Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 130 (2010)

Locally Testable vs. Locally Decodable Codes

Tali Kaufman[*]
Dept. of Computer Science and Applied Math.
Weizmann Institute
Rehovot, Israel
kaufmant@mit.edu

Michael Viderman[†]
Dept. of Computer Science
Technion
Haifa 32000, Israel
viderman@cs.technion.ac.il

August 23, 2010

Abstract

We study the relation between locally testable and locally decodable codes. Locally testable codes (LTCs) are error-correcting codes for which membership of a given word in the code can be tested probabilistically by examining it in very few locations. Locally decodable codes (LDCs) allow to recover each message entry with high probability by reading only a few entries of a slightly corrupted codeword. A linear code $\mathcal{C} \subseteq \mathbb{F}_2^n$ is called sparse if $n \geq 2^{\Omega(\dim(\mathcal{C}))}$.

It is well-known that LTCs do not imply LDCs and that there is an intersection between these two families. E.g. the Hadamard code is both LDC and LTC. However, it was not known whether LDC implies LTC. We show the following results.

- Two-transitive codes with a local constraint imply LDCs, while they do not imply LTCs.
- Every non-sparse LDC contains a large subcode which is not LTC, while every subcode of an LDC remains LDC. Hence, every non-sparse LDC contains a subcode that is LDC but is not LTC.

The above results demonstrate inherent differences between LDCs and LTCs, in particular, they imply that LDCs do not imply LTCs.

[*]This research was partially supported by the Koshland Fellowship
[†]This research was partially supported by the Israeli Science Foundation (grant No. 679/06).

ISSN 1433-8092

1 Introduction

A linear code over a finite field $\mathbb{F}$ is a linear subspace $\mathcal{C} \subseteq \mathbb{F}^n$. The dimension of $\mathcal{C}$ is its dimension as a vector space, and its rate is the ratio of its dimension to n. The distance of $\mathcal{C}$ is the minimal Hamming distance between two different codewords. Typically, we are interested in codes whose distance is a linear to the block length n, i.e., $\Omega(n)$.

Locally testable codes (LTCs) are error-correcting codes for which membership of a given word in the code can be tested probabilistically by examining it in very few locations. More precisely such a code has a *tester*, which is a randomized algorithm with oracle access to the received word x. The tester reads at most q symbols from x and based on this local view decides if $x \in \mathcal{C}$ or not. It should accept codewords with probability one, and reject words that are far (in Hamming distance) from the code with noticeable probability.

In recent years, starting with the work of Goldreich and Sudan [12], several surprising constructions of LTCs have been given (see [10] for an extensive survey of some of these constructions). The principal challenge is to understand the largest asymptotic rate possible for LTCs, and to construct LTCs approaching this limit. We now know constructions of LTCs of dimension $n/\log^{O(1)} n$ which can be tested with only three queries [5, 7], [21]. The main open question in the subject is whether there are asymptotically good LTCs, i.e., LTCs that have dimension $\Omega(n)$ and distance $\Omega(n)$.

The only negative results on LTCs concern binary codes testable with just 2-queries [2, 15] (which is a severe restriction), random LDPC codes [4], cyclic codes [1] [1], Solvable codes [19] and codes with small redundancy in the small weight dual words [3].

On the other hand, locally decodable codes (LDCs) allow to recover each message entry with high probability by reading only a few entries of the codeword even if a constant fraction of it is adversely corrupted.

The best construction of LDCs was initiated by the breakthrough results of Yekhanin [26] who showed a (conditional) subexponential construction of 3-query LDCs. Later Efremenko [9] showed unconditional subexponential construction of LDCs. Gopalan showed that these codes can be considered as a sub-family of Reed-Muller codes [13].

Katz and Trevisan [16] were first who defined formally LDCs and showed that LDCs have superlinear blocklength. Goldreich et al. [11] showed that linear 2-query LDCs have exponential blocklength. This result was generalized by Dvir and Shpilka [8] for all arbitrarily large fields. Obata [22] and then Shiowattana and Lokam [23] showed asymptotically tight (exponential) lower bounds on the blocklength of 2-query LDCs. Kerenidis and de Wolf [20] showed exponential lower bounds for 2-query LDC and improved superlinear lower bound for q-query LDCs, where $q \geq 3$. Then Woodruff [25] improved this result for odd q and showed that q-query LDCs ($q \geq 3$) with k message bits and blocklength n have $n \geq \Omega(k^{1+\frac{1}{\lceil q/2-1 \rceil}})/\log(k)$ and for 3-query linear LDCs showed that $n \geq \Omega(k^2/\log\log(k))$. The known lower bounds for q-query LDCs for $q \geq 3$ seems to be very far from tight.

LDCs are related to private information retrieval protocols, initiated by [6], while LTCs are related to PCPs [12]. Both these families of error correcting codes are explicitly studied, for survey see e.g. [24]. In spite of the fact, the distinction between the two families of the codes was not made. Namely, it is well-known that there is an intersection between the two families of codes, e.g. the famous Hadamard code is 3-query LTC and 2-query LDC. Moreover, it is well-known that LTCs do not imply LDCs, i.e.,

[1] The last result rules out asymptotically good *cyclic* LTCs; the existence of asymptotically good cyclic codes has been a long-standing open problem, and the result shows the "intersection" of these questions concerning LTCs and cyclic codes has a negative answer.

there are LTCs which are not LDCs. This follows simply by comparing the upper and lower bounds on the blocklength of these families of codes. If $\mathcal{C} \subseteq \mathbb{F}^n$ is a q-query LDCs then $n \geq \Omega(\dim(C)^{q/(q-1)})$ (by Katz and Trevisan [16]), while there exist (best known) LTCs s.t. $n \leq O(\dim(\mathcal{C}) \cdot (\mathit{poly}\log(\dim(\mathcal{C}))))$ [5, 7], [21]. However, the other direction, i.e., whether LDCs imply LTCs, was not known.

1.1 Our Results

We show that LDC does not imply LTC, and in fact there are inherent differences between LDCs and LTCs. Specifically we show the following results.

- In Theorem 3.1 we show that codes invariant under two-transitive groups that obey a local constraint are LDCs, while they are not necessarily LTCs. This provides a general proof to the local decodability of polynomial codes such as Hadamard code, Reed-Muller codes and dual-BCH codes. Combining this with a recent result of [14], we obtain an explicit family of linear codes which is locally decodable but is not locally testable.

- In Theorem 4.1 we show that every non-sparse code contains a large subcode which is not LTC, while every subcode of an LDC remains LDC (Corollary 4.4). Hence, every non-sparse LDC contains a subcode that is LDC but is not LTC. Moreover, we show (Theorem 4.6) that if we consider uniform-LTCs (for which a tester picks every possible local constraint with the same probability) then, in fact, *every* non-sparse LDC has *many* large subcodes which are not uniform-LTCs (but still LDCs).

