminus \{w\}} x_{C'} = 1$ also holds, since $x_w^2 = 1$. We shall call such a constraint a "2-chain" — it connects two constraints intersecting in one variable. We can think of such a 2-chain as a tuple (u, C, w, C', w'), where $C \cup \{w\} \in H_u$ and $C' \cup \{w'\} \in H_w$, and this yields the constraint $x_C x_{C'} x_{w'} = x_u$ (see Fig. 1).

Consider now the 2-chains $\cup_{i \in [k]} \mathcal{H}_i^{(2)}$, i.e., 2-chains of the form (i, C, w, C', w') where $i \in [k]$. Then, the constraints have the form $x_C x_{C'} x_{w'} = b_i$, so they decode the i-th independent bit b_i. We have thus formed a new set of constraints with "right hand side" b_i.

A heuristic calculation. Let us now do a heuristic calculation (that ignores the key issue of approximate regularity) to see if we improve the density at lower Kikuchi levels by taking the XOR instances corresponding to 2-chains. For any fixed "head" $i \in [k]$, there are $(3\delta n)^2$ 2-chains. This is because we have δn choices for $C \cup \{w\} \in H_i$, followed by 3 ways to choose w from $C \cup \{w\}$, and then similarly $3\delta n$ choices in total for (C', w'). Let $\mathcal{H}_i^{(2)}$ denote the set of 2-chains with head i. We have thus produced $\sim n^2$ constraints and each constraint has arity 5,[6] as $|C| = |C'| = 2$.

The Kikuchi matrix in Definition 3.2 only makes sense for even q, but let us still do a "pretend" calculation of the relative density for the arity 5 constraints we have produced. This can be made precise with a slightly more sophisticated Kikuchi matrix, so this is still a meaningful heuristic.

The density (i.e., average degree) expression for a Kikuchi matrix A_i is now $n^2(\ell/n)^{q/2} \sim n^2(\ell/n)^{5/2} \sim \ell^{2.5}/n^{0.5}$. This density is $\gg 1$ whenever $\ell \gg n^{1/5}$, so one might expect to obtain a bound of $k \lesssim n^{1/5}$ (beating the $n^{1/3}$ bound for the naive XOR instance [AGKM23]) when working with 2-chains — a construction that crucially relies on additional structure in 3-LCC! While there

---

[6]Some constraints may have additional variable cancellations and thus have arity < 5. However, as the density gets worse as the arity increases, this is only "better" for us.

are lot of details that we have simply ignored in doing this calculation, it does suggest that we are able to achieve a constant-density Kikuchi matrix A_i at a lower level ℓ. A similar calculation (that we will omit here) for chains of larger length, say r, shows that the smallest level ℓ at which we can obtain constant density Kikuchi matrices is $\ell \sim n^{1/2r}$, and this suggests that we might be able to obtain constant density at level $\ell = \text{polylog}(n)$ if we work with $r \sim \log n$ length chains.

In Section 4, as a warmup to our somewhat technical proof of the main theorem, we present a complete analysis of the 2-chains (with extended commentary) to obtain a $k \leq \tilde{O}(n^{1/4})$ bound (giving a polynomial improvement on the $\sim n^{1/3}$ lower bound on 3-LDCs already!) in order to illustrate (a simplified version of) the set of new tools that go into the analysis.

3.3 From the heuristic to a proof

In the remaining part of this overview, we briefly discuss the technical tools we develop to turn the above heuristic calculation into a full proof. We note that the actual parameters become rather delicate. For readers familiar with the literature on random CSP refutation (our setting resembles semirandom XOR refutation with complicated correlations in the right hand sides), this is similar to the analysis getting rather delicate when dealing with XOR instances with super-constant arity.

Setting up the Kikuchi matrix. The instances produced by forming r-chains yield XOR instances of (odd) arity $2r + 1$. We build a different Kikuchi matrix by first applying the "Cauchy–Schwarz" trick — a standard idea in CSP refutation also utilized in [AGKM23]. In our case, the XOR instance produced after this trick corresponds to constraints formed by joining two r-chains at their "tails" whenever the tails match. We choose a variant of the Kikuchi matrix for the "Cauchy–Schwarzed instance" except for the key difference that it is indexed by $2r$-tuples of sets of size ℓ (instead of a single set of size ℓ) in the sketch above. This choice is crucial in the analysis of row pruning, in particular, as we discuss below, in obtaining bounds that significantly beat those obtained by a blackbox application of low-degree polynomial concentration [KV00], see below.

Regularity decomposition. If $H_1, H_2, \ldots H_n$ are such that no pair of variables appears in more than one hyperedge ("no heavy pairs") across all the H_i's, then it turns out that the resulting Kikuchi matrices satisfy approximate regularity after pruning a negligible fraction of rows. This no-heavy-pair property holds, e.g., if H_i's are uniformly random and independent hypergraph matchings of size δn.

However, when the H_i's are arbitrary, and in particular when there are "heavy pairs" (i.e. pairs of variables that appear in $\gg \log n$ hyperedges across the H_i's), the resulting Kikuchi matrices are *far* from being approximately regular. Our key technical idea is a new *decomposition* procedure that operates directly on the chains. Such a decomposition procedure partitions the chains into $\sim r$ different groups such that each group admits a (different, appropriately defined) Kikuchi matrix that satisfies approximate regularity. Regularity decompositions were already used in early applications of the Kikuchi matrix method for proving hypergraph Moore bound and smoothed CSP refutation [GKM22, HKM23]. However, our notion of regularity is (necessarily) significantly weaker (we call it "contiguously regular" partitioning) that, unlike [GKM22], does not "by design" ensure approximate regularity of the Kikuchi matrices after removing only a negligible fraction of

rows. Instead, our argument for approximate regularity relies on combining the guarantees of the decomposition with (1) an appropriate choice of Kikuchi matrix for each piece in the partition, and (2) the structure in the chains arising by virtue of H_i's being matchings.

Polynomial concentration: bounding expected derivatives. Our main technical step (the subject of Section 8) is proving that our weak notion of regularity combined with the fact that H_i's are matchings is enough to control expected partial derivatives of the "degree-polynomial" that computes the degrees of nodes in the Kikuchi graph.

We note that off-the-shelf low-degree polynomial concentration inequalities (e.g., the Kim–Vu inequality [KV00] or the related inequality of Schudy and Sviridenko [SS12]) lose an exponential factor in the degree of the polynomial in the tail bound. This exponential factor is too costly for us as the arity of our constraints, and thus the degree of the polynomial, is $O(\log n)$ that eventually restricts us to only a quasi-polynomial instead of an exponential lower bound on 3-LCCs. Instead, we induce a special "partite" structure (i.e., there exists a partition of the variables so that the degree of the polynomial is 1 when restricted to any single piece in the partition) in the polynomial by setting up our Kikuchi matrix to be indexed by tuples of sets (instead of a single set). For such partite polynomials, we prove an analog[7] of the Kim–Vu inequality for partite that gives sharper bounds when its expected partial derivatives decay appropriately.

We note that the analysis of the expected partial derivatives of the "degree polynomial" (which we use to prove approximate regularity) and the interplay of these bounds with our decomposition of chains is the key technical part (and the focus of Section 8) of our proof. In order to illustrate this technical part in a "base" case that still captures some of the complications, we present the case of 2-chains as a warmup in the next section.

4   Warmup: An $n \geq \tilde{\Omega}(k^4)$ Lower Bound via 2-Chains

In this section, we give a detailed sketch of the proof of the following theorem, which is a weaker version of our main result. Notice that this theorem already improves the best known 3-LCC lower bound [AGKM23] by a polynomial factor in k.

Theorem 4.1 (Weak version of Theorem 1). *Let $\mathcal{L} : \{-1, 1\}^k \to \{-1, 1\}^n$ be a $(3, \delta)$-LCC in normal form with $\delta = O(1)$. Then, $n \geq \tilde{\Omega}(k^4)$.*

The theorem above obtains a lower bound of $n \gtrsim k^4$ — worse than the bound of $n \gtrsim k^5$ predicted by the heuristic but still beating $n \gtrsim k^3$ from [AGKM23]; we discuss the reason that we do not match the heuristic in Remark 4.2.

Proof. As before, we have 3-uniform hypergraph matchings $H_1, \ldots, H_n$, where for any $u \in [n]$ and $C \in H_u$, we have that for any $b \in \{-1, 1\}^k$, $x = \mathcal{L}(b)$ satisfies $x_C = x_u$. Following Section 3.2, we

---

[7]We did not find a reference to a known result so we include a proof in Lemma 2.9.

shall let $\mathcal{H}_i^{(2)}$ denote the set of 2-chains with head i. We define the 5-XOR instance $\Phi_b(x)$ as

$$\Phi_b(x) := \sum_{i=1}^{k} b_i \sum_{\vec{C}=(i,C_0,w_0,C_1,w_1)\in\mathcal{H}_i^{(2)}} x_{C_0} x_{C_1} x_{w_1} \ .$$

We note that $\mathrm{val}(\Phi_b) = k(3\delta n)^2$ for any $b \in \{-1,1\}^k$, as the instance is satisfiable and has $k(3\delta n)^2$ constraints in total. Following the strategy in Section 3.1, we shall use spectral refutation via Kikuchi matrices to bound $\mathrm{val}(\Phi_b)$ with high probability for a random $b \in \{-1,1\}^k$.

4.1   Step 1: the Cauchy–Schwarz trick

As we have observed, the basic Kikuchi matrices in Definition 3.2 are only defined for constraints of even arity, but the constraints in $\mathcal{H}_i^{(2)}$ have arity 5, i.e., odd arity. The standard way to handle odd arity XOR instances is to use the "Cauchy–Schwarz trick", which produces even arity instances as follows. Let $\vec{C} \in \mathcal{H}_i^{(2)}$ and $\vec{C}' \in \mathcal{H}_j^{(2)}$ for $i \neq j \in [k]$ be two constraints in our initial 5-XOR instance, where $\vec{C} = (i, C_0, w_0, C_1, w_1)$ and $\vec{C}' = (j, C_0', w_0', C_1', w_1')$ where $w_1 = w_1'$, i.e., the last element of both chains is the same. From this pair, we can "cancel" $w_1 = w_1'$, producing the derived constraint $x_{C_0} x_{C_1} x_{C_0'} x_{C_1'} = b_i b_j$, which has arity 8. We do this for all pairs of chains with the same "tail" vertex w. We note that this process produces at least $(k(3\delta n)^2)^2 / n \sim k^2 n^3$ constraints.

We now define the following "Cauchy–Schwarzed instance" polynomial:

$$f_b(x) = \sum_{i \neq j \in [k]} b_i b_j \sum_{w \in [n]} \sum_{\vec{C}\in\mathcal{H}_i^{(2)},\vec{C}'\in\mathcal{H}_j^{(2)}:w_1=w_1'=w} x_{C_0} x_{C_1} x_{C_0'} x_{C_1'} \ .$$

The phrase "Cauchy–Schwarz trick" refers to the fact that one can show $k^2 n^4 \sim \Phi_b(x)^2 \leq n \cdot f_b(x) + o(k^2 n^4)$ via a simple application of the Cauchy–Schwarz inequality and a bound on the "diagonal terms" where $i = j$. This reduces the task to bounding the cross-term polynomial f_b.

We now observe that the "right-hand sides" of the constraints in f_b are no longer independent, as they are of the form $b_i b_j$ for $i \neq j \in [k]$, and this will cause an issue "downstream" when we apply matrix concentration bounds, as the matrices will not be independent. To recover independence, we consider the polynomial $f_{M,b}(x)$ defined for a (directed) matching M on $[k]$:

$$f_{M,b}(x) = \sum_{(i,j)\in M} b_i b_j \sum_{w \in [n]} \sum_{\vec{C}\in\mathcal{H}_i^{(2)},\vec{C}'\in\mathcal{H}_j^{(2)}:w_1=w_1'=w} x_{C_0} x_{C_1} x_{C_0'} x_{C_1'} \ .$$

Because we now sum over a matching, we have that $b_i b_j$ and $b_{i'} b_{j'}$ are independent for different directed edges (i, j) and (i', j') in M. And, we can easily relate f_b and $f_{M,b}$, as $f_b(x) = 2(k - 1)\mathbb{E}_M f_{M,b}(x)$ when k is even, and $f_b(x) = 2k\mathbb{E}_M f_{M,b}(x)$ when k is odd, where the expectation is over a maximum matching M. This is because the chance that M contains a directed edge (i, j) is $\frac{1}{2(k-1)}$ if k is even and $\frac{1}{2k}$ if k is odd. In particular, there exists a maximum matching M such that $\mathrm{val}(f_{M,b}) \geq \frac{2}{k}\mathrm{val}(f_b) \sim kn^3$.

14

Remark 4.2. Restricting to a matching M loses a factor of k in the number of constraints. This leads to a factor k "loss" in the density of the corresponding Kikuchi matrix and is the main reason why we obtain weaker bound of $n \geq \tilde{O}(k^4)$ instead of k^5 suggested by our heuristic calculation in Section 3.2. A better bound could be obtained by instead following the setup in [AGKM23], where they split $[k]$ randomly into a left and right set L and R and only consider constraints where $i \in L$ and $j \in R$ (thereby losing only $\sim 1/2$ of the constraints instead of a factor k). This careful setup is necessary in [AGKM23] for their goal of obtaining a cubic (as opposed to the known quadratic) bound, but this makes the "row pruning" step (i.e., arguing approximate regularity of Kikuchi graphs after removing a negligible fraction of constraints) significantly more challenging. In our case, the effect of this loss on the final lower bound diminishes as the length of the chain r grows and when $r \sim \log n$, disappears asymptotically, and so we pick a matching M to make the row pruning easier.

4.2   Step 2: spectral refutation via Kikuchi matrices

Let us now bound $\mathrm{val}(f_{M,b})$ (with high probability over $b \in \{-1, 1\}^k$) for any maximum matching M. We introduce our Kikuchi matrices:

Definition 4.3. For $i \neq j \in [k]$ and $\vec{C} = (i, C_0, w_0, C_1, w_1)$ and $\vec{C}' = (j, C_0', w_0', C_1', w_1')$ with $w_1 = w_1'$, we define the matrix $A_{i,j}^{(\vec{C}, \vec{C}')}$ as follows. The rows/columns of the matrix $A_{i,j}^{(\vec{C}, \vec{C}')}$ are indexed by a 4-tuple of sets (S_0, S_1, S_0', S_1'), each in $\binom{[n]}{\ell}$, and the $((S_0, S_1, S_0', S_1'), (T_0, T_1, T_0', T_1'))$-th entry is 1 if $S_0 \oplus T_0 = C_0$, $S_1 \oplus T_1 = C_1$, $S_0' \oplus T_0' = C_0'$, $S_1' \oplus T_1' = C_1'$, and is 0 otherwise.

We let $A_{i,j} = \sum_{\vec{C} \in \mathcal{H}_i^{(2)}, \vec{C}' \in \mathcal{H}_j^{(2)} : w_1 = w_1'} A_{i,j}^{(\vec{C}, \vec{C}')}$ and $A = \sum_{(i,j) \in M} b_i b_j A_{i,j}$.

We now observe that each matrix $A_{i,j}^{(\vec{C}, \vec{C}')}$ has exactly D^4 nonzero entries, where $D = 2 \cdot \binom{n-2}{\ell-1}$, and the matrix has N^4 rows/columns, where $N = \binom{n}{\ell}$. We note that $D/N \sim \ell/n$, and so the average number of nonzero entries per row (or column), i.e., the density, is $(D/N)^4 \sim (\ell/n)^4 = (\ell/n)^{q/2}$, as the arity of the constraints is 8.

We also observe that for any $x \in \{-1, 1\}^n$, $D^4 f_{M,b}(x) = x'^\top A x'$, where x' is the vector with (S_0, S_1, S_0', S_1')-th entry equal to $\prod_{v \in S_0} x_v \prod_{v \in S_1} x_v \prod_{v \in S_0'} x_v \prod_{v \in S_1'} x_v$. We thus have that

$$ kn^3 \cdot D^4 \leq D^4 \cdot \mathrm{val}(f_{M,b}) \leq \|A\|_{\infty \to 1} \leq N^4 \|A\|_2 \ . $$

For any $i \neq j$, the matrix $A_{i,j}$ has density $\sim m_{i,j}(D/N)^4 \sim (\ell/n)^4$, where $m_{i,j}$ is the number of the constraints in f_b with right-hand side $b_i b_j$. Let us now argue that each $m_{i,j}$ is at most $O(n^3)$. Indeed, $m_{i,j}$ is the number of pairs of 2-chains $(i, C_0, w_0, C_1, w_1) \in \mathcal{H}_i^{(2)}$ and $(j, C_0', w_0', C_1', w_1') \in \mathcal{H}_j^{(2)}$ where $w_1 = w_1'$. To show that $m_{i,j} \leq O(n^3)$, we pick w_0, w_1 and w_0', for a total of n^3 choices, and observe that this completely determines both chains. Indeed, because H_i is a matching, there is at most one constraint C in H_i that contains w_0, and then C_0 must be $C \setminus \{w\}$. This similarly shows that we have at most one choice of C_1 and also C_0'. Finally, because $w_1' = w_1$, and we know w_1, we thus know w_1' as well, which by similar reasoning gives us at most one choice for C_1', and we have determined

15

the entire chain. We note that we have a lower bound of $\sim kn^3$ on the total number of constraints $\sum_{(i,j)\in M} m_{i,j}$, so this calculation also shows that no $m_{i,j}$ can be much larger than the average.

Returning to the density calculation, we have shown that $A_{i,j}$ has density at most $n^3(\ell/n)^4 = \ell^4/n$. Again, following the blueprint in Section 3.1, we will set $\ell = n^{1/4} \cdot \text{polylog}(n)$, and we want to show that the matrices $A_{i,j}$ satisfy the approximate regularity condition, i.e., the number of rows/columns with more than $\Delta = \ell^4 \cdot \text{polylog}(n)/n$ nonzero entries is at most $N^4/\text{poly}(n)$. Let us finish the proof, assuming that this holds.

Proof assuming approximate regularity. Let $\mathcal{B}$ denote the set of rows/columns that are "bad" for some pair (i,j), i.e., the matrix $A_{i,j}$ has more than Δ nonzero entries in that row. Let $B_{i,j}$ be the matrix where the rows and columns in $\mathcal{B}$ have been all set to 0. Let $B = \sum_{(i,j)\in M} b_i b_j B_{i,j}$. We have that B is the sum of mean 0 independent matrices, each with spectral norm $\|B_{i,j}\|_2 \le \Delta$. Therefore, by matrix Khintchine (Fact 2.8), we have that with high probability over b, $\|B\|_2 \le O(\Delta\sqrt{k\log(N^4)}) = O(\Delta\sqrt{k\ell\log n})$.

Now, we observe that $\|A - B\|_{\infty\to 1} \le o(N)$. This is because the number of nonzero entries that we have removed from A to produce B is at most $k \cdot n^3 \cdot N^4/\text{poly}(n) = o(N^4)$ (there are k edges (i,j) in the matching M, each has $m_{i,j} \le n^3$ constraints, and each row of $A_{i,j}$ has at most $m_{i,j} \le n^3$ nonzero entries) provided that the $\text{poly}(n)$ factor is large enough. We thus conclude that

$$kn^3 \cdot D^4 \le D^4 \cdot \text{val}(f_{M,b}) \le \|A - B\|_{\infty\to 1} + N^4 \|B\|_2 \le o(N^4) + N^4 O(\Delta\sqrt{k\ell\log n}) \ .$$

Substituting the value for Δ and rearranging, we conclude that $k \le \ell \cdot \text{polylog}(n) \le \tilde{O}(n^{1/4})$.

We remark that Sections 4.1 and 4.2 are fairly mechanical, and they justify the use of the heuristic calculation. The place where we had "freedom" is in the choice of constraints to use in the initial XOR instance, which we chose to be the 2-chains $\mathcal{H}_i^{(2)}$. It thus remains to bound the number of bad rows $\mathcal{B}$. This "row pruning" step is key to converting the heuristic into a full proof.

4.3 Step 3: row pruning, the key technical step

We want to understand if, after dropping a $1/\text{poly}(n)$ fraction of the rows, every Kikuchi graph $A_{i,j}$ satisfies approximate regularity. This is equivalent to showing that for every matrix $A_{i,j}$, with probability at least $1 - 1/\text{poly}(n)$ a uniformly random row (S_0, S_1, S_0', S_1'), has at most Δ nonzero entries in $A_{i,j}$ for $\Delta = \ell^4 \cdot \text{polylog}(n)/n = \Delta_{avg}\text{polylog}(n)$.

