THEORY OF COMPUTING, Volume 15 (12), 2019, pp. 1–24
www.theoryofcomputing.org

Outlaw Distributions and Locally Decodable Codes

Jop Briët* Zeev Dvir† Sivakanth Gopi‡

Received August 11, 2017; Revised June 20, 2019; Published October 29, 2019

Abstract: Locally decodable codes (LDCs) are error correcting codes that allow for decoding of a single message bit using a small number of queries to a corrupted encoding. Despite decades of study, the optimal trade-off between query complexity and codeword length is far from understood. In this work, we give a new characterization of LDCs using distributions over Boolean functions whose expectation is hard to approximate (in L_∞ norm) with a small number of samples. We coin the term "outlaw distributions" for such distributions since they "defy" the Law of Large Numbers. We show that the existence of outlaw distributions over sufficiently "smooth" functions implies the existence of constant query LDCs and vice versa. We give several candidates for outlaw distributions over smooth functions coming from finite field incidence geometry, additive combinatorics and hypergraph (non)expanders.

We also prove a useful lemma showing that (smooth) LDCs which are only required to work on average over a random message and a random message index can be turned into true LDCs at the cost of only constant factors in the parameters.

A conference version of this paper appeared in the Proc. of the 8th Innovations in Theoretical Computer Science Conf. ITCS 2017 [7]. The present version includes a new application of the results to a problem from additive combinatorics (Corollary 1.7) and corrects a minor error in the proof of Theorem 3.1.

*Supported by a VENI grant and the Gravitation-grant NETWORKS-024.002.003 from the Netherlands Organisation for Scientific Research (NWO).

†Research supported by NSF CAREER award DMS-1451191 and NSF grant CCF-1523816.

‡Research supported by NSF grant CCF-1523816 and by the Sloan foundation.

ACM Classification: G.2

AMS Classification: 68Q87, 68R05, 05D40, 68R10, 05E99, 05B25

Key words and phrases: locally decodable codes, Boolean functions, incidence geometry, pseudorandomness, influence, gaussian width, hypergraphs

 DOI: 10.4086/toc.2019.v015a012

1 Introduction

Error correcting codes (ECCs) solve the basic problem of communication over noisy channels by encoding a message into a codeword from which, even if the channel partially corrupts it, the message can later be retrieved. Pioneering work of Shannon [32] showed the existence of optimal (capacity-achieving) ECCs, giving one of the earliest applications of the probabilistic method. The problem of explicitly constructing such codes has fueled the development of coding theory ever since. Similarly, the exploration of many other fascinating structures, such as Ramsey graphs, expander graphs, two-source extractors, etc., began with a striking existence proof via the probabilistic method, only to be followed by decades of catch-up work on explicit constructions. Locally decodable codes (LDCs) are a special class of error correcting codes whose development has not followed this line. The defining feature of LDCs is that they allow for ultrafast decoding of single message bits, a property that typical ECCs lack, as their decoders must read an entire (possibly corrupted) codeword to achieve the same. They were first formally defined in the context of channel coding in [25], although they (and the closely related locally correctable codes) implicitly appeared in several previous works in other settings, such as computation and program checking [4, 5], probabilistically checkable proofs [3, 2] and private information retrieval schemes (PIRs) [11]. More recently, LDCs have even found applications in Banach-space geometry [9] and LDC-inspired objects called local reconstruction codes found applications in fault tolerant distributed storage systems [21]. See [40] for a survey of LDCs and some of the applications.

Despite their many applications, our knowledge of LDCs is very limited; the best constructions we know are far from what is currently known about their limits. Although standard random (linear) ECCs do allow for some weak local-decodability, they are outperformed by even the earliest explicit constructions [26]. All the known constructions of LDCs were obtained by explicitly designing such codes using some algebraic objects like low-degree polynomials or matching vectors [40].

In this paper, we give a characterization of LDCs in probabilistic and geometric terms, making them amenable to probabilistic constructions. On the flip side, these characterizations might also be easier to work with for the purpose of showing lower bounds. We will make this precise in the next section. Let us first give the formal definition of an LDC.

Definition 1.1 (Locally decodable code)**.** For positive integers k,n,q and parameters $\eta,\delta \in (0,1/2]$, a map $C : \{0,1\}^k \to \{0,1\}^n$ is a (q,δ,η)-*locally decodable code* if, for every $i \in [k]$, there exists a randomized *decoder* (a probabilistic algorithm) $\mathcal{A}_i$ such that:

- For every message $x \in \{0,1\}^k$ and string $y \in \{0,1\}^n$ that differs from the codeword $C(x)$ in at most δn coordinates,
$$\Pr[\mathcal{A}_i(y) = x_i] \geq \frac{1}{2} + \eta. \tag{1.1}$$
- The decoder $\mathcal{A}_i$ non-adaptively[1] queries at most q coordinates of y.[2]

[1] Any adaptive q-query decoder can be converted into a (2^q-1)-query non-adaptive decoder. Since in this paper we only deal with constant q, we will assume that the decoders are always non-adaptive.

[2] We can assume that on input $y \in \{0,1\}^n$, the decoder $\mathcal{A}_i$ first samples a set $S \subseteq [n]$ of at most q coordinates according to a probability distribution depending on i only and then returns a random bit depending only on i, S and the values of y at S.

Known results. The main parameters of LDCs are the number of queries q and the length n of the encoding as a function of k and q, typically the parameters δ, η are fixed constants. The simplest example is the Hadamard code, which is a 2-query LDC with $n = 2^k$. The 2-query regime is the only nontrivial case where optimal lower bounds are known: it was shown in [27, 20] that exponential length is necessary. In general, Reed-Muller codes of degree $q-1$ are q-query LDCs of length $n = \exp(O(k^{1/q-1}))$. For a long time, these were the best constructions for constant q, until in breakthrough works, Yekhanin [39] and Efremenko [16] constructed 3-query LDCs with subexponential length $n = \exp(\exp(O(\sqrt{\log k})))$. More generally they constructed 2^r-query LDCs with length $n = \exp(\exp(O(\log^{1/r} k)))$. For $q \geq 3$, the best currently known lower bounds leave huge gaps, giving only polynomial bounds. A 3-query LDC must have length $n \geq \tilde{\Omega}(k^2)$, and more generally for a q-query LDC, $n \geq \tilde{\Omega}(k^{1+1/(\lceil q/2 \rceil - 1)})$ [27, 38].[3] LDCs where the codewords are over a large alphabet are studied because of their relation to private information retrieval schemes [11, 25]. In [15], 2-query LDCs of length $n = \exp(k^{o(1)})$ over an alphabet of size $\exp(k^{o(1)})$ were constructed. There is also some exciting recent work on LDCs where the number of queries grows with k, in which case there are explicit constructions with constant rate (that is, $n = O(k)$) and query complexity $q = \exp(O(\sqrt{\log n}))$; in fact we can even achieve the optimal rate-distance tradeoff of traditional error correcting codes [29, 28, 22]. We cannot yet rule out the exciting possibility that constant rate LDCs with polylogarithmic query complexity exist.

1.1 LDCs from distributions over smooth Boolean functions

Our main result shows that LDCs can be obtained from "outlaw" distributions over "smooth" functions. The term outlaw refers to the Law of Large Numbers, which says that the average of independent samples tends to the expectation of the distribution from which they are drawn. Roughly speaking, a probability distribution is an outlaw if many samples are needed for a good estimation of the expectation and a smooth function over the n-dimensional Boolean hypercube is one that has no influential variables. Paradoxically, while many instances of the probabilistic method use the fact that sample means of a small number of independent random variables tend to concentrate around the true mean, as captured for example by the Chernoff bound, our main result requires precisely the opposite. We show that if *at least* k samples from a distribution over smooth functions are needed to approximate the mean, then there exists an $O(1)$-query LDC sending $\{0,1\}^{\Omega(k)}$ to $\{0,1\}^n$, where the hidden constants depend only the smoothness and mean-estimation parameters.

To make this precise, we now formally define smooth functions and outlaw distributions. Given a function $f : \{-1,1\}^n \to \mathbb{R}$, its *spectral norm* is defined by

$$\|f\|_{\mathrm{sp}} = \sum_{S \subset [n]} |\widehat{f}(S)|,$$

where $\widehat{f}(S)$ are the Fourier coefficients of f (see Section 2 for basics on Fourier analysis). In words, the spectral norm of f is the ℓ_1-norm of its spectrum (i.e., its Fourier coefficients).[4] We also consider the supremum norm, $\|f\|_{L_\infty} = \sup\{|f(x)| : x \in \{-1,1\}^n\}$. It follows from the Fourier inversion formula that $\|f\|_{L_\infty} \leq \|f\|_{\mathrm{sp}}$. The *discrete derivative* of f in direction $i \in [n]$ is the function $(D_i f)(x) = (f(x) -$

[3] We use $\tilde{\Omega}(\,), \tilde{O}(\,), \tilde{\omega}(\,), \tilde{o}(\,)$ to hide polylogarithmic factors through out this paper.

[4] The spectral norm is also known as the *algebra norm* or *Wiener norm* and is also denoted by $\|f\|_A$ or $\|\hat{f}\|_1$ [23, 33].

$f(x^i))/2$, where x^i is the point that differs from x on the i-th coordinate. Smooth functions are functions whose discrete derivatives have small spectral norms.

Definition 1.2 (σ-smooth functions)**.** For $\sigma > 0$, a function $f : \{-1,1\}^n \to \mathbb{R}$ is σ-smooth if for every $i \in [n]$, we have $\|D_i f\|_{\mathrm{sp}} \le \sigma/n$.

Intuition for the above definition may be gained from the fact that smooth functions have no influential variables. The influences, $(\mathbb{E}_{x\in\{-1,1\}^n}[(D_i f)(x)^2])^{1/2}$, measure the extent to which changing the i-th coordinate of a randomly chosen point changes the value of f. Since $\|D_i f\|_{L_\infty} \le \|D_i f\|_{\mathrm{sp}}$, the directional derivatives of σ-smooth functions are uniformly bounded by σ/n, which is a much stronger condition than saying that the derivatives are small on average. Outlaws are defined as follows.

Definition 1.3 (Outlaw)**.** Let n be a positive integer and μ be a probability distribution over real-valued functions on $\{-1,1\}^n$. For a positive integer k and $\varepsilon > 0$, say that μ is a (k,ε)-*outlaw* if for independent random μ-distributed functions $f_1,\dots,f_k$ and $\bar{f} = \mathbb{E}_\mu[f]$,

$$\mathbb{E}\left[\left\|\frac{1}{k}\sum_{i=1}^{k}(f_i - \bar{f})\right\|_{L_\infty}\right] \ge \varepsilon.$$

Denote by $\kappa_\mu(\varepsilon)$ the largest integer k such that μ is a (k,ε)-outlaw.

