Noisy Decoding by Shallow Circuits with Parities: Classical and Quantum (Extended Abstract)

Jop Briët
Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

Harry Buhrman
Centrum Wiskunde & Informatica, Amsterdam, The Netherlands
University of Amsterdam, The Netherlands

Davi Castro-Silva
Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

Niels M. P. Neumann
Centrum Wiskunde & Informatica, Amsterdam, The Netherlands
The Netherlands Organisation for Applied Scientific Research (TNO), Den Haag, The Netherlands

Abstract

We consider the problem of decoding corrupted error correcting codes with $\mathrm{NC}^0[\oplus]$ circuits in the classical and quantum settings. We show that any such classical circuit can correctly recover only a vanishingly small fraction of messages, if the codewords are sent over a noisy channel with positive error rate. Previously this was known only for linear codes with large dual distance, whereas our result applies to any code. By contrast, we give a simple quantum circuit that correctly decodes the Hadamard code with probability $\Omega(\varepsilon^2)$ even if a $(1/2 - \varepsilon)$-fraction of a codeword is adversarially corrupted.

Our classical hardness result is based on an equidistribution phenomenon for multivariate polynomials over a finite field under biased input-distributions. This is proved using a structure-versus-randomness strategy based on a new notion of rank for high-dimensional polynomial maps that may be of independent interest.

Our quantum circuit is inspired by a non-local version of the Bernstein-Vazirani problem, a technique to generate "poor man's cat states" by Watts et al., and a constant-depth quantum circuit for the OR function by Takahashi and Tani.

2012 ACM Subject Classification Mathematics of computing → Coding theory; Theory of computation → Circuit complexity

Keywords and phrases Coding theory, circuit complexity, quantum complexity theory, higher-order Fourier analysis, non-local games

Digital Object Identifier 10.4230/LIPIcs.ITCS.2024.21

Category Extended Abstract

Related Version *Full Version*: https://arxiv.org/abs/2302.02870

Funding This work was supported by the Dutch Research Council (NWO/OCW), as part of the Quantum Software Consortium programme (project number 024.003.037), the NETWORKS programme (grant no. 024.002.003), and the Quantum Delta NL program.

Acknowledgements We thank Richard Cleve for helpful discussions on the proof of Theorem 2, Emanuele Viola for helpful pointers to the literature and comments on an earlier version of this manuscript, Tamar Ziegler for encouragement to explore the high-characteristic setting, as well as anonymous referees for helpful comments.

15th Innovations in Theoretical Computer Science Conference (ITCS 2024).
Editor: Venkatesan Guruswami; Article No. 21; pp. 21:1–21:11
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

Error correcting codes (ECCs), formally introduced in Shannon's celebrated work [28], protect digital signals from noise. An ECC is a map $E : \Sigma^k \to \Sigma^n$, for a finite alphabet Σ and positive integers $n \geq k$, with the property that any message $x \in \Sigma^k$ can be decoded from the codeword $E(x)$ even if the codeword is partially corrupted. If too many errors occur, however, recovering the original message may become impossible. In such cases one can instead resort to *list decoding*, an influential idea proposed in seminal works of Elias [10] and Wozencraft [39], which aims to give a small list of messages whose codewords are close to the received (corrupted) codeword. Complexity considerations appear naturally in this context, as encoding and decoding ideally allow for reliable communication with limited computational resources; they also appear because of the fundamental role played by ECCs in computational complexity itself (see e.g., [34] for a survey).

1.1 Error models

In the error model considered by Shannon [28], a codeword is corrupted according to some random process. A natural such process is given by the *symmetric channel*: for each coordinate of the codeword independently, the channel either transmits it unchanged with some probability ρ, or replaces it with a uniformly random element of Σ with probability $1-\rho$. We refer to ρ as the *bias* of the channel.[1] If $Z \in \Sigma^n$ is distributed according to the random outcome of the symmetric channel with bias ρ applied to a codeword $E(x)$, we write $Z \sim \mathcal{N}_\rho\big(E(x)\big)$. In this model the goal is to correctly decode a corrupted codeword with good probability over the noise.