The heavy pair degree. We now make a key observation. Whether the above approximate regularity property holds for a given collection of matchings $H_1, H_2, \ldots, H_n$ is governed by a single parameter that we call the *heavy pair degree d*. This is the maximum, over all pairs $\{v, v'\} \subseteq [n]$, of the number of hyperedges across the H_i's that contain $\{v, v'\}$. We will prove that if d is small enough then approximate regularity holds for every $A_{i,j}$ after dropping a $1/\text{poly}(n)$-fraction of rows. When d is large, this property will not hold for the $A_{i,j}$'s from Definition 4.3. Instead, we will define a *different* collection of Kikuchi matrices that have high density and for which row pruning succeeds.

Lemma 4.4 (Row pruning for 2-chains with no heavy pairs). *Let $H_1, \ldots, H_n$ be 3-uniform hypergraph matchings of size δn, and let d be the maximum, over all pairs $\{v, v'\}$ of vertices, of the number of pairs*

16

(u, C) with $u \in [n]$ and $C \in H_u$ where $\{v, v'\} \subseteq C$. Fix $i \neq j \in [k]$, and let $A_{i,j}$ be the matrix defined in [Definition 4.3](#) at level $\ell \in \mathbb{N}$.

Suppose that $d \leq \ell^2$. Then, the number of rows (S_0, S_1, S_0', S_1') of $A_{i,j}$ with more than $\Delta = \ell^4 \cdot \text{polylog}(n)/n$ nonzero entries is at most $N^4/\text{poly}(n)$.

We note that if the matchings $H_1, \ldots, H_n$ are *random*, then we have $d \leq \text{polylog}(n)$ with high probability, and so random matchings satisfy the "small heavy-pair degree" assumption with high probability. We can thus think of $d \leq \text{polylog}(n)$ as a pseudorandom property of a collection $H_1, \ldots, H_n$ of matchings. We now sketch a proof of [Lemma 4.4](#).

The degree polynomial and its partial derivatives. As the first step in the proof of [Lemma 4.4](#), we define a degree 4 polynomial $\text{Deg}_{i,j} : \{0, 1\}^{4n} \to \mathbb{N}$, where we think of the $4n$ variables as split into 4 groups of n variables $s^{(0)}, s^{(1)}, s'^{(0)}, s'^{(1)}$, which are indicator variables of the 4 sets S_0, S_1, S_0', S_1', respectively. This polynomial $\text{Deg}_{i,j}(s^{(0)}, s^{(1)}, s'^{(0)}, s'^{(1)})$ upper bounds the number of nonzero entries in the (S_0, S_1, S_0', S_1')-th row in the matrix $A_{i,j}$ in [Definition 4.3](#).

Formally, let $\mathcal{T}_{i,j}$ denote the (multi)-set of 4-tuples (u_0, u_1, v_0, v_1) such that there exists $\vec{C} = (i, C_0, w_0, C_1, w_1) \in \mathcal{H}_i^{(2)}$ and $\vec{C}' = (j, C_0', w_0', C_1', w_1') \in \mathcal{H}_j^{(2)}$ with $w_1 = w_1'$ such that $u_0 \in C_0, u_1 \in C_1, v_0 \in C_0', v_1 \in C_1'$; if there are multiple such pairs $(\vec{C}, \vec{C}')$ that produce the same (u_0, u_1, v_0, v_1), then we add this tuple multiple times. Then, we set

$$\text{Deg}_{i,j}(s^{(0)}, s^{(1)}, s'^{(0)}, s'^{(1)}) := \sum_{(u_0, u_1, v_0, v_1) \in \mathcal{T}_{i,j}} s^{(0)}_{u_0} s^{(1)}_{u_1} s'^{(0)}_{v_0} s'^{(1)}_{v_1} \ .$$

Note that $\text{Deg}_{i,j}$ is a polynomial with non-negative coefficients. We are interested in the probability that $\text{Deg}_{i,j}$, on uniform draws of 4-tuples of ℓ-size sets, takes a value that deviates from its expectation μ by some multiplicative factor. It turns out (see [Lemma 8.3](#)) that we can pass on to independent p-biased product distribution on $\{0, 1\}^{4n}$ for $p \sim \ell/n$ without much loss. This is helpful because the tail behavior of low-degree polynomials with non-negative coefficients on product distributions is determined by a bound on its expected partial derivatives. Namely, variants of the Kim-Vu inequality (see [Lemma 2.9](#)) show the following: *if the expectation of every partial derivative of* $\text{Deg}_{i,j}$ *is at most* μ, *then* $\text{Deg}_{i,j}(S_0, S_1, S_0', S_1') \leq O(\mu \log n)$ *with probability at least* $1 - 1/\text{poly}(n)$.

Let us now examine the expected partial derivatives of $\text{Deg}_{i,j}(s)$. We start by introducing notation to refer to them. Let $Z = (z_0, z_1, z_0', z_1') \in ([n] \cup \{\star\})^4$ be an ordered tuple of length 4, with entries either in n or set to $\star$, which we think of as an "unfixed" value. Then, Z encodes partial derivatives with respect to any subset of variables that use at most one variable in each of the groups $s^{(0)}, s^{(1)}, s'^{(0)}, s'^{(1)}$. All other partial derivatives of $\text{Deg}_{i,j}$ are 0 since $\text{Deg}_{i,j}$ has degree 1 in each of the 4 groups of variables (i.e., $\text{Deg}_{i,j}$ is 4-partite). We know that $\mathbb{E}[\text{Deg}_{i,j}(s)] = \mu_{(\star, \star, \star, \star)} \leq 2^4 (\ell/n)^4 \cdot n^3 = O(1) \cdot \ell^4/n$; the factor of 2^4 comes from the fact that each pair $(\vec{C}, \vec{C}')$ adds 2^4 different tuples to $\mathcal{T}_{i,j}$. Now, [Lemma 2.9](#) implies that the chance that $\text{Deg}_{i,j}$ takes a value larger than $\mu \cdot \text{polylog}(n)$ is at most $1/\text{poly}(n)$ if $\mu_Z \leq \mu$ for all Z.

Computing expected partial derivatives. To help bound the expected partial derivatives μ_Z, let us relate these parameters to combinatorial quantities of the hypergraphs $H_1, H_2, \ldots, H_n$. Notice that when we take partial derivatives with respect to some Z, the only monomials that "survive"

are ones that "contain" Z, and furthermore the expectation of the partial derivative is simply $(\ell/n)^{\# \text{ of } \star \text{ entries in } Z}$ times the number of such monomials. Formally, let $\deg_{i,j}(Z)$ be the number of pairs $(\vec{C}, \vec{C}') \in \mathcal{H}_i^{(2)} \times \mathcal{H}_j^{(2)}$ where $w_1 = w_1'$ and $z_0 \in C_0, z_1 \in C_1, z_0' \in C_0', z_1' \in C_1'$, where for the symbol $\star$, we say that $\star \in C$ always holds — we say that such a pair $(\vec{C}, \vec{C}')$ *contains* Z. Then, the expected partial derivative at Z is $\mu_Z = 2^{4-|Z|}(\ell/n)^{4-|Z|}\deg_{i,j}(Z)$, where $|Z|$ is the number of non-$\star$ entries in Z.[8] For example, $Z = (\star, \star, \star, \star)$ is contained in all such pairs of 2-chains, and so $\deg_{i,j}(\star, \star, \star, \star) = m_{i,j} \leq O(n^3)$ and $\mu_Z = \mu = 16(\ell/n)^4 m_{i,j}$. Let us use the shorthand $\mu_t = \max_{Z:|Z|=t} \mu_Z$.

Let Z be an arbitrary 4-tuple with at least one non-$\star$ entry. As explained above, estimating μ_Z is, up to scaling, equivalent to counting $\deg_{i,j}(Z)$, the number of pairs $(\vec{C}, \vec{C}')$ that contain Z. We next observe that if Z has no $\star$ entries, then the number of 2-chains $(\vec{C}, \vec{C}')$ containing Z is an absolute constant. This is because there is at most one constraint $C_0 \cup \{w_0\}$ that contains z_0 in H_i. Given this constraint, there are 2 choices for w_0, as $w_0 \in C_0 \cup \{w_0\} \setminus \{z_0\}$. Given w_0, there is at most one constraint $C_1 \cup \{w_1\}$ in H_1 that contains z_1, and then at most 2 choices for w_1. We can similarly use the knowledge of (z_0', z_1') to bound the number of choices for C_0', C_1'. All in all, we have at most $16 = O(1)$ choices for the pair $(\vec{C}, \vec{C}')$ given Z with no $\star$ entries. This immediately shows that for Z such that $|Z| = 4$, $\mu_Z \leq O(1) \leq \mu$.

Let us now deal with Z's with at least one $\star$ entry by breaking up into cases depending on $|Z|$. We will view the counting of $\deg_{i,j}(Z)$ as a procedure that makes a bounded number of choices to decode the pair $(\vec{C}, \vec{C}')$.

Let us deal with the case when $|Z| = 1$. By swapping the roles of i and j if needed, without loss of generality we can assume that one of z_0 or z_1 is non-$\star$, and all other entries in Z are $\star$. There are at most n choices for z_0 (if $z_1 \neq \star$) or z_1 (if $z_0 \neq \star$). We now have n choices for z_0', which again determines C_0' and w_0' up to 2 choices. We now observe that (C_1', w_1') is uniquely determined. Indeed, this is because we know w_1', as it equals w_1 (the two 2-chains must have matching tails), and therefore this determines the hyperedge $C_1' \cup \{w_1'\} \in H_{w_0'}$ uniquely. We have thus shown that for Z with $|Z| = 1$, we have $\deg_{i,j}(Z) \leq O(n^2)$, and so $\mu_Z \leq (\ell/n)^3 \cdot O(n^2) \leq O(\ell^3/n) \leq O(\ell^4/n)$.

Let us now handle the case when $|Z| = 2$. Similar arguments as above show that $\mathrm{Deg}_{i,j}(Z) \leq O(n)$ holds for all Z except when the non-$\star$ entries of Z look like $Z = (\star, z_1, \star, z_1')$ where $z_1, z_1' \neq \star$, and thus $\mu_Z \leq (\ell/n)^2 \cdot O(n) \leq O(\ell^4/n)$ for these Z's. To count $\deg_{i,j}(Z)$ for $Z = (\star, z_1, \star, z_1')$ where $z_1, z_1' \neq \star$, we pay a factor of n to determine z_0, and then this determines (up to an $O(1)$ factor) C_0 and C_1 as well. Now, we know w_1' (because it is equal to w_1) and z_1' which is in C_1'. Thus, the hyperedge $C_1' \cup \{w_1'\}$ must contain the pair $\{z_1', w_1'\}$. Using the heavy pair degree, there are at most d choices for the pair $(w_0', C_1' \cup \{w_1'\})$, and after learning w_0' we also know C_0'. Hence, we have paid a total of $O(nd)$ choices, which implies that $\mu_2 \leq (\ell/n)^2 \cdot O(nd) = O(\ell^2 d/n)$. For $|Z| = 3$, a similar issue arises and gives a bound of $\mu_3 \leq O(\ell d/n)$.

We can now finish the proof of Lemma 4.4.

---

[8]The extra factor of $2^{4-|Z|}$ comes from the fact that for every Z and pair $(\vec{C}, \vec{C}')$ containing Z, the pair $(\vec{C}, \vec{C}')$ produces $2^{4-|Z|}$ tuples (u_0, u_1, v_0, v_1) in $\mathcal{T}_{i,j}$ that contain Z. In this case, this is just a constant factor, so we can ignore it.

Proof of Lemma 4.4. Notice that if $d \leq \ell^2$ then $\mu_t \leq \mu$ for every t. Applying Lemma 2.9 now yields that the probability that $\text{Deg}_{i,j} > \mu \cdot \text{polylog}(n)$ is at most $1/\text{poly}(n)$. Taking a union bound on $k < n$ yields that the fraction of bad rows $|\mathcal{B}|/N$ is at most $1/\text{poly}(n)$, as desired. $\square$

4.4 Step 4: hypergraph decomposition to handle large heavy pair degree

We will handle the case when the heavy pair degree is high by designing a *different* Kikuchi matrix. To do this, we we will construct the cross term polynomial (obtained by applying the Cauchy–Schwarz inequality) slightly differently. Our current Kikuchi matrix is built from the XOR instance obtained by pairing up chains that agree on their tails and thus "cancel" (i.e., square out) one variable. When the heavy pair degree is large, we will build chains by cancelling a pair of variables instead. The number of pairs of chains that agree in a pair of variables instead of just their tails, i.e., the new number of "Cauchy–Schwarzed" constraints, will of course be smaller than before. On the other hand, since we cancel a pair of variables instead of just the tail, the arity of the resulting XOR instance will be smaller: 6 instead of 8. The punchline is that the density vs. arity trade-off (i.e., our key heuristic discussed in Section 3.2) breaks in our favor, *provided that there are many "heavy pairs".*

To formally implement this argument, we *decompose* the set of chains by "labeling" each chain by the heavy pair contained within, if one exists. Intuitively, this is the pair of variables in the chain that we intend to cancel in the Cauchy–Schwarz trick. If the chain does not contain any heavy pair, then we label it by its tail variable w, which we will cancel in the Cauchy–Schwarz trick as done before in Section 4.1. We let $\mathcal{H}_Q$ denote the set of chains labeled by the heavy pair Q, and $\mathcal{H}_w$ denote the set of chains labeled by the tail variable w. For technical reasons (that will become relevant when we do the row pruning argument for the different, yet-to-be-defined Kikuchi matrices), our decomposition will produce multiple pieces labeled by the *same* heavy pair Q, i.e., $\mathcal{H}_{Q,1}, \mathcal{H}_{Q,2}$, etc., and for two chains labeled by the same Q, we shall only cancel the pair Q if these two chains lie within the *same* piece $\mathcal{H}_{Q,p}$.

Formally, our hypergraph decomposition is as follows. Given the collection $\mathcal{H}^{(1)} = \{(u, C, w) : u \in [n], C \cup \{w\} \in H_u\}$ of 1-chains, we perform the following greedy algorithm: if there exists an ordered pair $Q = (Q_1, Q_2)$ such that there are more than $d := \ell^2$ 1-chains (u, C, w) in $\mathcal{H}^{(1)}$ with $Q_1 \in C$ and $Q_2 = w$, i.e., Q is a heavy pair contained in the chain (u, C, w), then we choose an arbitrary set of *exactly* d such 1-chains, remove them from $\mathcal{H}^{(1)}$, and place them in a new "partition" $\mathcal{H}_{Q,p}$; here, $p \in \mathbb{N}$ denotes the "label" of the partition, as we may be producing multiple partitions with the same Q, and so we will denote these different pieces of the partition by $\mathcal{H}_{Q,1}, \mathcal{H}_{Q,2}$, etc. Finally, if there is no such heavy pair Q, then we create partitions $\mathcal{H}_w$ for each $w \in [n]$, and add all remaining 1-chains with "tail w", i.e., 1-chains of the form (u, C, w), to $\mathcal{H}_w$.

This decomposition has the following properties:

(1) $\mathcal{H}^{(1)} = (\cup_w \mathcal{H}_w) \bigcup (\cup_{(Q,p)} \mathcal{H}_{Q,p})$ is a disjoint partition of $\mathcal{H}^{(1)}$;

(2) For each $Q = (Q_1, Q_2)$ and $p \in \mathbb{N}$, $\mathcal{H}_{Q,p}$ is a set of 1-chains that "contain" the tuple Q, i.e., each (u, C, w) in $\mathcal{H}_{Q,p}$ has $w = Q_2$ and $C \ni Q_1$;

(3) For each Q and $p \in \mathbb{N}$, $|\mathcal{H}_{Q,p}| = d$;

(4) For each $w \in [n]$, there is only one partition $\mathcal{H}_w$;

(5) The total number of partitions $\mathcal{H}_{Q,p}$ is at most $O(n^2/d)$, as there are at most $O(n^2)$ 1-chains, and each $\mathcal{H}_{Q,p}$ has exactly d 1-chains.

We stress that the decomposition is only on 1-chains, *not* the set of 2-chains $\cup_{i\in[k]}\mathcal{H}_i^{(2)}$ that are the constraints in the XOR instance! At a high level, this is because, e.g., the 2-chains in $\mathcal{H}_i^{(2)}$ (or $\mathcal{H}_j^{(2)}$) are formed by taking a 1-chain and *prepending* it with a hyperedge in H_i (or H_j), and so "first link" in each 2-chain is specific to the choice of $i \in [k]$, but the "second link" is an arbitrary 1-chain, and so it is "shared" across the $\mathcal{H}_i^{(2)}$'s in some informal sense.[9] This property turns out to be important when it comes time to bound the expected partial derivatives.

Now, we define $\mathcal{H}_{i,Q,p}^{(2)}$ to be the set of 2-chains (i, C_0, w_0, C_1, w_1) where the "second link" (w_0, C_1, w_1) is in $\mathcal{H}_{Q,p}$. Using the decomposition, we now define the following polynomials:

$$\Phi_b(x) := \sum_{i=1}^{k} b_i \sum_{\vec{C}=(i,C_0,w_0,C_1,w_1)\in\mathcal{H}_i^{(2)}} x_{C_0} x_{C_1} x_{w_1} \ ,$$

$$\Psi_{i,w}(x) := \sum_{C_0,w_0:C_0\cup\{w_0\}\in H_i} \sum_{(w_0,C_1,w_1)\in\mathcal{H}_w} x_{C_0} x_{C_1} \ ,$$

$$\Psi_{i,Q,p}(x) := \sum_{(i,C_0,w_0,C_1,w_1)\in\mathcal{H}_{i,Q,p}^{(2)}} x_{C_0} x_{C_1\setminus Q_1} \ ,$$

$$\Psi_b^{(0)}(x,y) := \sum_{i=1}^{k} \sum_{w\in[n]} b_i y_w \Psi_{i,w}(x) \ ,$$

$$\Psi_b^{(1)}(x,y) := \sum_{i=1}^{k} \sum_{(Q,p)} b_i y_{Q,p} \Psi_{i,Q,p}(x) \ ,$$

where above $y_{Q,p}$ and y_w are new variables. By definition, if we set $y_w = x_w$ and $y_Q = x_{Q_1} x_{Q_2}$, then we have that $\Phi_b(x) = \Psi^{(0)}(x,y) + \Psi^{(1)}(x,y)$. Indeed, all we have done is partitioned the constraints into these two polynomials and removed the "$x_{Q_1} x_{Q_2}$ term" from each monomial, replacing it with the new variable $y_{Q,p}$.

We now refute the two polynomials $\Psi^{(0)}(x,y)$ and $\Psi^{(1)}(x,y)$ separately using the machinery in Sections 4.1 to 4.3. In fact, Sections 4.1 to 4.3 immediately show that we can successfully refute the polynomial $\Psi^{(0)}(x,y)$. Indeed, the only issue that we encountered was in Section 4.3, where the row pruning failed if there was a pair $\{v,v'\}$ that appeared in more than ℓ^2 1-chains in $\mathcal{H}^{(1)}$. However, this cannot happen, as otherwise our decomposition algorithm would not have terminated.

It thus remains to handle the second polynomial, $\Psi^{(1)}(x,y)$. Applying the "Cauchy–Schwarz trick" of Section 4.1, we can reduce this to the case of bounding the polynomial:

$$f_{M,b}(x) = \sum_{(i,j)\in M} b_i b_j \sum_{(Q,p)} \Psi_{i,Q,p}(x) \Psi_{j,Q,p}(x) \ ,$$

---

[9]For this reason, in Section 6, the length of the chains defining the XOR constraints is $r + 1$, but we only decompose length r chains.
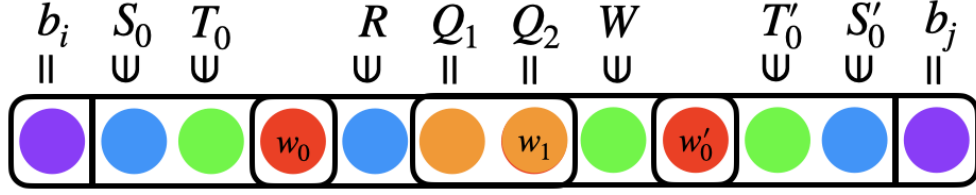
Figure 2: A pair of 2-chains $\vec{C} = (i, C_0, w_0, C_1, w_1) \in \mathcal{H}^{(2)}_{i,Q,p}$, $\vec{C}' = (j, C'_0, w'_0, C'_1, w'_1) \in \mathcal{H}^{(2)}_{j,Q,p}$. The blue vertices appear in the sets (S_0, R, S'_0) for the rows of the matrix $A^{(\vec{C}, \vec{C}')}_{i,j,Q,p}$, and the green vertices appear in the columns. The orange elements are the elements of Q that are canceled via the Cauchy–Schwarz operation.