1.2 On Sparse codes vs. Non-sparse codes

Recall that a code $\mathcal{C} \subseteq \mathbb{F}_2^n$ is called sparse if $n \geq 2^{\Omega(\dim(\mathcal{C}))}$, otherwise the code is non-sparse. A sparse code $\mathcal{C}$ is called unbiased if all nonzero codewords $c \in \mathcal{C}$ have relative weight ranging in $(\frac{1}{2} - n^{-\gamma}, \frac{1}{2} + n^{-\gamma})$ for some constant $\gamma > 0$. Kaufman and Sudan [17] showed that *all* sparse unbiased codes are LTCs and LDCs. Since every subcode of a sparse unbiased code is a sparse unbiased code we conclude that it is an LTC and LDC. However, sparse codes have exponential blocklength.

Our Theorem 4.1 shows that *every* non-sparse LDC contains a large subcode which is not LTC. Hence, every non-sparse LDC contains a subcode that is LDC but is not LTC. This demonstrates an inherent difference between sparse and non-sparse codes. In sparse codes local testability is preserved in subcodes, while in non-sparse codes local testability is not preserved in subcodes. In contrast to local testability, local decodability of all codes is always preserved in their subcodes (Corollary 4.4).

1.3 Paper Organization

We start with some definitions in Section 2. In Section 3 we provide the proof to our first theorem. In Section 4 we prove our second theorem, moreover we also prove a stronger version of our second theorem that applies only for a special type of LTCs known as uniform-LTCs.

2 Preliminaries

Let $\mathbb{F}$ be a finite field and $[n]$ be the set $\{1, \ldots, n\}$. In this work, we consider only linear codes. We start with a few definitions.

Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code over $\mathbb{F}$. For $w \in \mathbb{F}^n$, let $\mathrm{supp}(w) = \{i \in [n] \mid w_i \neq 0\}$ and $|w| = |\mathrm{supp}(w)|$. We define the *distance* between two words $x, y \in \mathbb{F}^n$ to be $\Delta(x, y) = |\{i \mid x_i \neq y_i\}|$ and the relative distance to be $\delta(x,y) = \frac{\Delta(x,y)}{n}$. The distance of a code is denoted by $\Delta(\mathcal{C})$ and defined to be the minimal value of $\Delta(x,y)$ for two distinct codewords $x,y \in \mathcal{C}$. Similarly, the relative distance of the code is denoted $\delta(\mathcal{C}) = \frac{\Delta(\mathcal{C})}{n}$. For $x \in \mathbb{F}^n$ and $\mathcal{C} \subseteq \mathbb{F}^n$, let $\delta(x,\mathcal{C}) = \min\limits_{y \in \mathcal{C}} \{\delta(x,y)\}$ denote the relative distance of x from the code $\mathcal{C}$. We note that $\Delta(\mathcal{C}) = \min\limits_{c \in \mathcal{C} \setminus \{0\}} \{|c|\}$. If $\delta(x,\mathcal{C}) \geq \epsilon$, we say that x is ϵ-far from $\mathcal{C}$ and otherwise x is ϵ-close to $\mathcal{C}$. Let $\dim(\mathcal{C})$ be the dimension of $\mathcal{C}$. The vector inner product between u_1 and u_2 is denoted by $\langle u_1, u_2 \rangle$. The dual code $\mathcal{C}^\perp$ is defined as $\mathcal{C}^\perp = \{u \in \mathbb{F}^n \mid \forall c \in \mathcal{C} : \langle u, c \rangle = 0\}$. In a similar way we define $\mathcal{C}^\perp_{\leq t} = \left\{u \in \mathcal{C}^\perp \mid |u| \leq t\right\}$ and $\mathcal{C}^\perp_t = \left\{u \in \mathcal{C}^\perp \mid |u| = t\right\}$.

For $w \in F^n$ and $S = \{j_1, j_2, \ldots, j_m\} \subseteq [n]$, where $j_1 < j_2 < \ldots < j_m$, let $w|_S = (w_{j_1}, w_{j_2}, \ldots, w_{j_m})$ be the *restriction* of w to the subset S. Let $\mathcal{C}|_S = \{c|_S \mid c \in \mathcal{C}\}$ denote the restriction of the code $\mathcal{C}$ to the subset S. For $T \subseteq \mathbb{F}^n$ and $w \in \mathbb{F}^n$ we say that $w \perp T$ if for all $t \in T$ we have $\langle w, t \rangle = 0$.

2.1 Codes invariant under groups

Let G be a group of permutations over $[n]$. For $\pi \in G$ and $w = (w_1, w_2, ..., w_n) \in F^n$ with some abuse of notation we let $\pi(w) = (w_{\pi^{-1}(1)}, ..., w_{\pi^{-1}(n)})$ be a π-permuted word. Note that since G is a group and $\pi \in G$ we have $\pi^{-1} \in G$. A linear code $\mathcal{C}$ is invariant under G if for every $\pi \in G$ and $c \in \mathcal{C}$ we have $\pi(c) \in \mathcal{C}$. Note that if $\mathcal{C}$ is invariant under G then also $\mathcal{C}^\perp$ is invariant under G. G is called 2-transitive if for all $i \neq j \in [n]$ and $i' \neq j' \in [n]$ we have $\pi \in G$ such that $\pi(i) = i'$ and $\pi(j) = j'$. A linear code $\mathcal{C}$ is 2-transitive if it is invariant under some 2-transitive permutation group G.

2.2 LTCs, LDCs and LCCs

In this section, we define LTCs, LDCs and LCCs formally and recall a few concepts that will be used later in this paper. We define LTCs following [3].

Definition 2.1 (LTCs and Testers)**.** Let $\mathcal{C} \in \mathbb{F}^n$ be a linear code. Given a distribution $\mathcal{D}$ over set $\mathcal{C}^\perp$, we define the support of $\mathcal{D}$ over $\mathcal{C}^\perp$ as $\mathcal{D}_S = \left\{u \in \mathcal{C}^\perp \mid \mathcal{D}(u) > 0\right\}$. We say that $\mathcal{D}$ is a (q,ϵ,δ)-distribution for the code $\mathcal{C}$, if the following conditions are satisfied:

- $\mathcal{D}_S \subseteq \mathcal{C}^\perp_{\leq q}$.
- For all $x \in \mathbb{F}^n$ such that $\delta(x,\mathcal{C}) \geq \delta$ it holds that $\mathop{\mathbf{Pr}}\limits_{u \sim \mathcal{D}}[\langle u, x \rangle \neq 0] \geq \epsilon$.

We say that $C \subseteq \mathbb{F}^n$ is a (q,ϵ,δ)-LTC if it has a (q,ϵ,δ)-distribution $\mathcal{D}$. If $\mathcal{D}$ is uniform over $\mathcal{C}^\perp_{\leq q}$ we say that $\mathcal{C}$ is a (q,ϵ,δ)-uniform LTC.

The parameter q is known as query complexity, ϵ is the rejection probability and δ is the distance threshold.

Note that if $\mathcal{C}$ is a (q,ϵ,δ)-LTC then $\mathcal{C}$ is also a (q,ϵ,δ')-LTC for all $\delta' \geq \delta$.

