

# Outlaw distributions and locally decodable codes

Jop Briët\*      Zeev Dvir†      Sivakanth Gopi‡

September 20, 2016

## 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  $L_\infty$  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 also give several candidates for outlaw distributions over smooth functions coming from finite field incidence geometry and from 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.

## 1 Introduction

Error correcting codes (ECCs) solve the basic problem of communication over noisy channels. They encode a message into a codeword from which, even if the channel partially corrupts it, the message can later be retrieved. With one of the earliest applications of the *probabilistic method*, formally introduced by Erdős in 1947, pioneering work of Shannon [Sha48] showed the existence of optimal (capacity-achieving) ECCs. 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

---

\*CWI, Science Park 123, 1098 XG Amsterdam, the Netherlands. Supported by a VENI grant from the Netherlands Organisation for Scientific Research (NWO). Email: [j.briet@cwi.nl](mailto:j.briet@cwi.nl).

†Department of Computer Science and Department of Mathematics, Princeton University, Princeton, NJ 08540, USA. Research supported by NSF grant CCF-1523816 and by the Sloan foundation. Email: [zeev.dvir@gmail.com](mailto:zeev.dvir@gmail.com).

‡Department of Computer Science, Princeton University, Princeton, NJ 08540, USA. Research supported by NSF grant CCF-1523816 and by the Sloan foundation. Email: [sgopi@cs.princeton.edu](mailto:sgopi@cs.princeton.edu).

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 ultra fast 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 [KT00], although they (and the closely related locally correctable codes) implicitly appeared in several previous works in other settings, such as program checking [BK95], probabilistically checkable proofs [AS98, ALM<sup>+</sup>98] and private information retrieval schemes (PIRs) [CKGS98]. More recently, LDCs have even found applications in Banach-space geometry [BNR12] and LDC-inspired objects called local reconstruction codes found applications in fault tolerant distributed storage systems [GHSY12]. See [Yek12] for a survey of LDCs and some of the applications.

Despite their many applications, our knowledge of LDCs is very limited; the best-known constructions 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 [KS07]. All the known constructions of LDCs were obtained by explicitly designing such codes using some algebraic objects like low-degree polynomials or matching vectors [Yek12].

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  $\eta, \delta \in (0, 1/2]$ , a map  $C : \{0, 1\}^k \rightarrow \{0, 1\}^n$  is a  $(q, \delta, \eta)$ -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  $\delta n$  coordinates,

$$\Pr[\mathcal{A}_i(y) = x_i] \geq \frac{1}{2} + \eta. \tag{1}$$

- The decoder  $\mathcal{A}_i$  (non-adaptively) queries at most  $q$  coordinates of  $y$ .<sup>1</sup>

**Known results.** The main parameters of LDCs are the number of queries  $q$  and the length of the encoding  $n$  as a function of  $k$  and  $q$ , typically the parameters  $\delta, \eta$  are some 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

---

<sup>1</sup> 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  $S$ .

are known: it was shown in [KW04, GKST02] 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 a breakthrough work by [Yek07, Efr09], 3-query LDCs were constructed 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-known lower bounds leave huge gaps, giving only polynomial bounds. Any 3-query LDC must have length  $n \geq \tilde{\Omega}(k^2)$  [Woo12], and more generally any  $q$ -query LDC must have length  $n \geq \tilde{\Omega}(k^{1+1/(\lceil q/2 \rceil - 1)})$  [KT00, Woo07]. LDCs where the codewords are over a large alphabet are also studied because of their relation to private information retrieval schemes [CKGS98, KT00]. In [DG15], 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 when the number of queries can also grow 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 [KSY10, KMRS16, GKdO<sup>+</sup>16]. 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  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ , its *spectral norm* (also known as the algebra norm or Wiener norm) is defined as

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

where  $\hat{f}(S)$  are the Fourier coefficients of  $f$  (see Section 2 for some preliminaries in Fourier analysis). 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  $\|f\|_{L_\infty} \leq \|f\|_{\text{sp}}$ . The  *$i$ th discrete derivative* of  $f$  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. Smooth functions are functions whose discrete derivatives have small spectral norms.

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

Intuition for the above definition may be gained from the fact that smooth functions have no influential variables. The influence of the  $i$ th variable,  $(\mathbb{E}_{x \in \{-1, 1\}^n} [(D_i f)(x)^2])^{1/2}$ , measures the extent to which changing the  $i$ th coordinate of a randomly chosen point changes the value of  $f$ . Since  $\|D_i f\|_{L_\infty} \leq \|D_i f\|_{\text{sp}}$ , the directional derivatives of  $\sigma$ -smooth functions are uniformly bounded by  $\sigma/n$ , 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  $\mu$  be a probability distribution over real-valued functions on  $\{-1, 1\}^n$ . For a positive integer  $k$  and  $\varepsilon > 0$ , say that  $\mu$  is a  $(k, \varepsilon)$ -outlaw if for independent random  $\mu$ -distributed functions  $f_1, \dots, f_k$  and  $\bar{f} = \mathbb{E}_\mu[f]$ ,