---

where M is a maximum matching, as before. Notice that the constraints in $f_{M,b}$ have arity 6 (see Fig. 2). Following the blueprint of Section 4.2, we define the following Kikuchi matrices.

Definition 4.5. For $i \neq j \in [k]$, (Q, p), and $\vec{C} = (i, C_0, w_0, C_1, w_1) \in \mathcal{H}^{(2)}_{i,Q,p}$, $\vec{C}' = (j, C'_0, w'_0, C'_1, w'_1) \in \mathcal{H}^{(2)}_{j,Q,p}$, we define the matrix $A^{(\vec{C}, \vec{C}')}_{i,j,Q,p}$ as follows. The matrix $A^{(\vec{C}, \vec{C}')}_{i,j,Q,p}$ is indexed by a 3-tuple of sets (S_0, R, S'_0), each in $\binom{[n]}{\ell}$, and the $(S_0, R, S'_0), (T_0, W, T'_0)$-th entry is 1 if $S_0 \oplus T_0 = C_0$, $S'_0 \oplus T'_0 = C'_0$, and $R = \{u\} \cup U$, $W = \{v\} \cup V$, where $C_1 = \{u, Q_1\}$, $C'_1 = \{v, Q_1\}$, and $U \subseteq [n]$ is a set of size $\ell - 1$ where $u, v \notin U$.

We let $A_{i,j} = \sum_{Q,p} \sum_{\vec{C} \in \mathcal{H}^{(2)}_{i,Q,p}, \vec{C}' \in \mathcal{H}^{(2)}_{j,Q,p}} A^{(\vec{C}, \vec{C}')}_{i,j,Q,p}$ and $A = \sum_{(i,j) \in M} b_i b_j A_{i,j}$.

Notice that for $\vec{C} = (i, C_0, w_0, C_1, w_1) \in \mathcal{H}^{(2)}_{i,Q,p}$ and $\vec{C}' = (j, C'_0, w'_0, C'_1, w'_1) \in \mathcal{H}^{(2)}_{j,Q,p}$, the split of the elements in the constraint across the row (S_0, R, S'_0) and the column (T_0, W, T'_0) is asymmetric: see Fig. 2.

Applying the same machinery in Section 4.2 to the matrices in Definition 4.5 will yield the correct lower bound provided that the row pruning step succeeds. It thus remains to bound the number of rows in $A_{i,j}$ for a fixed pair (i, j) with a number of nonzero entries exceeding the average by a polylog(n) factor.

We now apply Lemma 2.9. As before, we define a similar degree polynomial $\text{Deg}_{i,j}$, and the tail bound boils down to computing the expected partial derivatives μ_Z, where $Z = (z_0, r, z'_0) \in ([n] \cup \{\star\})^3$ is now a tuple of length 3, and $\mu_Z = (\ell/n)^{3-|Z|} \deg_{i,j}(Z)$, as the constraints have arity 3. We observe that $\deg_{i,j}(\star, \star, \star) \leq O(n^2 d)$, as we have $O(n^2)$ choices for $\vec{C} = (i, C_0, w_0, C_1, w_1) \in \mathcal{H}^{(2)}_i$ (which then determines (Q, p)), followed by $O(d)$ choices for (w'_0, C'_1, w'_1) (because this must be in $\mathcal{H}^{(1)}_{Q,p}$, which has size d), and then a unique choice for C_0. Therefore, $\mu_0 \leq (\ell/n)^3 \cdot O(n^2 d) = O(\ell^3 d/n)$.

Bounding μ_1 is straightforward, and we omit the calculations. We obtain a bound of $\mu_1 \leq (\ell/n)^2 \cdot O(nd) = O(\ell^2 d/n)$. Bounding μ_2 can be done with a trivial bound of $\deg_{i,j}(Z) \leq O(n)$,

yielding $\mu_2 \leq (\ell/n) \cdot O(n) = O(\ell)$. Finally, it is simple to bound $\deg_{i,j}(Z) \leq O(1)$ when $|Z| = 3$, and so we obtain $\mu_3 \leq O(1)$.

We notice that $\mu_0 \geq \mu_1$ and $\mu_2 \geq \mu_3$ always hold. So, either μ_0 or μ_2 must be the maximum. Because $d = \ell^2$, we have $\mu_0 = O(\ell^3 d/n) \sim \ell^5/n \gg \ell \sim \mu_2$ because $\ell^4 \gg n$, by choice of ℓ. Thus, $\mu_0 \gg \mu_2$, and so the row pruning argument, etc., will all succeed. This, combined with the refutation argument for $\Psi_b^{(0)}(x)$, implies that our heuristic calculation succeeds and we get a bound of $k \leq \tilde{O}(\ell)$, where ℓ is chosen to be $\tilde{O}(n^{1/4})$. Thus, we obtain a lower bound of $k \leq \tilde{O}(n^{1/4})$. $\qquad \square$

4.5   Preview: extending the warmup to a proof of Theorem 1

We now give a brief overview of how we shall extend the ideas used in this warmup to prove Theorem 1. First, we observe that in the argument we presented in Sections 4.1 to 4.4, there were only two crucial moments in the proof where we had a lot of freedom: (1) the choice of the constraints in the initial XOR instance (in this warmup, we chose the set of 2-chains with head $i \in [k]$), and (2) the choice of the hypergraph decomposition in Section 4.4 — the rest of the proof was fairly mechanical, and boiled down to computing the expected partial derivatives μ_Z. Namely, if we can choose the constraints and the decomposition so that the row pruning succeeds for all the resulting Kikuchi matrices, i.e., the expected partial derivatives of the degree polynomials are appropriately bounded, then the general machinery in Sections 4.1 to 4.3 succeeds in proving the lower bound predicted by the heuristic calculation in Section 3.2 (up to a small loss, see Remark 4.2).

As discussed in Section 3.2, we shall define the XOR instance using $(r + 1)$-chains for a parameter $r = O(\log n)$, and the heuristic calculation predicts that this will yield an exponential lower bound. Thus, the key technical component of the proof is to choose the decomposition of the $(r + 1)$-chains so that the degree polynomials of the resulting Kikuchi matrices all satisfy the bounded expected partial derivatives condition. In Section 4.4, we showed how to do this for the case when $r = 1$.

We now wish to point out the following crucial observation: the decomposition in Section 4.4 is "informed" by the row pruning calculation for the *undecomposed chains* done in Section 4.3. Specifically, in Section 4.3, we argued that if there is a violating partial derivative for the undecomposed chains, then there is some combinatorial structure in the chains (namely, a heavy pair) that is the "cause" of the large expected partial derivative, and this combinatorial structure is exactly the criteria that we use to decompose the hypergraph. In some sense, the hypergraph decomposition (along with the modified Cauchy–Schwarz trick and Kikuchi matrices) can be thought of as a precise way to "fix" this high expected partial derivative. For longer chains, there is once again an intimate relationship between the existence of a violating expected partial derivative and a certain "denser-than-anticipated" combinatorial structure (analogous to heavy pairs) being present in the chains we construct. For larger chains, this structure is a more complicated to describe, but an analogous chain decomposition for this structure accomplishes the same job.

More precisely, we generalize the decomposition of Section 4.4 as follows. As done in Section 4.4, we shall think of an $(r + 1)$-chain in $\mathcal{H}_i^{(r+1)}$ as being split into two subchains, the "first link" in H_i and then the rest of the chain, which is an r-chain. As before, our decomposition shall decompose the r-chain part only, and this induces a decomposition of the $(r + 1)$-chains in $\mathcal{H}_i^{(r+1)}$. Recall that in Section 4.4, we decomposed a 1-chain (u, C, w) by picking a Q where $Q_1 \in C$ and $Q_2 = w$.

Notice that Q only contains one element of the hyperedge C; there was no need to do a further decomposition to handle, e.g., heavy triples $Q = (Q_1, Q_1', Q_2)$ where $\{Q_1, Q_1'\} = C$ and $Q_2 = w$.

Now, we have r-chains $(u, C_1, w_1, \ldots, C_r, w_r)$, and we shall decompose if there is a $Q = (Q_1, \ldots, Q_{r+1}) \in ([n] \cup \{\star\})^r \times [n]$ such that (1) Q is heavy, i.e., is contained in many r-chains, meaning that (a) $Q_{h+1} = w_r$, and so in particular $Q_{h+1} \neq \star$, and (b) $Q_h \in C_h$ for $h = 1, \ldots, r$; and (2) Q is *contiguous*, meaning that if $h \in [r+1]$ is the minimal h such that $Q_h \neq \star$, then $Q_{h'} \neq \star$ for all $h' \geq h$, i.e., Q has $\star$'s followed by only non-$\star$ entries.

Condition (1) above is a somewhat natural extension of the decomposition method in Section 4.4, but condition (2) is trickier. It turns out (in a somewhat subtle way) that because the H_i's are matchings, if there is a violating expected partial derivative, then not only is there a heavy Q, but there must be a heavy *contiguous* Q. In a sense (that can be made precise), the contiguous Q's are *irreducible* violations and thus it is enough to only handle them.

5 Proof of Theorem 1: From LCCs to XOR Formulas

We now present the proof of Theorem 1 for the case of $\mathbb{F} = \mathbb{F}_2$. The proof is spread over Sections 5 to 8 and follows the steps in the warmup. In the current section, we define r-chains and the family of XOR instances associated to the LCC that we wish to refute. Then, in Section 6, we decompose the r-chains, and thereby decompose the $(r+1)$-chains forming the constraints in the XOR instance. Then, in Section 7, we define the Kikuchi matrices and finish the argument up to the proof of the row pruning lemma, Lemma 7.4, an analogue of Lemma 4.4 that is the key technical lemma. Finally, in Section 8, we prove Lemma 7.4.

Let $\mathcal{L} \colon \mathbb{F}_2^k \to \mathbb{F}_2^n$ be $(3, \delta, \varepsilon)$-locally correctable. Without loss of generality, by Fact 2.4 we can assume that $\mathcal{L}$ is $(3, \delta')$-normally decodable, where $\delta' \geq \delta/6$ and $n' = 2n$. For the remainder of the proof, we will redefine δ to be δ', and n to be $2n$. We shall also think of the code $\mathcal{L} \colon \mathbb{F}_2^k \to \mathbb{F}_2^n$ as a map $\mathcal{L} \colon \{-1, 1\}^k \to \{-1, 1\}^n$.

We will now define satisfiable XOR formulas Φ associated with the linear code $\mathcal{L}$. Let $\mathcal{L} \colon \{-1, 1\}^k \to \{-1, 1\}^n$ be a linear $(3, \delta)$-normally correctable code. Recall that without loss of generality, $\mathcal{L}$ is systematic, meaning that the first k bits of $\mathcal{L}$ are the message bits. In particular, for every $b \in \{-1, 1\}^k$, there is a unique $x \in \mathcal{L}$ such that $x|_{[k]} = b$. We can thus generate $x \leftarrow \mathcal{L}$ uniformly at random by first choosing $b \leftarrow \{-1, 1\}^k$ uniformly at random, and then setting x to be the unique extension of b.

Since $\mathcal{L}$ is a linear $(3, \delta)$-normally correctable code, there exist 3-uniform hypergraph matchings $H_1, \ldots, H_n$, each of size exactly δn, such that every $x \in \mathcal{L}$ satisfies the following system of 4-XOR constraints, i.e., each constraint has arity 4:

$$\forall u \in [n], C \in H_u, \ x_C x_u = 1. \tag{3}$$

We will construct an XOR formula by *long chain* derivations. Intuitively, a long chain derivation starts from the natural XOR constraints (3) and derives new ones by chaining together t constraints with an appropriate combinatorial structure. Below, we formalize the set of constraints in this formula as a family of hypergraphs built from the H_u's.
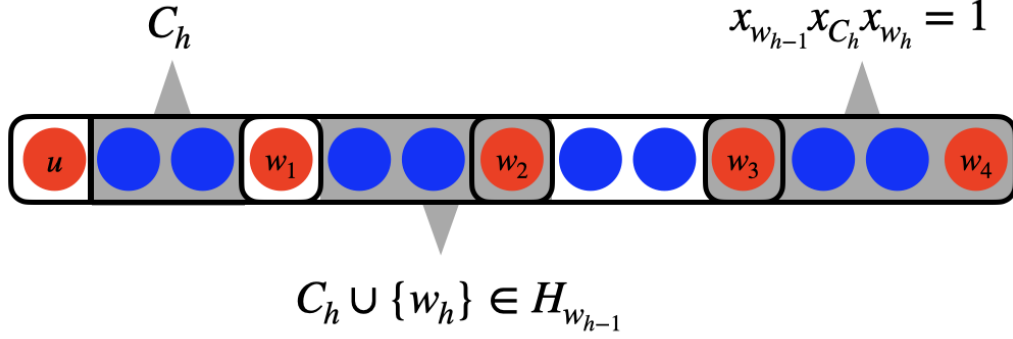
$$C_h \qquad\qquad\qquad\qquad x_{w_{h-1}} x_{C_h} x_{w_h} = 1$$

$$C_h \cup \{w_h\} \in H_{w_{h-1}}$$

Figure 3: A 4-chain. The pairs of blue vertices are the C_h's, and the red vertices are the w_h's. Note that for any $x \in \mathcal{L}$, we have $x_{w_{h-1}} x_{C_h} x_{w_h} = 1$.

Definition 5.1 (*t*-chain hypergraph $\mathcal{H}^{(t)}$). Let $t \geq 1$ be an integer. For any $u \in [n]$, let $\mathcal{H}_u^{(t)}$ denote the set of tuples of the form $(u, C_1, w_1, C_2, w_2, \ldots, C_t, w_t)$, where each $C_h \in \binom{[n]}{2}$, $w_h \in [n]$, and it holds that for all $1 \leq h \leq t$, $C_h \cup \{w_h\} \in H_{w_{h-1}}$ where we set $w_0 := u$. We call u the head, w_h's the *pivots* for $1 \leq h \leq t-1$ and w_t the *tail* in such a chain. We let $\mathcal{H}^{(t)} = \cup_{u \in [n]} \mathcal{H}_u^{(t)}$ denote the set of all t-chains, where $\mathcal{H}_u^{(t)}$ is the set of t-chains with head u.

The following simple observation helps us understand the combinatorial structure in the chains.

Observation 5.2. Let $x = \mathcal{L}(b)$ for a linear LCC over $\mathbb{F}_2$ with $\{H_u\}_{u \in [n]}$ being the associated matchings. Then, for any t-chain $(u, C_1, w_1, C_2, w_2, \ldots, C_t, w_t)$, x satisfies $x_u x_{w_t} \prod_{h=1}^{t} x_{C_h} = 1$.

Proof. We know that x satisfies $x_{w_h} x_{C_{h+1}} x_{w_{h+1}} = 1$ for every $0 \leq h \leq t$ where we define $w_0 = u$. Taking products of the left-hand sides of each of these t equations, we observe that for every $1 \leq h \leq t-1$, x_{w_h} is "squared out" (since $x_v^2 = 1$ for every $v \in [n]$), and this finishes the proof. $\qquad\square$

Building chains iteratively. It is useful to think of t-chains as being built by extending smaller chains by iteratively adding hyperedges to the head (i.e. to the left). The following notation and observation formalizes this.

Definition 5.3 (Extending Chains). For the t-chain hypergraph $\mathcal{H}^{(t)}$ built from 3-matchings $H_1, H_2, \ldots, H_n$ on $[n]$, we define $H_u \circ \mathcal{H}^{(t+1)}$ as:

$$H_u \circ \mathcal{H}^{(t)} = \cup_{w_0 \in [n]} \left\{ (u, C_0, \vec{C} \mid \vec{C} \in \mathcal{H}_{w_0}^{(t)}, \{C_0 \cup \{w_0\} \in H_u \right\} \ .$$

Observation 5.4. For $t \geq 1$, let $\mathcal{H}^{(t)}$ be the t-chain hypergraph built from 3-matchings $H_1, H_2, \ldots, H_n$ on $[n]$. Then, $\mathcal{H}^{(t+1)} = \cup_{u \in [n]} H_u \circ \mathcal{H}^{(t)} = \cup_{u \in [n]} \mathcal{H}_u^{(t')} \circ \mathcal{H}^{(t-t')}$ for any $0 < t' < t$.

Chains that fix some positions. We will often refer to the set of chains where some of the C_h's are forced to contain some $v_h \in [n]$. Towards this, we introduce the following terminology.

Definition 5.5 (Chains containing Q). For any $Q = (Q_1, \ldots, Q_t, Q_{t+1}) \in \{[n] \cup \star\}^t \times [n]$, we say that a chain $(u, C_1, w_1, \ldots, C_t, w_t) \in \mathcal{H}^{(t)}$ *contains* Q if $Q_{t+1} = w_t$ and for $1 \le h \le t$, if $Q_h \ne \star$, then $Q_h \in C_h$. We say that a Q is *contiguous* if there exists $s \le t$ such that $Q_h \ne \star$ for every $h \ge s+1$ and $Q_h = \star$ for every $1 \le h \le s$, i.e., the first s entries are $\star$, and the remaining entries are non-$\star$. We note that by definition, $Q_{t+1} \ne \star$ always.

We say that Q is *complete* if Q does not contain any $\star$. We say that $Q' \supseteq Q$ if whenever $Q_h \ne \star$, $Q'_h = Q_h$. We define the size $|Q|$ to be the number of coordinates in Q that do not equal $\star$.

We write $\mathcal{H}_Q^{(t)}$ to denote the set of all t-chains that contain Q, and for $u \in [n]$, we write $\mathcal{H}_{u,Q}^{(t)}$ to denote the set of t-chains with head u that contain Q.

We caution the reader that $\mathcal{H}_u^{(t)}$ and $\mathcal{H}_Q^{(t)}$, are different sets of chains. In context, it shall be easy to distinguish between the two cases as the type of u and Q are different: namely, we have $u \in [n]$ and $Q \in \{[n] \cup \star\}^{t+1}$.

XOR Formulas from r-chains. Next, we define XOR formulas associated with $\mathcal{H}^{(r+1)}$ that are guaranteed to be satisfiable. The length of the chain depends on a parameter r, which we shall set later.

Definition 5.6 (The XOR Formula Φ). Fix $r \in \mathbb{N}$.

For any $b = (b_1, \ldots, b_k) \in \{-1, 1\}^k$, define the polynomial Φ_b:

$$\Phi_b(x) = \sum_{i=1}^{k} b_i \sum_{(i, C_0, w_0, C_1, w_1 \ldots, C_r, w_r) \in \mathcal{H}^{(r+1)}} x_{w_r} \prod_{h=0}^{r} x_{C_h} = \sum_{i=1}^{k} b_i \sum_{C_0, w_0 : C_0 \cup \{w_0\} \in H_i} x_{C_0} \sum_{(w_0, C_1, w_1, \ldots, C_r, w_r) \in \mathcal{H}^{(r)}} x_{w_r} \prod_{h=1}^{r} x_{C_h} .$$

We will drop the subscript b when it is clear from the context.

We note the equality holds above as we are simply thinking of the chain $(i, C_0, w_0, \ldots, C_r, w_r)$ as being split into two parts, the 1-chain (i, C_0, w_0), followed by the r-chain $(w_0, C_1, w_1, \ldots, C_r, w_r)$. We write the polynomial in this form because for much of the proof, we shall wish to think of the r-chain as separate from the 1-chain (i, C_0, w_0).

We now observe that $\Phi_b(x)$ is satisfiable and thus has a high value.

Lemma 5.7. *For every $b \in \{-1, 1\}^k$, Φ_b is satisfied by $x = \mathcal{L}(b)$ and thus, $\mathrm{val}(\Phi_b) = k(3\delta n)^{r+1}$.*

Proof. Observe that Φ_b is a sum of monomials corresponding to a r-chain each of which is satisfied by $x = \mathcal{L}(b)$ by Observation 5.2. Thus, $\mathrm{val}(\Phi_b)$ equals the total number of chains of length r with head in $[k]$, which we next count.