To approximate the true mean of an outlaw μ to within ε on average in the L_∞-distance, one thus needs $\kappa_\mu(\varepsilon)+1$ samples. Note that if μ is a distribution over σ-smooth functions, then the distribution $\tilde{\mu}$ obtained by scaling functions in the support of μ by $1/\sigma$ is a distribution over 1-smooth functions and $\kappa_{\tilde{\mu}}(\varepsilon/\sigma) = \kappa_\mu(\varepsilon)$.

Our main result is then as follows.

Theorem 1.4 (Main theorem)**.** *Let n be a positive integer and $\varepsilon > 0$. Let μ be a probability distribution over 1-smooth functions on $\{-1,1\}^n$ and $k = \kappa_\mu(\varepsilon)$. Then, there exists a (q,δ,η)-LDC sending $\{0,1\}^l$ to $\{0,1\}^{2n}$ where $l = \Omega(\varepsilon^2 k/\log(1/\varepsilon))$, $q = O(1/\varepsilon)$, $\delta = \Omega(\varepsilon)$ and $\eta = \Omega(\varepsilon)$. Additionally, if μ is supported on degree-d polynomials, then we can take $q = d$.*

Note that the smoothness requirement is essential. For example the uniform distribution over the n dictator functions $f_i(x) = x_i$ for $i \in [n]$ is an $(n/2,1)$-outlaw, but it cannot imply constant rate since constant query LDCs which we know do not exist. In fact we establish a converse to Theorem 1.4, showing that its hypothesis is essentially equivalent to the existence of LDCs in the small query complexity regime.

Theorem 1.5. *If there exists a $C : \{0,1\}^k \to \{0,1\}^n$ is a (q,δ,η)-LDC, then there exists a probability distribution μ over 1-smooth degree-q polynomials on $\{-1,1\}^n$ such that*

$$\kappa_\mu(\varepsilon) \ge \eta k$$

where $\varepsilon = \eta\delta/(q2^{q/2})$.

Theorem 1.5 can in turn convert the problem of proving lower bounds on the length of LDCs to a problem on Banach space geometry. In particular, for a distribution μ over 1-smooth degree-q polynomials on $\{0,1\}^n$, one can upper bound $\kappa_\mu(\varepsilon)$ in terms of type constants of the space formed by the q-fold injective tensor product of ℓ_q^n with itself (or the space of q-tensors of size n viewed as multilinear forms on ℓ_q^n) [6]. Unfortunately, little is known about these constants, however.

Candidate outlaws. One scenario in which outlaw distributions can be obtained is using incidence geometry in finite fields. In particular, the following result can be derived from our main theorem (stated a bit informally here, see Section 6.1 for the formal version).

Corollary 1.6. *Let $p > 2$ be a fixed prime. Suppose that for a random set of directions, $D \subseteq \mathbb{F}_p^n$, of size $|D| \leq k$, with probability at least $1/2$, there exists a set $B \subset \mathbb{F}_p^n$ of size $|B| \geq \Omega(p^n)$ which does not contain any lines with direction in D. Then, there exists a p-query LDC sending $\{0,1\}^{\Omega(k)}$ to $\{0,1\}^{2p^n}$.*

The assumption in Corollary 1.6 that D be random is essential for it to be potentially interesting for LDCs. If we instead ask that every set D of directions satisfy the condition—as we did in the conference version of this paper—then letting D be a subspace shows that k must be smaller than a constant depending only on p and ε by Theorem 6.2 below.

The analogue of Corollary 1.6 in $\mathbb{Z}/N\mathbb{Z}$ where lines correspond to arithmetic progressions and directions correspond to common differences can also be used to construct LDCs. This setting was studied in [17], where it is shown that if D is a random subset of $\mathbb{Z}/N\mathbb{Z}$ of size $\omega(\log N)$, then almost surely every dense subset of $\mathbb{Z}/N\mathbb{Z}$ contains a 3-term arithmetic progression with common difference in D. In [18], the same authors showed that for every fixed integer $q \geq 4$, the same holds for q-term arithmetic progressions if D has of size $\omega(N^{1-1/q})$.[5] Our main result, together with the best currently known lower bounds on LDCs [27, 38] can be used to improve the bounds of [18] on random differences in Szemerédi's theorem as shown in the following corollary.[6]

Corollary 1.7. *Let $q \geq 3$ be an integer and $\alpha \in (0,1]$. Let $D \subseteq \mathbb{Z}/N\mathbb{Z}$ be a uniformly random subset of size $\tilde{\omega}(N^{1-1/\lceil q/2 \rceil})$. Then, almost surely as $N \longrightarrow \infty$, it holds that every subset $A \subset \mathbb{Z}/N\mathbb{Z}$ of size at least αN contains a q-term arithmetic progression with common difference in the set D.*

Another setting in which our approach leads to interesting open problems is in relation to pseudorandom hypergraphs. Consider a partition of the complete bipartite graph $K_{n,n}$ into n perfect matchings. It is known that picking $k = O(\log n)$ of these matchings at random will give us a pseudorandom (expander) graph (of degree k). For some particular partitions (e.g., given by an Abelian group) this bound is tight. The questions arising from our approach can be briefly summarized as follows: Can one find an n-vertex hypergraph H (say three uniform to be precise) and a partition of H into matchings so that, to get a pseudorandom hypergraph (defined appropriately) one needs at least k random matchings. This would give a code sending $\Omega(k)$-bit messages with encoding length $O(n)$ and so, becomes interesting when k is super-polylogarithmic in n. We elaborate on this in Section 6.2.

1.2 Techniques

Our proof of Theorem 1.4 proceeds in two steps. The first step consists of turning an outlaw over smooth functions into a seemingly crude type of LDC that is only required to work on average over a uniformly distributed message and a uniformly distributed message index. We call such codes *average-case smooth*

[5] Their results are actually stated slightly differently. Instead, their random set D is formed by putting each element of $\mathbb{Z}/N\mathbb{Z}$ in it independently with probability $p = \omega(N^{-1/q})$. However, standard arguments show that the two statements are equivalent.

[6] In [8], the first and third authors derive Corollary 1.7 in addition to another result in probabilistic combinatorics from first principles (which is to say, not through LDCs).

codes (see below). The second step consists of showing that such codes are in fact not much weaker than honest LDCs.

From outlaws to average-case smooth codes. The key ingredient for the first step is *symmetrization*, a basic technique from probability theory. We briefly sketch how this is used (we refer to Section 3 for the full proof). Suppose that $f_1,\dots,f_k$ are independent samples distributed according to a (k,ε)-outlaw with expectation $\bar{f}$ and supported on 1-smooth functions. We introduce an independent copy[7] f_i' of f_i for each $i \in [k]$ and consider the symmetrically distributed random functions $f_i - f_i'$. Since $\bar{f} = \mathbb{E}[f_i']$ for each $i \in [k]$, Jensen's inequality and Definition 1.3 imply that

$$\mathbb{E}\big[\|(f_1 - f_1') + \cdots + (f_k - f_k')\|_{L_\infty}\big] \geq \mathbb{E}\big[\|(f_1 - \mathbb{E}[f_1']) + \cdots + (f_k - \mathbb{E}[f_k'])\|_{L_\infty}\big] \geq \varepsilon k.$$

Since the random functions $f_i - f_i'$ are independent and symmetric, we get that for independent uniformly random signs $x_1,\dots,x_k \in \{-1,1\}$, the above left-hand side equals

$$\mathbb{E}\big[\|x_1(f_1 - f_1') + \cdots + x_k(f_k - f_k')\|_{L_\infty}\big].$$

The triangle inequality and the Averaging Principle then give that there exist *fixed* smooth functions $f_1^\star,\dots,f_k^\star$ such that on average over the random signs, we have

$$\mathbb{E}\big[\|x_1 f_1^\star + \cdots + x_k f_k^\star\|_{L_\infty}\big] \geq \varepsilon k/2. \tag{1.2}$$

To get an average-case smooth code out of this, we view each sequence $x = (x_1,\dots,x_k)$ as a k-bit message and choose an arbitrary n-bit string for which the L_∞-norm in (1.2) is achieved to be the its encoding, $C(x)$. This gives a map $C : \{-1,1\}^k \to \{0,1\}^n$ satisfying

$$\mathbb{E}\big[x_1 f_1^\star(C(x)) + \cdots x_k f_k^\star(C(x))\big] \geq \varepsilon k/2.$$

Equivalently, for uniform x and i, we have

$$\Pr[\mathcal{A}(f_i^\star(C(x))) = x_i] \geq \frac{1}{2} + \frac{\varepsilon}{4}$$

where $\mathcal{A}(\alpha)$ (for $\alpha \in [-1,1]$) is a ± 1-valued random variable with expected value α.[8] Finally, we use the smoothness property to transform the $f_i^\star$ into decoders with the desired properties. This is done in Section 3. It is in the application of the Averaging Principle where the probabilistic method appears in our construction of LDCs.

[7] in this context sometimes referred to as a "ghost copy" as it will later disappear again

[8] Note that a 1-smooth function is bounded by 1 since $\|f\|_{L_\infty} \leq \|f\|_{\mathrm{sp}} \leq \sum_{i=1}^n \|D_i f\|_{\mathrm{sp}} \leq n \cdot (1/n) \leq 1$.

Average-case smooth codes are LDCs. Katz and Trevisan [25] observed that LDC decoders must have the property that they select their queries according to distributions that do not favor any particular coordinate. The intuition for this is that if they did favor a certain coordinate, then corrupting that coordinate would cause the decoder to err with too high a probability. If instead, queries are sampled according to a "smooth" distribution, they will all fall on uncorrupted coordinates with good probability provided the fraction δ of corrupted coordinates and query complexity q aren't too large. The following definition allows us to make this intuition precise.

Definition 1.8 (Smooth code)**.** For positive integers k,n,q and parameters $\eta \in (0,1/2]$ and $c > 0$, a map $C : \{0,1\}^k \to \{0,1\}^n$ is a (q,c,η)-*smooth code* if, for every $i \in [k]$, there exists a randomized decoder $\mathcal{A}_i : \{0,1\}^n \to \{0,1\}$ such that:

- For every $x \in \{0,1\}^k$,
$$\Pr\big[x_i = \mathcal{A}_i\big(C(x)\big)\big] \geq \frac{1}{2} + \eta. \tag{1.3}$$
- The decoder $\mathcal{A}_i$ non-adaptively queries at most q coordinates of $C(x)$.
- For each $j \in [n]$, the probability that $\mathcal{A}_i$ queries the coordinate $j \in [n]$ is at most c/n.