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

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

To approximate the true mean of an outlaw  $\mu$  to within  $\varepsilon$  on average in the  $L_\infty$ -distance, one thus needs  $\kappa_\mu(\varepsilon) + 1$  samples. Note that if  $\mu$  is a distribution over  $\sigma$ -smooth functions, then the distribution  $\tilde{\mu}$  obtained by scaling functions in the support of  $\mu$  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  $\mu$  be a probability distribution over 1-smooth functions on  $\{-1, 1\}^n$  and  $k = \kappa_\mu(\varepsilon)$ . Then, there exists a  $(q, \delta, \eta)$ -LDC sending  $\{0, 1\}^l$  to  $\{0, 1\}^n$  where  $l = \Omega(\varepsilon^2 k / \log(1/\varepsilon))$ ,  $q = O(1/\varepsilon)$ ,  $\delta = \Omega(\varepsilon)$  and  $\eta = \Omega(\varepsilon)$ . Additionally, if  $\mu$  is supported by degree- $d$  functions, 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, 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  $C : \{0, 1\}^k \rightarrow \{0, 1\}^n$  is a  $(q, \delta, \eta)$ -LDC, then there exists a probability distribution  $\mu$  over 1-smooth degree- $q$  functions on  $\{-1, 1\}^n$  such that*

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

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

Let us remark in passing that 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, it can be shown that for a distribution  $\mu$  over 1-smooth degree- $q$  functions on  $\{0, 1\}^n$ , one can upper bound  $\kappa_\mu(\varepsilon)$  in terms of type constants of the space of  $q$ -linear forms on  $\ell_q^{n+1}$  [Bri15].

**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 every set of directions  $D \subset \mathbb{F}_p^n$  of size  $|D| \leq k$ , 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  $\{0, 1\}^{p^n}$ .*

Another setting in which our approach leads to interesting open problems is in relation to expansion in 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 an 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 an expander (defined appropriately) one needs at least  $k$  random matchings. This would give a code sending  $k$  bit messages with encoding length  $n$  and so, becomes interesting when  $k$  is super poly-logarithmic 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 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 high-dimensional probability. 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 smooth functions distributed according to a  $(k, \varepsilon)$ -outlaw with expectation  $\bar{f}$ . We introduce an independent copy<sup>2</sup>  $f'_i$  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

$$\mathbb{E}[\|(f_1 - f'_1) + \dots + (f_k - f'_k)\|_{L_\infty}] \geq \mathbb{E}[\|(f_1 - \mathbb{E}[f'_1]) + \dots + (f_k - \mathbb{E}[f'_k])\|_{L_\infty}] \geq \varepsilon k.$$

Now, using the fact that the random functions  $f_i - f'_i$  are independent and symmetric, for independent uniformly distributed random signs  $x_1, \dots, x_k \in \{-1, 1\}$ , the above left-hand side equals

$$\mathbb{E}[\|x_1(f_1 - f'_1) + \dots + x_k(f_k - f'_k)\|_{L_\infty}].$$

---

<sup>2</sup>in this context sometimes referred to as a “ghost copy” as it will later disappear again

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

$$\mathbb{E}[\|x_1 f_1^* + \dots + x_k f_k^*\|_{L_\infty}] \geq \varepsilon k/2. \quad (2)$$

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  $L_\infty$ -norm in (2) is achieved to be the its encoding,  $C(x)$ . This gives a map  $C : \{-1, 1\}^k \rightarrow \{0, 1\}^n$  satisfying

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

Equivalently, for uniform  $x$  and  $i$ , we have  $\Pr[f_i^*(C(x)) = x_i] \geq \frac{1}{2} + \frac{\varepsilon}{4}$ . Finally, we use the smoothness property to transform the  $f_i^*$  into decoders with the desired properties. This is done in Section 3. Let us point out that it is in the application of the Averaging Principle where the probabilistic method appears in our construction of LDCs.

**Average-case smooth codes are LDCs.** Katz and Trevisan [KT00] 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 of corrupted coordinates  $\delta$  and query complexity  $q$  aren’t too large. The following definition allows us to make this intuition precise.

**Definition 1.7** (Smooth code). For positive integers  $k, n, q$  and parameters  $\eta \in (0, 1/2]$  and  $c > 0$ , a map  $C : \{0, 1\}^k \rightarrow \{0, 1\}^n$  is a  $(q, c, \eta)$ -smooth code if, for every  $i \in [k]$ , there exists a randomized decoder  $\mathcal{A}_i : \{0, 1\}^n \rightarrow \{0, 1\}$  such that

- For every  $x \in \{0, 1\}^k$ ,

$$\Pr[x_i = \mathcal{A}_i(C(x))] \geq \frac{1}{2} + \eta. \quad (3)$$

- 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 1.8** (Katz–Trevisan). *Let  $C : \{0, 1\}^k \rightarrow \{0, 1\}^n$  be a  $(q, \delta, \eta)$ -LDC, then  $C$  is also a  $(q, q/\delta, \eta)$ -smooth code. Conversely, if  $C : \{0, 1\}^k \rightarrow \{0, 1\}^n$  is a  $(q, c, \eta)$ -smooth code, then  $C$  is also 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 in Theorem 1.8 showing 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 1.9** (Average-case smooth code). A code as in Definition 1.7 is a  $(q, c, \eta)$ -average-case smooth code if instead of the first item, (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

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

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

**Lemma 1.10.** *Let  $C : \{0, 1\}^k \rightarrow \{0, 1\}^n$  be a  $(q, c, \eta)$ -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 1.10 is as follows. We first switch the message and codeword alphabets to  $\{-1, 1\}$  and let  $f_i : \{-1, 1\}^k \rightarrow [-1, 1]$  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  $T = \{(f_1(z), \dots, f_k(z)) : z \in \{-1, 1\}^k\}$  has large *Gaussian width*, in particular it holds that for a standard  $k$ -dimensional Gaussian vector  $g$ , we have  $\mathbb{E}[\sup_{t \in T} \langle g, t \rangle] \gtrsim \varepsilon k$ .<sup>3</sup> Next, we employ a powerful result of [MV03] showing that  $T$  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  $\varepsilon$  and  $c$ . Finally, we obtain an LDC via an application of Theorem 1.8. 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 1.10 to convert them to LDCs. In Section 4, we prove Lemma 1.10 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 Cayley hypergraphs.