Define $w_{-1} := i$. Given w_{h-1} for $h \ge 0$, there are δn choices for the set $C_h \cup \{w_h\} \in H_h$ and for each such choice, there are 3 choices for the next pivot w_h. Thus, the number of $(r+1)$-chains with head i is $(3\delta n)^{r+1}$. Summing over the k possible heads completes the proof. $\square$

6 Contiguously Regular Partition of Chains

In this section, we partition the r-chain hypergraph $\mathcal{H}^{(r)}$ into buckets that satisfy a useful regularity property. We first abstract out the relevant properties of the partition below and then show how it

can be using a simple greedy partitioning algorithm. This partitioning will be key to setting up and analyzing our spectral refutation in the next section.

Definition 6.1 (Contiguously regular partition). For $\delta > 0$ and $r \in \mathbb{N}$, let $\mathcal{H}^{(r)}$ be the r-chain hypergraph built form 3-matchings $H_1, H_2, \ldots, H_n$ on $[n]$ of size δn each. Let $\mathcal{H}^{(r)} = \cup_{Q,p} \mathcal{H}^{(r)}_{Q,p}$ be a disjoint partition of $\mathcal{H}^{(r)}$ indexed by $Q \in ([n] \cup \{\star\})^r \times [n]$ and $p \in [m]$ for some large enough $m \in \mathbb{N}$. We say that such a partition is d-contiguously regular if the following conditions hold:

(1) for every $Q \in ([n] \cup \{\star\})^r \times [n]$ and $p \in [m]$, $\mathcal{H}^{(r)}_{Q,p} \subseteq \mathcal{H}^{(r)}_Q$,

(2) for every (Q, p) such that $\mathcal{H}^{(r)}_{Q,p} \neq \emptyset$, Q is contiguous,

(3) if $|Q| = 1$, then $\mathcal{H}^{(r)}_{Q,p} = \emptyset$ whenever $p > 1$,

(4) for every contiguous Q, Q' such that $Q' \supseteq Q$,

$$\left| \left\{ \vec{C}' \in \mathcal{H}^{|Q'|-1} \mid \vec{C}' \text{ contains } Q', \text{ and } \exists \vec{C} \text{ extending } \vec{C}', \vec{C} \in \mathcal{H}^{(r)}_{Q,p} \right\} \right| \leq d^{|Q'|-1} .$$

(5) For every t, the set P_t of all (Q, p) such that $\mathcal{H}^{(r)}_{Q,p} \neq \emptyset$ and $|Q| = t + 1$ satisfies $|P_t| d^t \leq n(3\delta n)^t$. Observe that $|P_0| \leq n$ is forced by (3).

We now give a bit of intuition for the definition. A contiguously regular partition takes the set of r-chains $\mathcal{H}^{(r)}$ and decomposes it into pieces, where the pieces are intuitively indexed by Q; however, for technical reasons, we will want to have multiple pieces assigned to the *same* Q, and so we disambiguate these pieces using the label p, i.e., we can have pieces $\mathcal{H}^{(r)}_{Q,1}, \mathcal{H}^{(r)}_{Q,2}, \mathcal{H}^{(r)}_{Q,3}$, etc.

Condition (1) says that the chains in the piece $\mathcal{H}^{(r)}_{Q,p}$ in the decomposition are all chains that contain Q, hence why we view them as indexed by Q. Condition (2) says that the only nonempty pieces have a *contiguous* Q, hence the name "contiguously regular partition". Condition (3) says that if $|Q| = 1$, then there is only one piece with this Q. Recall that when $|Q| \geq 2$, we can have pieces $\mathcal{H}^{(r)}_{Q,1}, \mathcal{H}^{(r)}_{Q,2}, \mathcal{H}^{(r)}_{Q,3}$, etc.; we have asserted that when $|Q| = 1$, this does not happen. Condition (4) is a *regularity* condition saying that chains in $\mathcal{H}^{(r)}_{Q,p}$ appear in this piece because the tuple Q is "maximal". Condition (5) asserts that the number of pieces with $|Q|$ of a given size is not too large.

We now make the following observation.

Observation 6.2. Items (1), (2), and (4) imply that $|\mathcal{H}^{(r)}_{Q,p}| \leq n(3\delta n)^{r-|Q|} d^{|Q|-1}$ for all (Q, p).

Moreover, let $\mathcal{H}^{(r)}_{u,Q,p}$ denote the set of chains in $\mathcal{H}^{(r)}_{Q,p}$ with head u. Then, Items (1), (2) and (4) imply that $|\mathcal{H}^{(r)}_{u,Q,p}| \leq (3\delta n)^{r-|Q|} d^{|Q|-1}$ when $|Q| \leq r$.

Proof. To see this, we apply item (4) with $Q' = Q$, and we now count the chains in $\mathcal{H}^{(r)}_{Q,p}$ by (1) first choosing a suffix $\vec{C}' \in \mathcal{H}^{|Q|-1}_{|Q|}$, and then (2) completing the chain. By Item (4), we have at most $d^{|Q|-1}$ choices for the suffix. Once the suffix is fixed, we now complete the chain as follows. If $|Q| = r + 1$, then we have chosen the entire chain and are done. Otherwise, we do the following. First, we choose

w_0, which has n choices. Then, we choose $C_1 \cup \{w_1\} \in H_{w_0}$, which has δn choices, followed by $w_1 \in C_1 \cup \{w_1\}$, which has 3 choices. We repeat this until we reach the point in the chain where we have determined $w_{r-|Q|}$. Because we also know the suffix $\vec{C}'$, we have already determined $w_{r+1-|Q|}$. Because $H_{w_{r-|Q|}}$ is matching, there is at most one $C_{r+1-|Q|}$ such that $C_{r+1-|Q|} \cup \{w_{r+1-|Q|}\} \in H_{w_{r-|Q|}}$, and so we have determined the entire chain. We have thus made at most $n(3\delta n)^{r-|Q|}d^{|Q|-1}$ choices when $|Q| \le r$, and d^r choices if $|Q| = r + 1$. In both cases, this is at most $n(3\delta n)^{r-|Q|}d^{|Q|-1}$.

Finally, we note that for $|Q| \le r$, the above argument also bounds $|\mathcal{H}^{(r)}_{u,Q,p}|$. We simply save a factor of n because w_0 must be equal to u. □

We now give an algorithm that, given $\mathcal{H}^{(r)}$ and d, outputs a d-contiguously regular partition of $\mathcal{H}^{(r)}$ using a simple iterative greedy scheme.

Lemma 6.3 (Contiguously regular partition of chains). *For $\delta > 0$, let $H_1, H_2, \ldots, H_n$ be arbitrary δn size 3-matchings on $[n]$. For $r \in \mathbb{N}$, let $\mathcal{H}^{(r)}$ be the r-chain hypergraph built from $H_1, H_2, \ldots, H_n$. Then, for every $d \in \mathbb{N}$, there exists a d-contiguously regular partition $\mathcal{H}^{(r)} = \cup_{Q,p}\mathcal{H}^{(r)}_{Q,p}$.*

Proof. The greedy algorithm that computes the decomposition is given below.

Algorithm 6.4.

Given: An r-chain hypergraph $\mathcal{H}^{(r)}$.

Output: A contiguously d-regular partition $\mathcal{H}^{(r)} = \cup_{Q,p}\mathcal{H}^{(r)}_{Q,p}$.

Operation:

1. **Initialize:** For $Q = (w)$ for each $w \in [n]$, let $\mathcal{H}^{(0)}_{Q,1} = \{(w)\}$, i.e., the set of 0-chains with tail w.

2. **Iterative Greedy Fixing:** For $t = 1, \ldots, r$, do:
   (a) Initialize $\mathcal{H}^{(t)}_{(\star,Q),p} = \cup_{u \in [n]}H_u \circ \mathcal{H}^{(t-1)}_{Q,p}$ for every $Q \in ([n] \cup \{\star\})^{t-1} \times [n]$.
   (b) For every $Q' = (u, Q)$ for $Q \in [n]^{t-1} \times [n]$, initialize $p' = 1$ and do:
      i. Let $\mathcal{R} = \mathcal{H}^{(t)}_{(\star,Q),p} \cap \{\vec{C} \in \mathcal{H}^{(t)} \mid \vec{C} \text{ contains } Q'\}$. If $|\mathcal{R}| \le d^{|Q'|-1}$, end.
      ii. Otherwise, select exactly $d^{|Q'|-1}$ t-chains from $\mathcal{R}$, remove them from $\mathcal{H}^{(t)}_{(\star,Q),p}$, and put them in a new piece $\mathcal{H}^{(t)}_{Q',p'}$.
      iii. Set $p' = p' + 1$.

We now verify that our decomposition satisfies the properties required of a contiguously d-regular partition. The key observation is that the algorithm iterates over $t = 1, \ldots, r$, and computes, after the t-th iteration, a d-contiguously regular partition of $\mathcal{H}^{(t)}$. Indeed, we prove this by induction. For $t = 0$ this trivially holds.

We now show the induction step. We observe that properties (1) and (2) are trivial. Property (3) holds because of the following. We observe that the pieces in the decomposition of $\mathcal{H}^{(t)}$ are either obtained by extending "old" pieces to get $\mathcal{H}^{(t)}_{(\star,Q),p} = \cup_{u \in [n]}H_u \circ \mathcal{H}^{t-1}_{Q,p}$, or by adding "new" pieces

produced in step (2b). We note that we only produce new pieces for Q with $|Q| \geq 2$, so we cannot violate property (3). Property (4) follows because the loop in step (2bi) finished.

Finally, to check property (5), we need to bound $|P_{t'}|$ for $0 \leq t' \leq t$. As $|P_0| \leq n$ always holds, it remains to bound $|P_{t'}|$ for $t' \leq t$. We note that all the "new" pieces have a Q where $|Q| = t + 1$. Hence, for $t' \leq t - 1$, $|P_{t'}|$ satisfies property (5) by the induction hypothesis. To bound $|P_t|$, we observe that each new partition contains $d^{|Q'|} = d^t$ chains. As $\mathcal{H}^{(t)}$ has at most $n(3\delta n)^t$ chains (see, e.g., the proof of Lemma 5.7), the bound on $|P_t|$ follows. $\qquad \square$

Every d-contiguously regular partition of $\mathcal{H}^{(r)}$ naturally relates to a "bipartite" polynomial Ψ (i.e., Ψ has additional variables $y_{Q,p}$ corresponding to labels of the buckets in the partition in addition to the original variables $x \in \{-1, 1\}^n$) such that $\mathrm{val}(\Psi)$ upper bounds $\mathrm{val}(\Phi)$. Our main technical argument will construct a spectral refutation to upper bound $\mathrm{val}(\Psi)$ for a d-contiguously regular partition of $\mathcal{H}^{(r)}$ for an appropriate choice of d.

Definition 6.5 (Bipartite XOR Formulas from a contiguously regular partition). Fix $r, d \in \mathbb{N}$ and for the r-chain hypergraph $\mathcal{H}^{(r)}$ built from 3-matchings $H_1, H_2, \ldots, H_n$ on $[n]$ of size δn each, let $\mathcal{H}^{(r)} = \cup_{p \in P, Q \in ([n] \cup \{\star\})^r \times [n]} \mathcal{H}^{(r)}_{Q,p}$ be a contiguously d-regular partition. For each nontrivial piece $\mathcal{H}^{(r)}_{Q,p}$, we define $\Psi_{i,Q,p}$ as the following XOR formula with terms corresponding $(r+1)$-chains obtained by (1) taking r-chains from a single piece $\mathcal{H}^{(r)}_{Q,p}$ with x_Q "modded out" from the corresponding monomial and (2) *joining* with a 1-chain (i, C_0, w_0). Namely,

$$\Psi_{i,Q,p}(x) = \sum_{C_0,w_0 : C_0 \cup \{w_0\} \in H_i} \quad \sum_{(w_0, C_1, w_1, C_2, w_2, \ldots, C_r, w_r) \in \mathcal{H}^{(r)}_{Q,p}} x_{C_0} x_{w_r \setminus Q_{r+1}} \prod_{h=1}^{r} x_{C_h \setminus Q_h} \ .$$

Here, we use the convention that if $Q_h = \star$, then $C_h \setminus Q_h := C_h$. We note that because $w_r = Q_{r+1}$, we have $x_{w_r \setminus Q_{r+1}} = 1$.

For each $0 \leq t \leq r$, let $\Psi^{(t)}(x, y) = \sum_{i=1}^{k} \sum_{(Q,p) \in P_t} b_i y_{Q,p} \Psi_{i,Q,p}(x)$. Finally, we let $\Psi(x, y) = \sum_{0 \leq t \leq r} \Psi^{(t)}(x, y)$; here, for every piece $\mathcal{H}^{(r)}_{Q,p}$ in the contiguously regular partition, we introduce a new variable $y_{Q,p}$.

We next observe that Ψ is satisfiable and thus has a large value for every $b \in \{-1, 1\}^k$. Indeed, the observation is that we have replaced the monomial x_Q in Φ with a new variable $y_{Q,p}$ for each (Q, p).

Lemma 6.6. *Let $\mathcal{H}^{(r)} = \cup_{p \in P, Q} \mathcal{H}^{(r)}_{Q,p}$ be a contiguously d-regular partition. Fix $b \in \{-1, 1\}^k$ and $x \in \{-1, 1\}^n$. Then, there is a y such that $\Psi(x, y) = \Phi(x)$. In particular, setting $x = \mathcal{L}(b)$, we have that $\mathrm{val}(\Psi(x, y)) \geq \mathrm{val}(\Phi(x)) \geq k(3\delta n)^{r+1}$.*

We note that the system of equations in $\Phi(x)$ is satisfiable, and so $\mathrm{val}(\Phi)$ is simply the number of constraints in the instance.

Proof. Set $y_{Q,p} = x_Q$ for every (Q, p), where $x_Q := \prod_{h : Q_h \neq \star} x_{Q_h}$. $\qquad \square$

28

For intuition, we observe that for random matchings $H_1, \ldots, H_n$, the $O(\log n)$-contiguously regular partition is the trivial one.

Lemma 6.7 (Trivial partition is regular for random matchings). *Let $H_1, H_2, \ldots, H_n$ be uniformly random and independent 3-matchings on $[n]$. Then, the trivial partition of the associated r-chain hypergraph $\mathcal{H}^{(r)}$, where we set $\mathcal{H}^{(r)}_{Q,p}$ to be the set of all chains with tail w if $Q = (\star, \ldots, \star, w)$ and $p = 1$, and empty otherwise, is $O(\log n)$-regular with probability at least $1 - 1/n$ over the draw of H_i's.*

Remark 6.8. Eventually (in Lemma 7.4 and Sections 7.5 and 7.6), we will set the parameter d to be constant. However, if the matchings $H_1, \ldots, H_n$ are random, then with high probability the trivial partition will *not* be $O(1)$-regular. However, if we run Algorithm 6.4 to decompose the r-chains, then with high probability over the draw of $H_1, \ldots, H_n$, only a $o(1)$-fraction of the r-chains will be placed in a "non-trivial component" of the decomposition, i.e., in a piece $\mathcal{H}^{(r)}_{Q,p}$ where $|Q| \geq 2$. Phrased differently, if we discard a $o(1)$-fraction of hyperedges from the random matchings, then the trivial partition of the r-chain hypergraph of the remaining hyperedges will be $O(1)$-regular. This fact is somewhat analogous to the fact that sparse random graphs are not, e.g., triangle-free with high probability, but can be made triangle-free by removing a very small number of edges.

Proof. We claim that the trivial partition of $\mathcal{H}^{(r)}$ is d-regular for $d = O(\log n)$ with probability at least $1 - 1/n$. In the trivial refinement, as defined above, we partition $\mathcal{H}^{(r)}$ by simply placing a chain in $\mathcal{H}^{(r)}_{Q,p}$ if $Q = (\star, \ldots, \star, w)$, $p = 1$, and the tail of the chain is $Q_{r+1} = w$.

Towards this, we first prove that for every pair $u, v \subseteq [n]$, the number of hyperedges in the multiset $\cup_{u \in [n]} H_i$ that contain u and v is at most $O(\log n)$. To see this, observe that the chance that there are some u, v that co-occur in a hyperedge in at least b different H_i's is at most $n^2 \binom{n}{b}(3/n)^b \leq 3^b/b! \leq 1/n$ if $b = c \log_2 n$ for some large enough $c > 0$. We will now set $d_2 = 2c \log_2 n$ and confirm d-regularity of the trivial refinement.

Now take any contiguous $Q \in ([n] \cup \{\star\})^r \times [n]$ of size $|Q| = t + 1$. Consider the chains $\vec{C'} = (w_0, C_1, w_1, \ldots, C_t, w_t) \in \mathcal{H}^{(t)}$ that contain Q. We now iteratively choose $(w_{t-1}, C_t, w_t), (w_{t-2}, C_{t-1}, w_{t-1}), \ldots, (w_0, C_1, w_1)$. Assuming we have made the first h choices in the list, we have determined w_{t-h-1}. There are at most $c \log_2 n$ choices for a hyperedge in any of H_is that contains Q_{t-h-1} and w_{t-h-1} and given this choice, at most 2 choices for w_{t-h}. So in total, we have at most $(2c \log_2 n)^t = d^t$ choices. □

7 Spectral Refutation via Kikuchi Matrices

In Section 6, we defined polynomials $\Psi^{(t)}(x, y)$ such that $\mathbb{E}_b[\text{val}(\Phi)] \leq \sum_{t=0}^r \mathbb{E}_b[\text{val}(\Psi^{(t)})]$. Thus, to prove Theorem 1, we need to upper bound $\mathbb{E}_b[\text{val}(\Psi^{(t)})]$ for each t. In this section, we will define, for each $0 \leq t \leq r$, a Kikuchi matrix $A^{(t)}$ such that $\mathbb{E}_b[\text{val}(\Psi^{(t)})^2] \leq \|A^{(t)}\|_{\infty \to 1}$. Then, in Section 8 we shall bound $\|A^{(t)}\|_{\infty \to 1}$ and finish the proof.

7.1 Step 1: the Cauchy–Schwarz trick

First, we show that we can relate $\Psi^{(t)}(x, y)$ to a certain "cross-term" polynomial f_M obtained via applying the Cauchy–Schwarz inequality.

Lemma 7.1 (Cauchy–Schwarz trick). *Let M be a maximum directed matching[10][11] of $[k]$ and let f_M be the cross-term polynomial defined as*

$$f_M(x) = f_M^{(t)} = \sum_{\{i,j\} \in M} b_i b_j \sum_{(Q,p) \in P_t} \Psi_{i,Q,p}(x) \Psi_{j,Q,p}(x).$$

Then,

$$\mathbb{E}_{b \leftarrow \{-1,1\}^k} \mathrm{val}(\Psi^{(t)}) \leq k \left(|P_t| (3\delta n)^{r-t} d^t \right)^2 + |P_t| 2k \mathbb{E}_M \mathbb{E}_{b \leftarrow \{-1,1\}^k} [\mathrm{val}(f_M^{(t)})],$$

where the expectation $\mathbb{E}_M$ is over a uniformly random maximum matching M.

Proof. We will first apply the Cauchy–Schwarz inequality to eliminate the y variables:

$$\Psi^{(t)}(x,y)^2 = \left(\sum_{(Q,p) \in P_t} y_{Q,p} \left(\sum_{i=1}^{k} b_i \Psi_{i,Q,p} \right) \right)^2$$

$$\leq \left(\sum_{(Q,p) \in P_t} y_{Q,p}^2 \right) \left(\sum_{(Q,p) \in P_t} \left(\sum_{i=1}^{k} b_i \Psi_{i,Q,p} \right)^2 \right)$$

$$= |P_t| \left(\sum_{(Q,p) \in P_t} \sum_{i=1}^{k} \Psi_{i,Q,p}^2 + \sum_{(Q,p) \in P_t} \sum_{i \neq j \in [k]} b_i b_j \Psi_{i,Q,p}(x) \Psi_{j,Q,p}(x) \right).$$

Observe that $|\Psi_{i,Q,p}(x)|$ is at most the number of $(w_0, C_1, w_1, \ldots, C_r, w_r) \in \mathcal{H}_{Q,p}^{(r)}$ and $C_0 \in \binom{[n]}{2}$ such that $C \cup \{w_0\} \in H_i$. If $|Q| = r + 1$, i.e., $t = r$, then we observe that by Observation 6.2, we have $|\mathcal{H}_{Q,p}^{(r)}| \leq d^{|Q|-1}$, and for each choice of $(w_0, C_1, w_1, \ldots, C_r, w_r) \in \mathcal{H}_{Q,p}^{(r)}$, there is at most one choice of C_0. If $|Q| \leq r$, then we have at most $(3\delta n)$ choices for (C_0, w_0), and for each w_0, we have by Observation 6.2 that $|\mathcal{H}_{w_0,Q,p}^{(r)}| \leq (3\delta n)^{r-|Q|} d^{|Q|-1}$, giving us $(3\delta n)^{r-|Q|+1} d^{|Q|-1}$ choices in total. We thus have that $|\Psi_{i,Q,p}(x)| \leq (3\delta n)^{r-|Q|+1} d^{|Q|-1}$, regardless of $|Q| = t + 1$.