The formal version of Katz and Trevisan's observation is as follows.

Theorem 1.9 (Katz–Trevisan)**.** *If $C : \{0,1\}^k \to \{0,1\}^n$ is a (q,δ,η)-LDC, then C is a $(q,q/\delta,\eta)$-smooth code. Conversely, if $C : \{0,1\}^k \to \{0,1\}^n$ is a (q,c,η)-smooth code, then C is a $(q,\eta/2c,\eta/2)$-LDC.*

Our second step in the proof of Theorem 1.4 is a stronger form of the converse part of Theorem 1.9. We show that even smooth codes that are only required to work *on average* can be turned into LDCs, losing only a constant factor in the rate and success probability.

Definition 1.10 (Average-case smooth code)**.** A code as in Definition 1.8 is a (q,c,η)-*average-case* smooth code if instead of the first item, (1.3) is required to hold only on average over uniformly distributed $x \in \{0,1\}^k$ and uniformly distributed $i \in [k]$, which is to say that

$$\Pr\big[x_i = \mathcal{A}_i\big(C(x)\big)\big] \geq \frac{1}{2} + \eta,$$

where the probability is taken over x, i and the randomness used by $\mathcal{A}_i$.

Lemma 1.11. *Let $C : \{0,1\}^k \to \{0,1\}^n$ be a (q,c,η)-average-case smooth code. Then, there exists an $(q,\Omega(\eta/c),\Omega(\eta))$-LDC sending $\{0,1\}^l$ to $\{0,1\}^n$ where $l = \Omega(\eta^2 k/\log(1/\eta))$.*

The idea behind the proof of Lemma 1.11 is as follows. We first switch the message and codeword alphabets to $\{-1,1\}$ and let $f_i : \{-1,1\}^k \to [-1,1]$ be the expected decoding function $f_i(z) = \mathbb{E}[\mathcal{A}_i(z)]$. The properties of C then easily imply that the set $T \subseteq [-1,1]^k$ given by $T = \{(f_1(z),\ldots,f_k(z)) : z \in \{-1,1\}^k\}$ has large *Gaussian width*, in particular it holds that for a standard k-dimensional Gaussian vector g, we have $\mathbb{E}[\sup_{t\in T}\langle g,t\rangle] \gtrsim \varepsilon k$.[9] Next, we employ a powerful result of [31] showing that T

[9]We write $A \gtrsim B$ and $A = \Omega(B)$ interchangeably to mean that $A \geq cB$ for some absolute constant $c > 0$ independent of all parameters involved.

contains an l-dimensional hypercube-like structure with edge length some absolute constant $c \in (0,1]$, for $l \gtrsim k$. Roughly speaking, this implies that C is a smooth code on $\{-1,1\}^l$ whose decoding probability depends on ε and c. Finally, we obtain an LDC via an application of Theorem 1.9. The full proof is given in Section 4.

1.3 Organization

Section 2 contains some preliminaries in Fourier analysis over the Boolean cube. In Section 3, we prove our main theorem (Theorem 1.4) by first showing that outlaw distributions over smooth functions imply existence of average-case smooth codes and using Lemma 1.11 to convert them to LDCs. In Section 4, we prove Lemma 1.11 showing how to convert average-case smooth-codes to LDCs. In Section 5, we show the converse to our main theorem (Theorem 1.5) showing how to get outlaw distributions over smooth functions from LDCs. Finally in Section 6, we give some candidate constructions of outlaw distributions over smooth functions using incidence geometry and hypergraph pseudorandomness.

2 Preliminaries

We recall a few basic definitions and facts from analysis over the n-dimensional Boolean hypercube $\{-1,1\}^n$. Equipped with the coordinate-wise multiplication operation, the hypercube forms an Abelian group whose group of characters is formed by the functions $\chi_S(x) = \prod_{i\in S} x_i$ for all $S \subseteq [n]$. The characters form a complete orthonormal basis for the space of real-valued functions on $\{-1,1\}^n$ endowed with the inner product

$$\langle f, g\rangle = \mathbb{E}_{x\in\{-1,1\}^n}[f(x)g(x)],$$

where we use the notation $\mathbb{E}_{a\in S}$ to denote the expectation with respect to a uniformly distributed element a over a set S. The *Fourier transform* of a function $f : \{-1,1\}^n \to \mathbb{R}$ is the function $\widehat{f} : 2^{[n]} \to \mathbb{R}$ defined by $\widehat{f}(S) = \langle f, \chi_S\rangle$. The Fourier inversion formula (which follows from orthonormality of the character functions) asserts that

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

Parseval's Identity relates the L_2-norms of f and its Fourier transform by

$$\left(\mathbb{E}_{x\in\{-1,1\}^n}[f(x)^2]\right)^{1/2} = \Big(\sum_{S\subseteq[n]} |\widehat{f}(S)|^2\Big)^{1/2}.$$

A function f has *degree* q if $\widehat{f}(S) = 0$ when $|S| > q$ and the *degree-q truncation* of f, denoted $f^{\leq q}$, is the degree-q polynomial defined by

$$f^{\leq q} = \sum_{|S|\leq q} \widehat{f}(S)\chi_S.$$

A function f is a *q-junta* if it depends only on a subset of q of its variables, or equivalently, if there exists a subset $T \subseteq [n]$ of size $|T| \leq q$ such that $\widehat{f}(S) = 0$ for every $S \not\subseteq T$. The *i-th discrete derivative* $D_i f$ is the function $(D_i f)(x) = (f(x) - f(x^i))/2$, where x^i is the point that differs from x on the i-th coordinate. It is easy to show that the i-th discrete derivative in of a function f is given by

$$D_i f = \sum_{S \ni i} \widehat{f}(S)\chi_S.$$

Hence, it follows that $\|D_i f\|_{\mathrm{Spec}} = \sum_{S \ni i} |\widehat{f}(S)|$.

3 From outlaws to average-case smooth codes

In this section we prove Theorem 1.4. For convenience, in the remainder of this paper, we switch the message and codeword alphabets of all codes from $\{0,1\}^n$ to $\{-1,1\}^n$. We begin by showing that outlaw distributions over degree-q polynomials give q-query average-case smooth codes. Combined with Lemma 1.11, this implies the second part of Theorem 1.4.

Theorem 3.1. *Let μ be a probability distribution on 1-smooth degree-q polynomials on $\{-1,1\}^n$, let $\varepsilon \in (0,1]$ and let $k = \kappa_\mu(\varepsilon)$. Then, there exists a $(q,2,\varepsilon/4)$-average-case smooth code sending $\{-1,1\}^k$ to $\{-1,1\}^{2n}$.*

Proof. The proof uses a symmetrization argument. Let $\mathcal{F} = (f_1,\dots,f_k)$ and $\mathcal{F}' = (f_1',\dots,f_k')$ be two k-tuples of independent μ-distributed random variables and let $\bar{f} = \mathbb{E}_\mu[f]$. Then, by definition of $\kappa_\mu(\varepsilon)$ and triangle inequality for L_∞ norm,

$$\begin{aligned}
\varepsilon &\le \mathbb{E}_{\mathcal{F}}\Big[\Big\|\frac{1}{k}\sum_{i=1}^{k}(f_i - \bar{f})\Big\|_{L_\infty}\Big] \\
&= \mathbb{E}_{\mathcal{F}}\Big[\Big\|\frac{1}{k}\sum_{i=1}^{k}\big(f_i - \mathbb{E}_{\mathcal{F}'}[f_i']\big)\Big\|_{L_\infty}\Big] \\
&\le \mathbb{E}_{\mathcal{F},\mathcal{F}'}\Big[\Big\|\frac{1}{k}\sum_{i=1}^{k}(f_i - f_i')\Big\|_{L_\infty}\Big].
\end{aligned}$$

The random variables $f_i - f_i'$ are symmetrically distributed, which is to say that they have the same distribution as their negations $f_i' - f_i$. Since they are independent, it follows that for every $x \in \{-1,1\}^k$, the random variable $x_1(f_1 - f_1') + \dots + x_k(f_k - f_k')$ has the same distribution as $(f_1 - f_1') + \dots + (f_k - f_k')$. Therefore,

$$\begin{aligned}
\mathbb{E}_{\mathcal{F},\mathcal{F}'}\Big[\Big\|\frac{1}{k}\sum_{i=1}^{k}(f_i - f_i')\Big\|_{L_\infty}\Big] &= \mathbb{E}_{x\in\{-1,1\}^k}\Big[\mathbb{E}_{\mathcal{F},\mathcal{F}'}\Big[\Big\|\frac{1}{k}\sum_{i=1}^{k}x_i(f_i - f_i')\Big\|_{L_\infty}\Big]\Big] \\
&\le 2\mathbb{E}_{\mathcal{F}}\Big[\mathbb{E}_{x\in\{-1,1\}^k}\Big[\Big\|\frac{1}{k}\sum_{i=1}^{k}x_i f_i\Big\|_{L_\infty}\Big]\Big].
\end{aligned}$$

Applying the Averaging Principle to the outer expectation, we find that there exist 1-smooth degree-q polynomials $f_1^\star,\dots,f_k^\star : \{-1,1\}^n \to \mathbb{R}$ such that

$$\mathbb{E}_{x\in\{-1,1\}^k}\Big[\Big\|\frac{1}{k}\sum_{i=1}^{k}x_i f_i^\star\Big\|_{L_\infty}\Big] \ge \frac{\varepsilon}{2}. \tag{3.1}$$

Let $\tilde{C}:\{-1,1\}^k\to\{-1,1\}^n$ and $\sigma:\{-1,1\}^k\to\{-1,1\}$ be such that for each $x\in\{-1,1\}^k$, we have

$$\sigma(x)\frac{1}{k}\sum_{i=1}^k x_i f_i^\star\big(\tilde{C}(x)\big)=\Big|\frac{1}{k}\sum_{i=1}^k x_i f_i^\star\big(\tilde{C}(x)\big)\Big|=\Big\|\frac{1}{k}\sum_{i=1}^k x_i f_i^\star\Big\|_{L_\infty}. \tag{3.2}$$