## 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  $\langle f, g \rangle =$

---

<sup>3</sup>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.

$\mathbb{E}_{x \in \{-1, 1\}^n} [f(x)g(x)]$ , 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  $f : \{-1, 1\}^n \rightarrow \mathbb{R}$  is the function  $\widehat{f} : 2^{[n]} \rightarrow \mathbb{R}$  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

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

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- $q$  function 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  $\|D_i f\|_{\text{Spec}} = \sum_{S \ni i} |\widehat{f}(S)|$ .

### 3 From outlaws to LDCs

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- $q$  functions give  $q$ -query average-case smooth codes. Combined with Lemma 1.10, this implies the second part of Theorem 1.4.

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

*Proof:* The proof uses a symmetrization argument. Let  $\mathcal{F} = (f_1, \dots, f_k)$ ,  $\mathcal{F}' = (f'_1, \dots, f'_k)$  be two  $k$ -tuples of independent  $\mu$ -distributed random variables and let  $\bar{f} = \mathbb{E}_\mu[f]$ . Then, by



definition of  $\kappa_\mu(\varepsilon)$  and Jensen's inequality,

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

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,

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

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

$$\mathbb{E}_{x \in \{-1, 1\}^k} \left[ \left\| \frac{1}{k} \sum_{i=1}^k x_i f_i^* \right\|_{L_\infty} \right] \geq \frac{\varepsilon}{2}. \quad (4)$$

Define the code  $C : \{-1, 1\}^k \rightarrow \{-1, 1\}^n$  such that for each  $x \in \{-1, 1\}^k$ , we have

$$\frac{1}{k} \sum_{i=1}^k x_i f_i^*(C(x)) = \left\| \frac{1}{k} \sum_{i=1}^k x_i f_i^* \right\|_{L_\infty}. \quad (5)$$

For each  $i \in [k]$ , define the decoder  $\mathcal{A}_i$  as follows. Let  $\nu_i : 2^{[n]} \rightarrow [0, 1]$  be the probability distribution defined by  $\nu_i(S) = |\widehat{f}_i^*(S)| / \|\widehat{f}_i^*\|_{\text{sp}}$ . Given a string  $z \in \{-1, 1\}^n$ , with probability  $1 - \|\widehat{f}_i^*\|_{\text{sp}}$ , the decoder  $\mathcal{A}_i$  returns a uniformly random sign, and with probability  $\|\widehat{f}_i^*\|_{\text{sp}}$ , it samples a set  $S \subseteq [n]$  according to  $\nu_i$  and returns  $\chi_S(z)$ . This is a valid probability distribution since for any 1-smooth function  $f$ , we have

$$\|f\|_{\text{sp}} = \sum_{S \subseteq [n]} |\widehat{f}(S)| \leq \sum_{S \subseteq [n]} |S| |\widehat{f}(S)| = \sum_{i=1}^n \sum_{S \ni i} |\widehat{f}(S)| \leq n \cdot \frac{1}{n} = 1.$$

Then,  $\mathcal{A}_i$  queries at most  $q$  coordinates of  $z$  and since  $f_i^*$  is 1-smooth, the probability that it queries any coordinate  $j \in [n]$  is at most  $\|D_j f_i^*\|_{\text{sp}} \leq 1/n$ . We also have  $\mathbb{E}[\mathcal{A}_i(z)] = f_i^*(z)$ . Therefore, by (4) and (5), we have

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

Hence,  $\mathcal{C}$  is a  $(q, 1, \varepsilon/4)$ -average-case smooth code.  $\square$

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

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

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

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

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

$$\|f - f^{\leq q}\|_{L_\infty} \leq \sum_{|S| > q} |\hat{f}(S)| \leq \frac{1}{q} = \frac{\varepsilon}{4}. \quad (6)$$

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

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

Hence, by the triangle inequality and (6), we have

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

giving the claim.  $\square$

*Proof of Theorem 1.4:* By applying Lemma 3.2 to  $\mu$ , we get a distribution  $\tilde{\mu}$  over 1-smooth degree  $q = O(1/\epsilon)$  functions with  $k' = \kappa_{\tilde{\mu}}(\epsilon/2) \geq \kappa_{\mu}(\epsilon) = k$ . By Theorem 3.1, we get a  $(q, 1, \Omega(\epsilon))$ -average-case smooth code  $C' : \{-1, 1\}^{k'} \rightarrow \{-1, 1\}^n$ . Finally we use Lemma 1.10 to convert  $C'$  to a  $(q, \Omega(\epsilon), \Omega(\epsilon))$ -LDC  $C : \{-1, 1\}^{\ell} \rightarrow \{-1, 1\}^n$  where  $\ell = \Omega(\epsilon^2 k' / \log(1/\epsilon))$ . For the last part of the theorem we can simply apply Theorem 3.1 directly.  $\square$

## 4 From average-case smooth codes to LDCs

In this section, we prove Lemma 1.10. For this, we need the notion of the Vapnik–Chervonenkis dimension (VC-dimension).

**Definition 4.1** (VC-dimension). For  $T \subset [-1, 1]^k$  and  $w > 0$ ,  $\text{vc}(T, w)$  is defined as the size of the largest subset  $\sigma \subset [k]$  such that there exists a shift  $s \in [-1, 1]^k$  satisfying the following: for every  $x \in \{-1, 1\}^{\sigma}$ , there exists  $t \in T$  such that for every  $i \in \sigma$ ,  $(t_i - s_i)x_i \geq w/2$ .