We say that a family of codes $\left\{\mathcal{C}^{(n)} \;\middle|\; n \in \mathbb{Z}\right\}$ is locally testable if there exist constants $q,\epsilon,\delta > 0$ such that for infinitely many n it holds that $\mathcal{C}^{(n)} \subseteq \mathbb{F}^n$ is a (q,ϵ,δ)-LTC, where $\delta \leq \delta(\mathcal{C}^{(n)})/3$.

Note that every perfect code $\mathcal{C}$ is $(0,1,\delta > \delta(\mathcal{C})/2)$-LTC, i.e., the code is locally testable with 0 queries and highest possible rejection probability when the distance threshold is $\delta > \delta(\mathcal{C})/2$ since there are no words which are δ-far from the code. Hence, to avoid trivial cases we must require the distance threshold δ to be at most $\delta(\mathcal{C})/2$. Moreover, in the area of locally testable codes we usually require even less distance threshold, at most $\delta \leq \delta(\mathcal{C})/3$. E.g., all known constructions of LTCs satisfy this requirement (see e.g., [12, 17, 18, 21, 7]). From the other side, if for all constants $q, \epsilon > 0$ the code $\mathcal{C}$ is not $(q, \epsilon, \delta(\mathcal{C})/3)$-LTC we say that $\mathcal{C}$ is not locally testable (see e.g., [1, 3, 14]).

Remark 2.2. Our proofs will follow also if we define "uniform LTC" as LTC with a uniform distribution over $\mathcal{C}_q^{\perp}$.

We note that sometimes (e.g. [3]) uniform LTCs mean that the associated distribution is uniform over its support and not over all $\mathcal{C}_{\leq q}^{\perp}$, which is a less restrictive assumption.

Now we define Locally Decodable Codes (LDCs).

Definition 2.3 (LDCs)**.** Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code of dimension k. Let $E_{\mathcal{C}}$ be the encoding function, i.e., $\mathcal{C} = \left\{E_{\mathcal{C}}(x) \mid x \in \mathbb{F}^k\right\}$. Then $\mathcal{C}$ is a (q, ϵ, δ)-LDC if there exists a randomized decoder ($\mathbb{D}$) that reads at most q entries and the following condition holds:

- For all $x \in \mathbb{F}^k$, $i \in [k]$ and $\hat{c} \in \mathbb{F}^n$ such that $\Delta(E_{\mathcal{C}}(x), \hat{c}) \leq \delta n$ we have $\mathbf{Pr}\left[\mathbb{D}^{\hat{c}}[i] = x_i\right] \geq \dfrac{1}{|\mathbb{F}|} + \epsilon$, i.e., with probability at least $\frac{1}{|\mathbb{F}|} + \epsilon$ entry x_i will be recovered correctly.

Note that definition implies that $\delta < \delta(\mathcal{C})/2$. We say that a family of codes $\left\{C^{(n)} \,\middle|\, n \in \mathbb{Z}\right\}$ is locally decodable if there exist constants $q, \epsilon, \delta > 0$ such that for infinitely many n it holds that $C^{(n)} \subseteq \mathbb{F}^n$ is a (q, ϵ, δ)-LDC.

Now we define locally self-correctable codes (LCCs).

Definition 2.4 (LCCs)**.** Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code of dimension k. Then $\mathcal{C}$ is a (q, ϵ, δ)-LCC if there exists a self-corrector ($\mathbb{SC}$) that reads at most q entries and the following condition holds:

- For all $c \in \mathcal{C}$, $i \in [n]$ and $\hat{c} \in F^n$ such that $\Delta(c, \hat{c}) \leq \delta n$ we have $\mathbf{Pr}\left[\mathbb{SC}^{\hat{c}}[i] = c_i\right] \geq \dfrac{1}{|\mathbb{F}|} + \epsilon$, i.e., with probability at least $\frac{1}{|\mathbb{F}|} + \epsilon$ entry c_i will be recovered correctly.

We say that a code $\mathcal{C}$ is locally self-correctable when $q, \epsilon, \delta > 0$ are constants. Note that the definition implies that $\delta < \delta(\mathcal{C})/2$.

The following folklore claim says that LCCs imply LDCs with the same parameters.

Claim 2.5. *If $\mathcal{C} \subseteq \mathbb{F}^n$ is a (q, ϵ, δ)-LCC then $\mathcal{C}$ is a (q, ϵ, δ)-LDC.*

Proof. Let $k = \dim(\mathcal{C})$. We pick a generator matrix $G \in \mathbb{F}^{n \times k}$ for $\mathcal{C}$, i.e., $\mathcal{C} = \left\{Gm \mid m \in \mathbb{F}^k\right\}$ such that the first k rows of G form identity matrix[2]. Hence the first k symbols of the code are message symbols, i.e., for all $m \in \mathbb{F}^k$ we have $(Gm)|_{[k]} = m$.

Let $\mathbb{SC}$ be a self-corrector for a code $\mathcal{C}$ that for every $i \in [n]$ reads at most q symbols and recovers the symbol i with probability at least $\frac{1}{|\mathbb{F}|} + \epsilon$ even if at most δ-fraction of the symbols was adversely corrupted. In particular, $\mathbb{SC}$ recovers with probability at least $\frac{1}{|\mathbb{F}|} + \epsilon$ every coordinate $i \in [k]$, i.e., every message symbol. We conclude that $\mathcal{C}$ is a (q, ϵ, δ)-LDC. □

We stress that LDCs do not imply LCCs. To see this, let $C \subseteq \mathbb{F}_2^n$ be an LDC. Append to it one entry (with coordinate $(n+1)$) obtaining $C' \subseteq \mathbb{F}_2^{(n+1)}$, such that this entry will not be involved in low-weight constraints of C' and thus could not be recovered with constant query complexity after the codeword will be corrupted. However, the extended code remains LDC. Claim 2.6 provides a formal proof to the above intuition.

[2] $\mathcal{C}$ need not be systematic but it can be easily converted into one as was stated.

Claim 2.6. *There exist constants $q, \epsilon, \delta > 0$ and a code $C' \subseteq \mathbb{F}_2^{n+1}$ s.t. C' is a (q, ϵ, δ)-LDC, but for any constants $q', \epsilon', \delta' > 0$ it holds that C' is not a (q', ϵ', δ')-LCC.*

Proof. Let $C \subseteq \mathbb{F}_2^n$ be a (q, ϵ, δ)-LDC with $\dim(C) = \omega(\log(n))$ (such codes exist, e.g. [9]). Claim 4.5 implies that there exists a word $u \in \mathbb{F}_2^n$ such that $\Delta(u, C^{\perp}) \geq \omega(1)$. Let $C' \subseteq \mathbb{F}_2^{n+1}$ be s.t. $C'|_{[n]} = C$ and for every $c' \in C'$ we have $(c')_{(n+1)} = \langle u, c'|_{[n]} \rangle$, i.e., the first n coordinates of the code C' are identical to the code C and the last bit of the code C' is a sum of the bits indexed by $\mathrm{supp}(u)$.