The combinatorial worst-case error model of Hamming [18] instead assumes that the codeword is corrupted arbitrarily on at most some $\delta \in [0, 1)$ fraction of coordinates. We will refer to δ as the *error parameter*. In this setting, the number of errors that can be tolerated depends on the minimal Hamming distance between any pair of distinct codewords, or *minimal distance* of the code, denoted d_E. Since the Hamming ball of diameter $d_E - 1$ around any point $y \in \Sigma^n$ contains at most one codeword, a message can be retrieved if fewer than $d_E/2$ errors have occurred.

If more errors occur, faithful decoding is no longer possible and list decoding enters the picture. For $\delta \in [0, 1)$ and positive integer L, a code is (δ, L)*-list decodable* if for any point $y \in \Sigma^n$, the Hamming ball of radius δn centered around y contains at most L codewords. It is well known that any (δ, L)-list decodable code satisfies $L \geq \Omega(1/\varepsilon^2)$ when $\delta = (1-\varepsilon)(1-|\Sigma|^{-1})$ [17]. If fewer than a δ-fraction of codeword coordinates are corrupted, then a random element from this list will give the correct message with probability at least $1/L$.

1.2 Circuits

A well-studied problem is that of decoding corrupted ECCs by constant-depth circuits with n inputs, k outputs and size $\mathrm{poly}(n)$, for example in the context of black-box hardness amplification [31, 35, 36]. Two classes of such circuits are AC^0, consisting of unbounded-fan-in AND, OR and NOT gates, and the class NC^0, consisting of arbitrary bounded-fan-in gates; without loss of generality, we may assume that the fan-in of any gate in NC^0 is at most two.

[1] In this model, each coordinate is thus corrupted with probability $(1-\rho)(1-|\Sigma|^{-1})$, which is usually referred to as the *error rate*. For our purposes, however, the bias will be a more convenient parameterization.

The extensions of these classes where unbounded-fan-in parity gates are added to the gate sets are denoted by $\mathrm{AC}^0[\oplus]$ and $\mathrm{NC}^0[\oplus]$, respectively. These are proper extensions since parity cannot be computed by AC^0 circuits and NC^0 is a proper subset of AC^0 (see [1]). An important distinction is that the outputs of NC^0 circuits depend on only a constant number of coordinates of the input, whereas the outputs of $\mathrm{NC}^0[\oplus]$ circuits can depend on the whole input. The classes AC^0 and $\mathrm{NC}^0[\oplus]$ are incomparable since $\mathrm{NC}^0[\oplus]$ cannot compute the n-bit AND function; indeed, $\mathrm{NC}^0[\oplus]$ circuits can compute only constant-degree polynomials over $\mathbb{F}_2$ (see Section 3), whereas AND has degree n.

We also consider the quantum counterparts of the above circuit classes, denoted QX, where X is one of the classes discussed above; these classes were first introduced by Moore [22] and Moore and Nilsson [23]. Thus, QNC^0 is the class of constant-depth quantum circuits containing arbitrary one- and two-qubit gates, while $\mathrm{QNC}^0[\oplus]$ includes unbounded-fan-in parity gates acting on superpositions. In contrast with their classical analogues, the classes $\mathrm{QNC}^0[\oplus]$ and QAC^0 are known to be equivalent [15, 20, 22].[2]

1.3 Quantum advantage

The above-mentioned classes of quantum circuits recently enjoyed renewed interest in the context of provable separations between quantum and classical complexity classes.

One of the principal challenges in quantum computing is to determine for which types of problems quantum computers offer a significant advantage over classical ones. Celebrated examples of practical importance, such as Shor's algorithm for integer factoring [29], require quantum computers of a vastly larger scale than currently available. Moreover, formally proving classical hardness of factoring appears to be beyond the scope of currently-available techniques. Constant-depth circuits form an attractive computational model, as they will likely be easier to implement in practice and, from the perspective of complexity theory, provide one of the few settings currently amenable to provable lower bounds.