Observe that if  $T$  is convex, then  $\text{vc}(T, w)$  is the maximum dimension of a shifted hypercube with edge lengths at least  $w$  contained in  $T$ .

**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 is easy to see that a large VC-dimension implies a 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** ([MV03]). *Let  $T \subset [-1, 1]^k$ . Then, the Gaussian width of  $T$  is bounded as*

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

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 1.10:* The proof works by showing that the average-case smooth code property implies that the image of the (average) decoding functions should have large Gaussian width. We then use Theorem 4.3 to find a hypercube like structure inside the image, which we use to construct a smooth code. Finally we use Theorem 1.8 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  $f_i : \{-1, 1\}^k \rightarrow [-1, 1]$  be the expected decoding function  $f_i(z) = \mathbb{E}[\mathcal{A}_i(z)]$ . Let  $g$  be a standard  $k$ -dimensional Gaussian vector and  $T = \{(f_1(z), \dots, f_k(z)) : z \in \{-1, 1\}^k\}$ . By the definition of average-case smooth code we have

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

(See for instance [Tal14, Lemma 3.2.10] for the last inequality.) By Theorem 4.3, for some constant  $\alpha > 0$ , we have

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

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

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

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

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  $i \in \sigma$ ,

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

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 1.8,  $C'$  is also a  $(q, \Omega(\eta/c), \Omega(\eta))$ -LDC.  $\square$

## 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 1.8,  $C : \{-1, 1\}^k \rightarrow \{-1, 1\}^n$  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  $\nu_i : 2^{[n]} \rightarrow [0, 1]$  be the probability distribution used by  $\mathcal{B}_i$  to sample a set  $S \subseteq [n]$  of at most  $q$  coordinates and let  $f_{i,S} : \{-1, 1\}^n \rightarrow [-1, 1]$  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  $f_i : \{-1, 1\}^n \rightarrow [-1, 1]$  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  $\delta/(q2^{q/2})$ -smooth. Since the functions  $f_{i,S} : \{-1, 1\}^n \rightarrow \{-1, 1\}$  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

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

which gives the claim. By (3), it holds for every  $x \in \{-1, 1\}^k$  and every  $i \in [k]$  that

$$x_i f_i(C'(x)) \geq 2\eta. \tag{7}$$

Define the distribution  $\mu$  to correspond to the process of sampling  $i \in [k]$  uniformly at random and returning  $f_i$ . Let  $\bar{g} = (f_1 + \dots + f_k)/k$  be the mean of  $\mu$ . We show that  $\kappa_\mu(\eta) \geq \eta k$ . To this end, let  $l = \eta k$ , let  $\sigma : [l] \rightarrow [k]$  be an arbitrary map and define the functions  $g_1, \dots, g_l$  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 (7) that  $f_{\sigma(i)}(C(x)) \in [2\eta, 1]$  for every  $i \in [l]$  and that  $f_i(C(x)) \leq 0$  for every other  $i \in [k]$ . Hence,

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

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

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

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

## 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  $x, y \in \mathbb{F}_p^n$ , the *line* with origin  $x$  in direction  $y$ , denoted  $\ell_{x,y}$ , is the sequence  $(x + \lambda y)_{\lambda \in \mathbb{F}_p}$ . A line is nontrivial if  $y \neq 0$ .

**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  $n \geq n_1(p, \varepsilon)$  and  $k$  be positive integers. Assume that for every set  $A \subseteq \mathbb{F}_p^n$  of size  $|A| \leq k$ , there exists a set  $B \subseteq \mathbb{F}_p^n$  of size  $\varepsilon p^n$  such that every nontrivial line through  $A$  contains at most  $p - 2$  points of  $B$ . Then, there exists a  $(p - 1, c, c)$ -LDC sending  $\{0, 1\}^l$  to  $\{0, 1\}^{p^n}$ , where  $l = \Omega(c^2 k / \log(1/c))$ .*

The proof uses the following version of Szemerédi’s Theorem [Tao12, Theorem 1.5.4] and its standard “Varnavides-type” corollary (see for example [TV06, Exercise 10.1.9]).

**Theorem 6.2** (Szemerédi’s theorem). *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  $n \geq n_0(p, \varepsilon)$  and let  $S \subseteq \mathbb{F}_p^n$  be a set of size  $|S| \geq \varepsilon p^n$ . Then,  $S$  contains a nontrivial 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  $c(p, \varepsilon)p^{2n}$  nontrivial lines, that is,*

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

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

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

$$F_x(f) = \mathbb{E}_{y \in \mathbb{F}_p^n \setminus \{0\}} \left[ \prod_{\lambda \in \mathbb{F}_p^*} \phi(f)(x + \lambda y) \right]. \quad (8)$$

Then, for a set  $B \subseteq \mathbb{F}_p^n$ , the value  $F_x(\phi^{-1}(1_B))$  equals the fraction of all nontrivial 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  $\varepsilon p^n$ , it thus 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 (8) involve disjoint sets of variables and can be expanded as

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

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

Let  $\mu$  be the uniform probability distribution over all  $F_x$ . We claim that  $\kappa_\mu(c(p, \varepsilon)) \geq k$ , which implies the result by Theorem 1.4 since  $\mu$  is supported by degree  $(p - 1)$ -functions. For every set  $A \subseteq \mathbb{F}_p^n$  of size  $|A| \leq k$ , let  $B \subseteq \mathbb{F}_p^n$  be an arbitrary set as in the assumption of the corollary and let  $f_A = \phi^{-1}(1_B)$ . 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  $\mu$ -distributed random functions and since every nontrivial line through  $A$  meets  $B$  in at most  $p - 2$  points, we have  $F_{z_i}(f_A) = 0$  for every  $i \in [k]$ . Hence,