Let $C:\{-1,1\}^k\to\{-1,1\}^{2n}$ be the code given by $C(x)=(\tilde{C}(x),\sigma(x)\tilde{C}(x))$, that is, the concatenation of $\tilde{C}(x)$ and either an unchanged or a coordinate-wise negated copy of $\tilde{C}(x)$ depending on the value of $\sigma(x)$. For each $i\in[k]$, define the decoder $\mathcal{A}_i$ as follows. Let $\nu_i:2^{[n]}\to[0,1]$ be the probability distribution defined by $\nu_i(S)=|\widehat{f_i^\star}(S)|/\|f_i^\star\|_{\mathrm{sp}}$. Given a string $(z,z')\in\{-1,1\}^{2n}$, with probability $1-\|f_i^\star\|_{\mathrm{sp}}$, the decoder $\mathcal{A}_i$ returns a uniformly random sign, and with probability $\|f_i^\star\|_{\mathrm{sp}}$, it samples a set $S\subseteq[n]$ according to ν_i and a uniformly random $s\in S$, and returns $\chi_{S\setminus\{s\}}(z)z'_s$. To see that $\|f_i^\star\|_{\mathrm{sp}}\le 1$, observe that for any 1-smooth function f, we have

$$\|f\|_{\mathrm{sp}}=\sum_{S\subset[n]}|\widehat{f}(S)|\le\sum_{S\subset[n]}|S||\widehat{f}(S)|=\sum_{i=1}^n\sum_{S\ni i}|\widehat{f}(S)|\le n\frac{1}{n}=1.$$

Then, $\mathcal{A}_i$ queries at most q coordinates of (z,z') and since $f_i^\star$ is 1-smooth, the probability that it queries any coordinate $j\in[2n]$ is at most $\|D_j f_i^\star\|_{\mathrm{sp}}\le 1/n$. For each $x\in\{-1,1\}^k$, the expected output of $\mathcal{A}_i$ on input $C(x)$ satisfies $\mathbb{E}[\mathcal{A}_i(C(x))]=\sigma(x)f_i^\star\big(\tilde{C}(x)\big)$. Therefore, by (3.1) and (3.2), we have

$$\begin{aligned}\mathbb{E}_{x\in\{-1,1\}^k,i\in[k]}\left[\Pr[x_i=\mathcal{A}_i(C(x))]\right]&=\frac{1}{2}+\frac{1}{2}\mathbb{E}_{x\in\{-1,1\}^k,i\in[k]}\left[x_i\mathbb{E}[\mathcal{A}_i(C(x))]\right]\\&=\frac{1}{2}+\frac{1}{2}\mathbb{E}_{x\in\{-1,1\}^k,i\in[k]}\left[x_i\,\sigma(x)\,f_i^\star\big(\tilde{C}(x)\big)\right]\\&=\frac{1}{2}+\frac{1}{2}\mathbb{E}_{x\in\{-1,1\}^k}\left[\Big\|\frac{1}{k}\sum_{i=1}^k x_i f_i^\star\Big\|_{L_\infty}\right]\\&\ge\frac{1}{2}+\frac{\varepsilon}{4}.\end{aligned}$$

Hence, C is a $(q,1/2,\varepsilon/4)$-average-case smooth code. □

The final step before the proof of Theorem 1.4 is to show that for any distribution μ over smooth functions, there exists a distribution $\tilde{\mu}$ over smooth functions of bounded degree that is not much more concentrated than μ.

Lemma 3.2. *Let μ be a probability distribution over 1-smooth functions and let $\varepsilon>0$. Then, there exists a probability distribution $\tilde{\mu}$ over 1-smooth functions of degree $q=4/\varepsilon$ such that $\kappa_{\tilde{\mu}}(\varepsilon/2)\ge\kappa_\mu(\varepsilon)$.*

Proof. We first establish that smooth functions have low-degree approximations in the supremum norm. If $f:\{-1,1\}^n\to\mathbb{R}$ is 1-smooth, then

$$q\sum_{|S|>q}|\widehat{f}(S)|\le\sum_{S\subset[n]}|S||\widehat{f}(S)|=\sum_{i=1}^n\sum_{S\ni i}|\widehat{f}(S)|=\sum_{i=1}^n\|D_i f\|_{\mathrm{sp}}\le 1.$$

It follows that the degree-q truncation $f^{\le q}$ satisfies

$$\left\|f - f^{\le q}\right\|_{L_\infty} \le \sum_{|S|>q} |\widehat{f}(S)| \le \frac{1}{q} = \frac{\varepsilon}{4}. \tag{3.3}$$

Define $\tilde{\mu}$ as follows: sample f according to μ and output $f^{\le q}$. Clearly, $\tilde{\mu}$ is also a distribution over 1-smooth functions. For $k = \kappa_\mu(\varepsilon)$, we have

$$\mathbb{E}_{f_1,\dots,f_k\sim\mu}\left[\left\|\frac{1}{k}\sum_{i=1}^{k}\left(f_i - \mathbb{E}[f_i]\right)\right\|_{L_\infty}\right] \ge \varepsilon.$$

Hence, by the triangle inequality and (3.3), we have

$$\mathbb{E}_{f_1,\dots,f_k\sim\tilde{\mu}}\left[\left\|\frac{1}{k}\sum_{i=1}^{k}\left(f_i - \mathbb{E}[f_i]\right)\right\|_{L_\infty}\right] \ge \frac{\varepsilon}{2},$$

giving the claim. □

Proof of Theorem 1.4. By applying Lemma 3.2 to μ, we get a distribution $\tilde{\mu}$ over 1-smooth degree $q = O(1/\varepsilon)$ polynomials with $k' = \kappa_{\tilde{\mu}}(\varepsilon/2) \ge \kappa_\mu(\varepsilon) = k$. By Theorem 3.1, we get a $(q,2,\Omega(\varepsilon))$-average-case smooth code $C' : \{-1,1\}^{k'} \to \{-1,1\}^{2n}$. Finally we use Lemma 1.11 to convert C' to a $(q,\Omega(\varepsilon),\Omega(\varepsilon))$-LDC $C : \{-1,1\}^{\ell} \to \{-1,1\}^{2n}$ where $\ell = \Omega(\varepsilon^2 k'/\log(1/\varepsilon))$. For the last part of the theorem we can apply Theorem 3.1 and Lemma 1.11 directly. □

4 From average-case smooth codes to LDCs

In this section, we prove Lemma 1.11. For this, we need the notion of the Vapnik–Chervonenkis dimension (VC-dimension).

Definition 4.1 (VC-dimension [31])**.** Let $T \subset [-1,1]^k$ and $w > 0$. Then, $\mathrm{vc}(T,w)$ is defined as the size of the largest subset $\sigma \subset [k]$ such that for some shift $s \in [-1,1]^k$, it holds that for every $x \in \{-1,1\}^\sigma$, there exists $t \in T$ such that for every $i \in \sigma$, $(t_i - s_i)x_i \ge w/2$.

Observe that if T is convex, then $\mathrm{vc}(T,w)$ is the maximum dimension of a shifted hypercube with edge lengths at least w contained in T. Note that usually VC-dimension for a subset $T \subset \{0,1\}^n$ is defined as the size of the largest subset $\sigma \subset [n]$ which is fully shattered by T (i.e., $T|_\sigma = \{0,1\}^\sigma$). This coincides with Definition 4.1 for any $w \in (0,1]$ (by the taking the shift $s = (1/2,1/2,\dots,1/2)$). Thus Definition 4.1 can be seen as a generalization of the usual definition to more general sets.

Definition 4.2 (Gaussian width)**.** Let g be a k-dimensional standard Gaussian vector, with independent standard normal distributed entries. The *Gaussian width* of a set $T \subseteq \mathbb{R}^k$ is defined as

$$E(T) = \mathbb{E}_g[\sup_{t\in T}\langle g,t\rangle].$$

It can be shown that large VC-dimension implies large Gaussian width. The following theorem shows the converse: containing a hypercube-like structure is the only way to have large Gaussian width.

Theorem 4.3 ([31])**.** *Let $T \subset [-1,1]^k$. Then, the Gaussian width of T is bounded as*

$$E(T) \lesssim \sqrt{k} \int_{\alpha E(T)/k}^{1} \sqrt{\mathrm{vc}(T,w)\log(1/w)}dw$$

for some absolute constant $\alpha > 0$.

Finally, we use that fact that, as for LDCs, we can assume that on input $y \in \{0,1\}^n$, the decoder $\mathcal{A}_i$ of a smooth code first samples a set $S \subseteq [n]$ of at most q coordinates according to a probability distribution that depends on i only and then returns a random sign depending only on i, S and the values of y at S.

Proof of Lemma 1.11. For the proof we show that the average-case smoothness property implies that the image of the (average) decoding functions has large Gaussian width. Then, Theorem 4.3 gives a hypercube like structure inside the image, which we use to construct a smooth code. Finally we use Theorem 1.9 to convert the smooth code to an LDC.

Recall the switch of the message and codeword alphabets to $\{-1,1\}$. For each $i \in [k]$, let $f_i : \{-1,1\}^n \to [-1,1]$ be the expected decoding function $f_i(z) = \mathbb{E}[\mathcal{A}_i(z)]$. Let g be a standard k-dimensional Gaussian vector and $T = \{(f_1(z),\dots,f_k(z)) : z \in \{-1,1\}^n\}$. By the definition of average-case smooth code we have

$$2\eta k \le \mathbb{E}_{x\in\{-1,1\}^k}\left[\sum_{i=1}^{k} x_i f_i(C(x))\right] \le \mathbb{E}_{x\in\{-1,1\}^k}\left[\sup_{t\in T}\langle x,t\rangle\right] \lesssim \mathbb{E}_g\left[\sup_{t\in T}\langle g,t\rangle\right].$$

(See for instance [34, Lemma 3.2.10] for the last inequality.) By Theorem 4.3, for some constant $\alpha > 0$, we have

$$\eta k \lesssim \sqrt{k}\int_{\alpha\eta}^{1}\sqrt{\mathrm{vc}(T,w)\log(1/w)}dt \le \sqrt{k}\cdot\sqrt{\mathrm{vc}(T,\alpha\eta)\log(1/\alpha\eta)}$$

where we used the fact that $\mathrm{vc}(T,w)$ is decreasing in w. So for $\tau = \alpha\eta$, we have $\mathrm{vc}(T,\tau) \gtrsim \eta^2 k/\log(1/\eta)$. By the definition of VC-dimension, there exists a subset $\sigma \subset [k]$ of size $|\sigma| \ge \mathrm{vc}(T,\tau)$ and a shift $s \in [-1,1]^k$ such that for every $x \in \{-1,1\}^\sigma$ there exists $t \in T$ such that $(t_i - s_i)x_i \ge \tau/2$ for every $i \in \sigma$.