A recent series of works, starting with a breakthrough of Bravyi, Gosset and König [4], considered the relative power of *shallow quantum circuits*. For instance:

- The 2D-Hidden Linear Function problem can be solved exactly in QNC^0 while any AC^0 circuit succeeds with exponentially small probability under a certain input distribution [2]; this strengthened the main result of [4] showing that this problem separates QNC^0 from NC^0 in the worst case.
- The Relaxed Parity Halving problem can be solved exactly in QNC^0 while any AC^0 circuit succeeds with probability at most $\frac{1}{2} + \exp(-n^{\varepsilon})$ under the uniform distribution [2].
- The Parallel Parity Bending problem can be solved with probability $1 - o(1)$ by a $\mathrm{QNC}^0/\mathsf{qpoly}$ circuit while any $\mathrm{AC}^0[\oplus]/\mathsf{rpoly}$ succeeds with probability at most $O(n^{-\varepsilon})$ [2].
- The problem of simulating correlations obtained from measuring graph states QNC^0 and NC^0, even in the average-case [11].
- The 1D-Magic Square problem separates noisy QNC^0 circuits from NC^0 [5].

Similar separations based on other relational and sampling-based problems were proven in [9, 16, 38]. A common feature of all these problems is that they were specifically designed to prove separations between shallow quantum and classical circuits.

[2] In addition, the works of Moore showed that these classes are all equivalent to $\mathrm{QAC}^0[q]$, the class QAC^0 with additional modulo-q gates. For an integer $q > 1$, a modulo-q gate evaluates to 1 if the sum of its inputs equals 0 mod q and evaluates to 0 otherwise. Classically, the classes $\mathrm{AC}^0[p]$ and $\mathrm{AC}^0[q]$ are incomparable if p and q are powers of distinct primes [26, 30].

We instead consider the problem of decoding a corrupted error-correcting code, which arises naturally in computer science. This problem is well studied in the context of classical complexity theory, where shallow circuits endowed with parity gates are also considered; see Sections 2.2 and 2.3 for further discussion.

1.4 The Hadamard code

A basic but important example of an ECC is the *Hadamard code*, which encodes k-bit messages into codewords of length $n = 2^k$ and is given by the $\mathbb{F}_2$-linear map $H(x) = (\langle x, y\rangle)_{y\in\mathbb{F}_2^k}$, where $\langle x, y\rangle = y^{\mathsf{T}}x$. This code has minimal distance $n/2$ and is $(1/2 - \varepsilon, O(1/\varepsilon^2))$-list decodable for any $\varepsilon \in (0, 1/2]$, which is known to be optimal for any code [17].

Under the symmetric channel, the Chernoff bound implies that unique decoding of the Hadamard code is possible with high probability for any constant bias $\rho > 0$.[3] This is due to the fact that, with high probability, the Hamming ball of radius $(1/4 - \rho/4)n$ around a corrupted version of a codeword C contains no other codewords than C itself.

For the worst-case Hamming model, Goldreich and Levin [12] famously gave an efficient list decoding algorithm for the Hadamard code that runs in time $\mathrm{poly}(k, 1/\varepsilon)$, for error parameter $\delta = 1/2 - \varepsilon$. For fixed $\varepsilon > 0$, their algorithm gives a probabilistic AC^0 circuit that, on input length n, correctly returns the original message with probability $\Omega(1)$.

2 Our results

Here we consider the following problem. Let $E : \mathbb{F}_2^k \to \mathbb{F}_2^n$ be a (binary) error correcting code. Given a map $\phi : \mathbb{F}_2^n \to \mathbb{F}_2^k$ representing some decoding procedure, we wish to bound the probability of correct message retrieval:

$$\Pr\big[\phi\big(E(x) + Z\big) = x\big], \tag{1}$$

where $x \in \mathbb{F}_2^k$ is some message and $Z \in \mathbb{F}_2^n$ is an error string. We consider two scenarios, one classical and one quantum.