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

which gives the claim.  $\square$

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,  $\mathbb{P}\mathbb{F}_p^{n-1}$  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 every set of directions  $D \subset \mathbb{P}\mathbb{F}_p^{n-1}$  of size  $|D| \leq k$ , 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  $D$ . Then, there exists a  $(p, c, c)$ -LDC sending  $\{0, 1\}^l$  to  $\{0, 1\}^{p^n}$ , where  $l = \Omega(c^2 k / \log(1/c))$ .*

**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  $\varepsilon$  are constant with respect to  $n$ . 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 [Tao14]), 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  $\lambda$ . Since  $f$  is homogeneous, the coefficient of  $\lambda^d$ , which must be zero, equals  $f(y)$ , giving the result.  $\square$

The following lemma is essentially contained in [BR].

**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  $g(\lambda) = a_0 + a_1\lambda + \dots + a_d\lambda^d = f(x + \lambda y) \in \mathbb{F}_p[\lambda]$  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  $\lambda_1, \dots, \lambda_d \in \mathbb{F}_p^*$  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  $\lambda = 0$ .  $\square$

The final ingredient for the proof of Theorem 6.5 is the DeMillo–Lipton–Schwartz–Zippel Lemma as it appears in [CT14].

**Lemma 6.9** (DeMillo–Lipton–Schwartz–Zippel). *Let  $f \in \mathbb{F}_p[x_1, \dots, x_n]$  be a nonzero polynomial of degree  $d$  and denote  $r = |\mathbb{F}_p|$ . Then,*

$$|Z(f)| \leq \left(1 - \frac{1}{r^{d/(r-1)}}\right)r^n.$$

*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  $f \in \mathbb{F}_p[x_1, \dots, x_n]$  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.  $\square$

## 6.2 Uniformity of random Cayley hypergraphs

A second candidate for constructing outlaws comes from quasirandom properties of Cayley graphs and hypergraphs.



**Random Cayley graphs and 2-query LDCs.** For graphs, an important quasirandom property is spectral expansion. If for a regular graph  $G$ , we let  $1 = \lambda_1(G) \geq \lambda_2(G) \geq \dots \geq \lambda_n(G) \geq -1$  denote the eigenvalues of the normalized adjacency matrix, then the second eigenvalue is defined as  $\lambda(G) = \max_{i \in \{2, \dots, n\}} |\lambda_i(G)|$ . The importance of this parameter stems from the fact that if it is small, then every large subset of vertices is connected to its complement by a large number of edges [Tan84, AM85], a property that sparse random graphs have with high probability (we refer to [HLW06] for a survey on expander graphs).

A famous result of Alon and Roichman [AR94] asserts that random Cayley graphs require relatively low degree to be spectral expanders with good probability. For a finite group  $\Gamma$  and an element  $g \in \Gamma$ , the Cayley graph  $\text{Cay}(\Gamma, \{g\})$  is the 2-regular graph with vertex set  $\Gamma$  and edge set  $\{\{u, gu\} : u \in \Gamma\}$ , where in case  $g^2 = 1$ , all edges are doubled (parallel edges are thus allowed). For a multiset  $S = \{g_1, \dots, g_k\} \subseteq \Gamma$ , the Cayley graph  $\text{Cay}(\Gamma, S)$  is the  $2k$ -regular graph formed by the disjoint union of the graphs  $\text{Cay}(\Gamma, \{g_1\}), \dots, \text{Cay}(\Gamma, \{g_k\})$ .

**Theorem 6.10** (Alon–Roichman Theorem). *For any  $\varepsilon \in (0, 1)$  there exists a  $c(\varepsilon) \in (0, \infty)$  such that the following holds. Let  $\Gamma$  be a finite group of cardinality  $n$ . Let  $k \geq c(\varepsilon) \log n$  be an integer and let  $g_1, \dots, g_k$  be independent uniformly distributed elements from  $\Gamma$ . Then, with probability at least  $1/2$ , the Cayley graph  $G = \text{Cay}(\Gamma, \{g_1, \dots, g_k\})$  satisfies  $\lambda(G) \leq \varepsilon$ .*

The link with Theorem 1.4 follows from the fact that the above result can equivalently be phrased as saying that the normalized adjacency matrix  $A_G$  of the random graph  $G$  as in Theorem 6.10 *concentrates* around its expectation. The normalized adjacency matrix of  $\text{Cay}(\Gamma, \{g_i\})$ , denoted  $A_i$ , has expectation  $J/n$ , where  $J$  is the all-ones matrix, and  $A_G$  is the average  $(A_1 + \dots + A_k)/k$ . The Schatten- $\infty$  norm (also known as the spectral norm or operator norm) of a matrix  $B$  is given by  $\|B\|_{S_\infty} = \sup\{x^\top B y : \|x\|_{\ell_2} \leq 1, \|y\|_{\ell_2} \leq 1\}$ . Due to the characterization  $\lambda(G) = \|A_G - J/n\|_{S_\infty}$ , Theorem 6.10 says that the value

$$\left\| \frac{1}{k} \sum_{i=1}^k (A_i - J/n) \right\|_{S_\infty}$$

is small with good probability provided  $k$  is large enough. Good concentration is thus good for expansion. Our Theorem 1.4 implies that *poor* concentration is good for LDCs!