Let $q', \epsilon', \delta' > 0$ be constants. We argue that there is no $u' \in (C')^{\perp}$ s.t. $n + 1 \in \mathrm{supp}(u')$ and $|u'| \leq q' + 1 = O(1)$. Assume the contrary, and let $u' \in (C')^{\perp}$ be such that $n + 1 \in \mathrm{supp}(u')$ and $|u'| \leq q' + 1$. This implies that the last symbol $(n + 1)$ of the code C' is a sum of only q' symbols of $C'|_{[n]}$. Recall that $C'|_{(n+1)}$ was defined as a sum of symbols indexed by $\mathrm{supp}(u)$. We conclude that for all $c' \in C'|_{[n]}$ we have $\langle u, c' \rangle = \langle u', c' \rangle$ which implies that $\langle u + u', c' \rangle = 0$. Since $C'|_{[n]} = C$ we get that $u + u' \in C^{\perp}$. This means $\Delta(u, C^{\perp}) \leq |u'| \leq q' + 1$ contradicting our assumption that $\Delta(u, C^{\perp}) \geq \omega(1) > q' + 1$.

We conclude that the last bit of C' is not involved in the constraints of weight at most $q' + 1$. Hence any "local view" which contains only q' queries but does not contain the last bit has no information about the last bit of C' and thus C' is not (q', ϵ', δ')-LCC. However, C' is a $(q, \epsilon, \delta - \frac{1}{n+1})$-LDC because if $(\delta - \frac{1}{n+1}) \cdot (n + 1) \leq \delta n$ symbols are corrupted then at most δn symbols from the first n symbols are corrupted. Recall that $C'|_{[n]} = C$ and C is a (q, ϵ, δ)-LDC. So, every message bit can be recovered with probability at least $\frac{1}{2} + \epsilon$ and only q queries. □

3 Two Transitivity with a Local Constraint implies Local Correction

In this section we show (Theorem 3.1) that 2-transitive codes with local constraints imply LCCs and hence also LDCs. However, there exists a family of two-transitive codes with local constraints which is not locally testable, due to [14]. We conclude in Corollary 3.2 that a family of codes $\{\mathcal{C}^{(n)}\}_{n \in \mathbb{Z}}$ (explicitly) shown in [14] is LCC (and LDC) but is not LTC.

Theorem 3.1 (2-transitivity implies LCCs)**.** *If $\mathcal{C} \subseteq \mathbb{F}^n$ is a 2-transitive code such that $\mathcal{C}_q^{\perp} \neq \emptyset$ then $\mathcal{C}$ is a $(q - 1, \frac{1}{6}, \frac{1}{3q})$-LCC (LDC).*

Moreover, there exists a family of 2-transitive codes $\{\mathcal{C}^{(n)}\}_{n \in \mathbb{Z}}$, where $\mathcal{C}^{(n)} \subseteq \mathbb{F}^n$ and $(\mathcal{C}^{(n)})_8^{\perp} \neq \emptyset$, which is not $(q', \epsilon', 1/7)$-LTC for all constants $q', \epsilon' > 0$.

The following corollary follows immediately from Theorem 3.1.

Corollary 3.2. *There exists a family of linear codes $\{\mathcal{C}_n\}_{n \in \mathbb{N}}$, where $\mathcal{C}_n \subseteq \mathbb{F}^n$, which is a $(7, \frac{1}{6}, \frac{1}{24})$-LCC (LDC) but is not $(q', \epsilon', 1/7)$-LTC for all constants $q', \epsilon' > 0$.*

Since by Claim 2.5 (q, ϵ, δ)-LCC is also a (q, ϵ, δ)-LDC then Theorem 3.1 and the lower bound on the blocklength of LDCs by Kerenidis and de Wolf [20] imply the next corollary.

Corollary 3.3. *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a 2-transitive linear code and $k = \dim(\mathcal{C})$. If $\mathcal{C}_q^{\perp} \neq \emptyset$ then $n \geq \Omega(k/\log(k))^{1+1/(\lceil \frac{q}{2} - 1 \rceil)}$.*

Notice that under the famous conjecture that LDCs have superpolynomial blocklength we have that 2-transitive codes with constant weight duals have superpolynomial blocklength.

Proof of Theorem 3.1. Assume $\mathcal{C}$ (and thus $\mathcal{C}^{\perp}$) is invariant under a 2-transitive permutations group G (note that $G \neq \emptyset$, e.g., G contains the identity permutation). Let $u \in \mathcal{C}_q^{\perp}$ (note $\mathcal{C}_q^{\perp} \neq \emptyset$) and let $\operatorname{supp}(u) = \{i_1, i_2, \ldots, i_q\}$. Hence for every $i \in [n]$ there exists $u' \in \mathcal{C}_q^{\perp}$ such that $i \in \operatorname{supp}(u')$, e.g. pick $g \in G$ s.t. $g(i_1) = i$ and let $u' = g(u)$.

We define the self-corrector of entry $i \in [n]$ ($\mathbb{SC}_i$) which on word w

- picks random $g \in G$ such that $g(i) = i$
- queries all entries of $w|_{\operatorname{supp}(g(u')\setminus\{i\})}$
- and recovers the entry $w|_{(i)}$ by $\dfrac{-\sum_{j \in (\operatorname{supp}(g(u'))\setminus\{i\})} w|_j \cdot g(u')|_j}{u'_{(i)}}$.

This self-corrector queries only $q-1$ entries and has perfect completeness, i.e., for all $c = (c_1, \ldots, c_n) \in \mathcal{C}$ and $i \in [n]$ it holds that $\mathbb{SC}_i[c]$ returns c_i. Assume the self-corrector $\mathbb{SC}_i$ is given a word w such that for some $c \in C$ we have $\delta(w, c) \leq \frac{1}{3q}$. Let $I = \operatorname{supp}(w - c)$ and note that $|I| \leq \frac{n}{3q}$. Think of I as a set of corrupted coordinates. Notice that if $\mathbb{SC}_i$ picks $g \in G$, $g(i) = i$ such that $(\operatorname{supp}(g(u')) \setminus \{i\}) \cap I = \emptyset$ then $\mathbb{SC}_i$ recovers correctly the entry w_i, i.e., $SC_i[w] = c_i$. This is true because

$$\frac{\sum_{j \in (\operatorname{supp}(g(u'))\setminus\{i\})} w|_j \cdot g(u')|_j}{u'_{(i)}} = \frac{\sum_{j \in (\operatorname{supp}(g(u'))\setminus\{i\})} c|_j \cdot g(u')|_j}{u'_{(i)}} = c_i,$$

where the last equality follows because $\langle c, g(u') \rangle = 0$. In other words, whenever all the coordinates of $g(u')$ are correct but may be the i's coordinate, $\mathbb{SC}_i$ recovers correctly the entry w_i. Proposition 3.4 implies that for $j \neq i$ and random $g \in G$ such that $g(i) = i$ we have that $g(j)$ is uniformly distributed in $[n] \setminus \{i\}$. We conclude that

$$\mathop{\mathbf{Pr}}_{g \in G, g(i)=i}\left[(\operatorname{supp}(g(u')) \cap I) \setminus \{i\} \neq \emptyset\right] \leq \frac{(q-1)|I|}{n-1} \leq \frac{q|I|}{n}.$$

It follows that the probability that $\mathbb{SC}_i$ picks $g \in G$, $g(i) = i$ such that $\operatorname{supp}(g(u')) \cap I \subseteq \{i\}$ is at least $1 - \frac{q}{3q} = \frac{2}{3}$. So, with probability at least $2/3$ the self-corrector $\mathbb{SC}_i$ picks $g \in G$ such that $g(i) = i$ and $|\operatorname{supp}(g(u')) \setminus \{i\} \cap I| = \emptyset$ and the correction succeeds.