2.1 Classical setting

In the first scenario, ϕ represents an $\mathrm{NC}^0[\oplus]$ circuit, x is uniformly distributed and $Z \sim \mathcal{N}_\rho(0)$, so that $E(x) + Z$ is a random codeword corrupted according to the binary symmetric channel with bias ρ. Our main result in this setting says that (1) tends to zero, for any $\rho \in [0, 1)$ and any code:

▶ **Theorem 1** (Impossibility of decoding by $\mathrm{NC}^0[\oplus]$). *For any $\rho \in [0, 1)$, $d \in \mathbb{N}$ and $\varepsilon \in (0, 1]$, there is a $k_0 = k_0(d, \rho, \varepsilon) \in \mathbb{N}$ such that the following holds. Let $k \geq k_0$ and n be positive integers, $E : \mathbb{F}_2^k \to \mathbb{F}_2^n$ be any map and $\phi : \mathbb{F}_2^n \to \mathbb{F}_2^k$ be a map computable by an $\mathrm{NC}^0[\oplus]$ circuit of depth at most d. Then, for a uniformly distributed $x \in \mathbb{F}_2^k$ and $Z \sim \mathcal{N}_\rho(0)$, we have that*

$$\Pr\big[\phi\big(E(x) + Z\big) = x\big] < \varepsilon.$$

In particular, this theorem shows that no $\mathrm{NC}^0[\oplus]$ circuit can correctly decode more than an ε-fraction of codewords with probability higher than ε over the noise distribution, if the messages are long enough depending on ε, the error rate $(1 - \rho)/2 > 0$ and the

[3] This even holds for any code over a large enough alphabet, as shown in [27].

depth of the circuit. As a consequence of Yao's minimax principle [41] and the Chernoff bound, it follows that any probabilistic $\mathrm{NC}^0[\oplus]$ circuit will also fail (with high probability) to correctly decode any binary ECC in the worst-case Hamming model, for any constant error parameter $\delta \in (0, 1/2]$.

We note that the decay we obtain on the probability (1) of correct message retrieval as a function of the message length is extremely slow, making Theorem 1 a qualitative result rather than quantitative. Nevertheless, we conjecture that the true decay of this probability is exponential in the message length k; this would clearly be optimal, as can be seen by taking a constant map ϕ which always returns some fixed message. In Section 7 of the full version [7] we provide some evidence to support this conjecture.

2.2 Quantum setting

In the second scenario, we consider the worst-case Hamming model with constant-depth quantum circuits. Our main result in this setting is an explicit $\mathrm{QNC}^0[\oplus]$ circuit capable of decoding the Hadamard code.

▶ **Theorem 2** (Decoding Hadamard with $\mathrm{QNC}^0[\oplus]$). *There is a family of* $\mathrm{QNC}^0[\oplus]$ *circuits* $(\mathcal{C}_n)_{n\in\mathbb{N}}$ *such that the following holds. Let* $k \in \mathbb{N}$, $n = 2^k$ *and* $\varepsilon \in (0, 1/2]$. *Then, for any* $y \in \mathbb{F}_2^n$ *and any* $x \in \mathbb{F}_2^k$ *satisfying* $d\big(y, H(x)\big) \le (\frac{1}{2} - \varepsilon)n$, *on input* y *the circuit* $\mathcal{C}_n$ *returns* x *with probability* $\Omega(\varepsilon^2)$.

We note that the bound $\Omega(\varepsilon^2)$ obtained in the theorem is optimal, since in general there can be $\Theta(\varepsilon^{-2})$ messages $x \in \mathbb{F}_2^k$ satisfying $d\big(y, H(x)\big) \le (\frac{1}{2} - \varepsilon)n$. This bound is non-trivial only when $\varepsilon = \Omega(1/\sqrt{n})$, as there are n possible messages.

As a simple corollary of Theorem 2, we obtain a similar result for the problem of list decoding the Hadamard code.

