High-Entropy Dual Functions and Locally Decodable Codes

Jop Briët
CWI, Science Park 123, 1098 XG Amsterdam, The Netherlands
j.briet@cwi.nl

Farrokh Labib
CWI, Science Park 123, 1098 XG Amsterdam, The Netherlands
labib@cwi.nl

Abstract

Locally decodable codes (LDCs) allow any single encoded message symbol to be retrieved from a codeword with good probability by reading only a tiny number of codeword symbols, even if the codeword is partially corrupted. LDCs have surprisingly many applications in computer science and mathematics (we refer to [13, 10] for extensive surveys). But despite their ubiquity, they are poorly understood. Of particular interest is the tradeoff between the codeword length N as a function of message length k when the *query complexity*—the number of probed codeword symbols—and alphabet size are constant. The Hadamard code is a 2-query LDC of length $N = 2^{O(k)}$ and this length is optimal in the 2-query regime [11]. For $q \geq 3$, near-exponential gaps persist between the best-known upper and lower bounds. The family of Reed-Muller codes, which generalize the Hadamard code, were for a long time the best-known examples, giving q-query LDCs of length $\exp(O(k^{1/(q-1)}))$, until breakthrough constructions of *matching vector* LDCs of Yekhanin and Efremenko [12, 6].

In contrast with other combinatorial objects such as expander graphs, the probabilistic method has so far not been successfully used to beat the best explicit LDC constructions. In [3], a probabilistic framework was given that could in principle yield best-possible LDCs, albeit non-constructively. A special instance of this framework connects LDCs with a probabilistic version of Szemerédi's theorem. The setup for this is as follows: For a finite abelian group G of size $N = |G|$, let $D \subseteq G$ be a random subset where each element is present with probability ρ independently of all others. For $k \geq 3$ and $\varepsilon \in (0, 1)$, let E be the event that every subset $A \subseteq G$ of size $|A| \geq \varepsilon|G|$ contains a proper k-term arithmetic progression with common difference in D. For fixed $\varepsilon > 0$ and sufficiently large N, it is an open problem to determine the smallest value of ρ — denoted ρ_k — such that $\Pr[E] \geq \frac{1}{2}$. In [3] it is shown that there exist k-query LDCs of message length $\Omega(\rho_k N)$ and codeword length $O(N)$. As such, Szemerédi's theorem with random differences, in particular lower bounds on ρ_k, can be used to show the existence of LDCs. Conversely, this connection indirectly implies the best-known upper bounds on ρ_k for all $k \geq 3$ [8, 4]. However, a conjecture from [9] states that over $\mathbb{Z}_N$ we have $\rho_k \leq O_k(N^{-1} \log N)$ for all k, which would be best-possible. Truth of this conjecture would imply that over this group, Szemerédi's theorem with random differences cannot give LDCs better than the Hadamard code. For finite fields, Altman [1] showed that this is false. In particular, over $\mathbb{F}_p^n$ for p odd, he proved that $\rho_3 \geq \Omega(p^{-n}\, n^2)$; generally, $\rho_k \geq \Omega(p^{-n}\, n^{k-1})$ holds when $p \geq k + 1$ [2]. In turn, these bounds are conjectured to be optimal for the finite-field setting, which would imply that over finite fields, Szemerédi's theorem with random differences cannot give LDCs better than Reed-Muller codes.

The finite-field conjecture is motivated mainly by the possibility that so-called *dual functions* can be approximated well by *polynomial phases*, functions of the form $e^{2\pi i P(x)/p}$ where P is a multivariate polynomial over $\mathbb{F}_p$. We show that this is false. Using Yekhanin's matching-vector-code construction, we give dual functions of order k over $\mathbb{F}_p^n$ that cannot be approximated in L_∞-distance by polynomial phases of degree $k - 1$. This answers in the negative a natural finite-field analog of a problem of Frantzikinakis over $\mathbb{N}$ [7, Problem 1].

2012 ACM Subject Classification Theory of computation

Keywords and phrases Higher-order Fourier analysis, dual functions, finite fields, additive combinatorics, coding theory

12th Innovations in Theoretical Computer Science Conference (ITCS 2021).
Editor: James R. Lee; Article No. 76; pp. 76:1–76:2
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Digital Object Identifier 10.4230/LIPIcs.ITCS.2021.76

Category Extended Abstract

Related Version This is an extended abstract of [5], `https://arxiv.org/abs/2010.14956`.

Funding *Jop Briët*: Supported by the Gravitation grant NETWORKS-024.002.003 from the Dutch Research Council (NWO)
Farrokh Labib: Supported by the Gravitation grant NETWORKS-024.002.003 from the Dutch Research Council (NWO)

References

1 Daniel Altman. On Szemerédi's theorem with differences from a random set. *Acta Arith*, 195:97–108, 2020. doi:10.4064/aa190531-25-10.

2 Jop Briët. Subspaces of tensors with high analytic rank. *Online Journal of Analytic Combinatorics*, 2020. To appear. Available at arXiv: 1908.04169.

3 Jop Briët, Zeev Dvir, and Sivakanth Gopi. Outlaw distributions and locally decodable codes. *Theory of Computing*, 15(12):1–24, 2019. Preliminary version in ITCS'17. doi:10.4086/toc.2019.v015a012.

4 Jop Briët and Sivakanth Gopi. Gaussian width bounds with applications to arithmetic progressions in random settings. *International Mathematics Research Notices*, page rny238, 2018. doi:10.1093/imrn/rny238.

5 Jop Briët and Farrokh Labib. High-entropy dual functions over finite fields and locally decodable codes. *arXiv preprint*, 2020. arXiv:2010.14956.

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

7 Nikos Frantzikinakis. Some open problems on multiple ergodic averages. *Bull. Hellenic Math. Soc.*, 60:41–90, 2016. doi:10.1109/tac.2015.2392613.

8 Nikos Frantzikinakis, Emmanuel Lesigne, and Mate Wierdl. Random sequences and pointwise convergence of multiple ergodic averages. *Indiana Univ. Math. J.*, pages 585–617, 2012.

9 Nikos Frantzikinakis, Emmanuel Lesigne, and Mate Wierdl. Random differences in Szemerédi's theorem and related results. *J. Anal. Math.*, 130:91–133, 2016. doi:10.1007/s11854-016-0030-z.

10 Sivakanth Gopi. *Locality in coding theory.* PhD thesis, Princeton University, 2018.

11 Iordanis Kerenidis and Ronald de Wolf. Exponential lower bound for 2-query locally decodable codes via a quantum argument. *J. Comput. System Sci.*, 69(3):395–420, 2004. Earlier version in STOC'03. doi:10.1016/j.jcss.2004.04.007.

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

13 Sergey Yekhanin. Locally decodable codes. *Found. Trends Theor. Comput. Sci.*, 6(3):139–255, 2012. doi:10.1561/0400000030.