Thus, for $|Q| = t + 1$, $\sum_{Q,p} \sum_{i=1}^{k} \Psi_{i,Q,p}^2 \leq k|P_t| \left((3\delta n)^{r-t} d^t \right)^2$. This gives us an upper bound of $k|P_t| \left((3\delta n)^{r-t} d^t \right)^2$ on the first term.

Let's now analyze the second term. Since a uniformly random maximum matching on $[k]$ contains a (directed) edge (i,j) with probability exactly $\frac{1}{2(k-1)}$ if k is even, and $\frac{1}{2k}$, if k is odd, we have:

$$\sum_{i \neq j \in [k]} b_i b_j \sum_{(Q,p) \in P_t} \Psi_{i,Q,p}(x) \Psi_{j,Q,p}(x) \leq 2k \mathbb{E}_M \sum_{(i,j) \in M} b_i b_j \sum_{(Q,p) \in P_t} \Psi_{i,Q,p}(x) \Psi_{j,Q,p}(x) = 2k \mathbb{E}_M [f_M^{(t)}].$$

Using that $\mathrm{val}(\mathbb{E}_M[f_M^{(t)}]) \leq \mathbb{E}_M[\mathrm{val}(f_M^{(t)})]$ completes the proof. $\square$

---

[10] A directed matching is a matching, only the edges are additionally directed

[11] This is a perfect matching if k is even, and will leave one element of $[k]$ unmatched if k is odd.

7.2 Step 2: defining the Kikuchi matrices

It thus remains to bound $\mathrm{val}(f_M)$ for an arbitrary directed maximum matching M.

For $i \in [k]$ and (Q,p), we let $\mathcal{H}_{i,Q,p}^{(r+1)}$ denote the set of chains in $\mathcal{H}^{(r+1)}$ of the form $(i, C_0, w_0, C_1, w_1, \ldots, C_r, w_r)$ where $(w_0, C_1, w_1, \ldots, C_r, w_r) \in \mathcal{H}_{Q,p}^{(r)}$. We define the Kikuchi matrices that we consider below.

Definition 7.2 (Kikuchi matrices for a fixed t). Let $i, j \in [k]$.

Let $\vec{C} = (i, C_0, w_0, C_1, w_1, C_2, w_2, \ldots, C_r, w_r) \in \mathcal{H}_{i,Q,z}^{(r+1)}$ and $\vec{C}' = (j, C_0', w_0', C_1', w_1', C_2', w_2', \ldots, C_r', w_r') \in \mathcal{H}_{j,Q,z}^{(r+1)}$. For Q of size $|Q| = t+1$, we let $A_{i,j,Q,p}^{(\vec{C},\vec{C}')} \in \{0,1\}^{\binom{[n]}{\ell}^{2r+2-t}}$ be the matrix with rows and columns by indexed by $(2r+2-t)$-tuples of sets $(S_0, \ldots, S_{r-t}, S_0', \ldots, S_{r-t}', R_1, \ldots, R_t)$ of size exactly ℓ. Note that when $t = 0$, we do not have any "R_h's" in the row/column index tuples.

We set $A_{i,j,Q,p}^{(\vec{C},\vec{C}')}((S_0, \ldots, S_{r-t}, S_0', \ldots, S_{r-t}', R_1, \ldots, R_t), (T_0, \ldots, T_{r-t}, T_0', \ldots T_{r-t}', W_1, \ldots, W_t))$ equal to 1 if the following holds, and otherwise we set this entry to be 0.

1. For $h = 0, \ldots, r - t$, we have $S_h \oplus T_h = C_h$,

2. For $h = 0, \ldots, r - t$, we have $S_h' \oplus T_h' = C_h'$,

3. For $h = 1, \ldots, t$, we have $R_h = \{u\} \cup U$, $W_h = \{v\} \cup U$, where $C_{r-t+h} = \{u, Q_h\}$, $C_{r-t+h}' = \{v, Q_h\}$, and $U \subseteq [n]$ is a set of size $\ell - 1$ with $u, v \notin U$.

We let $A_{i,j} = \sum_{(Q,p)\in P_t} \sum_{\vec{C}\in\mathcal{H}_{i,Q,p}^{(r+1)}, \vec{C}'\in\mathcal{H}_{j,Q,p}^{(r+1)}} A_{i,j,Q,p}^{(\vec{C},\vec{C}')}$, and for any matching M on $[k]$, let $A_M = \sum_{(i,j)\in M} b_i b_j A_{i,j}$.

7.3 Step 3: relating the "Cauchy–Schwarzed" polynomial f_M and the Kikuchi matrix A

The following lemma shows that we can express $f_M(x)$ as a (scaling of a) quadratic form on the matrix A.

Lemma 7.3. Let $x \in \{-1,1\}^n$, and let $x' \in \{-1,1\}^N$, where $N = \binom{n}{\ell}^{2r+2-t}$, denote the vector where the $(S_0, S_1, \ldots, S_{r-t}, S_0', S_1', \ldots, S_{r-t}', R_1, \ldots, R_t)$-th entry of x' is $\prod_{h=0}^{r-t} x_{S_h} x_{S_h'} \prod_{h=1}^{t} x_{R_h}$. Then, $x'^\top A x' = D f_M(x)$, where $D = 2^{2r+2-2t} \binom{n-2}{\ell-1}^{2r+2-t}$. Note that $D/N = 2^{2r+2-2t} \left(\frac{\ell(n-\ell)}{n(n-1)}\right)^{2r+2-t}$. In particular, $\mathrm{val}(f_M) \leq \frac{1}{D} \|A\|_{\infty\to 1}$.

Proof. Expanding definitions, we have

$$x'^\top A x' = \sum_{(i,j)\in M} b_i b_j \sum_{(Q,p)\in P_t} \sum_{\vec{C}\in\mathcal{H}_{i,Q,p}^{(r+1)}, \vec{C}'\in\mathcal{H}_{j,Q,p}^{(r+1)}} x'^\top A_{i,j,Q,p}^{(\vec{C},\vec{C}')} x' \ ,$$

$$f_M(x) = \sum_{(i,j)\in M} b_i b_j \sum_{(Q,p)\in P_t} \Psi_{i,Q,p}(x) \Psi_{j,Q,p}(x) \ ,$$

where we recall that

$$\Psi_{i,Q,p}(x) = \sum_{C_0, w_0 : C_0 \cup \{w_0\} \in H_i} x_{C_0} \sum_{(w_0, C_1, w_1, C_2, w_2, \ldots, C_r, w_r)\in\mathcal{H}_{Q,p}^{(r)}} \prod_{h=1}^{r-t} x_{C_h} \prod_{h=1}^{t} x_{C_{r-t+h}\setminus Q_h}$$

31

$$= \sum_{\vec{C} \in \mathcal{H}_{i,Q,p}^{(r+1)}} \prod_{h=0}^{r-t} x_{C_h} \prod_{h=1}^{t} x_{C_{r-t+h} \setminus Q_h}$$

Thus, it suffices to show that

$$x'^{\top} A_{i,j,Q,p}^{(\vec{C},\vec{C}')} x' = D \cdot \prod_{h=0}^{r-t} x_{C_h} \prod_{h=1}^{t} x_{C_{r-t+h} \setminus Q_h} \prod_{h=0}^{r-t} x_{C'_h} \prod_{h=1}^{t} x_{C'_{r-t+h} \setminus Q_h} \ .$$

Let $\vec{S} = (S_0, S_1, \ldots, S_{r-t}, S'_0, S'_1, \ldots, S'_{r-t}, R_1, \ldots, R_t)$ and $\vec{T} = (T_0, \ldots, T_{r-t}, T'_0, \ldots T'_{r-t}, W_1, \ldots, W_t)$ be such that $A_{i,j,Q,p}^{(\vec{C},\vec{C}')}(\vec{S}, \vec{T}) = 1$. Then, we have that

$$x'_{\vec{S}} x'_{\vec{T}} = \prod_{h=0}^{r-t} x_{S_h} x_{T_h} x_{S'_h} x_{T'_h} \prod_{h=1}^{t} x_{R_h} x_{W_h} = \prod_{h=0}^{r-t} x_{S_h \oplus T_h} x_{S'_h \oplus T'_h} \prod_{h=1}^{t} x_{R_h \oplus W_h}$$

$$= \prod_{h=0}^{r-t} x_{C_h} x_{C'_h} \prod_{h=1}^{t} x_{C_{r-t+h} \setminus Q_h} x_{C'_{r-t+h} \setminus Q_h} \ ,$$

where we use that this entry of $A_{i,j,Q,p}^{(\vec{C},\vec{C}')}(\vec{S}, \vec{T})$ is nonzero and $x \in \{-1,1\}^n$ so that $x_u^2 = 1$ for any $u \in [n]$.

Next, we prove that there are exactly D pairs $(\vec{S}, \vec{T})$ where $A_{i,j,Q,p}^{(\vec{C},\vec{C}')}(\vec{S}, \vec{T})$ is nonzero. We observe that, for each $h = 0, \ldots, r-t+1$, there are exactly $2\binom{n-2}{\ell-1}$ pairs (S_h, T_h) such that $S_h \oplus T_h = C_h$. Indeed, this is because $C_h = \{u, v\}$ has size exactly 2, so $S_h \oplus T_h = C_h$ implies that $S_h = \{u\} \cup U$ and $T_h = \{v\} \cup U$. There are 2 choices for $u \in C_h$ to assign to S_h, and then afterward there are $\binom{n-2}{\ell-1}$ choices for the set U, which is a set of size $\ell - 1$ not containing either of u, v. For $h = 1, \ldots, t$, there are exactly $\binom{n-2}{\ell-1}$ pairs (R_h, W_h) satisfying the condition, as this is the number of choices for U. Here, note that we do not have the additional factor of 2 because the u coming from C_{r-t+h} must be in R_h. Combining, we see that $D = 2^{2(r-t+1)} \binom{n-2}{\ell-1}^{2(r-t+1)+t}$, as required. The "in particular" follows by noting $\mathrm{val}(f_M) = \max_{x \in \{-1,1\}^n} f_M(x)$ and $\|A\|_{\infty \to 1} = \max_{x', y' \in \{-1,1\}^N} x'^{\top} A y'$. $\qquad \square$

7.4 Step 4: bounding the $\infty \to 1$-norm of A via row pruning

By Lemma 7.3, in order to upper bound $\mathbb{E}_b[\mathrm{val}(f_M)]$, it suffices to bound $\mathbb{E}_b[\|A\|_{\infty \to 1}]$.

We always have $\|A\|_{\infty \to 1} \leq N\|A\|_2$. It turns out that $\|A\|_2$ is governed by the maximum degree (relative to the average) of any of $A_{i,j}$, where by degree we mean the number of nonzero entries in a row/column. However, the $A_{i,j}$'s can have rows of degree significantly larger the the average, and this prohibits the spectral norm from giving a good bound on $\mathrm{val}(f_M)$.

The key observation is that, for a certain choice of our parameters ℓ, r, d, the fraction of these "bad rows" is very small and does not noticeably affect $\|A\|_{\infty \to 1}$. We can thus "zero out" these bad rows and then use the spectral certificate on the "pruned matrix" to bound $\mathrm{val}(f_M)$. Establishing this combinatorial fact is the crux of our proof and is captured in the following lemma. The next section (Section 8) is dedicated to the proof of this lemma, and constitutes the key component of

the proof. In fact, at a high level all the steps done so far in the proof are somewhat generic, and our key innovation is choosing the decomposition step carefully to ensure that this row pruning step succeeds.

Lemma 7.4 (Row pruning). *Fix $r \geq 1$, and let $\mathcal{H}^{(r)}$ denote the r-chain hypergraph. Let $\cup_{Q,p} \mathcal{H}_{Q,p}^{(r)}$ be a d-contiguously regular partition of $\mathcal{H}^{(r)}$. Fix $0 \leq t \leq r$ and a maximum directed matching M on $[k]$. Let A be the Kikuchi matrix defined in [Definition 7.2](#), which depends on r, t, the pieces $\cup_{(Q,p) \in P_t} \mathcal{H}_{Q,p}^{(r)}$ of the refinement, and the matching M.*

Let $\Gamma > 0$ be a constant, and let $\Delta = 9 \cdot 2^{2r+2-2t} (\ell/n)^{2r+2-t} d^t (3\delta n)^{2r+1-t}$. Let $\mathcal{B}$ denote the set of rows $\vec{S}$ such that there exists $i \neq j \in [k]$ where the $\vec{S}$-th row/column of $A_{i,j}$ has more than Δ nonzero entries.

Suppose there is $\gamma \in (0,1)$ such that

(1) $\gamma \leq \frac{1}{c\Gamma r^3 \log_2 n}$, where c is a sufficiently large absolute constant;

(2) $\frac{2}{3\delta\ell} \leq \gamma$;

(3) $\left(\frac{3\gamma\delta\ell}{4}\right)^{r+1} \geq n$;

(4) $(2r+2)\exp(-\ell/64r^2) \leq \ell^{-\Gamma r}$;

(5) $d = 3\delta\ell\gamma$.

Then, the number of bad rows is $|\mathcal{B}| \leq 2\ell^{-\Gamma r} N$.

7.5 Step 5: finishing the proof

Let Γ be a sufficiently large constant, $r = O(\log n)$, $\gamma = 1/O(\log^4 n)$, and $\ell = O(\log^4 n/\delta)$ for a sufficiently large constant. This choice of parameters satisfies all the conditions in [Lemma 7.4](#) and furthermore they satisfy

$$(2\ell^{-\Gamma r}) \cdot (2\ell)^{2r+2-t} \leq 1/n^2 .$$

We note that by our choice of parameters, d is constant.

Applying [Lemma 6.3](#), we can construct a contiguous d-regular refinement of $\mathcal{H}^{(r)}$, given by $\mathcal{H}^{(r)} = \cup_{Q,p} \mathcal{H}_{Q,p}^{(r)}$ and polynomials $\Psi^{(t)}$ for $0 \leq t \leq r$ such that $k(3\delta n)^{r+1} \leq \mathbb{E}_b[\text{val}(\Phi)] \leq \sum_{t=0}^{r} \mathbb{E}_b[\text{val}(\Psi^{(t)})]$.

By [Lemma 7.1](#), we have

$$(\mathbb{E}_b[\text{val}(\Psi^{(t)})])^2 \leq \mathbb{E}_b[\text{val}(\Psi^{(t)})^2] = \mathbb{E}_b[\text{val}((\Psi^{(t)})^2)] \leq k \left(|P_t|(3\delta n)^{r-t} d^t \right)^2 + |P_t| 2k \mathbb{E}_M \mathbb{E}_{b \leftarrow \{-1,1\}^k}[\text{val}(f_M^{(t)})] .$$

By [Lemma 7.3](#), we have for any maximum directed matching M,

$$\mathbb{E}_{b \leftarrow \{-1,1\}^k}[\text{val}(f_M^{(t)})] \leq \frac{1}{D} \|A\|_{\infty \to 1} .$$

Let $B_{i,j}$ denote the matrix $A_{i,j}$ after we zero out all rows and columns in $\mathcal{B}$. We observe that $\|B_{i,j}\|_2 \leq \Delta$ as every row and *column* in $B_{i,j}$ has at most Δ nonzero entries; the fact about the

columns of $B_{i,j}$ follows because $A_{i,j} = A_{j,i}^\top$, so the set $\mathcal{B}$ contains all the bad columns as well. Thus, $\|A_{i,j} - B_{i,j}\|_{\infty \to 1} \le |\mathcal{B}| \cdot (2\ell)^{2r+2-t} \le 2\ell^{-\Gamma r} N \cdot (2\ell)^{2r+2-t} \le N/n^2$, by our choice of parameters. Here, we used that every row of $A_{i,j}$ can (crudely) have at most $(2\ell)^{2r+2-t}$ nonzero entries.

Now, let $B = \sum_{(i,j) \in M} b_i b_j B_{i,j}$. The random matrix B is a Rademacher series with k terms in $\mathbb{R}^{N \times N}$. By the Matrix Khintchine (Fact 2.8) inequality, we have that $\mathbb{E}_b[\|B\|_2] \le O(\Delta \sqrt{k \log N}) = O(\Delta \sqrt{kr\ell \log n})$. Note that here we use that M is a matching, so $b_i b_j$ and $b_{i'} b_{j'}$ are independent Rademacher random variables for distinct edges (i, j) and (i', j') in the matching.

Hence, we have that

$$D\mathbb{E}_b[\mathrm{val}(f_M)] \le \mathbb{E}_b[\|A\|_{\infty \to 1}] \le \mathbb{E}_b[\|B\|_{\infty \to 1} + \|A - B\|_{\infty \to 1}]$$
$$\le \mathbb{E}_b[N\|B\|_2] + kN/n^2 \le NO(\Delta\sqrt{kr\ell \log n}) + o(N).$$

Thus, $\mathbb{E}_b[\mathrm{val}(f_M)] \le \frac{N}{D}O(\Delta\sqrt{kr\ell \log n})$. Using the bound on Δ from Lemma 7.4, we have that

$$\frac{N}{D}\Delta \le 2^{-2r-2+2t}\left(\frac{n(n-1)}{\ell(n-\ell)}\right)^{2r+2-t} \cdot 9 \cdot 2^{2r+2-2t}(\ell/n)^{2r+2-t}d^t(3\delta n)^{2r+1-t}$$
$$= \left(\frac{n-1}{n-\ell}\right)^{2r+2-t} \cdot 9 \cdot d^t(3\delta n)^{2r+1-t} \le e^{\frac{O(\ell r)}{n}} \cdot 9 \cdot d^t(3\delta n)^{2r+1-t} \le O(1) \cdot d^t(3\delta n)^{2r+1-t},$$

and so we conclude that $\mathbb{E}_b[\mathrm{val}(f_M)] \le d^t(3\delta n)^{2r+1-t}O(\sqrt{kr\ell \log n})$, where we use that $\ell r \le n$.

We thus have

$$\mathbb{E}_b[\mathrm{val}(\Psi^{(t)})] \le \sqrt{k\left(|P_t|(3\delta n)^{r-t}d^t\right)^2 + |P_t|2k \cdot d^t(3\delta n)^{2r+1-t}O(\sqrt{kr\ell \log n})}.$$

Next, we note that we have $|P_t|d^t \le |\mathcal{H}^{(t)}| = n(3\delta n)^t$, and so

$$\mathbb{E}_b[\mathrm{val}(\Psi^{(t)})] \le \sqrt{k\left(n(3\delta n)^t \cdot (3\delta n)^{r-t}\right)^2 + n(3\delta n)^t \cdot 2k \cdot (3\delta n)^{2r+1-t}O(\sqrt{kr\ell \log n})}$$
$$\le O(1) \cdot \sqrt{k}\frac{1}{3\delta}(3\delta n)^{r+1} + O(1) \cdot (3\delta n)^{r+1}\sqrt{\frac{k}{3\delta}}(kr\ell \log n)^{1/4}$$
$$\le O(1) \cdot \sqrt{k} \cdot (3\delta n)^{r+1}\left(\frac{1}{3\delta} + \sqrt{\frac{1}{3\delta}}(kr\ell \log n)^{1/4}\right)$$
$$\le O(1) \cdot \sqrt{k} \cdot (3\delta n)^{r+1}\sqrt{\frac{1}{3\delta}}(kr\ell \log n)^{1/4},$$

assuming that $\delta^{-2} \le O(kr\ell \log n) = O(k \log^6 n/\delta)$.[12] Thus,

$$k(3\delta n)^{r+1} \le \mathbb{E}_b[\mathrm{val}(\Phi)] \le \sum_{t=0}^{r} \mathbb{E}_b[\mathrm{val}(\Psi^{(t)})] \le (r+1) \cdot O(1) \cdot \sqrt{k} \cdot (3\delta n)^{r+1}\sqrt{\frac{1}{3\delta}}(kr\ell \log n)^{1/4}$$
$$\implies k \le \frac{1}{9\delta^2} \cdot O(r^5\ell \log n) = O(\log^{10} n/\delta^3),$$

---

[12]When we optimize the $\log n$ factor in the next step, we will no longer need this assumption, which is why it does not appear in Theorem 1.