**Corollary 6.11.** *Let  $\Gamma$  be a finite group of cardinality  $n$  and let  $\varepsilon \in (0, 1)$ . Let  $k$  be the largest positive integer such that with probability at least  $1/2$ , for independent uniformly distributed elements  $g_1, \dots, g_k \in \Gamma$ , the graph  $G = \text{Cay}(\Gamma, \{g_1, \dots, g_k\})$  satisfies  $\lambda(G) > \varepsilon$ . Then, there exists a  $(2, 1, \Omega(\varepsilon))$ -LDC sending  $\{0, 1\}^l$  to  $\{0, 1\}^n$ , where  $l = \Omega(\varepsilon^2 k / \log(1/\varepsilon))$ .*

It is not hard to show that if  $\Gamma$  is Abelian, then any Cayley graph  $G = \text{Cay}(\Gamma, S)$  generated by  $|S| \leq (\log |\Gamma|)/3$  group elements satisfies  $\lambda(G) \geq 1/2$  [HLW06, Proposition 11.5]. Together with Corollary 6.11, this fact implies the existence of 2-query LDCs of exponential length; arguably the most round-about way to prove this!

Below, we prove a more general version of Corollary 6.11, giving the existence of  $q$ -query LDCs from lower bounds on the required degree for uniformity of random  $q$ -uniform Cayley hypergraphs. For this, we gather the following definitions.

**Regular hypergraphs and uniformity.** A  $q$ -uniform hypergraph  $H = (V, E)$  with vertex set  $V$  has as edge set  $E$  a family of unordered  $q$ -element multisets with possible parallel edges. For  $u_1, \dots, u_q \in V$  let  $e_H(u_1, \dots, u_q)$  denote the number of edges equal to  $\{u_1, \dots, u_q\}$ .<sup>4</sup> The *adjacency form* of  $H$  is the  $q$ -linear form  $\bar{A}_H : \mathbb{R}^V \times \dots \times \mathbb{R}^V \rightarrow \mathbb{R}$  given by  $\bar{A}_H(1_{\{u_1\}}, \dots, 1_{\{u_q\}}) = e_H(u_1, \dots, u_q)$ . The degree of a vertex  $v \in V$  is defined by  $\bar{A}_H(1_{\{v\}}, 1_V, \dots, 1_V)$  and  $H$  is  $k$ -regular if every vertex has degree exactly  $k$ , in which case its normalized adjacency form is  $A_H = \bar{A}_H/k$ . Observe that we obtain the usual definition of a regular graph for the case  $q = 2$ .

Given a regular hypergraph  $H = (V, E)$  and a regular sub-hypergraph  $J = (V, E')$  of  $H$  based on a multiset  $E' \subseteq E$ , we define the *uniformity of  $J$  relative to  $H$*  by

$$\Delta_H(J) = \max_{T_1, \dots, T_q \subseteq V} \frac{1}{|V|} \left| A_J(1_{T_1}, \dots, 1_{T_q}) - A_H(1_{T_1}, \dots, 1_{T_q}) \right|. \quad (9)$$

Observe that we get the usual notion of uniformity for graphs if  $H$  is the complete graph with all self-loops. We say that  $J$  is  $\varepsilon$ -uniform relative to  $H$  if  $\Delta_H(J) \leq \varepsilon$ .

**Cayley hypergraphs.** A Cayley hypergraph over a finite group  $\Gamma$  is a disjoint union of hypergraphs with edge sets of the form  $\{\pi_1(v), \dots, \pi_q(v)\}$ ,  $v \in \Gamma$ , where  $\pi_1, \dots, \pi_q$  are permutations on  $\Gamma$ , and where the edge multiplicities are set such that

$$e_H(u_1, \dots, u_q) = \sum_{\sigma \in S_q} \sum_{v \in \Gamma} 1_{\{u_1\}}(\pi_{\sigma(1)}(v)) \cdots 1_{\{u_q\}}(\pi_{\sigma(q)}(v)). \quad (10)$$

This choice of multiplicities ensures that we always get a  $(q!)$ -regular hypergraph regardless of the choice of permutations. The permutations defining Cayley hypergraphs are given in terms of an integer vector  $r \in (\mathbb{Z} \setminus \{0\})^q$  such that  $\Gamma$  has no elements of order  $r_j$  for every  $j \in [q]$ , so that for every  $g \in \Gamma$  and  $j \in [q]$ , the map  $u \mapsto u^{r_j} g$  is a permutation. For a  $q$ -tuple  $\mathbf{g} = (g[1], \dots, g[q]) \in \Gamma^q$ , we then define  $\text{Cay}^{(q)}(\Gamma, r, \mathbf{g})$  to be the hypergraph as above with  $\pi_j(u) = u^{r_j} g[j]$ . For a multiset  $S = \{\mathbf{g}_1, \dots, \mathbf{g}_k\} \subseteq \Gamma^q$ , we let  $\text{Cay}^{(q)}(\Gamma, r, S)$  be the  $(q!)k$ -regular hypergraph given by the disjoint union of  $\text{Cay}^{(q)}(\Gamma, r, \{\mathbf{g}_i\})$  for  $i \in [k]$ .