Now we will define the code $C' : \{-1,1\}^\sigma \to \{-1,1\}^n$. Given $x \in \{-1,1\}^\sigma$, there exists $t(x) \in T$ such that $(t(x)_i - s_i)x_i \ge \tau/2$ for every $i \in \sigma$. Define $C'(x) \in \{-1,1\}^n$ to be one of the preimages of $t(x)$ under f, that is,

$$\big(f_1(C'(x)),\dots,f_k(C'(x))\big) = t(x).$$

Let W_p denote a $\{-1,1\}$-valued random variable with mean p. The decoding algorithms $\mathcal{A}'_i(y)$ run $\mathcal{A}_i(y)$ internally and give their output as follows:

$$\mathcal{A}'_i(y) = \begin{cases} \text{Output } W_{(1-s_i)/2} & \text{if } \mathcal{A}_i(y) \text{ returns } 1,\\ \text{Output } -W_{(1+s_i)/2} & \text{if } \mathcal{A}_i(y) \text{ returns } -1.\end{cases}$$

Therefore, for every $x \in \{-1,1\}^{\sigma}$ and for every $i \in \sigma$,

$$\begin{aligned}
x_i \mathbb{E}[\mathcal{A}'_i(C'(x))] &= x_i \mathbb{E}\left[\frac{(1+\mathcal{A}_i(C'(x)))}{2} W_{(1-s_i)/2} - \frac{(1-\mathcal{A}_i(C'(x)))}{2} W_{(1+s_i)/2}\right] \\
&= \frac{x_i}{2} \mathbb{E}\left[\mathcal{A}_i(C'(x)) - s_i\right] \\
&= \frac{x_i}{2} (f_i(C'(x)) - s_i) \\
&= \frac{x_i}{2} (t(x)_i - s_i) \\
&\geq \frac{\tau}{4} \gtrsim \eta.
\end{aligned}$$

Since the probability that $\mathcal{A}'_i(C'(x))$ queries any particular location of $C'(x)$ is still at most c/n, it follows that C' is a $(q,c,\Omega(\eta))$-smooth code. By Theorem 1.9, C' is also a $(q,\Omega(\eta/c),\Omega(\eta))$-LDC. □

5 From LDCs to outlaws

In this section we prove Theorem 1.5, the converse of our main result.

Proof of Theorem 1.5. By Theorem 1.9, the map $C : \{-1,1\}^k \to \{-1,1\}^n$ is also a $(q,q/\delta,\eta)$-smooth code. For each $i \in [k]$, let $\mathcal{B}_i$ be its decoder for the i-th index. Let $\nu_i : 2^{[n]} \to [0,1]$ be the probability distribution used by $\mathcal{B}_i$ to sample a set $S \subseteq [n]$ of at most q coordinates and let $f_{i,S} : \{-1,1\}^n \to [-1,1]$ be function whose value at $y \in \{-1,1\}^n$ is the expectation of the random sign returned by $\mathcal{B}_i(y)$ conditioned on the event that it samples S. Since this value depends only on the coordinates in S, the function $f_{i,S}$ is a q-junta.

Fix an $i \in [k]$ and let $f_i : \{-1,1\}^n \to [-1,1]$ be the function given by $f_i = \mathbb{E}_{S\sim\nu_i}[f_{i,S}]$. Then, since a q-junta has degree at most q, so does f_i. We claim that f_i is $(q2^{q/2})/\delta$-smooth. Since the functions $f_{i,S} : \{-1,1\}^n \to \{-1,1\}$ are q-juntas, it follows from Parseval's identity that they have spectral norm at most $2^{q/2}$. Moreover, for each $j \in [n]$, we have $\Pr_{S\sim\nu_i}[j \in S] \leq q/(\delta n)$. Hence, since $f_{i,S}$ depends only on the coordinates in S, we have

$$\left\|D_j f_i\right\|_{\mathrm{sp}} \leq \sum_{S \ni j} \nu_i(S) \left\|f_{i,S}\right\|_{\mathrm{sp}} \leq \frac{q2^{q/2}}{\delta n},$$

which gives the claim. By (1.3), it holds for every $x \in \{-1,1\}^k$ and every $i \in [k]$ that

$$x_i f_i\big(C'(x)\big) \geq 2\eta. \tag{5.1}$$

Define the distribution μ to correspond to the process of sampling $i \in [k]$ uniformly at random and returning f_i. Let $\bar{g} = (f_1 + \cdots + f_k)/k$ be the mean of μ. We show that $\kappa_\mu(\eta) \geq \eta k$. To this end, let $l = \eta k$, let $\sigma : [l] \to [k]$ be an arbitrary map and define the functions $g_1, \ldots, g_l$ by $g_i = f_{\sigma(i)}$. Let $x \in \{-1,1\}^k$ be such that for each $i \in [l]$, we have $x_{\sigma(i)} = 1$ and $x_j = -1$ elsewhere. It follows from (5.1) that $f_{\sigma(i)}\big(C(x)\big) \in [2\eta, 1]$ for every $i \in [l]$ and that $f_i\big(C(x)\big) \le 0$ for every other $i \in [k]$. Hence,

$$\begin{aligned}\Big\|\frac{1}{l}\sum_{i=1}^{l}(g_i-\bar{g})\Big\|_{L_\infty} &\ge \Big(\frac{1}{l}\sum_{i=1}^{l}(g_i-\bar{g})\Big)\big(C(x)\big)\\ &= \frac{1}{l}\sum_{i=1}^{l} f_{\sigma(i)}\big(C(x)\big) - \frac{1}{k}\sum_{i=1}^{k} f_i\big(C(x)\big)\\ &\ge 2\eta - \frac{l}{k} = \eta.\end{aligned}$$

If σ maps each element in $[l]$ to a uniformly random element in $[k]$, then $g_1,\dots,g_l$ are independent, μ-distributed and satisfy

$$\mathbb{E}\left[\Big\|\frac{1}{l}\sum_{i=1}^{l}(g_i-\bar{g})\Big\|_{L_\infty}\right] \ge \eta,$$

which shows that $\kappa_\mu(\eta) \ge l$. Finally we can scale all the functions in μ to make them 1-smooth, and get a distribution $\tilde{\mu}$ over 1-smooth functions with $\kappa_{\tilde{\mu}}(\eta\delta/(q2^{q/2})) \ge \eta k$. ☐

6 Candidate outlaws

In this section we elaborate on the candidate outlaws mentioned in the introduction.

6.1 Incidence geometry

We begin by describing a variant of Corollary 1.6 based on a slightly different assumption and show conditions under which this assumption holds. Let p be an odd prime, let $\mathbb{F}_p$ be a finite field with p elements and let n be a positive integer. For $x,y \in \mathbb{F}_p^n$ and $y \neq 0$, the *line* with origin x in direction y, denoted $\ell_{x,y}$, is the sequence $(x+\lambda y)_{\lambda\in\mathbb{F}_p}$.

Corollary 6.1. *For every odd prime p and $\varepsilon \in (0,1]$, there exist a positive integer $n_1(p,\varepsilon)$ and a $c = c(p,\varepsilon) \in (0,1/2]$ such that the following holds. Let $n \ge n_1(p,\varepsilon)$ and k be positive integers. Assume that for independent uniformly distributed elements $z_1,\dots,z_k \in \mathbb{F}_p^n$, with probability at least $1/2$, there exists a set $B \subseteq \mathbb{F}_p^n$ of size εp^n such that every line passing through some point of the set $\{z_1,\dots,z_k\}$ contains at most $p-2$ points of B. Then, there exists a $(p-1,c,c)$-LDC sending $\{0,1\}^l$ to $\{0,1\}^{2p^n}$, where $l = \Omega(c^2k/\log(1/c))$.*

The proof uses the following version of Szemerédi's Theorem, which follows easily from the density Hales–Jewett theorem [19] (see also [35, Theorem 1.5.4]), and its standard "Varnavides-type" corollary (see for example [37, Exercise 10.1.9]).

Theorem 6.2 (Szemerédi's theorem for finite vector spaces)**.** *For every odd prime p and any $\varepsilon \in (0,1]$, there exists a positive integer $n_0(p,\varepsilon)$ such that the following holds. Let $n \ge n_0(p,\varepsilon)$ and let $S \subseteq \mathbb{F}_p^n$ be a set of size $|S| \ge \varepsilon p^n$. Then, S contains a line.*

Corollary 6.3. *For every odd prime p and any $\varepsilon \in (0,1]$, there exists a positive integer $n_1(p,\varepsilon)$ and a $c(p,\varepsilon) \in (0,1]$ such that the following holds. Let $n \geq n_1(p,\varepsilon)$ and let $S \subseteq \mathbb{F}_p^n$ be a set of size $|S| \geq \varepsilon p^n$. Then, S contains at least $c(p,\varepsilon)p^{2n}$ lines, that is,*

$$\Pr_{x\in\mathbb{F}_p^n, y\in\mathbb{F}_p^n\setminus\{0\}}\left[\{(x+\lambda y)_{\lambda=0}^{p-1}\} \subset S\right] \geq c(p,\varepsilon).$$

Proof of Corollary 6.1. Abusing notation, we identify functions $f : \mathbb{F}_p^n \to \{-1,1\}$ with vectors in $\{-1,1\}^{\mathbb{F}_p^n}$. Let $\phi : \{-1,1\} \to \{0,1\}$ be the map $\phi(\alpha) = (\alpha+1)/2$. For a function $f : \mathbb{F}_p^n \to \{-1,1\}$, let $\phi(f) : \mathbb{F}_p^n \to \{0,1\}$ be the function $\phi(f)(x) = \phi(f(x))$ and for $f : \mathbb{F}_p^n \to \{0,1\}$, define $\phi^{-1}(f) : \mathbb{F}_p^n \to \{-1,1\}$ analogously.

For every $x \in \mathbb{F}_p^n$, let $F_x : \{-1,1\}^{\mathbb{F}_p^n} \to \mathbb{R}$ be the degree-$(p-1)$ polynomial

$$F_x(f) = \mathbb{E}_{y\in\mathbb{F}_p^n\setminus\{0\}}\left[\prod_{\lambda\in\mathbb{F}_p^*} \phi(f)(x+\lambda y)\right]. \tag{6.1}$$

Then, for a set $B \subseteq \mathbb{F}_p^n$, the value $F_x(\phi^{-1}(1_B))$ equals the fraction of all lines $\ell_{x,y}$ through x of which B contains the $p-1$ points $\{x+\lambda y : \lambda \in \mathbb{F}_p^*\}$. If B has size at least εp^n, it follows from Corollary 6.3 that $\mathbb{E}_{x\in\mathbb{F}_p^n}[F_x(\phi^{-1}(1_B))] \geq c(p,\varepsilon)$. Moreover, since the monomials in the expectation of (6.1) can be expanded as

$$\prod_{\lambda\in\mathbb{F}_p^*} \phi(f)(x+\lambda y) = \frac{1}{2^{p-1}} \sum_{S\subseteq\mathbb{F}_p^*} \prod_{\lambda\in S} f(x+\lambda y),$$

it follows that each F_x is $\frac{p-1}{2(1-p^{-n})}$-smooth.