34

by our choice of r, ℓ. We note that, up to the proof of Lemma 7.4, this *almost* finishes the proof of Theorem 1. The issue is that we have lost an additional $\log^2 n$-factor. In the next and final step, we shall save this factor by reformulating the above proof as a reduction to a 2-LDC and applying a off-the-shelf bound on linear 2-LDCs instead of a spectral refutation to finish.

7.6 Step 6: optimizing the $\log n$ factor

We shall now reformulate the arguments in Section 7.5 to give us a reduction from the 3-LCC $\mathcal{L}$ to a 2-LDC $\mathcal{L}'$. Instead of bounding $\text{val}(\Psi^{(t)})$ using the $\infty \to 1$ norm of the Kikuchi matrices, we shall instead use the Kikuchi matrices to give a *reduction* to a linear 2-LDC, and then we apply the lower bound of [GKST06] (Fact 2.5). The difference between Section 7.5 and this subsection is similar to the difference between the main proof and the proof in Appendix B in [AGKM23], which also saves some additional $\log n$ factors in the setting of 3-LDC lower bounds.

The reason for the savings is that, in the case of 2-query linear codes, Fact 2.5 shows a lower bound of $2 \log_2 n \geq \delta k$, which saves a factor of δ over the lower bound from spectral refutation of $O(\log n) \geq \delta^2 k$ for general codes. In our reduction, we shall produce a 2-LDC with $\delta' \sim \delta/(\log^2 n)$, so this optimization saves us a $O(\log^2 n)$ factor. As a result, we get a final lower bound of $k \leq O(\log^8 n)$, as opposed to the lower bound of $k \leq O(\log^{10} n)$ that we obtained in Section 7.5.

We proceed similarly to Section 7.5. Let Γ be a sufficiently large constant, $r = O(\log n)$, $\gamma = 1/O(\log^4 n)$, and $\ell = O(\log^4 n/\delta)$ for a sufficiently large constant. We note that this choice of parameters satisfies all the conditions in Lemma 7.4, and furthermore they satisfy

$$(2\ell^{-\Gamma r}) \cdot (2\ell)^{2r+2-t} \leq 1/n^2 \ .$$

Applying Lemma 6.3, we can construct a contiguous d-regular refinement of $\mathcal{H}^{(r)}$, given by $\mathcal{H}^{(r)} = \cup_{Q,p}\mathcal{H}^{(r)}_{Q,p}$ and polynomials $\Psi^{(t)}$ for $0 \leq t \leq r$ such that $k(3\delta n)^{r+1} \leq \mathbb{E}_b[\text{val}(\Phi)] \leq \sum_{t=0}^{r} \mathbb{E}_b[\text{val}(\Psi^{(t)})]$.

Now, we observe that there exists $t \in \{0, \dots, r\}$ such that $k(3\delta n)^{r+1}/(r+1) \leq \mathbb{E}_b[\text{val}(\Psi^{(t)})]$. In particular, $\Psi^{(t)}$ has at least $k(3\delta n)^{r+1}/(r+1)$ constraints. For the remainder of the proof, we let t be this particular value in $\{0, \dots, r\}$.

By Lemma 7.1, we have

$$(\mathbb{E}_b[\text{val}(\Psi^{(t)})])^2 \leq \mathbb{E}_b[\text{val}(\Psi^{(t)})^2] = \mathbb{E}_b[\text{val}((\Psi^{(t)})^2)] \leq k\left(|P_t|(3\delta n)^{r-t}d^t\right)^2 + |P_t|2k\mathbb{E}_M\mathbb{E}_{b\leftarrow\{-1,1\}^k}[\text{val}(f_M^{(t)})] \ .$$

Therefore, there exists a maximum directed matching M on $[k]$ such that

$$\frac{1}{k|P_t|}(\mathbb{E}_b[\text{val}(\Psi^{(t)})])^2 - |P_t|\left((3\delta n)^{r-t}d^t\right)^2 \leq 2\mathbb{E}_{b\leftarrow\{-1,1\}^k}[\text{val}(f_M^{(t)})] \ .$$

For the remainder of the proof, we let M be this particular directed matching.

Let $L = \{i : (i, j) \in M\}$ denote the "left halves" of the edges in the matching M. We note that $k' := |L| \geq \frac{k-1}{2}$. Let $\mathcal{L}': \{-1,1\}^L \to \{-1,1\}^{2N}$, where $N := \binom{n}{\ell}^{2r+2-t}$, be the linear code defined from $\mathcal{L}$ as follows. For each $b \in \{-1,1\}^L$, we first extend b to be in $\{-1,1\}^k$ by setting $b_j = 1$ for all $j \notin L$ (for $b \in \{-1,1\}^L$, we shall abuse notation and think of b as in $\{-1,1\}^k$ using this

trivial extension). Then, we let $x = \mathcal{L}(b)$, and finally we let $x' := \mathcal{L}'(b)$ be the vector with $2N$ coordinates, one for each row/column of A, where the $\vec{S}$-th entry (similarly $\vec{T}$-th entry) is given by $x'_{\vec{S}} = \prod_{h=0}^{r-t} x_{S_h} x_{S'_h} \prod_{h=1}^{t} x_{R_h}$.

We make the following observations. First, we note that $\mathcal{L}'$ is clearly a linear map. Secondly, following Lemma 7.3, we note that for every $b \in \{-1,1\}^L$, every $(i,j) \in M$ (which implies that $i \in L$ and $j \notin L$), and row $\vec{S}$ and column $\vec{T}$ where $A_{i,j}(\vec{S}, \vec{T}) = 1$, we have that $x' = \mathcal{L}'(b)$ satisfies $x'_{\vec{S}} x'_{\vec{T}} = b_i$.

We now show that $\mathcal{L}'$ is a $(2, \delta)$-LDC for $\delta' = \Omega(\delta/r^2)$. Formally, we shall show that for each $(i,j) \in M$, there exists a matching $G''_{i,j}$ on $[2N]$ such that for every $b \in \{-1,1\}^L$, each edge $(\vec{S}, \vec{T})$ in $G''_{i,j}$, we have $x'_{\vec{S}} x'_{\vec{T}} = b_i$ where $x' = \mathcal{L}'(b)$, and furthermore $\frac{1}{k'} \sum_{(i,j)\in M} |G''_{i,j}| \geq \delta' \cdot 2N$, where we recall that $k' := |L|$ is the dimension of $\mathcal{L}'$.

We have already argued that for every edge $(\vec{S}, \vec{T})$ in the bipartite graph $G_{i,j}$ defined by the adjacency matrix $A_{i,j}$ (where the rows and columns form the left and right sets of vertices), we have $x'_{\vec{S}} x'_{\vec{T}} = b_i$. It thus remains to show that $G_{i,j}$ has a matching of size $\delta' N$.

As before, let $B_{i,j}$ denote the matrix $A_{i,j}$ after we zero out all rows and columns in $\mathcal{B}$. We observe that every row and *column* in $B_{i,j}$ has at most Δ nonzero entries; the fact about the columns of $B_{i,j}$ follows because $A_{i,j} = A_{j,i}^\top$, so the set $\mathcal{B}$ contains all the bad columns as well. Thus, the bipartite graph $G'_{i,j}$ defined by the adjacency matrix $B_{i,j}$ has maximum (left or right) degree at most Δ and therefore has a matching $G''_{i,j}$ of size at least $|G'_{i,j}|/\Delta$.

Now, the number of edges removed is at most $|\mathcal{B}| \cdot (2\ell)^{2r+2-t} \leq 2\ell^{-\lceil r\rceil} N \cdot (2\ell)^{2r+2-t} \leq N/n^2$, by our choice of parameters. Here, we used that every row of $A_{i,j}$ can (crudely) have at most $(2\ell)^{2r+2-t}$ nonzero entries. Thus, we have that $|E(G'_{i,j})| \geq |E(G_{i,j})| - N/n^2$.

In order to finish the reduction, we need to lower bound $\frac{1}{N} \sum_{(i,j)\in M} |E(G''_{i,j})|$. We have that

$$
\frac{1}{N} \sum_{(i,j)\in M} |E(G''_{i,j})| \geq \sum_{(i,j)\in M)} \frac{1}{N\Delta} |E(G'_{i,j})| \geq \sum_{(i,j)\in M)} \frac{1}{N\Delta} \left(|E(G_{i,j})| - \frac{N}{n^2} \right) \geq \frac{1}{N\Delta} \left(D \cdot \mathbb{E}_{b\leftarrow\{-1,1\}^k}[\mathrm{val}(f_M^{(t)})] - \frac{kN}{n^2} \right)
$$

$$
\geq \frac{D}{2N\Delta} \left(\frac{1}{k|P_t|} (\mathbb{E}_b[\mathrm{val}(\Psi^{(t)})])^2 - |P_t| \left((3\delta n)^{r-t} d^t \right)^2 \right) - \frac{k}{\Delta n^2}
$$

$$
\geq \frac{D}{2N\Delta} \left(\frac{1}{k|P_t|} \cdot \frac{k^2 (3\delta n)^{2r+2}}{(r+1)^2} - |P_t| \left((3\delta n)^{r-t} d^t \right)^2 \right) - \frac{k}{\Delta n^2} .
$$

Using the bound on Δ from Lemma 7.4, we have that

$$
\frac{N\Delta}{D} \leq 2^{-2r-2+2t} \left(\frac{n(n-1)}{\ell(n-\ell)} \right)^{2r+2-t} \cdot 9 \cdot 2^{2r+2-2t} (\ell/n)^{2r+2-t} d^t (3\delta n)^{2r+1-t}
$$

$$
= \left(\frac{n-1}{n-\ell} \right)^{2r+2-t} \cdot 9 \cdot d^t (3\delta n)^{2r+1-t} \leq e^{\frac{O(\ell r)}{n}} \cdot 9 \cdot d^t (3\delta n)^{2r+1-t} \leq O(1) \cdot d^t (3\delta n)^{2r+1-t} ,
$$

which implies that

$$
\frac{1}{N} \sum_{(i,j)\in M} |E(G''_{i,j})| \geq \frac{1}{O(1)d^t (3\delta n)^{2r+1-t}} \left(\frac{1}{k|P_t|} \cdot \frac{k^2 (3\delta n)^{2r+2}}{(r+1)^2} - |P_t| \left((3\delta n)^{r-t} d^t \right)^2 \right) - \frac{k}{\Delta n^2} .
$$

Next, we note that we have $|P_t|d^t \le |\mathcal{H}^{(t)}| = n(3\delta n)^t$, and so

$$\frac{1}{N} \sum_{(i,j)\in M} |E(G''_{i,j})| \ge \frac{1}{O(1)d^t(3\delta n)^{2r+1-t}} \left(\frac{1}{k|P_t|} \cdot \frac{k^2(3\delta n)^{2r+2}}{(r+1)^2} - |P_t| \left((3\delta n)^{r-t}d^t \right)^2 \right) - \frac{k}{\Delta n^2}$$

$$\ge \frac{1}{O(1)} \left(\frac{k(3\delta)}{(r+1)^2} - \frac{1}{3\delta} \right) - \frac{k}{\Delta n^2} \ge \frac{1}{O(1)} \frac{k(3\delta)}{(r+1)^2} \ ,$$

if we assume that $k \ge \Omega(r^2/\delta^2) = \Omega(\log^2 n/\delta^2)$. Note that if not, then we have that $\delta^2 k \le O(\log^2 n)$, which is a better lower bound than Theorem 1.

Therefore, we have shown that

$$\frac{1}{k'} \sum_{(i,j)\in M} |E(G''_{i,j})| \ge \delta' N \ ,$$

where $k' = |L| \ge \frac{k-1}{2}$ and $\delta' = \Omega(\delta/r^2)$. Hence, by Fact 2.5, it follows that

$$O(\ell r \log n) \ge 2\log_2 N \ge \delta' k \ge \Omega(\delta k/r^2)$$
$$\implies k \le O(\ell r^3 \log n/\delta) \le O(\log^8 n/\delta^2) \ ,$$

i.e., $2^{\Omega((\delta^2 k)^{1/8})} \le n$. This finishes the proof of Theorem 1 for the case of $\mathbb{F} = \mathbb{F}_2$, up to the proof of Lemma 7.4.

8   Row Pruning: Proof of Lemma 7.4

In this section, we prove Lemma 7.4, restated below, which is the main technical component in the proof of Theorem 1.

Lemma 8.1 (Row pruning). *Fix $r \ge 1$, and let $\mathcal{H}^{(r)}$ denote the r-chain hypergraph. Let $\cup_{Q,p}\mathcal{H}^{(r)}_{Q,p}$ be a d-contiguously regular partition of $\mathcal{H}^{(r)}$. Fix $0 \le t \le r$ and a maximum directed matching M on $[k]$. Let A be the Kikuchi matrix defined in Definition 7.2, which depends on r, t, the pieces $\cup_{(Q,p)\in P_t}\mathcal{H}^{(r)}_{Q,p}$ of the refinement, and the matching M.*

Let $\Gamma > 0$ be a constant, and let $\Delta = 9 \cdot 2^{2r+2-2t}(\ell/n)^{2r+2-t}d^t(3\delta n)^{2r+1-t}$. Let $\mathcal{B}$ denote the set of rows $\vec{S}$ such that there exists $i \ne j \in [k]$ where the $\vec{S}$-th row/column of $A_{i,j}$ has more than Δ nonzero entries.

Suppose there is $\gamma \in (0,1)$ such that

(1) $\gamma \le \frac{1}{c\Gamma r^3 \log_2 n}$, *where c is a sufficiently large absolute constant;*

(2) $\frac{2}{3\delta\ell} \le \gamma$;

(3) $\left(\frac{3\gamma\delta\ell}{4} \right)^{r+1} \ge n$;

(4) $(2r+2)\exp(-\ell/64r^2) \le \ell^{-\Gamma r}$;

(5) $d = 3\delta\ell\gamma$.

Then, the number of bad rows is $|\mathcal{B}| \leq 2\ell^{-\Gamma r} N$.

For $i \neq j \in [k]$ and a row $\vec{S}$, let $\deg_{i,j}(\vec{S})$ denote the number of nonzero entries in the $\vec{S}$-th row. The main idea of the proof is to observe that for any i, j, $\deg_{i,j}(\vec{S})$ is upper-bounded by a $(2r + 2 - t)$-partite polynomial $\mathrm{Deg}_{i,j}(s^{(1)}, s^{(2)}, \ldots, s^{(2r+2-t)})$ in $n(2r + 2 - t)$ variables $s_u^{(h)}$ for $1 \leq h \leq 2r + 2 - t$ and $u \in [n]$ that define $\vec{S}$, i.e., $s^{(h)}$ (in $\{0, 1\}^n$) represents the $0 - 1$ indicator vector of S_h (or S_h' or R_h, depending on the value of h). The contiguous regularity property allows us to control the expected partial derivatives of $\mathrm{Deg}_{i,j}$ and thus apply the tail bounds for partite polynomials in Lemma 2.9.

Let us first set up the polynomial $\mathrm{Deg}_{i,j}$ formally.

For $\vec{C} \in \mathcal{H}_{i,Q,p}^{(r+1)}$ and $\vec{C}' \in \mathcal{H}_{j,Q,p}^{(r+1)}$, let $\mathcal{T}_{i,j,Q,p}^{(\vec{C},\vec{C}')}$ denote the set of $(2r + 2 - t)$-tuples $(u_0, \ldots, u_{r-t}, u_{r-t+1}, \ldots, u_r, v_0, \ldots, v_{r-t})$ such that for $h = 0, \ldots, r - t$, $u_h \in C_h$ and $v_h \in C_h'$, and for $h = 1, \ldots, t$, we have $u_{r-t+h} \in C_h \setminus Q_h$. For a row $\vec{S}$ and a tuple $U = (u_0, \ldots, u_{r-t}, u_{r-t+1}, \ldots, u_r, v_0, \ldots, v_{r-t})$, we write $U \in \vec{S}$ to mean that $u_h \in S_h, v_h \in S_h'$ for $h = 0, \ldots, r - t$ and $u_{r-t+h} \in R_h$ for $h = 1, \ldots, t$.

We next make an easy observation about the structure of the matrices $A_{i,j,Q,p}^{(\vec{C},\vec{C}')}$.

Observation 8.2. Every row of $A_{i,j,Q,p}^{(\vec{C},\vec{C}')}$ has at most 1 non-zero entry. Further, for every non-zero row $\vec{S}$, there is a unique $(2r + 2 - t)$-tuple $U \in \mathcal{T}_{i,j,Q,p}^{\vec{C},\vec{C}'}$ such that $U \in \vec{S}$. Finally, $U \in \vec{S}$ does not guarantee a non-zero entry.

Let $\mathcal{T}_{i,j} = \bigcup_{(Q,p) \in P_t} \bigcup_{\vec{C} \in \mathcal{H}_{i,Q,p}^{(r+1)}, \vec{C}' \in \mathcal{H}_{j,Q,p}^{(r+1)}} \mathcal{T}_{i,j,Q,p}$. Then, by the above observation, the number of nonzero entries in the $\vec{S}$-th row of $A_{i,j}$ is upper bounded by the number of tuples $U \in \mathcal{T}_{i,j}$ with $U \in \vec{S}$. Define the following polynomial $\mathrm{Deg}_{i,j}$ that counts this latter quantity, as follows.

Let $s = (s^{(0)}, \ldots, s^{(r-t)}, s^{(r-t+1)}, \ldots, s^{(r)}, s'^{(0)}, \ldots, s'^{(r-t)})$ be a partitioned set of $0, 1$-valued variables where each $s^{(h)}, s'^{(h)}$ is an n-tuple $(s_u^{(h)})_{u \in [n]}$. We view s as the tuple of 0-1 indicators for the sets appearing in a $\vec{S}$. Formally, we have

$$\mathrm{Deg}_{i,j}(s) := \sum_{U \in \mathcal{T}_{i,j}} \prod_{h=0}^{r} s_{u_h}^{(h)} \prod_{h=0}^{r-t} s_{v_h}'^{(h)} .$$

Let $\mathcal{D}$ denote the uniform distribution over the rows $\vec{S}$, i.e., each $s^{(h)}$ is drawn independently and uniformly at random from $\{0, 1\}^n$ conditioned on $\|s^{(h)}\|_1 = \ell$. Thus, to bound the fraction of rows with a large number of nonzero entries, it suffices to prove bounds on the tail probability of $\mathrm{Deg}_{i,j}$ on $\mathcal{D}$. As $\mathcal{D}$ is not quite a product distribution, we cannot directly apply Lemma 2.9. Nonetheless, we shall show that its tail bounds behave like those for a product distribution, via the following coupling lemma.

Let $\mathcal{D}'$ denote the distribution where each $s_u^{(h)}, s_v'^{(h)}$ are chosen independently as a p-biased Bernoulli random variable where $p = (1 + \beta)\ell/n$ independently for $\beta = \frac{1}{4r}$. The following lemma relates tail bounds for $\mathrm{Deg}_{i,j}$ on $\mathcal{D}'$ with those on $\mathcal{D}$.

Lemma 8.3 (Coupling). *We have* $\Pr_{s \leftarrow \mathcal{D}}[\mathrm{Deg}_{i,j}(s) \geq \Delta] \leq \Pr_{s \leftarrow \mathcal{D}'}[\mathrm{Deg}_{i,j}(s) \geq \Delta] + (2r + 2) \exp(-\ell/64r^2)$.

Proof of Lemma 8.3. To relate the two probabilities, we will couple $\mathcal{D}'$ with $\mathcal{D}$ as follows. First, sample $s \leftarrow \mathcal{D}'$. Then, for each $h = 0, \ldots, r$, set $s^{(h)}$ to be a uniformly random subset of $s^{(h)}$ (if one exists), and similarly for $h = 0, \ldots, r - t$, set $s'^{(h)}$ to be a uniformly random subset of $s'^{(h)}$ of size ℓ also. If one of the sets S_h or S'_h has size $< \ell$, i.e., $\|s^{(h)}\|_1 < \ell$ for some $h \in \{0, \ldots, r\}$ or $\|s'^{(h)}\|_1 < \ell$ for some $h \in \{0, \ldots, r - t\}$, then the coupling fails and we abort. Let $\mathcal{J}$ be the joint distribution induced by this coupling.