**Examples.** Notice that  $\text{Cay}^{(2)}(\Gamma, (1, 1), (g, h))$  is the 2-regular Cayley graph generated by the element  $g^{-1}h$ . If  $\Gamma = \mathbb{F}_p^n$  for some odd prime  $p$ , if  $\mathbf{1} \in \mathbb{Z}^p$  is the all-ones vector and if  $S = \{(0, y, 2y, \dots, (p-1)y) : y \in \mathbb{F}_p^n\}$ , then the edges of  $H = \text{Cay}^{(p)}(\Gamma, \mathbf{1}, S)$  are all the affine lines in  $\mathbb{F}_p^n$ . Moreover, if  $D \subseteq \mathbb{F}_p^n$  and  $S' = \{(0, y, 2y, \dots, (p-1)y) : y \in D\}$ , then the edges  $J = \text{Cay}^{(p)}(\Gamma, \mathbf{1}, S')$  are all the affine lines whose direction lies in  $D$ . Finally, if  $\Delta_H(J) \leq \varepsilon$ , then for every set  $T \subseteq \mathbb{F}_p^n$ , the fraction of lines in  $T$  with direction in  $D$  is within  $\varepsilon$  of the fraction of all lines in  $T$ . In [BR] it is shown that if  $n \geq p^2$ , then there is an absolute constant  $c(p) \in (0, 1)$  depending on  $p$  only such that if  $H$  and  $J$  are as above, then  $\Delta_H(J) < c(p)$  implies that  $|D| \geq \Omega(n^{p-1})$ .<sup>5</sup>

<sup>4</sup>Curly brackets delimit *multisets*: unordered lists that may contain repeated elements.

<sup>5</sup>Their proof, however, does not show that the assumption of Corollary 6.4 holds for  $|D| \geq \Omega(n^{p-1})$ , as it relies on the construction of a combinatorial rectangle  $T_1 \times \dots \times T_p$  consisting of different sets.

**Random  $q$ -uniform Cayley hypergraphs and  $q$ -query LDCs.** Theorem 1.4 implies the following link between LDCs and random Cayley hypergraphs.

**Corollary 6.12.** *Let  $\Gamma$  be a finite group of cardinality  $n$  and let  $H = \text{Cay}^{(q)}(\Gamma, r, S)$  be a  $q$ -uniform Cayley hypergraph. Let  $k$  be the largest positive integer such that with probability at least  $1/2$ , if  $\mathbf{g}_1, \dots, \mathbf{g}_k$  are independent uniformly distributed random elements over  $S$ , the random Cayley hypergraph  $J = \text{Cay}^{(q)}(\Gamma, r, \{\mathbf{g}_1, \dots, \mathbf{g}_k\})$  satisfies  $\Delta_H(J) \geq \varepsilon$ . Then, there exists a  $(q, 1, \Omega(\varepsilon))$ -LDC sending  $\{0, 1\}^l$  to  $\{0, 1\}^n$ , where  $l = \Omega(\varepsilon^2 k / \log(1/\varepsilon))$ .*

*Proof:* Associate with every  $q$ -linear form  $A : \mathbb{R}^\Gamma \times \dots \times \mathbb{R}^\Gamma \rightarrow \mathbb{R}$  a homogeneous  $q|\Gamma|$ -variate degree- $q$  function  $f : \{-1, 1\}^{q|\Gamma|} \rightarrow \mathbb{R}$  in the obvious way. It follows from (10) and the normalization of the adjacency forms that for every fixed  $g \in \Gamma$ , the functions  $f_g$  associated with the adjacency form of the hypergraph  $\text{Cay}^{(q)}(\Gamma, r, \{g\})$  are 1-smooth.

Moreover, for any  $q$ -linear form  $A$  on  $\mathbb{R}^n$ , it holds that

$$\frac{\max \{|A(x[1], \dots, x[q])| : x[1], \dots, x[q] \in \{-1, 1\}^n\}}{\max \{|A(1_{T_1}, \dots, 1_{T_q})| : T_1, \dots, T_q \subseteq [n]\}} \geq 1.$$

The form  $A_J$  is the average of  $k$  independent identically distributed adjacency forms  $A_1, \dots, A_k$  that have expectation  $A_H$ . Letting  $f_1, \dots, f_k$  and  $\bar{f}$  be the functions associated with  $A_1, \dots, A_k$  and  $A_H$ , respectively, it follows that

$$\begin{aligned} \mathbb{E} \left[ \left\| \frac{1}{k} \sum_{i=1}^k (f_i - \bar{f}) \right\|_{L_\infty} \right] &= \mathbb{E} \left[ \max \{|(A_J - A_H)(x[1], \dots, x[q])| : x[1], \dots, x[q] \in \{-1, 1\}^\Gamma\} \right] \\ &\geq \mathbb{E}[\Delta_H(J)] \\ &\geq \frac{\varepsilon}{2}, \end{aligned}$$

where the last line follows from the fact that uniformity is nonnegative. The result now follows from Theorem 1.4.  $\square$

**Spectral expansion and uniformity.** For regular *graphs*, the famous Expander Mixing Lemma [HLW06] shows that  $\Delta(G) \leq \lambda(G)$  holds in general. Whereas the reverse inequality does not hold in general [CZ16], the reason why Corollary 6.11 could be stated in terms of the second eigenvalue is that for Cayley graphs, a near-reverse inequality does hold: for some absolute constant  $c \in (0, \infty)$ , we have  $\Delta(G) \geq c\lambda(G)$  [KRS16, CZ16]. A second eigenvalue for hypergraphs, analogous to uniformity for hypergraphs, was defined and studied in [BR], where first steps to generalize the Alon–Roichman Theorem were taken. While the Expander Mixing Lemma easily generalizes, it is unknown whether the analogue of [KRS16, CZ16] holds for Cayley hypergraphs.