Let μ be the uniform distribution over F_x. We claim that $\kappa_\mu(c(p,\varepsilon)) \geq k$, which implies the result by Theorem 1.4 since μ is supported on degree $(p-1)$-polynomials. For every set $A \subseteq \mathbb{F}_p^n$, let $B_A \subseteq \mathbb{F}_p^n$ be a set of maximum size such that every line through A contains at most $p-2$ points of B_A, and let $f_A = \phi^{-1}(1_{B_A})$. Let z be a uniformly distributed random variable over $\mathbb{F}_p^n$, let $z_1,\dots,z_k$ be independent copies of z and let $A = \{z_1,\dots,z_k\}$. Then, $F_{z_1},\dots,F_{z_k}$ are independent μ-distributed random functions. Moreover, in the event that $|B_A| \geq \varepsilon p^n$, we have

$$(\mathbb{E}_z[F_z]-F_{z_i})(f_A) = \mathbb{E}_z\big[F_z(\phi^{-1}(1_{B_A}))\big] - F_{z_i}(\phi^{-1}(1_{B_A})) = \mathbb{E}_z\big[F_z(\phi^{-1}(1_{B_A}))\big] \geq c(p,\varepsilon)$$

for every $i \in [k]$. Since this event happens with probability at least $1/2$, we have

$$\mathbb{E}\Big[\Big\|\frac{1}{k}\sum_{i=1}^{k}\big(F_{z_i}-\mathbb{E}[F_z]\big)\Big\|_{L_\infty}\Big] \geq \mathbb{E}\Big[\Big|\frac{1}{k}\Big(\sum_{i=1}^{k}\big(F_{z_i}-\mathbb{E}[F_z]\big)\Big)(f_A)\Big|\Big] \geq \frac{c(p,\varepsilon)}{2},$$

which gives the claim. □

The proof of the formal version of Corollary 1.6 (given below) is similar to that of Corollary 6.1, so we omit it. In the following, $\mathbb{PF}_p^{n-1}$ is the projective space of dimension $n-1$, which is the space of directions in $\mathbb{F}_p^n$. The formal version of Corollary 1.6 is then as follows.

Corollary 6.4. *For every odd prime p and $\varepsilon \in (0,1]$, there exist a positive integer $n_1(p,\varepsilon)$ and a $c = c(p,\varepsilon) \in (0,1/2]$ such that the following holds. Let $n \geq n_1(p,\varepsilon)$ and k be positive integers. Suppose that for independent uniformly distributed elements $z_1,\dots,z_k \in \mathbb{PF}_p^{n-1}$, with probability at least $1/2$, there exists a set $B \subset \mathbb{F}_p^n$ of size $|B| \geq \varepsilon p^n$ which does not contain any lines with direction in $\{z_1,\dots,z_k\}$. Then, there exists a (p,c,c)-LDC sending $\{0,1\}^l$ to $\{0,1\}^{2p^n}$, where $l = \Omega(c^2k/\log(1/c))$.*

Feasible parameters for Corollary 6.1. Proving lower bounds on k for which the assumption of Corollary 6.1 holds true thus allows one to infer the existence of $(p-1)$-query LDCs with rate $\Omega(k/N)$ for $N = p^n$, provided p and ε are constants independent of n. We establish the following bounds, which imply the (well-known) existence of $(p-1)$-query LDCs with message length $k = \Omega((\log N)^{p-2})$.

Theorem 6.5. *For every odd prime p there exists an $\varepsilon(p) \in (0,1]$ such that the following holds. For every set $A \subseteq \mathbb{F}_p^n$ of size $|A| \leq \binom{n+p-3}{p-2} - 1$, there exists a set $B \subseteq \mathbb{F}_p^n$ of size $\varepsilon(p)p^n$ such that every line through A contains at most $p-2$ points of B.*

The proof uses some basic properties of polynomials over finite fields. For an n-variate polynomial $f \in \mathbb{F}_p[x_1,\dots,x_n]$ denote $Z(f) = \{x \in \mathbb{F}_p^n : f(x) = 0\}$. The starting point of the proof is the following standard result (see for example [36]), showing that small sets can be "captured" by zero-sets of nonzero, homogeneous polynomials of low degree.

Lemma 6.6 (Homogeneous Interpolation)**.** *For every $A \subseteq \mathbb{F}_p^n$ of size $|A| \leq \binom{n+d-1}{d} - 1$, there exists a nonzero homogeneous polynomial $f \in \mathbb{F}_p[x_1,\dots,x_n]$ of degree d such that $A \subseteq Z(f)$.*

The next two lemmas show that if f is nonzero, homogeneous and degree d, and if $a \in \mathbb{F}_p^*$ is such that $f^{-1}(a)$ is nonempty, then lines through $Z(f)$ meet $f^{-1}(a)$ in at most d points.

Lemma 6.7. *Let $f \in \mathbb{F}_p[x_1,\dots,x_n]$ be a nonzero homogeneous polynomial of degree d. Let $a \in \mathbb{F}_p^*$ be such that the set $f^{-1}(a)$ is nonempty. Then, every line that meets $f^{-1}(a)$ in $d+1$ points must have direction in $Z(f)$.*

Proof. The univariate polynomial $g(\lambda) = f(x+\lambda y)$ formed by the restriction of f to a line $\ell_{x,y}$ has degree at most d. By the Factor Theorem, such a polynomial must be the constant polynomial $g(\lambda) = a$ to assume the value a for $d+1$ values of λ. Since f is homogeneous, the coefficient of λ^d, which must be zero, equals $f(y)$, giving the result. □

The following lemma, which is inspired by the solution of the finite-field Kakeya problem in [14], is essentially contained in [10].

Lemma 6.8 (Briët–Rao)**.** *Let $f \in \mathbb{F}_p[x_1,\dots,x_n]$ be a nonzero homogeneous polynomial of degree d. Let $a \in \mathbb{F}_p^*$ be such that $f^{-1}(a)$ is nonempty. Then, there exists no line that intersects $Z(f)$, meets $f^{-1}(a)$ in at least d points and has direction in $Z(f)$.*

Proof. For a contradiction, suppose there exists a line $\ell_{x,y}$ through $Z(f)$ that meets $f^{-1}(a)$ in d points and has direction $y \in Z(f)$. Observe that for every $\lambda \in \mathbb{F}_p$, the shifted line $\ell_{x+\lambda y,y}$ also meets $f^{-1}(a)$ in d points. Hence, without loss of generality we may assume that the line starts in $Z(f)$, that is $x \in Z(f)$.

Let $g(\lambda) = a_0 + a_1\lambda + \cdots + a_d\lambda^d = f(x+\lambda y) \in \mathbb{F}_p[\lambda]$ be the restriction of f to $\ell_{x,y}$. It follows that $a_0 = g(0) = f(x) = 0$ and, since f is homogeneous, that $a_d = f(y) = 0$. Moreover, there exist distinct elements $\lambda_1, \ldots, \lambda_d \in \mathbb{F}_p^*$ such that $g(\lambda_i) = f(x+\lambda_i y) = a$ for every $i \in [d]$. Then $g(\lambda) - a$ is a degree $d-1$ polynomial with d distinct roots. But it cannot be the zero polynomial since it takes value $-a$ when $\lambda = 0$. □

The final ingredient for the proof of Theorem 6.5 is a variant of the DeMillo–Lipton–Schwartz–Zippel Lemma; see for instance [12, Appendix C] for a proof.

Lemma 6.9. *Let $f \in \mathbb{F}_p[x_1, \ldots, x_n]$ be a nonzero polynomial of degree d. Then,*

$$|Z(f)| \leq \Big(1 - \frac{1}{p^{d/(p-1)}}\Big) p^n.$$

Proof of Theorem 6.5. Let $A \subseteq \mathbb{F}_p^n$ be a set of size $|A| \leq \binom{n+p-3}{p-2} - 1$. Let $f \in \mathbb{F}_p[x_1, \ldots, x_n]$ be a nonzero degree-$(p-2)$ homogeneous polynomial such that $A \subseteq Z(f)$, as promised to exist by Lemma 6.6. By Lemma 6.9, there exists an $a \in \mathbb{F}_p^*$ such that the set $B = f^{-1}(a)$ has size at least $|B| \geq p^n/p^{(2p-3)/(p-1)}$. By Lemma 6.7, every line that meets B in $p-1$ points must have direction in $Z(f)$, but by Lemma 6.8 no such line can pass through $Z(f)$. Hence, every line through A meets B in at most $p-2$ points. □

6.2 Hypergraph pseudorandomness

A second candidate for constructing outlaws comes from special types of hypergraphs. A hypergraph $H = (V,E)$ is a pair consisting of a finite vertex set V and an edge set E of subsets of V that allows for parallel (repeated) edges. A hypergraph is t-uniform if all its edges have size t. For subsets $W_1, \ldots, W_t \subseteq V$, define the induced edge count by

$$e_H(W_1, \ldots, W_t) = \sum_{v_1 \in W_1} \cdots \sum_{v_t \in W_t} 1_E(\{v_1, \ldots, v_t\}).$$

A perfect matching in a t-uniform hypergraph is a family of vertex-disjoint edges that intersects every vertex. We shall use the following notion of pseudorandomness.

Definition 6.10 (Relative pseudorandomness)**.** Let $H = (V,E)$, $J = (V,E')$ be t-uniform hypergraphs with identical vertex sets. Then J is ε-pseudorandom relative to H if for all $W_1, \ldots, W_t \subseteq V$, we have

$$\left| \frac{e_J(W_1, \ldots, W_t)}{|E'|} - \frac{e_H(W_1, \ldots, W_t)}{|E|} \right| < \varepsilon. \tag{6.2}$$

The left-hand side of (6.2) compares the fraction of edges that the sets $W_1, \ldots, W_t$ induce in J with the fraction of edges they induce in H. Standard concentration arguments show that if $|E| \geq |V|$, then a random hypergraph J whose edge set E' is formed by independently putting each edge of E in E' with probability $p = p(\varepsilon, t)$, is ε-pseudorandom relative to H with high probability. A deterministic hypergraph J is thus pseudorandom relative to H if it mimics this property of truly random sub-hypergraphs. For graphs, relative ε-pseudorandomness turns into a common notion sometimes referred to as ε-uniformity when H is the complete graph with all loops, in which case (6.2) says that the number of edges induced by a pair of vertex-subsets W_1, W_2 is roughly equal to the product of their densities $(|W_1|/|V|)(|W_2|/|V|)$. Uniformity in graphs is closely connected to the perhaps better-known notion of spectral expansion [24]. These two notions were recently shown to be equivalent (up-to universal constants) for all vertex-transitive graphs [13].