Fix $h \in \{0, \ldots, r\}$. By Chernoff bound, we have for every $\delta \in [0, 1]$ and for any h,

$$\Pr_{s \sim \mathcal{D}'}[\|s^{(h)}\|_1 < (1 - \delta)(1 + \beta)\ell] \leq \exp\left(\frac{\delta^2 \ell(1 + \beta)}{2}\right) .$$

Setting $\delta = 1 - \frac{1}{1+\beta}$ and noting that $\beta = \frac{1}{4r} < 1$, we see that $\left(\frac{\delta^2 \ell(1+\beta)}{2}\right) = \frac{\ell}{2(4r+16r^2)} \geq \frac{\ell}{64r^2}$. Hence, the probability that $\mathcal{J}$ aborts is at most $(2r + 2) \exp(-\ell/64r^2) \leq \ell^{-\Gamma r}$. Here, we use Item 4 in the assumptions of the parameters in Lemma 7.4.

We also observe that $\mathrm{Deg}_{i,j}$ is monotone, that is, $\mathrm{Deg}_{i,j}(s) \geq \mathrm{Deg}_{i,j}(s')$ for *any* Boolean variables s, s' where $s' \leq s$ coordinate-wise. In particular, if we first sample $s \leftarrow \mathcal{D}'$ and it holds that $\mathrm{Deg}_{i,j}(s) \leq \Delta$, then it also holds that $\mathrm{Deg}_{i,j}(s') \leq \Delta$ also, regardless of the choice of s' made by the coupling $\mathcal{J}$. We thus have

$$\Pr_{s' \leftarrow \mathcal{D}}[\mathrm{Deg}_{i,j}(s') > \Delta] \leq \Pr_{(s,s') \sim \mathcal{J}}[\mathrm{Deg}_{i,j}(s) > \Delta \mid \mathcal{J} \text{ does not abort}] \leq \Pr_{s \leftarrow \mathcal{D}'}[\mathrm{Deg}_{i,j}(s) > \Delta] + \ell^{-\Gamma r}.$$

This completes the proof. $\qquad\square$

We finally obtain a tail bound on $\mathrm{Deg}_{i,j}$ for the product distribution $\mathcal{D}'$ to complete the proof.

Lemma 8.4. *For $\Delta = 9 \cdot 2^{2r+2-2t}(\ell/n)^{2r+2-t} \cdot d^t(3\delta n)^{2r+1-t}$, we have $\Pr_{s \leftarrow \mathcal{D}'}[\mathrm{Deg}_{i,j}(s) \geq \Delta] \leq \ell^{-\Gamma r}$.*

Proof of Lemma 8.4. We will apply Lemma 2.9 to bound $\Pr_{s \leftarrow \mathcal{D}'}[\mathrm{Deg}_{i,j}(s) \geq \Delta]$. Note that $\mathrm{Deg}_{i,j}$ is homogeneous, multilinear, $(2r + 2 - t)$-partite polynomial. To apply Lemma 2.9, we will now bound the expected partial derivatives μ_Z of $\mathrm{Deg}_{i,j}$ for each tuple $Z \in \{[n] \cup \{\star\})^{2r+2-t}$.

Let

$$\mu = 3 \cdot 2^{2r+2-2t}(\ell/n)^{2r+2-t} \cdot d^t(3\delta n)^{2r+1-t} = 3 \cdot 2^{2r+2-2t}\frac{\ell}{n}(3\delta\ell)^{2r+1-t}d^t . \tag{4}$$

Claim 8.5 (Bounding Expected Partials). Let $\gamma = \frac{c}{\Gamma r^3 \log_2 n}$, $\ell \geq \frac{4n^{1/r}}{3\delta\gamma^4}$, $d = 3\gamma\delta\ell$. Then, for any $h = 0, \ldots, 2r + 2 - t$, we have $\mu_Z := \mu_Z(\mathrm{Deg}_{i,j}) \leq \mu \cdot \gamma^{|Z|}$.

We postpone the proof of Claim 8.5, and now finish the proof of Lemma 7.4 by using Lemma 2.9. Applying of Lemma 2.9 with $\beta = 1/(2r + 2)$ and γ, we see that

$$\Pr_{x \leftarrow \mathcal{D}'}[P_{i,j}(x) \geq 3\mu] \leq r(n + 1)^r \exp\left(-\frac{\frac{1}{2(2r+2)^2}}{2\gamma + \frac{1}{3(2r+2)}\gamma}\right)$$

$$\leq r(n + 1)^r \exp\left(-\frac{1}{24\gamma(r + 1)^2}\right) \leq \ell^{-\Gamma r} ,$$

as $3 \geq (1 + \frac{1}{2r+2})^{2r+2} \geq (1 + \frac{1}{2r+2})^{2r+2-t}$ and $\gamma \leq \frac{1}{c\Gamma r^3 \log_2 n}$ by Item 1 of our parameter assumptions. $\qquad\square$

It thus remains to prove Claim 8.5.

Proof of Claim 8.5. For $U \in \mathcal{T}_{i,j}$, we say $Z \subseteq U$ if Z and U agree on all non-$\star$ entries of Z. We let $\deg_{i,j}(Z)$ denote the number of tuples $U \in \mathcal{T}_{i,j}$ where $Z \subseteq U$. Note that $\mu_Z = p^{2r+2-t-|Z|} \deg_{i,j}(Z)$. Let's now estimate $\deg_{i,j}(Z)$ – which equals the number of triples $(U, \vec{C}, \vec{C}')$ where $U \in \mathcal{T}_{i,j}^{(\vec{C},\vec{C}')}$ and $Z \subseteq U$.

Fix a Z and let Z_1 denote the first $r+1$ entries, and Z_2 denote the last $r+1-t$ entries. First, we argue that there are at most $2^{|Z_1|}(3\delta n)^{r+1-|Z_1|}$ choices for $\vec{C} \in \cup_{(Q,p) \in P_t} \mathcal{H}_{i,Q,p}^{(r+1)}$ for which Z_1 is contained in $\vec{C}$. To see why, consider choosing $\vec{C}$ iteratively. Given the first $h-1$ choices, let's now consider the h-th choice. If $(Z_1)_h$ is a $\star$, then there are δn choices for the hyperedge $C_h \cup \{w_h\} \in H_{w_{h-1}}$, as we already know w_{h-1} (when $h = 0$, $w_{-1} := i$ is fixed). Then, there are 3 choices for w_h within this hyperedge. If $(Z_1)_h = u_h \neq \star$, then, there is a unique hyperedge in $H_{w_{h-1}}$ containing u_h. This hyperedge has two other vertices that could be chosen as w_h. Hence, we have $2^{|Z_1|}(3\delta n)^{r+1-|Z_1|}$ choices in total. Observe that once we have chosen $\vec{C}$, we also know the index (Q,p) of the partition in the refinement that $\vec{C}$ comes from.

Next, let's count the number of partial tuples $U = (u_0, \ldots, u_r)$ that we can produce from this $\vec{C}$. For each non-$\star$ entry of Z_1, we know $u_h = (Z_1)_h$. For each h where $(Z_1)_h = \star$, if $h \in \{0, \ldots, r-t\}$, then we only know $u \in C_h$, which gives us $|C_h| = 2$ choices for u_h. If $h \in \{r-t+1, \ldots, r\}$, then we know that u_h must equal $C_h \setminus Q_{h-(r-t)}$ – a unique choice. We thus pay an additional $2^{r+1-t-|Z_1'|}$, where Z_1' is the partial tuple $((Z_1)_0, \ldots, (Z_1)_{r-t})$, to determine $(u_0, \ldots, u_r)$.

We now have two cases.

(1) Case 1: Z_2 has no $\star$ entries, i.e., $|Z_2| = r+1-t$. This implies that $(v_0, \ldots, v_{r-t}) = ((Z_2)_0, \ldots, (Z_2)_{r-t})$, and so we have uniquely determined U. By an argument similar to above, we also have at most 2^{r+1} choices for $\vec{C}' \in \mathcal{H}_{j,Q,p}^{(r+1)}$ (recall that we already know Q, which determines $v_{r+1-t}, \ldots, v_r$ up to 2^t choices). Hence, we have argued in this case that $\deg_{i,j}(Z) \leq 2^{2r+2}(3\delta n)^{r+1-|Z_1|}$, where we use that $|Z_1| \leq t + |Z_1'|$.

It then follows that

$$\mu_Z = p^{2r+2-|Z|} \deg_{i,j}(Z) \leq p^{2r+2-t-(r+1-t)-|Z_1|} 2^{2r+2} (3\delta n)^{r+1-|Z_1|}$$

$$\leq (1+\beta)^{2r+2} \cdot (\ell/n)^{r+1-|Z_1|} 2^{2r+2} (3\delta n)^{r+1-|Z_1|}$$

$$\leq \left(1 + \frac{1}{4r}\right)^{4r} 2^{2r+2} (3\delta \ell)^{r+1-|Z_1|}$$

$$\leq 3 \cdot 2^{2r+2} \cdot (3\delta \ell)^{r+1-|Z_1|} \ .$$

Now, we observe that since $\mu = 3 \cdot 2^{2r+2-2t} \frac{\ell}{n} (3\delta \ell)^{2r+1-t} d^t$ and $d = 3\delta \ell \gamma$, we have that

$$\frac{\mu_Z}{\gamma^{|Z|}\mu} = \frac{3 \cdot 2^{2r+2} \cdot (3\delta \ell)^{r+1-|Z_1|}}{\gamma^{|Z_1|+r+1-t} \cdot 3 \cdot 2^{2r+2-2t} \frac{\ell}{n} (3\delta \ell)^{2r+1-t} d^t} = \frac{2^{2t} n}{\gamma^{|Z_1|+r+1} \cdot \ell (3\delta \ell)^{r+|Z_1|}}$$

$$\leq \frac{2^{2(r+1)} n}{\gamma^{|Z_1|+r+1} \cdot (3\delta \ell)^{r+1+|Z_1|}} \leq 1 \ ,$$

provided that $3\delta\ell \geq 1$ and $(3\delta\ell\gamma/4)^{r+1} \geq n$, which hold by Item 3 of the parameter assumptions.

(2) Case 2: Z_2 has at least one $\star$ entry. In this case, let us write $Z_2 = (Z_2', \star, Z_2'')$, where Z_2'' does not contain any $\star$. Note that Z_2'' may be empty, i.e., have length 0.

We observe that there are at most $2^{|Z_2'|}(3\delta n)^{r+1-t-(|Z_2''|+1)-|Z_2'|}$ choices for the partial chain $(j, C_0', w_0, \ldots, C_{h'}', w_{h'}')$ where $h' = r - t - (|Z_2''| + 1)$ (i.e., the number of entries in Z_2'). As in Case 1, we argue inductively and consider the step when we have chosen $j, C_0', w_0, \ldots, C_h', w_h$ for some $0 \leq h < h'$. If $(Z_2')_h = \star$, then, there are δn choices for choosing the next hyperedge and 3 choices for deciding the w_{h+1}' within it giving a total of $3\delta n$ choices. If $(Z_2')_h \neq \star$, then there is at most one hyperedge (so no choice to be made) in $H_{w_h'}$ that could appear as the next link and, given the hyperedge, there are 2 choices for the w_{h+1}'.

Given the first h' links in the partial chain, we have at most $2^{r+1-t-(|Z_2''|+1)-|Z_2'|}$ choices for the partial tuple $(v_0, \ldots, v_{h'})$. So in total, we have $2^{r+1-t-(|Z_2''|+1)}(3\delta n)^{r+1-t-(|Z_2''|+1)-|Z_2'|}$ choices for the partial chain $(j, C_0', w_0, \ldots, C_{h'}', w_{h'}')$ and the partial tuple $(v_0, \ldots, v_{h'})$.

To count the number of ways to complete the chain, we break our analysis into two subcases.

(a) Subcase 1: Z_2'' is empty, and so $|Z_2| = |Z_2'|$. In this case, $h' = r - t - 1$. Since we have already chosen (Q, p), the number of different choices for the partial chain $(w_{r-t}', C_{r-t+1}', w_{r-t+1}', \ldots, C_r', w_r')$ must be at most d^t, by Item (4) in Definition 6.1 with $Q' = Q$. Given this choice, w_{r-t}' is fixed so there is at most one choice for a hyperedge in $H_{w_{r-t-1}'}$ that contains w_{r-t}' and given that choice, there are two possible ways to choose v_{r-t}. In total, we have made at most $2d^t$ choices.

In the case that $t = 0$, the partial chain is the "0-chain" given by $Q_{r+1} = w_r'$, and as we have $w_r' = Q_{r+1} = w_r$, this gives a unique choice for the "chain", i.e., $d^0 = 1$ choices.

(b) Subcase 2: Z_2'' is nonempty. We observe that for $h' = r - t - (|Z_2''| + 1)$, the partial chain $(w_{h'+1}', C_{h'+2}', w_{h'+2}', \ldots, C_r', w_r')$ must contain the *complete* tuple $Z_2'' \| Q_z^{(t)}$, where $\cdot\|\cdot$ denotes concatenation. Thus, by d-regularity, there are at most $d^{t+|Z_2''|}$ choices of such tuples. Given the choice of this partial chain, there are $2^{r-t-h'}$ choices for $(v_{h'+1}, \ldots, v_{r-t})$. Hence, in total we have made $2^{|Z_2''|+1}d^{t+|Z_2''|}$ choices.

We note that in either subcase, we make at most $2^{|Z_2''|+1}d^{t+|Z_2''|}$ choices to pick $\vec{C}' \in \mathcal{H}_{j,Q,p}^{(r+1)}$ and $(v_{h'+1}, \ldots, v_{r-t})$, where we can have $|Z_2''| = 0$. Thus, the total number of choices of $\vec{C}'$ and $(v_0, \ldots, v_{r-1})$ is at most

$$2^{r+1-t-(|Z_2''|+1)}(3\delta n)^{r+1-t-(|Z_2''|+1)-|Z_2'|}2^{|Z_2''|+1}d^{t+|Z_2''|} = 2^{r+1-t}(3\delta n)^{r-t-|Z_2|} \cdot d^{t+|Z_2''|},$$

and thus, the total number of triples $(U, \vec{C}, \vec{C}')$ that contribute to $\deg_{i,j}(Z)$ is at most $2^{r+1-t-|Z_1'|+|Z_1|}(3\delta n)^{r+1-|Z_1|}2^{r+1-t}(3\delta n)^{r-t-|Z_2|} \cdot d^{t+|Z_2''|}$.

Thus,

$$\mu_Z \leq p^{2r+2-t-|Z_1|-|Z_2|}2^{r+1-t-|Z_1'|+|Z_1|}(3\delta n)^{r+1-|Z_1|}2^{r+1-t}(3\delta n)^{r-t-|Z_2|} \cdot d^{t+|Z_2''|}$$

41

$$\leq (1+\beta)^{2r+2}2^{2r+2-2t+|Z_1|}(\ell/n)^{2r+2-t-|Z_1|-|Z_2|}(3\delta n)^{2r+1-t-|Z_1|-|Z_2|} \cdot d^{t+|Z_2''|}$$

$$\leq (1+\beta)^{2r+2}2^{2r+2-2t+|Z_1|}(\ell/n)(3\delta\ell)^{2r+1-t-|Z_1|-|Z_2|} \cdot d^{t+|Z_2''|}$$

$$\leq \mu\frac{1}{3}(1+\beta)^{2r+2}2^{|Z_1|}(3\delta\ell)^{-|Z_1|-|Z_2|} \cdot d^{|Z_2''|}$$

$$\leq \mu\left(\frac{2}{3\delta\ell}\right)^{|Z_1|}(3\delta\ell)^{-|Z_2'|}\left(\frac{d}{3\delta\ell}\right)^{|Z_2''|}$$

$$\leq \mu\left(\frac{2}{3\delta\ell}\right)^{|Z_1|}(3\delta\ell)^{-|Z_2'|}\gamma^{|Z_2''|}\ ,$$

using that $(1+\beta)^{2r+2} \leq 3$ and $d = 3\gamma\delta\ell$. We thus have

$$\mu_Z \leq \mu\max\{\frac{2}{3\delta\ell},\gamma\}^{|Z|} \leq \gamma^{|Z|}\mu\,,$$

where we use Item 2 in the parameter assumptions.

This finishes the proof of Claim 8.5. $\qquad\square$

9   Discussion

We conclude with some remarks on the proof of Theorem 1, possible strengthenings, and extensions.

(1) **Non-linear codes.** The lower bound in Theorem 1 applies only to linear codes. However, we note that we only use linearity of the code to argue a lower bound on $\mathrm{val}(\Phi_b)$, the XOR instance polynomial for $(r+1)$-chains. For the natural XOR instances (i.e., when $r = 0$), a lower bound on $\mathrm{val}(\Phi_b)$ easily follows even for non-linear codes. This is the reason why the 3-LDC lower bounds in [AGKM23] apply to non-linear codes. The issue (that nevertheless appears surmountable) that prevents us from obtaining a similar lower bound on $\mathrm{val}(\Phi_b)$ for XOR instances with chains of length > 1 is the following: for non-linear codes, we are only guaranteed that each constraint is satisfied for a non-trivial constant fraction of codewords. That is, $\mathbb{E}_{x\leftarrow\mathcal{L}}[x_C x_u] \geq \varepsilon$ for some constant $\varepsilon > 0$ (for linear codes, we instead obtain $x_C x_u = 1$ for all $x \in \mathcal{L}$). In particular, it is not clear that $\mathbb{E}_b[\Phi_b(\mathcal{L}(b))]$ is non-trivially lower-bounded.

(2) **LCCs with more queries.** While our approach can likely improve the lower bounds (beyond those known for LDCs) even for $q > 3$, the improvements based on natural generalizations of our approach are likely to only yield a polynomial factor improvement. Our explanation is rooted in the heuristic calculation based on the density of the Kikuchi matrices explained earlier in Section 3.2. For larger q, the number of length $(r+1)$-chains with head $i \in [k]$ is still $k(3\delta n)^{r+1}$. The arity of the derived constraints, however, is now $(q-1)(r+1)+1$. This means that the density (i.e., average degree of the natural Kikuchi matrix) at level ℓ is $(3\delta n)^{r+1}(\ell/n)^{\frac{(q-1)(r+1)+1}{2}} \to \left(n(\ell/n)^{\frac{q-1}{2}}\right)^{r+1}$ for large r. Thus, the optimal ℓ turns out to be $n^{1-\frac{2}{q-1}}$, and so we can only hope to achieve a lower bound of $k \leq \tilde{O}(n^{1-\frac{2}{q-1}})$. This nevertheless would

yield an improvement on the current best-known lower bound of $k \leq \tilde{O}(n^{1-\frac{2}{q}})$, inherited from q-LDCs, by a polynomial factor via long chains.

(3) **Optimality of Reed–Muller codes?** Our main result Theorem 1 comes close to showing that Reed–Muller codes, which achieve a blocklength of $n = 2^{O(\sqrt{k})}$, are optimal linear 3-LCCs — a longstanding goal in understanding LCCs. Closing the gap between our result and the blocklength of Reed–Muller codes relates to optimizing the polylog(n) factors in our analysis. Let us now explain each of the log n factors that we "lose" with an eye for the losses that appear naturally surmountable and ones that appear rather inherent.

First, we note that we must take the chain length r to be $\geq O(\log n)$ and $\ell \geq 1/\delta$ for the heuristic calculation in Section 3.2 to work. Second, we note that the application of matrix Khintchine (Fact 2.8) loses a $\sqrt{\log N} = \sqrt{\ell r \log n}$ factor. Thus, in the ideal case, our method could potentially yield that $k \leq O(\log N)$ where $\ell \sim 1/\delta$ and $r = O(\log n)$. This would yield a bound of $k \leq O(\log^2 n)$, or in other words $n \geq 2^{\Omega(\sqrt{k})}$, matching the blocklength of Reed–Muller codes up to constant factors in the exponent.

However, our argument currently loses additional log n factors that appear improvable. First, the hypergraph decomposition step loses a factor of r in the "density" because we need to refute at least one of the $\sim r$ subinstances produced each of which may only have $1/(r+1)$-fraction of all the $(r+1)$-chains. This loses us $O(r^2)$ factor in the density once we use the Cauchy–Schwarz trick. Second, we cannot take ℓ to be as small as $1/\delta$, i.e., a constant. Currently, we need to take $\ell \geq O(\log^4 n)$ for the tail bounds used in the proof of Lemma 7.4 to be effective.