## References

- [ALM<sup>+</sup>98] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, 1998.
- [AM85] N. Alon and V. D. Milman.  $\lambda_1$ , isoperimetric inequalities for graphs, and super-concentrators. *Journal of Combinatorial Theory. Series B*, 38(1):73–88, 1985.
- [AR94] N. Alon and Y. Roichman. Random Cayley graphs and expanders. *Random Structures & Algorithms*, 5, 1994.
- [AS98] S. Arora and S. Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM*, 45(1):70–122, 1998.
- [BK95] M. Blum and S. Kannan. Designing programs that check their work. *Journal of the ACM*, 42(1):269–291, 1995.
- [BNR12] J. Briët, A. Naor, and O. Regev. Locally decodable codes and the failure of cotype for projective tensor products. *Electronic Research Announcements in Mathematical Sciences (ERA-MS)*, 19:120–130, 2012.
- [BR] J. Briët and S. Rao. To appear.
- [Bri15] J. Briët. On embeddings of  $\ell_1^k$  from locally decodable codes. *Electronic Colloquium on Computational Complexity (ECCC)*, (86), 2015.
- [CKGS98] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan. Private information retrieval. *Journal of the ACM*, 45(6):965–981, 1998.
- [CT14] G. Cohen and A. Tal. Two structural results for low degree polynomials and applications. *arXiv preprint arXiv:1404.0654*, 2014.
- [CZ16] D. Conlon and Y. Zhao. Quasirandom Cayley graphs. *preprint*, 2016. Available at arXiv:1603.03025 [math.CO].
- [DG15] Z. Dvir and S. Gopi. 2-server pir with sub-polynomial communication. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, pages 577–584. ACM, 2015.
- [Efr09] K. Efremenko. 3-query locally decodable codes of subexponential length. In *Proceedings of the forty-first annual ACM symposium on Theory of computing (STOC 2009)*, pages 39–44. 2009.
- [GHSY12] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin. On the locality of codeword symbols. *IEEE Trans. Information Theory*, 58(11):6925–6934, 2012. doi:10.1109/TIT.2012.2208937.

- [GKdO<sup>+</sup>16] S. Gopi, S. Kopparty, R. M. de Oliveira, N. Ron-Zewi, and S. Saraf. Locally testable and locally correctable codes approaching the gilbert-varshamov bound. *Electronic Colloquium on Computational Complexity (ECCC)*, 23:122, 2016.
- [GKST02] O. Goldreich, H. Karloff, L. J. Schulman, and L. Trevisan. Lower bounds for linear locally decodable codes and private information retrieval. In *Computational Complexity, 2002. Proceedings. 17th IEEE Annual Conference on*, pages 143–151. IEEE, 2002.
- [HLW06] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc.*, 43:439–561, 2006.
- [KMRS16] S. Kopparty, O. Meir, N. Ron-Zewi, and S. Saraf. High-rate locally-correctable and locally-testable codes with sub-polynomial query complexity. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 202–215. 2016. doi: 10.1145/2897518.2897523.
- [KRS16] Y. Kohayakawa, V. Rödl, and M. Schacht. Discrepancy and eigenvalues of Cayley graphs. *preprint*, 2016. Available at arXiv:1602.02291 [math.CO].
- [KS07] T. Kaufman and M. Sudan. Sparse random linear codes are locally decodable and testable. In *Foundations of Computer Science, 2007. FOCS'07. 48th Annual IEEE Symposium on*, pages 590–600. IEEE, 2007.
- [KSY10] S. Kopparty, S. Saraf, and S. Yekhanin. High-rate codes with sublinear-time decoding. *Electronic Colloquium on Computational Complexity (ECCC)*, 17:148, 2010.
- [KT00] J. Katz and L. Trevisan. On the efficiency of local decoding procedures for error-correcting codes. In *Proceedings of the 32nd annual ACM symposium on Theory of computing (STOC 2000)*, pages 80–86. ACM Press, 2000.
- [KW04] I. Kerenidis and R. d. Wolf. Exponential lower bound for 2-query locally decodable codes via a quantum argument. *J. of Computer and System Sciences*, 69:395–420, 2004. Preliminary version appeared in STOC'03.
- [MV03] S. Mendelson and R. Vershynin. Entropy and the combinatorial dimension. *Invent. Math.*, 152(1):37–55, 2003.
- [Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948. ISSN 1538-7305. doi:10.1002/j.1538-7305.1948.tb01338.x.
- [Tal14] M. Talagrand. *Upper and lower bounds for stochastic processes*, volume 60 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of*

*Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer, Heidelberg, 2014. Modern methods and classical problems.

- [Tan84] M. R. Tanner. Explicit concentrators from generalized  $n$ -gons. *SIAM Journal on Algebraic Discrete Methods*, 5(3):287–293, 1984.
- [Tao12] T. Tao. *Higher order Fourier analysis*, volume 142. American Mathematical Society, 2012.
- [Tao14] T. Tao. Algebraic combinatorial geometry: the polynomial method in arithmetic combinatorics, incidence combinatorics, and number theory. *EMS Surv. Math. Sci.*, 1:1–46, 2014.
- [TV06] T. Tao and V. Vu. *Additive Combinatorics*. Cambridge University Press, 2006.
- [Woo07] D. Woodruff. New lower bounds for general locally decodable codes. *Electronic Colloquium on Computational Complexity (ECCC)*, 14(006), 2007.
- [Woo12] D. P. Woodruff. A quadratic lower bound for three-query linear locally decodable codes over any field. *J. Comput. Sci. Technol.*, 27(4):678–686, 2012. doi:10.1007/s11390-012-1254-8.
- [Yek07] S. Yekhanin. Towards 3-query locally decodable codes of subexponential length. In *Proceedings of the 39th annual ACM symposium on Theory of computing (STOC 2007)*, pages 266–274. 2007.
- [Yek12] S. Yekhanin. Locally decodable codes. *Foundations and Trends in Theoretical Computer Science*, 6(3):139–255, 2012.