We shall be interested in hypergraphs whose edge set can be partitioned into a family of "blocks," such that randomly removing relatively few of the blocks likely leaves a hypergraph that is *not* pseudorandom relative to the original. (Think of a Jenga tower[10] that's already in a delicate balance, so that there are only few ways, or perhaps even no way, to remove many blocks without having it collapse.) Our blocks will be formed by perfect matchings. For technical reasons, the formal definition takes the view of building a new hypergraph out of randomly selected matchings, as opposed to obtaining one by randomly removing matchings.

Definition 6.11 (Jenga hypergraph)**.** A t-uniform hypergraph H is (k,ε)-jenga if its edge set can be partitioned into a family $\mathcal{M}$ of perfect matchings such that, with probability at least $1/2$, the disjoint union of k independent uniformly distributed matchings from $\mathcal{M}$ forms a hypergraph which is *not* ε-pseudorandom relative to H.

We have the following simple corollary to Theorem 1.4.

Corollary 6.12. *Let n,k,t be positive integers and $\varepsilon \in (0,1]$. Assume that there exists a t-uniform n-vertex hypergraph that is (k,ε)-jenga. Then, there exists a $(t,1,\Omega(\varepsilon/t^2))$-LDC sending $\{0,1\}^l$ to $\{0,1\}^{2tn}$, where $l = \Omega(\varepsilon^2 k/t^4 \log(t^2/\varepsilon))$.*

Proof. Let $H = (V,E)$ be a hypergraph as assumed in the corollary. Let $\mathcal{M}$ be a partition of E into perfect matchings such that if $M_1,\dots,M_k$ are independent and uniformly distributed over $\mathcal{M}$, then with probability at least $1/2$, the hypergraph $J = (V, M_1 \uplus \cdots \uplus M_k)$ is not ε-pseudorandom relative to H.[11]

Let $V_1,\dots,V_t$ be copies of V. For each $M \in \mathcal{M}$, define $f_M : \mathbb{R}^{V_1\cup\cdots\cup V_t} \to \mathbb{R}$ by

$$f_M(x[1],\dots,x[t]) = \frac{1}{|M|}\sum_{v_1\in V_1}\cdots\sum_{v_t\in V_t} 1_M(\{v_1,\dots,v_t\})\,x[1]_{v_1}\cdots x[t]_{v_t}, \quad x[i]\in\mathbb{R}^{V_i}.$$

Denote by $f_M^{\pm}$ the restriction of f_M to the Boolean hypercube $\{-1,1\}^{tn}$, which is a degree-t polynomial in tn variables. Then, for each variable $x[i]_{v_i}$ for $v_i \in V_i$, the spectral norm $\left\| D_{x[i]_{v_i}} f_M^{\pm} \right\|_{\mathrm{sp}}$ equals $|M|^{-1}$ times the number of monomials containing $x[i]_{v_i}$. Since every one of the tn variables appears in exactly one monomial and $|M| = n/t$, it follows that $f_M^{\pm}$ is σ-smooth for $\sigma = tn/|M| = t^2$. Moreover, for $J = (V,M)$ and $W_1,\dots,W_t \subseteq V$, we have

$$f_M(1_{W_1},\dots,1_{W_t}) = \frac{e_J(W_1,\dots,W_t)}{|M|}.$$

Let $M_1,\dots,M_k$ be independent uniformly distributed matchings from $\mathcal{M}$ and consider the random hypergraph $J = (V, M_1 \uplus \cdots \uplus M_k)$. Let $\bar{f} = \mathbb{E}[f_{M_1}]$ be the expectation of the random function f_{M_1} and note

[10] *Jenga*® is a game of dexterity in which players begin with a tower of wooden blocks and take turns trying to remove a block without making the tower collapse.

[11] $M_1 \uplus \cdots \uplus M_k$ is the disjoint union of the matchings (i.e., we repeat an edge r times if it occurs in r matchings).

that $\mathbb{E}[f_{M_i}] = \bar{f}$ for each $i \in [k]$. Let $\bar{f}^{\pm}$ denote the restriction of $\bar{f}$ to $\{-1,1\}^{tn}$ and note that $\bar{f}^{\pm} = \mathbb{E}[f^{\pm}_{M_i}]$. Then, since the functions $f_{M_i} - \bar{f}$ are multilinear,

$$\begin{aligned}
\mathbb{E}\Big[\Big\|\frac{1}{k}\sum_{i=1}^{k}(f^{\pm}_{M_i} - \bar{f}^{\pm})\Big\|_{L_\infty}\Big] &= \mathbb{E}\Big[\max_{x[1],\dots,x[t]\in\{-1,1\}^n}\Big|\frac{1}{k}\sum_{i=1}^{k}(f_{M_i} - \bar{f})(x[1],\dots,x[t])\Big|\Big] \\
&\geq \mathbb{E}\Big[\max_{W_1,\dots,W_t\subseteq V}\Big|\frac{1}{k}\sum_{i=1}^{k}(f_{M_i} - \bar{f})(1_{W_1},\dots,1_{W_t})\Big|\Big] \\
&= \mathbb{E}\Big[\max_{W_1,\dots,W_t\subseteq V}\Big|\frac{e_J(W_1,\dots,W_t)}{k|M|} - \frac{e_H(W_1,\dots,W_t)}{|E|}\Big|\Big] \\
&\geq \frac{\varepsilon}{2}.
\end{aligned}$$

The result now follows from Theorem 1.4. □

In the context of outlaws and LDCs, the relevant question concerning Jenga hypergraphs is the following. Let $\kappa^J(n,t,\varepsilon)$ denote the maximum integer k such that there exists an n-vertex t-uniform hypergraph that is (k,ε)-jenga.

Question 6.13. For integer $t \geq 2$ and parameter $\varepsilon \in (0,1]$, what is the growth rate of $\kappa^J(n,t,\varepsilon)$ as a function of $n \in \mathbb{N}$?

For $t = 2$ (graphs), the answer to Question 6.13 follows from a famous result of Alon and Roichman [1] on expansion of random Cayley graphs. In particular, the proof of this result, due to Landau and Russell [30], based on the matrix Chernoff bound of Ahlswede and Winter, implies that for constant $\varepsilon \in (0,1]$ we have $\kappa(n,2,\varepsilon) = \Theta(\log n)$. Indeed, the matrix Chernoff bound shows that any partitioning of the complete graph on n vertices into perfect matchings is $(O(\log n), O(1))$-Jenga. The lower bound follows for instance by partitioning the edge set of the complete graph with vertex set $V = \mathbb{F}_2^m$ into the collection of matchings of the form $M_y = \big\{\{x, x+y\} : x \in \mathbb{F}_2^m\big\}$ for each $y \in \mathbb{F}_2^m \setminus \{0\}$. Any $m-1$ of such matchings give a graph with two disconnected components of equal size, making it $(m-1, 1/4)$-jenga. Via Corollary 6.12, this arguably gives the most round-about way to prove the existence of 2-query LDCs matching the parameters of the Hadamard code! Generalizing the above example, [10] considered the p-uniform hypergraph on $\mathbb{F}_p^m$ whose edges are the (unordered) lines. It was shown that this hypergraph is (m^{p-1}, ε)-jenga for some $\varepsilon = \varepsilon(p)$ depending on p only, by partitioning the edge set according to the directions of the lines, that is, partitioning it with the matchings $M_y = \big\{\{x + \lambda y : \lambda \in \mathbb{F}_p\} : x \in \mathbb{F}_p^m\big\}$, $y \in \mathbb{F}_p^m \setminus \{0\}$. To the best of our knowledge, the best upper bounds on $\kappa^J(n,t,\varepsilon)$ for constant $t \geq 3$ and $\varepsilon \in (0,1]$ follow from upper bounds on LDCs, via Corollary 6.12.

We end with the following natural question concerning Jenga hypergraphs.

Question 6.14. Is $\kappa^J(n,t,\varepsilon)$ largest for the complete hypergraph?

Acknowledgements. J. B. and S. G. thank the Simons Institute for hosting them during the 2017 Pseudorandomness program, where part of this work was done. We thank Jacob Fox for pointing out to us that randomness is essential in Corollary 1.6, Quentin Guilmant for pointing out an error in an earlier version of the proof of Theorem 3.1 and the anonymous referees for helpful remarks that greatly improved the exposition of this paper.

References

[1] NOGA ALON AND YUVAL ROICHMAN: Random Cayley graphs and expanders. *Random Struct. Algorithms*, 5(2):271–284, 1994. [doi:10.1002/rsa.3240050203] 19

[2] SANJEEV ARORA, CARSTEN LUND, RAJEEV MOTWANI, MADHU SUDAN, AND MARIO SZEGEDY: Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998. Preliminary version in FOCS'92. [doi:10.1145/278298.278306] 2

[3] SANJEEV ARORA AND SHMUEL SAFRA: Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998. Preliminary version in FOCS'92. [doi:10.1145/273865.273901] 2

[4] LÁSZLÓ BABAI, LANCE FORTNOW, LEONID A. LEVIN, AND MARIO SZEGEDY: Checking computations in polylogarithmic time. In *Proc. 23rd STOC*, pp. 21–31. ACM Press, 1991. [doi:10.1145/103418.103428] 2

[5] MANUEL BLUM AND SAMPATH KANNAN: Designing programs that check their work. *J. ACM*, 42(1):269–291, 1995. Preliminary version in STOC'89. [doi:10.1145/200836.200880] 2

[6] JOP BRIËT: On embeddings of ℓ_1^k from locally decodable codes, 2016. [arXiv:1611.06385] 4

[7] JOP BRIËT, ZEEV DVIR, AND SIVAKANTH GOPI: Outlaw distributions and locally decodable codes. In *Proc. 8th Symp. Innovations in Theoret. Comp. Sci. (ITCS'17)*, pp. 20:1–20:19. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2017. [doi:10.4230/LIPIcs.ITCS.2017.20, arXiv:1609.06355] 1