These additional log n factors that we lose not appear to be inherent to our approach. To save these losses would require a sharper chain decomposition method (that does not lose a factor r in the density) and a sharper concentration bound than Lemma 2.9. While these appear technically challenging, it does look plausible that one remove these additional log n factors and obtain a lower bound that matches the blocklength of Reed–Muller codes up to absolute constant factors in the exponent.

Acknowledgements

References

[AGKM23] Omar Alrabiah, Venkatesan Guruswami, Pravesh K. Kothari, and Peter Manohar. A near-cubic lower bound for 3-query locally decodable codes from semirandom CSP refutation. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 1438–1448. ACM, 2023.

[ALM⁺98]   Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM (JACM)*, 45(3):501–555, 1998.

[AS98]   Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of np. *Journal of the ACM (JACM)*, 45(1):70–122, 1998.

[AS21]   Vahid R Asadi and Igor Shinkar. Relaxed locally correctable codes with improved parameters. In *48th International Colloquium on Automata, Languages, and Programming (ICALP 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2021.

[BBH23]   Afonso S Bandeira, March T Boedihardjo, and Ramon van Handel. Matrix concentration inequalities and free probability. *Inventiones mathematicae*, pages 1–69, 2023.

[BC23]   Jop Briët and Davi Castro-Silva. On the threshold for szemerédi's theorem with random differences. *arXiv preprint arXiv:2304.03234*, 2023.

[BDL13]   Abhishek Bhowmick, Zeev Dvir, and Shachar Lovett. New bounds for matching vector families. In *Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, June 1-4, 2013*, pages 823–832. ACM, 2013.

[BFNW93]   László Babai, Lance Fortnow, Noam Nisan, and Avi Wigderson. BPP has subexponential time simulations unless EXPTIME has publishable proofs. *Comput. Complex.*, 3:307–318, 1993.

[BGH⁺04]   Eli Ben-Sasson, Oded Goldreich, Prahladh Harsha, Madhu Sudan, and Salil P. Vadhan. Robust pcps of proximity, shorter pcps and applications to coding. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, Chicago, IL, USA, June 13-16, 2004*, pages 1–10. ACM, 2004.

[BGT17]   Arnab Bhattacharyya, Sivakanth Gopi, and Avishay Tal. Lower bounds for 2-query lccs over large alphabet. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

[BIW10]   Omer Barkol, Yuval Ishai, and Enav Weinreb. On locally decodable codes, self-correctable codes, and t-private pir. *Algorithmica*, 58(4):831–859, 2010.

[BK95]   Manuel Blum and Sampath Kannan. Designing programs that check their work. *Journal of the ACM (JACM)*, 42(1):269–291, 1995.

[BLR93]   Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. *Journal of computer and system sciences*, 47(3):549–595, 1993.

[CGS20]   Alessandro Chiesa, Tom Gur, and Igor Shinkar. Relaxed locally correctable codes with nearly-linear block length and constant query complexity. In *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SODA 2020, Salt Lake City, UT, USA, January 5-8, 2020*, pages 1395–1411. SIAM, 2020.

[CY23]   Gil Cohen and Tal Yankovitz. Asymptotically-good rlccs with $(\log{n})^{2+o(1)}$ queries. *Electron. Colloquium Comput. Complex.*, TR23-110, 2023.

[DGGW19]   Zeev Dvir, Sivakanth Gopi, Yuzhou Gu, and Avi Wigderson. Spanoids - an abstraction of spanning structures, and a barrier for lccs. In *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, volume 124 of *LIPIcs*, pages 32:1–32:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

[DSW14]   Zeev Dvir, Shubhangi Saraf, and Avi Wigderson. Breaking the quadratic barrier for 3-lcc's over the reals. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 784–793. ACM, 2014.

[Dvi10]   Zeev Dvir. On matrix rigidity and locally self-correctable codes. In *Proceedings of the 25th Annual IEEE Conference on Computational Complexity, CCC 2010, Cambridge, Massachusetts, USA, June 9-12, 2010*, pages 291–298. IEEE Computer Society, 2010.

[Dvi12]   Zeev Dvir. Incidence theorems and their applications. *CoRR*, abs/1208.5073, 2012.

[Dvi16]   Zeev Dvir. Lecture notes on linear locally decodable codes. https://www.cs.princeton.edu/~zdvir/LDCnotes/LDC8.pdf, Fall 2016.

[Efr09]   Klim Efremenko. 3-query locally decodable codes of subexponential length. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 39–44. ACM, 2009.

[Fei08]   Uriel Feige. Small linear dependencies for binary vectors of low weight. In *Building Bridges: Between Mathematics and Computer Science*, pages 283–307. Springer, 2008.

[GKM22]   Venkatesan Guruswami, Pravesh K. Kothari, and Peter Manohar. Algorithms and certificates for Boolean CSP refutation: smoothed is no harder than random. In *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*, pages 678–689. ACM, 2022.

[GKST06]   Oded Goldreich, Howard Karloff, Leonard J Schulman, and Luca Trevisan. Lower bounds for linear locally decodable codes and private information retrieval. *Computational Complexity*, 15(3):263–296, 2006.

[Gri01]   Dima Grigoriev. Linear lower bound on degrees of positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259(1):613–622, 2001.

[GRR20]   Tom Gur, Govind Ramnarayan, and Ron Rothblum. Relaxed locally correctable codes. *Theory of Computing*, 16(1):1–68, 2020.

[Ham73]   Noboru Hamada. On the p-rank of the incidence matrix of a balanced or partially balanced incomplete block design and its applications to error correcting codes. *Hiroshima Math. J.*, 3:153–226, 1973.

[HKM23]   Jun-Ting Hsieh, Pravesh K. Kothari, and Sidhanth Mohanty. A simple and sharper proof of the hypergraph Moore bound. In *Proceedings of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023, Florence, Italy, January 22-25, 2023*, pages 2324–2344. SIAM, 2023.

[IK99]   Yuval Ishai and Eyal Kushilevitz. Improved upper bounds on information-theoretic private information retrieval (extended abstract). In *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing, May 1-4, 1999, Atlanta, Georgia, USA*, pages 79–88. ACM, 1999.

[IS18]   Eran Iceland and Alex Samorodnitsky. On coset leader graphs of structured linear codes. *Electron. Colloquium Comput. Complex.*, TR18-023, 2018.

[KM23]   Vinayak M. Kumar and Geoffrey Mon. Relaxed local correctability from local testing. *CoRR*, abs/2306.17035, 2023.

[KSY14]   Swastik Kopparty, Shubhangi Saraf, and Sergey Yekhanin. High-rate codes with sublinear-time decoding. *Journal of the ACM (JACM)*, 61(5):1–20, 2014.

[KT00]   Jonathan Katz and Luca Trevisan. On the efficiency of local decoding procedures for error-correcting codes. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pages 80–86, 2000.

[KV00]   Jeong Han Kim and Van H Vu. Concentration of multivariate polynomials and its applications. *Combinatorica*, 20(3):417–434, 2000.

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

[LFKN90]   Carsten Lund, Lance Fortnow, Howard J. Karloff, and Noam Nisan. Algebraic methods for interactive proof systems. In *31st Annual Symposium on Foundations of Computer Science, St. Louis, Missouri, USA, October 22-24, 1990, Volume I*, pages 2–10. IEEE Computer Society, 1990.

[LP91]   Françoise Lust-Piquard and Gilles Pisier. Noncommutative Khintchine and Paley inequalities. *Ark. Mat.*, 29(2):241–260, 1991.

[Mek14]   Raghu Meka. Discrepancy and beating the union bound. https://windowsontheory.org/2014/02/07/discrepancy-and-beating-the-union-bound, February 2014.

[RS96]   Ronitt Rubinfeld and Madhu Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2):252–271, 1996.

[Sch08]   Grant Schoenebeck. Linear level lasserre lower bounds for certain k-csps. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 593–602. IEEE Computer Society, 2008.

[Sha90]   Adi Shamir. Ip=pspace. In *31st Annual Symposium on Foundations of Computer Science, St. Louis, Missouri, USA, October 22-24, 1990, Volume I*, pages 11–15. IEEE Computer Society, 1990.

[SS12]   Warren Schudy and Maxim Sviridenko. Concentration and moment inequalities for polynomials of independent random variables. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2012, Kyoto, Japan, January 17-19, 2012*, pages 437–446. SIAM, 2012.

[Ton11]   Vladimir D. Tonchev. Finite geometry designs, codes, and Hamada's conjecture. In *Information security, coding theory and related combinatorics*, volume 29 of *NATO Sci. Peace Secur. Ser. D Inf. Commun. Secur.*, pages 437–448. IOS, Amsterdam, 2011.

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

[Tro15]   Joel A. Tropp. An introduction to matrix concentration inequalities. *Found. Trends Mach. Learn.*, 8(1-2):1–230, 2015.

[Woo07]   David Woodruff. New lower bounds for general locally decodable codes. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 14, 2007.

[Woo10]   David P. Woodruff. A quadratic lower bound for three-query linear locally decodable codes over any field. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 13th International Workshop, APPROX 2010, and 14th International Workshop, RANDOM 2010, Barcelona, Spain, September 1-3, 2010. Proceedings*, volume 6302 of *Lecture Notes in Computer Science*, pages 766–779. Springer, 2010.

[Yek08]   Sergey Yekhanin. Towards 3-query locally decodable codes of subexponential length. *Journal of the ACM (JACM)*, 55(1):1–16, 2008.

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

[Zou12]   Anastasios Zouzias. A matrix hyperbolic cosine algorithm and applications. In *Automata, Languages, and Programming - 39th International Colloquium, ICALP 2012, Warwick, UK, July 9-13, 2012, Proceedings, Part I*, volume 7391 of *Lecture Notes in Computer Science*, pages 846–858. Springer, 2012.

A   Linear 3-LCC Lower Bounds over Larger Fields

In this section, we prove Theorem 1 in the case where the finite field $\mathbb{F}$ is not $\mathbb{F}_2$. The proof will be nearly identical to the proof in Sections 5 to 8 for the case of $\mathbb{F} = \mathbb{F}_2$, and so we shall only give a proof sketch and mainly focus on the parts of the proof where modifications are required.

To begin, we recall that by Fact 2.4, there exist 3-uniform hypergraph matchings $H_1, \ldots, H_n$, each of size at least δn, such that for each $u \in [n]$ and $C = \{v_1, v_2, v_3\} \in H_u$, there exists $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{F} \setminus \{0\}$ such that for every $x \in \mathcal{L}$, it holds that $\alpha_1 x_{v_1} + \alpha_2 x_{v_2} + \alpha_3 x_{v_3} = x_u$. Furthermore, without loss of generality we can assume that the code is systematic, i.e., for any $b \in \mathbb{F}^k$, $x = \mathcal{L}(b)$ satisfies $x_i = b_i$ for all $i \in [k]$.

Next, let us define a code $\mathcal{L}' \colon \{-1, 1\}^k \to \{-1, 1\}^{n(|\mathbb{F}|-1)}$ where, for each $u \in [n]$ and $\alpha \in \mathbb{F} \setminus \{0\}$, we set $\mathcal{L}'(b)_{(u,\alpha)} = \alpha \mathcal{L}(b)_u$. Let $n' = n(|\mathbb{F}| - 1)$, and associate $[n']$ with the set $[n] \times (\mathbb{F} \setminus \{0\})$. We now observe that $\mathcal{L}'$ is a 3-LCC in normal form with the additional property that the coefficients of all constraints can be taken to be 1 without loss of generality. Formally, there exist 3-uniform hypergraph matchings $H_1, \ldots, H_{n'}$ such that (1) each H_u has $|H_u| \geq \delta n'/(|\mathbb{F}| - 1)$, and (2) for each $u \in [n']$ and each $C = \{v_1, v_2, v_3\} \in H_{(u,\alpha)}$, every $x \in \mathcal{L}$ satisfies $x_u = x_{v_1} + x_{v_2} + x_{v_3}$.

Moreover, there is now a group action of $(\mathbb{F} \setminus \{0\}, \times)$ on the elements of $[n']$, namely for any $\alpha \in \mathbb{F} \setminus \{0\}$, this action maps $u \mapsto \alpha u$. We note that this action respects the constraints. Namely, for $C = \{v_1, v_2, v_3\} \in \binom{[n']}{3}$, if we define $\alpha C = \{\alpha v_1, \alpha v_2, \alpha v_3\}$, then we have that $H_{\alpha u} = \alpha H_u = \{\alpha C : C \in H_u\}$. For the proof, we will be using the fact that there is a negation action for $\alpha = -1$; this is because this transformation has made all coefficients in the constraints be equal to 1, so to cancel a variable x_u we shall only need x_{-u}.

We shall now abuse notation and redefine n' to be n, and we now simply assume that we have this group action on $[n]$. We have thus added this additional property to the code, and in doing so we have only decreased δ by a factor of $|\mathbb{F}| - 1$.

We now turn to the main part of the proof. Following Section 5, we define t-chains. The definition of t-chains now requires a small modification because in the original definition we formed longer chains by canceling a variable x_w via the operation $x_w + x_w = 0$, which was specific to the field $\mathbb{F}_2$. Now, we use the negation action on $[n]$ to cancel a variable.

Definition A.1 (*t*-chain hypergraph $\mathcal{H}^{(t)}$). Let $t \geq 1$ be an integer. For any $u \in [n]$, let $\mathcal{H}_u^{(t)}$ denote the set of tuples of the form $(u, C_1, w_1, C_2, w_2, \ldots, C_t, w_t)$, where each $C_h \in \binom{[n]}{2}$, $w_h \in [n]$, and it holds that for all $1 \leq h \leq t$, $C_h \cup \{w_h\} \in H_{-w_{h-1}}$ where we set $w_0 := u$.

Given any t-chain $(u, C_1, w_1, C_2, w_2, \ldots, C_t, w_t)$, we let the negation of the chain, denoted by $-(u, C_1, w_1, C_2, w_2, \ldots, C_t, w_t)$, be the chain $(-u, -C_1, -w_1, -C_2, -w_2, \ldots, -C_t, -w_t)$.

As before, we note that the linear equation defined by a t-chain or its negation is satisfied by any $x \in \mathcal{L}$.

In Section 5, we defined an instance polynomial Φ_b related to the system of linear constraints. This was natural over $\mathbb{F}_2$ as there is a group isomorphism between $(\mathbb{F}_2, +)$ and $\{-1, 1\} \in (\mathbb{R}, \times)$. Here, we can make a similar definition by using a nontrivial group homomorphism π from $(\mathbb{F}, +)$ to $(\mathbb{C}, \times)$ where the image of π is contained in the unit circle $\{z \in \mathbb{C} : |z| = 1\}$. However, the instance polynomial Φ_b (and the "decomposed polynomials" $\Psi_{i,Q,p}$ defined later) were only formally needed to discuss sets of linear constraints that are satisfied by the subspace $\mathcal{L}$. Thus, to avoid using the group homomorphism π, here we shall simply use these polynomials to refer to the underlying sets of constraints.

We now perform the hypergraph decomposition step as in Section 6, which is unchanged (once we use the updated definition of chain).[13] This produces the subinstances $\Psi^{(t)}(x, y)$, as before.

We now finish the proof following Section 7.6 in Section 7. We let t denote the value $0 \leq t \leq r$ such that $\Psi^{(t)}$ contains at least $k(3\delta n)^{r+1}/(r+1)$ constraints. Applying the Cauchy–Schwarz trick, we then have that there exists a maximum directed matching M on $[k]$ such that

$$\frac{1}{2k|P_t|} \left(\frac{k(3\delta n)^{r+1}}{r+1} \right)^2 - \frac{|P_t|}{2} \left((3\delta n)^{r-t} d^t \right)^2$$

is a lower bound on number of constraints in the system of linear equations given by:

$$b_i - b_j = \sum_{h=0}^{r-t} x_{C_h} + \sum_{h=1}^{t} x_{C_{r-t+h} \setminus Q_h} + \sum_{h=0}^{r-t} x_{-C'_h} + \sum_{h=1}^{t} x_{-C'_{r-t+h} \setminus -Q_h}$$

for every $(Q, p) \in P_t$, $(i, C_0, w_0, C_1, w_1, C_2, w_2, \ldots, C_r, w_r) \in \mathcal{H}_{i,Q,p}^{(r)}$, $(j, C'_0, w'_0, C'_1, w'_1, C'_2, w'_2, \ldots, C'_r, w'_r) \in \mathcal{H}_{j,Q,p}^{(r)}$. Here, we let $x_{C_h} := \sum_{v \in C_h} x_v$.

As before, the definition of the Kikuchi matrices Definition 7.2 is nearly identical: we merely swap b_j with $-b_j$. Because of this, the key technical part of the argument, namely the row pruning step Lemma 7.4, holds without any changes.

Now, we define the code $\mathcal{L}' \colon \mathbb{F}^k \to \mathbb{F}^{2N}$ identically as before. We let $L = \{i : (i, j) \in M\}$ denote the "left halves" of the edges in the matching M, and we define $\mathcal{L}' \colon \mathbb{F}^L \to \mathbb{F}^{2N}$ to be the same map as before; we simply replace sums with products, as we have not used the homomorphism π to embed $\mathbb{F}$ into $\mathbb{C}$. Namely, for $x' = \mathcal{L}'(b)$, the $\vec{S}$-th entry of x' is $x'_{\vec{S}} = \sum_{h=0}^{r-t} (x_{S_h} + x_{S'_h}) + \sum_{h=1}^{t} x_{R_h}$, and similarly for the $\vec{T}$-th entry, where $x = \mathcal{L}(b)$.

Now, the same calculation as before shows that $\mathcal{L}'$ is a $(2, \delta)$-LDC for $\delta' = \Omega(\delta/r^2)$ provided that $\delta^2 k \leq O(\log^2 n)$. Namely, there are matchings $G''_{i,j}$ on $[2N]$ such that (1) for any $(\vec{S}, \vec{T}) \in E(G''_{i,j})$,

---

[13] We note that the naive application of the decomposition step will produce partitions $\mathcal{H}_{Q,p}^{(r)}$ where $\mathcal{H}_{-Q,p}^{(r)}$ is not necessarily equal to $-\mathcal{H}_{Q,p}^{(r)}$. This turns out to not matter in the proof; as it turns out, we merely need that both decompositions $\cup_{Q,p} - \mathcal{H}_{Q,p}^{(r)}$ and $\cup_{Q,p} \mathcal{H}_{Q,p}^{(r)}$ are both contiguously regular partitions of $\mathcal{H}^{(r)}$, which obviously holds. Nonetheless, we note that one could also easily modify the decomposition step to respect this negation action.

it holds that $x'_{\vec{S}} - x'_{\vec{T}} = b_i - b_j = b_i$ (as $b_j = 0$ for $j \notin L$), and (2) $\frac{1}{k'} \sum_{(i,j) \in M} |E(G''_{i,j})| \geq \delta' \cdot 2N$, where $k' = |L| \geq \frac{k-1}{2}$.

As before, we now apply Fact 2.5. It follows that

$$O(\ell r \log n) \geq 2 \log_2 N \geq \delta' k \geq \Omega(\delta k / r^2)$$
$$\implies k \leq O(\ell r^3 \log n / \delta) \leq O(\log^8 n / \delta^2) \ .$$

Recall now that we had redefined n to be $n(|\mathbb{F}| - 1)$ and δ to be $\delta = \delta / (|\mathbb{F}| - 1)$. Thus, we have that for the original code, $\frac{k\delta^2}{(|\mathbb{F}|-1)^2} \leq O(\log^8 n)$ provided that $|\mathbb{F}| \leq n$. Note that if $|\mathbb{F}| \geq k$, then Theorem 1 becomes trivial, and so we can assume that $|\mathbb{F}| \leq k \leq n$ (as we always have $k \leq n$). Finally, we note that we have assumed (when we substitute back the original values of δ and n) that $\frac{\delta^2 k}{(|\mathbb{F}|-1)^2} \leq O(\log^2(n|\mathbb{F}|))$, which implies that $\frac{\delta^2 k}{(|\mathbb{F}|-1)^2} \leq O(\log^2 n)$, as we may again assume that $|\mathbb{F}| \leq n$. This is a stronger lower bound than Theorem 1, and so this finishes the proof of Theorem 1 for larger fields.