To see that two transitivity does not imply local testability we recall the main result of [14] (Theorem 4.9) that shows a family of 2-transitive codes $\left\{\mathcal{C}^{(n)}\right\}_n$ such that $\mathcal{C}_8^{\perp} \neq \emptyset$ which is not $(q', \epsilon', 1/7)$-LTC for all constants $q', \epsilon' > 0$. □

The proof of Proposition 3.4 is inspired by [1, Section 7].

Proposition 3.4. *Let G be a 2-transitive group and $G_{(i)} = \{g \in G \mid g(i) = i\}$. Then $G_{(i)}$ is a group of permutations s.t. for all $i' \neq i$ and $j' \neq i$ there exists $g \in G_{(i)}$ s.t. $g(i') = j'$. Furthermore, for any $i' \neq i$ and $j' \neq i$ we have $\mathop{\mathbf{Pr}}_{g \in G_{(i)}}\left[g(i') = j'\right] = \dfrac{1}{n-1}$.*

Proof. Let $id \in G$ be the identity permutation, i.e., for all $j \in [n]$ we have $id(j) = j$. We know that $id \in G_{(i)}$, for every $g \in G_{(i)}$ there exists $g^{-1} \in G_{(i)}$, and if $h_1, h_2 \in G_{(i)}$ then also $h_1 \circ h_2 \in G_{(i)}$. We conclude that $G_{(i)}$ is a group of permutations.

For any $i' \neq i$ and $j' \neq i$ there exists $g \in G$ s.t. $g(i) = i$ and $g(i') = j'$ because G is 2-transitive, moreover, $g \in G_{(i)}$.

We argue that for any $i' \neq i$ and $j' \neq i$ we have $\mathop{\mathbf{Pr}}_{g \in G_{(i)}}[g(i') = j'] = \frac{1}{n-1}$. It is sufficient to show that for any $i', j_1', j_2' \neq i$ we have $\mathop{\mathbf{Pr}}_{g \in G_{(i)}}[g(i') = j_1'] = \mathop{\mathbf{Pr}}_{g \in G_{(i)}}[g(i') = j_2']$.

Assume by a way of contradiction that $\mathop{\mathbf{Pr}}_{g \in G_{(i)}}[g(i') = j_1'] > \mathop{\mathbf{Pr}}_{g \in G_{(i)}}[g(i') = j_2']$. Let $h \in G_{(i)}$ s.t. $h(j_2') = j_1'$. Since $G_{(i)}$ is a group then random g is distributed in $G_{(i)}$ exactly as hg is distributed in $G_{(i)}$ and thus

$$\mathop{\mathbf{Pr}}_{g \in G_{(i)}}[g(i') = j_1'] > \mathop{\mathbf{Pr}}_{g \in G_{(i)}}[g(i') = j_2'] = \mathop{\mathbf{Pr}}_{g \in G_{(i)}}[h(g(i')) = j_1'] = \mathop{\mathbf{Pr}}_{g \in G_{(i)}}[g(i') = j_1'].$$

Contradiction. □

4 Non-sparse LDCs contain subcodes that are not LTCs

In this section we show (Theorem 4.1) that non-sparse LDCs contain subcodes that are not LTCs. This demonstrates an important difference between LTCs and LDCs. It turns out that reducing dimension of LDCs remains LDCs (Corollary 4.4), however LTCs are not stable to the dimension reduction. This leads to an interesting observation that every non-sparse LDC has a large subcode which is not LTC (but still LDC). Notice that non-sparse LDCs include Reed-Muller codes of low degree as well as subexponential LDCs that were recently discovered by Yekhanin [26] and Efremenko [9].

Theorem 4.1 (Non-sparse LDCs contain non-LTCs as subcodes)**.** *Let $q > 0$ and $0 < \epsilon, \delta < 1$ be constants. Then for every linear code $\mathcal{C} \subseteq \mathbb{F}^n$ that is a (q, ϵ, δ)-LDC with $\dim(\mathcal{C}) \geq \omega(\log(n))$ and any constants $q', \epsilon' > 0$ there exists a linear subcode $\mathcal{C}' \subset \mathcal{C}$ such that $\dim(\mathcal{C}') \geq \omega(\log(n))$, $\mathcal{C}'$ is a (q, ϵ, δ)-LDC but $\mathcal{C}'$ is not $(q', \epsilon', \delta(\mathcal{C}))$-LTC.*

In the following we show that every subcode of an LDC remains an LDC.

Claim 4.2. *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code and a (q, ϵ, δ)-LDC. Assume that $\mathcal{C}' \subset \mathcal{C}$ is a linear subcode of $\mathcal{C}$. Then $\mathcal{C}'$ is a (q, ϵ, δ)-LDC.*

Proof. Assume $\mathcal{C}$ has the message space $\mathbb{F}^k$ and has the decoder D. Let E_C be an encoding function for $\mathcal{C}$, i.e., $\mathcal{C} = \left\{E_C(x) \mid x \in \mathbb{F}^k\right\}$. Let $T \subset \mathbb{F}^k$ be a (linear) message space for $\mathcal{C}'$, i.e., $\mathcal{C}' = \{E_C(x') \mid x' \in T\}$.

For a linear subspace $M \subset \mathbb{F}^k$ and a subset $S \subseteq [k]$ we say that S is an M-independent subset if there is no $u \in M^{\perp} \setminus \left\{0^k\right\}$ such that $\mathrm{supp}(u) \subseteq S$. We say that $S \subseteq [k]$ is a maximal M-independent subset if S is M-independent and for all $i \in [k] \setminus S$ it holds that $S \cup \{i\}$ is not M-independent subset. Note that if S is a maximal M-independent subset then $|S| = \dim(M)$.