[8] JOP BRIËT AND SIVAKANTH GOPI: Gaussian width bounds with applications to arithmetic progressions in random settings. *Internat. Math. Res. Notices*, 2018. [doi:10.1093/imrn/rny238, arXiv:1711.05624] 5

[9] JOP BRIËT, ASSAF NAOR, AND ODED REGEV: Locally decodable codes and the failure of cotype for projective tensor products. *Electron. Res. Announc. Math. Sci.*, 19:120–130, 2012. [doi:10.3934/era.2012.19.120, arXiv:1208.0539] 2

[10] JOP BRIËT AND SHRAVAS RAO: Arithmetic expanders and deviation bounds for random tensors, 2016. [arXiv:1610.03428] 16, 19

[11] BENNY CHOR, ODED GOLDREICH, EYAL KUSHILEVITZ, AND MADHU SUDAN: Private information retrieval. *J. ACM*, 45(6):965–982, 1998. Preliminary version in FOCS'95. [doi:10.1145/293347.293350] 2, 3

[12] GIL COHEN AND AVISHAY TAL: Two structural results for low degree polynomials and applications. In *Proc. Internat. Workshop on Approximation, Randomization, and Combinat. Optim. (RANDOM'15)*, pp. 680–709. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2015. [doi:10.4230/LIPIcs.APPROX-RANDOM.2015.680, arXiv:1404.0654] 17

[13] DAVID CONLON AND YUFEI ZHAO: Quasirandom Cayley graphs. *Discrete Analysis*, pp. 6:1–6:14, 2017. [doi:10.19086/da.1294, arXiv:1603.03025] 18

[14] ZEEV DVIR: On the size of Kakeya sets in finite fields. *J. Amer. Math. Soc.*, 22(4):1093–1097, 2009. [doi:10.1090/S0894-0347-08-00607-3, arXiv:0803.2336] 16

[15] ZEEV DVIR AND SIVAKANTH GOPI: 2-Server PIR with subpolynomial communication. *J. ACM*, 63(4):39:1–39:15, 2016. Preliminary version in STOC'15. [doi:10.1145/2968443, arXiv:1407.6692] 3

[16] KLIM EFREMENKO: 3-Query locally decodable codes of subexponential length. *SIAM J. Comput.*, 41(6):1694–1703, 2012. Preliminary version in STOC'09. [doi:10.1137/090772721] 3

[17] NIKOS FRANTZIKINAKIS, EMMANUEL LESIGNE, AND MÁTÉ WIERDL: Random sequences and pointwise convergence of multiple ergodic averages. *Indiana Univ. Math. J.*, 61(2):585–617, 2012. [doi:10.1512/iumj.2012.61.4571] 5

[18] NIKOS FRANTZIKINAKIS, EMMANUEL LESIGNE, AND MÁTÉ WIERDL: Random differences in Szemerédi's theorem and related results. *J. d'Analyse Mathématique*, 130(1):91–133, 2016. [doi:10.1007/s11854-016-0030-z, arXiv:1307.1922] 5

[19] HILLEL FURSTENBERG AND YITZHAK KATZNELSON: A density version of the hales-jewett theorem. *J. d'Analyse Mathématique*, 57(1):64–119, 1991. [doi:10.1007/BF03041066] 14

[20] ODED GOLDREICH, HOWARD KARLOFF, LEONARD J. SCHULMAN, AND LUCA TREVISAN: Lower bounds for linear locally decodable codes and private information retrieval. *Comput. Complexity*, 15(3):263–296, 2006. Preliminary version in CCC'02. [doi:10.1007/s00037-006-0216-3] 3

[21] PARIKSHIT GOPALAN, CHENG HUANG, HUSEYIN SIMITCI, AND SERGEY YEKHANIN: On the locality of codeword symbols. *IEEE Trans. Inform. Theory*, 58(11):6925–6934, 2012. [doi:10.1109/TIT.2012.2208937, arXiv:1106.3625] 2

[22] SIVAKANTH GOPI, SWASTIK KOPPARTY, RAFAEL OLIVEIRA, NOGA RON-ZEWI, AND SHUBHANGI SARAF: Locally testable and locally correctable codes approaching the Gilbert-Varshamov bound. *IEEE Trans. Inform. Theory*, 64(8):5813–5831, 2018. Preliminary version in SODA'17. [doi:10.1109/TIT.2018.2809788] 3

[23] BEN GREEN AND TOM SANDERS: Boolean functions with small spectral norm. *Geom. Funct. Anal.*, 18(1):144–162, 2008. [doi:10.1007/s00039-008-0654-y, arXiv:math/0605524] 3

[24] SHLOMO HOORY, NATHAN LINIAL, AND AVI WIGDERSON: Expander graphs and their applications. *Bull. Amer. Math. Soc.*, 43(4):439–561, 2006. [doi:10.1090/S0273-0979-06-01126-8] 18

[25] JONATHAN KATZ AND LUCA TREVISAN: On the efficiency of local decoding procedures for error-correcting codes. In *Proc. 32nd STOC*, pp. 80–86. ACM Press, 2000. [doi:10.1145/335305.335315] 2, 3, 7

[26] TALI KAUFMAN AND MADHU SUDAN: Sparse random linear codes are locally decodable and testable. In *Proc. 48th FOCS*, pp. 590–600. IEEE Comp. Soc. Press, 2007. [doi:10.1109/FOCS.2007.8] 2

[27] IORDANIS KERENIDIS AND RONALD DE WOLF: Exponential lower bound for 2-query locally decodable codes via a quantum argument. *J. Comput. System Sci.*, 69(3):395–420, 2004. Preliminary version in STOC'03. [doi:10.1016/j.jcss.2004.04.007, arXiv:quant-ph/0208062] 3, 5

[28] SWASTIK KOPPARTY, OR MEIR, NOGA RON-ZEWI, AND SHUBHANGI SARAF: High-rate locally correctable and locally testable codes with sub-polynomial query complexity. *J. ACM*, 64(2):11:1–11:42, 2017. Preliminary version in STOC'16. [doi:10.1145/3051093, arXiv:1504.05653] 3

[29] SWASTIK KOPPARTY, SHUBHANGI SARAF, AND SERGEY YEKHANIN: High-rate codes with sublinear-time decoding. *J. ACM*, 61(5):28:1–28:20, 2014. Preliminary version in STOC'11. [doi:10.1145/2629416] 3

[30] ZEPH LANDAU AND ALEXANDER RUSSELL: Random Cayley graphs are expanders: A simple proof of the Alon–Roichman theorem. *Electron. J. Combin.*, 11(1):R62, 2004. 19

[31] SHAHAR MENDELSON AND ROMAN VERSHYNIN: Entropy and the combinatorial dimension. *Inventiones Math.*, 152(1):37–55, 2003. [doi:10.1007/s00222-002-0266-3, arXiv:math/0203275] 7, 11, 12

[32] CLAUDE E. SHANNON: A mathematical theory of communication. *Bell System Tech. J.*, 27(3):379–423, 623–656, 1948. [doi:10.1002/j.1538-7305.1948.tb01338.x] 2

[33] AMIR SHPILKA, AVISHAY TAL, AND BEN LEE VOLK: On the structure of Boolean functions with small spectral norm. *Comput. Complexity*, 26(1):229–273, 2017. Preliminary version in ITCS'14. [doi:10.1007/s00037-015-0110-y, arXiv:1304.0371] 3

[34] MICHEL TALAGRAND: *Upper and Lower bounds For Stochastic Processes*. Springer, 2014. [doi:10.1007/978-3-642-54075-2] 12

[35] TERENCE TAO: *Higher Order Fourier Analysis*. Amer. Math. Soc., 2012. [doi:10.1090/gsm/142] 14

[36] TERENCE TAO: Algebraic combinatorial geometry: The polynomial method in arithmetic combinatorics, incidence combinatorics, and number theory. *EMS Surveys in Math. Sciences*, 1(1):1–46, 2014. [doi:10.4171/EMSS/1, arXiv:1310.6482] 16

[37] TERENCE TAO AND VAN VU: *Additive Combinatorics*. Cambridge Univ. Press, 2006. [doi:10.1017/CBO9780511755149] 14

[38] DAVID WOODRUFF: New lower bounds for general locally decodable codes. *Electronic Colloq. Comput. Complexity (ECCC'07)*, 14(6), 2007. 3, 5

[39] SERGEY YEKHANIN: Towards 3-query locally decodable codes of subexponential length. *J. ACM*, 55(1):1:1–1:16, 2008. Preliminary version in STOC'07. [doi:10.1145/1326554.1326555] 3

[40] SERGEY YEKHANIN: Locally decodable codes. *Found. Trends Theor. Comput. Sci.*, 6(3):139–255, 2012. Preliminary version in CSR'11. [doi:10.1561/0400000030] 2

AUTHORS

Jop Briët
Senior researcher
Centrum Wiskunde & Informatica (CWI)
Amsterdam, The Netherlands
j.briet@cwi.nl
https://homepages.cwi.nl/~jop/

Zeev Dvir
Associate professor
Department of Computer Science and Department of Mathematics
Princeton University
zdvir@princeton.edu
www.cs.princeton.edu/~zdvir/

Sivakanth Gopi
Postdoctoral Researcher
Microsoft Redmond
sigopi@microsoft.com
https://www.microsoft.com/en-us/research/people/sigopi/

ABOUT THE AUTHORS

JOP BRIËT was born in Leiden, The Netherlands. He spent a year in the United States when he was one year old, but remembers nothing of it. He received his Ph. D. from the University of Amsterdam in 2011 and was advised by Harry Buhrman. He met his coauthors on this paper during a postdoc at the Courant Institute of Mathematical Sciences in New York. He is interested in the interplay between mathematics and theoretical computer science, especially for additive combinatorics, geometric function analysis, coding theory and quantum information.

ZEEV DVIR was born in Jerusalem, Israel. He received his Ph. D. from the Weizmann Institute in Israel in 2008. His advisors were Ran Raz and Amir Shpilka. He has a broad interest in theoretical computer science and mathematics and especially in computational complexity, pseudorandomness, coding theory and discrete mathematics.

SIVAKANTH GOPI (called "Gopi" by friends and colleagues) is a postdoctoral researcher in the algorithms research group at Microsoft Research, Redmond. He received his Ph. D. at Princeton University in 2018 under the supervision of Zeev Dvir; the title of his dissertation was "Locality in coding theory." His main research interests are in coding theory, complexity theory, and additive combinatorics. He is also part of the DNA Storage project and Project Laplace (differential privacy) at Microsoft. He enjoys spending leisure time learning new things, staying fit, exploring nature and talking to his family.