Recall that $T \subset \mathbb{F}^k$ is a linear subspace of $\mathbb{F}^k$. Let $T' \subseteq [k]$ be a maximal T-independent subset (clearly, such subset exists). Let $k' = \dim(T) = |T'|$ and without loss of generality (renaming of bits) we assume that $T' = [k']$. Note that $k' < k$. Let $E_T : \mathbb{F}^{k'} \to \mathbb{F}^k$ be a (linear) encoding function for the linear subspace T, i.e., $T = \left\{E_T(x') \mid x' \in \mathbb{F}^{k'}\right\}$. Note that for all $x' \in \mathbb{F}^{k'}$ we have $E_T(x')|_{[k']} = x'$, namely, the encoding function E_T preserves all "input" bits. Then the code $\mathcal{C}'$ has message space $\mathbb{F}^{T'} = \mathbb{F}^{k'}$ that means $\mathcal{C}' = \left\{E_C(E_T(x')) \mid x' \in \mathbb{F}^{k'}\right\}$. We conclude that for every $x' \in \mathbb{F}^{k'}$ there exists $x \in \mathbb{F}^k$ such that $x|_{[k']} = x'$ and $E_C(x) = E_C(E_T(x'))$.

Recall that $\mathcal{C} \subseteq \mathbb{F}^n$ is a (q, ϵ, δ)-LDC and has the decoder D. We argue that $\mathcal{C}'$ has the same decoder D. Let w be δ-close to $\mathcal{C}'$. But we know that $\mathcal{C}' \subset C$. That means there exists $x' \in \mathbb{F}^{k'}$ such that

$\delta(w, E_C(E_T(x'))) \leq \delta$, and moreover $\delta(w, E_C(x)) \leq \delta$, where $x = E_T(x') \in \mathbb{F}^k$. Note that for all $i \in [k']$ we have $x'_i = x_i$. The fact that $\mathcal{C}$ is a (q, ϵ, δ)-LDC implies that for all $i \in [k'] \subset [k]$ the decoder D recovers correctly the message entry ($x'_i = x_i$) with probability at least $\frac{1}{2} + \epsilon$. We conclude that $\mathcal{C}'$ is a (q, ϵ, δ)-LDC. □

Remark 4.3. The special case of reducing dimension is a removing of columns from the generator matrix. E.g., given a code $\mathcal{C} = \left\{Gx \mid x \in \mathbb{F}^k\right\}$, where $G \in \mathbb{F}^{n \times k}$ is a generator matrix for $\mathcal{C}$. Let $G' \in \mathbb{F}^{n \times (k-1)}$ be obtained by removing the last column of G. Then $\mathcal{C}' = \left\{G'x \mid x \in \mathbb{F}^{k-1}\right\}$ is a linear subcode of $\mathcal{C}$ and $\dim(\mathcal{C}') = \dim(\mathcal{C}) - 1$. In this case message space of $\mathcal{C}'$ is $\mathbb{F}^{k-1}$, while the message space of $\mathcal{C}$ is $\mathbb{F}^k$.

Corollary 4.4 (LDCs are stable for dimension reduction)**.** *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code and a (q, ϵ, δ)-LDC. Take any sequence of linear subcodes: $\mathcal{C}_1 \subset \mathcal{C}_2 \subset \ldots \subset \mathcal{C}_f = \mathcal{C}$. Then for all $i \in [f]$ it holds that $\mathcal{C}_i$ is a (q, ϵ, δ)-LDC.*

Now we show the auxiliary claim and then prove Theorem 4.1.

Claim 4.5. *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code such that $\dim(\mathcal{C}) = \omega(\log(n))$. Then there exists $w \in \mathbb{F}^n$ such that $\Delta(w, \mathcal{C}^\perp) \geq \omega(1)$.*

Proof. For integer R let $V(n, R) = \sum_{i=0}^{R} \binom{n}{i} \cdot (|\mathbb{F}| - 1)^i$ be the volume of a sphere in $\mathbb{F}^n$ of radius R. Let $k = \dim(C) \geq \omega(\log(n))$ and $S = C^\perp$. Then $\dim(S) = n - k$ and $|S| = |\mathbb{F}|^{n-k} = |\mathbb{F}|^n / |\mathbb{F}|^k$. Recall that a covering radius of a code S is $R_S = \max_{w \in \mathbb{F}^n} \Delta(w, S)$, i.e., the largest Hamming distance of any word in $\mathbb{F}^n$ from S. Note that if R_S is constant then $V(n, R_S)$ is polynomial in n and vice versa, if $V(n, R_S)$ is super-polynomial in n then R_S goes to infinity with n. Assume by a way of contradiction that there exists a constant $t > 0$ such that for all $w \in \mathbb{F}^n$ we have $\Delta(w, S) \leq t$, i.e., $R_S \leq t = O(1)$.

The covering radius bound[3] states that

$$|S| \cdot V(n, R_S) \geq |\mathbb{F}|^n.$$

But then $V(n, R_S) \geq |\mathbb{F}|^k$, where $k \geq \omega(\log(n))$. Hence $V(n, R_S)$ must be super-polynomial in n, and $R_S \geq \omega(1)$. Contradiction. □

Proof of Theorem 4.1. Claim 4.5 implies that there exists $u \in \mathbb{F}^n$ such that $\Delta(u, \mathcal{C}^\perp) \geq \omega(1) > q'$. Let $S = \mathrm{span}(\mathcal{C}^\perp \cup \{u\})$. Note that for all $u' \in S$ if $|u'| \leq q'$ then $u' \in \mathcal{C}^\perp$. Let $\mathcal{C}' = S^\perp$ and then $\mathcal{C}'^\perp = S$. We have $\mathcal{C}^\perp \subset \mathcal{C}'^\perp$, $\mathcal{C}' \subset \mathcal{C}$ and in particular, $\dim(\mathcal{C}') = \dim(\mathcal{C}) - 1$. We argue that $\mathcal{C}'$ is not $(q', \epsilon', \delta(\mathcal{C}))$-LTC. We have $c \in \mathcal{C} \setminus \mathcal{C}'$ since $C' \subset C$. However $c \perp (\mathcal{C}')^\perp_{\leq q'}$ because $(\mathcal{C}')^\perp_{\leq q'} \subseteq \mathcal{C}^\perp$ by construction. Hence c is $\delta(\mathcal{C})$-far from $\mathcal{C}'$ but will be accepted with probability 1 by any q'-query tester of $\mathcal{C}'$.

We conclude that $\mathcal{C}' \subset \mathcal{C}$ and $\dim(\mathcal{C}') = \dim(\mathcal{C}) - 1$ but $\mathcal{C}'$ is not $(q', \epsilon', \delta(\mathcal{C}))$-LTC. Claim 4.2 guarantees that $\mathcal{C}'$ is a (q, ϵ, δ)-LDC since $\mathcal{C}'$ is a linear subcode of $\mathcal{C}$. The Theorem follows. □

4.1 Non-sparse LDCs contain many subcodes which are not uniform-LTCs

Theorem 4.1 shows that every non-sparse LDC $\mathcal{C}$ contains a *single* sequence of linear subcodes $\mathcal{C}_1 \subsetneq \mathcal{C}_2 \subsetneq \mathcal{C}_3 \subsetneq \cdots \subsetneq \mathcal{C}$ which are all not LTCs. In the following (Theorem 4.6) we show that every non-sparse LDC and *every* long enough sequence of linear subcodes $\mathcal{C}_1 \subsetneq \mathcal{C}_2 \subsetneq C_3 \subsetneq \cdots \subsetneq \mathcal{C}_\ell = \mathcal{C}$ contains at least one subcode $\mathcal{C}_i$ which is not uniform LTC.

[3] For any code $\mathcal{C} \subseteq \mathbb{F}^n$ (whether linear or not) the covering bound states that the covering radius R of $\mathcal{C}$ relates to n and $|\mathcal{C}|$ by $|\mathcal{C}| \cdot V(n, R) \geq |\mathbb{F}|^n$.

Theorem 4.6. *Let $q, q', \epsilon, \epsilon', \delta > 0$ be constants. Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code such that $\dim(\mathcal{C}) \geq \omega(\log(n))$. Then* every *sequence of ℓ linear subcodes $\mathcal{C}_1 \subsetneq \mathcal{C}_2 \subsetneq \mathcal{C}_3 \subsetneq \cdots \subsetneq \mathcal{C}_\ell = \mathcal{C}$, where $\ell \geq (q' \log(n))/\epsilon'$, contains at least one code $\mathcal{C}_i$ which is not $(q', \epsilon', \delta(\mathcal{C})/2)$-uniform LTC. Moreover, if $\mathcal{C}$ is a (q, ϵ, δ)-LDC then all linear subcodes $\mathcal{C}_i$ in the sequence are (q, ϵ, δ)-LDCs.*

Note that if $\dim(\mathcal{C}) \geq \omega(\log(n))$ then $\mathcal{C}$ contains sequences of subcodes of length $\omega(\log(n))$. Now we prove two simple claims that will be useful in the proof of Theorem 4.6.

Claim 4.7. *Let $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code. Moreover, let a linear code $\mathcal{C}' \subsetneq \mathcal{C}$ be a $(q, \epsilon, \delta(\mathcal{C}))$-uniform LTC. Then $|\mathcal{C}^\perp_{\leq q}| \leq (1 - \epsilon)|\mathcal{C}'^\perp_{\leq q}|$.*

Proof. Let $\mathcal{D}$ be the uniform distribution over $\mathcal{C}'^\perp_{\leq q}$. We know that $\mathcal{C}^\perp_{\leq q} \subset \mathcal{C}'^\perp_{\leq q}$. Consider any $w \in \mathcal{C} \setminus \mathcal{C}'$ (note that $\mathcal{C} \setminus \mathcal{C}' \neq \emptyset$). Since $\mathcal{C}'$ is a $(q, \epsilon, \delta(\mathcal{C}))$-uniform LTC and $\delta(w, \mathcal{C}') \geq \delta(\mathcal{C})$ it holds that $\mathop{\mathbf{Pr}}\limits_{u \sim \mathcal{D}}[\langle u, w \rangle \neq 0] \geq \epsilon$. Notice that if for $u \in \mathcal{C}'^\perp$ it holds that $\langle u, w \rangle \neq 0$ then $u \notin \mathcal{C}^\perp$. So, there are at least $\epsilon|\mathcal{C}'^\perp_{\leq q}|$ words in $\mathcal{C}'^\perp_{\leq q}$ that are not in $\mathcal{C}^\perp_{\leq q}$. Thus we have $|\mathcal{C}^\perp_{\leq q}| \leq (1 - \epsilon)|\mathcal{C}'^\perp_{\leq q}|$. □

Claim 4.8. *Let $q', \epsilon' > 0$ be constants. Let ℓ be the minimal integer such that $\ell \geq \frac{q' \log n}{\epsilon'}$ and $\mathcal{C} \subseteq \mathbb{F}^n$ be a linear code such that $\dim(\mathcal{C}) > \frac{q' \log n}{\epsilon'}$. Then at least one of the codes in the sequence of the linear subcodes $\mathcal{C}_1 \subsetneq \mathcal{C}_2 \subsetneq \cdots \subsetneq \mathcal{C}_\ell = \mathcal{C}$ is not $(q', \epsilon', \delta(\mathcal{C})/2)$-uniform LTC.*

Proof. Note that for all $i \in [\ell - 1]$ we have $\dim(\mathcal{C}_i) < \dim(\mathcal{C}_{i+1})$. Assume that for all $i \in [\ell]$, $\mathcal{C}_i$ is a $(q', \epsilon', \delta(\mathcal{C})/2)$-uniform LTC. If $(\mathcal{C}_\ell)^\perp_{\leq q} = \emptyset$ then for any word $w \in \mathbb{F}^n$ such that $|\operatorname{supp}(w)| = \delta(\mathcal{C})/2$ (i.e., $\delta(w, \mathcal{C}_\ell) \geq \delta(\mathcal{C})/2$) it holds that $w \perp (\mathcal{C}_\ell)^\perp_{\leq q}$. Contradiction. We conclude that $|(\mathcal{C}_\ell)^\perp_{\leq q'}| \geq 1$.

Claim 4.7 implies that for all $i \in [\ell - 1]$ we have that $|(\mathcal{C}_{i+1})^\perp_{\leq q'}| \leq (1 - \epsilon') \cdot |(\mathcal{C}_i)^\perp_{\leq q}|$. Then it holds that

$$|(\mathcal{C}_\ell)^\perp_{\leq q'}| \leq (1 - \epsilon')^\ell \cdot |(\mathcal{C}_1)^\perp_{\leq q'}| < e^{-\epsilon' \ell} \cdot n^{q'} \leq 1,$$

where $\ell \geq \frac{q' \log n}{\epsilon'}$. We conclude that $(\mathcal{C}_\ell)^\perp_{\leq q'} = \emptyset$. Contradiction. □

Proof of Theorem 4.6. Assume $\mathcal{C}_1 \subsetneq \mathcal{C}_2 \subsetneq \ldots \subsetneq \mathcal{C}_\ell = \mathcal{C}$, where $\ell \geq (q' \log(n))/\epsilon'$. Claim 4.8 says that at least one of the codes in the sequence is not $(q', \epsilon', \delta(\mathcal{C})/2)$-uniform LTC. Corollary 4.4 implies that for all $i \in [\ell]$ the code $\mathcal{C}_i$ is a (q, ϵ, δ)-LDC. □

Acknowledgements.

We thank Eli Ben-Sasson, who refused to co-author this paper for many valuable discussions and ideas. We also thank Anna Gal for asking the question about the relation between LTCs and LDCs. We would like to thank Ronny Roth and Partha Mukhopadhyay for helpful discussions. We thank the anonymous referees for valuable comments on an earlier version of this article. Finally, we would like to thank Klim Efremenko and Eli Ben-Sasson for their helpful comments on Claim 4.2.

References

[1] L. Babai, A. Shpilka, and D. Stefankovic, "Locally testable cyclic codes," in *Proceedings: 44th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2003, 11–14 October 2003, Cambridge, Massachusetts*, IEEE, Ed. IEEE Computer Society Press, 2003, pp. 116–125.

[2] E. Ben-Sasson, O. Goldreich, and M. Sudan, "Bounds on 2-query codeword testing," in *RANDOM-APPROX*, ser. Lecture Notes in Computer Science, vol. 2764. Springer, 2003, pp. 216–227. [Online]. Available: http://springerlink.metapress.com/openurl.asp?genre=article&amp;issn=0302-9743&amp;volume=2764&amp;spage=216

[3] E. Ben-Sasson, V. Guruswami, T. Kaufman, M. Sudan, and M. Viderman, "Locally testable codes require redundant testers," in *IEEE Conference on Computational Complexity*. IEEE Computer Society, 2009, pp. 52–61. [Online]. Available: http://doi.ieeecomputersociety.org/10.1109/CCC.2009.6

[4] E. Ben-Sasson, P. Harsha, and S. Raskhodnikova, "Some 3CNF properties are hard to test," *SIAM Journal on Computing*, vol. 35, no. 1, pp. 1–21, 2005. [Online]. Available: http://epubs.siam.org/SICOMP/volume-35/art_44544.html

[5] E. Ben-Sasson and M. Sudan, "Simple PCPs with poly-log rate and query complexity," in *STOC*. ACM, 2005, pp. 266–275. [Online]. Available: http://doi.acm.org/10.1145/1060590.1060631

[6] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan, "Private information retrieval," *JACM: Journal of the ACM*, vol. 45, 1998.

[7] I. Dinur, "The PCP theorem by gap amplification," *Journal of the ACM*, vol. 54, no. 3, pp. 12:1–12:44, Jun. 2007.

[8] Z. Dvir and A. Shpilka, "Locally decodable codes with two queries and polynomial identity testing for depth 3 circuits," *SIAM J. Comput*, vol. 36, no. 5, pp. 1404–1434, 2007. [Online]. Available: http://dx.doi.org/10.1137/05063605X

[9] K. Efremenko, "3-query locally decodable codes of subexponential length," in *STOC*, M. Mitzenmacher, Ed. ACM, 2009, pp. 39–44. [Online]. Available: http://doi.acm.org/10.1145/1536414.1536422

[10] O. Goldreich, "Short locally testable codes and proofs (survey)," *Electronic Colloquium on Computational Complexity (ECCC)*, no. 014, 2005. [Online]. Available: http://eccc.hpi-web.de/eccc-reports/2005/TR05-014/index.html

[11] O. Goldreich, H. J. Karloff, L. J. Schulman, and L. Trevisan, "Lower bounds for linear locally decodable codes and private information retrieval," *Computational Complexity*, vol. 15, no. 3, pp. 263–296, 2006. [Online]. Available: http://dx.doi.org/10.1007/s00037-006-0216-3

[12] O. Goldreich and M. Sudan, "Locally testable codes and PCPs of almost-linear length," *Journal of the ACM*, vol. 53, no. 4, pp. 558–655, Jul. 2006.

[13] P. Gopalan, "A note on efremenko's locally decodable codes," *Electronic Colloquium on Computational Complexity (ECCC)*, no. 069, 2009. [Online]. Available: http://www.eccc.uni-trier.de/report/2009/069/

[14] E. Grigorescu, T. Kaufman, and M. Sudan, "2-transitivity is insufficient for local testability," in *IEEE Conference on Computational Complexity*. IEEE Computer Society, 2008, pp. 259–267. [Online]. Available: http://doi.ieeecomputersociety.org/10.1109/CCC.2008.31

[15] V. Guruswami, “On 2-query codeword testing with near-perfect completeness,” in *ISAAC*, ser. Lecture Notes in Computer Science, vol. 4288. Springer, 2006, pp. 267–276. [Online]. Available: http://dx.doi.org/10.1007/11940128_28

[16] J. Katz and L. Trevisan, “On the efficiency of local decoding procedures for error-correcting codes,” in *STOC*, 2000, pp. 80–86. [Online]. Available: http://doi.acm.org/10.1145/335305.335315

[17] T. Kaufman and M. Sudan, “Sparse random linear codes are locally decodable and testable,” in *FOCS*. IEEE Computer Society, 2007, pp. 590–600. [Online]. Available: http://doi.ieeecomputersociety.org/10.1109/FOCS.2007.65

[18] T. Kaufman and M. Sudan, “Algebraic property testing: the role of invariance,” in *STOC*. ACM, 2008, pp. 403–412. [Online]. Available: http://doi.acm.org/10.1145/1374376.1374434

[19] T. Kaufman and A. Wigderson, “Symmetric ldpc and local testing,” in *ICS*, 2010.

[20] I. Kerenidis and R. de Wolf, “Exponential lower bound for 2-query locally decodable codes,” *Electronic Colloquium on Computational Complexity (ECCC)*, no. 059, 2002. [Online]. Available: http://eccc.hpi-web.de/eccc-reports/2002/TR02-059/index.html

[21] O. Meir, “Combinatorial construction of locally testable codes,” in *STOC*. ACM, 2008, pp. 285–294. [Online]. Available: http://doi.acm.org/10.1145/1374376.1374419

[22] K. Obata, “Optimal lower bounds for 2-query locally decodable linear codes,” in *RANDOM*, ser. Lecture Notes in Computer Science, J. D. P. Rolim and S. P. Vadhan, Eds., vol. 2483. Springer, 2002, pp. 39–50. [Online]. Available: http://link.springer.de/link/service/series/0558/bibs/2483/24830039.htm

[23] Shiowattana and Lokam, “An optimal lower bound for 2-query locally decodable linear codes,” *IPL: Information Processing Letters*, vol. 97, 2006.

[24] L. Trevisan, “Some applications of coding theory in computational complexity,” Sep. 23 2004. [Online]. Available: http://arxiv.org/abs/cs/0409044

[25] Woodruff, “New lower bounds for general locally decodable codes,” in *ECCCTR: Electronic Colloquium on Computational Complexity, technical reports*, 2007.

[26] S. Yekhanin, “Towards 3-query locally decodable codes of subexponential length,” *J. ACM*, vol. 55, no. 1, 2008. [Online]. Available: http://doi.acm.org/10.1145/1326554.1326555

Appendix

For the sake of completeness we state the main result of Grigorescu et al. [14], who showed a 2-transitive code with dual codewords of weight 8 which is not LTC. For additional details see [14].

Let Tr be a trace function from $\mathbb{F}_{2^s}$ to $\mathbb{F}_2$. For positive integers $k < s$ let

$$\mathcal{F}^*_{k,s} = \left\{ f : \mathbb{F}_{2^s} \mapsto \mathbb{F}_2 \mid \exists \beta, \beta_0, \ldots, \beta_k \in \mathbb{F}_{2^n} \text{ s.t. } f(x) = \mathrm{Tr}(\beta + \beta_0 x + \sum_{i=1}^{k} \beta_i x^{2^i+1}) \right\}.$$

Theorem 4.9 ([14]). *Assume $k = \omega(1)$ and $s > 2k + 1$. Let $C = \mathcal{F}^*_{k,s}$ be a linear code. Then C is 2-transitive and $C^{\perp}_{8} \neq \emptyset$, but C is not $(q, \epsilon, 1/7)$-LTC for all constants $q, \epsilon > 0$.*

ECCC ISSN 1433-8092

http://eccc.hpi-web.de