▶ **Corollary 3.** *There is a family of* $\mathrm{QNC}^0[\oplus]$ *circuits* $(\mathcal{C}_n)_{n\in\mathbb{N}}$ *such that the following holds. Let* $k \in \mathbb{N}$, $n = 2^k$ *and* $\varepsilon \in [1/\sqrt{n}, 1/2]$. *Then, on any input* $y \in \mathbb{F}_2^n$, *with probability* $1 - \varepsilon$ *the circuit* $\mathcal{C}_n$ *returns a list* $L(y)$ *of size* $O(\varepsilon^{-2}\log(1/\varepsilon))$ *which contains every* $x \in \mathbb{F}_2^k$ *with* $d\big(y, H(x)\big) \le (\frac{1}{2} - \varepsilon)n$.

Proof. For a large enough constant $C > 0$, consider $C\varepsilon^{-2}\log(1/\varepsilon)$ parallel instances of the circuit from Theorem 2. This gives a list $L(y)$ of the claimed size such that any message $x \in \mathbb{F}_2^k$ satisfying $d\big(y, H(x)\big) \le (\frac{1}{2} - \varepsilon)n$ appears in $L(y)$ with probability at least $1 - \varepsilon^3$. Since there are at most $O(1/\varepsilon^2)$ such messages, it follows from the union bound that with probability at least $1 - O(\varepsilon)$ every such message appears in $L(y)$. ◀

▶ Remark 4. Note that the circuits obtained in this corollary also output several messages whose codewords differ from the input y in more than $(\frac{1}{2} - \varepsilon)n$ coordinates; this differs from the usual notion of the list decoding problem, which aims to output a list of all messages $x \in \mathbb{F}_2^k$ with $d\big(y, H(x)\big) \le (\frac{1}{2} - \varepsilon)n$ and none other. One can also solve the usual list decoding problem for the Hadamard code using $\mathrm{QNC}^0[\oplus]$ circuits, by making use of MAJORITY gates (and more general threshold gates) to prune the obtained list. We omit the details, as they are not so relevant for us.

As a consequence of Theorem 1 and Theorem 2, we conclude that the problem of list decoding the Hadamard code separates the complexity classes $\mathrm{NC}^0[\oplus]$ and $\mathrm{QNC}^0[\oplus]$; this holds for any positive error parameter $\delta > 0$. The task of proving quantum advantage for a natural problem such as list decoding was the original motivation for the present work.

In the high-error regime where the parameter δ approaches the information-theoretic limit of 1/2 (which is relevant for hardness amplification), a stronger separation follows by combining Theorem 2 with a result of Sudan showing hardness of noisy decoding by $\mathrm{AC}^0[\oplus]$ circuits (see Corollary 7 below).[4] To state this separation theorem precisely, we consider the following problem:

List-Hadamard problem. Let $\varepsilon : \mathbb{N} \to (0,1]$ be a function. For each dyadic number $n = 2^k$ we define the problem $\mathrm{LH}_n(\varepsilon)$ as follows: given $y \in \mathbb{F}_2^n$, output a list of at most $n/4$ elements in $\mathbb{F}_2^k$ containing every $x \in \mathbb{F}_2^k$ satisfying $d\big(y, H(x)\big) \leq \big(\frac{1}{2} - \varepsilon(n)\big)n$.

The most general form of our quantum advantage result is given by the following theorem:

▶ **Theorem 5** (Quantum-vs-classical separation). *For any constant $\delta \in (0, \frac{1}{2})$, list decoding the Hadamard code with error parameter δ separates $\mathrm{QNC}^0[\oplus]$ from $\mathrm{NC}^0[\oplus]$. Moreover, for any $(\log n)/\sqrt{n} \leq \varepsilon(n) \leq 1/(\log n)^{\omega(1)}$, the list-Hadamard problem $\mathrm{LH}_n(\varepsilon)$ separates $\mathrm{QNC}^0[\oplus]$ from $\mathrm{AC}^0[\oplus]$.*

2.3 Related results and discussion

Both the problem of decoding corrupted ECCs and the problem of proving quantum-versus-classical separations of complexity classes are well studied, and there are several results in the literature related to the results presented here.

The main strength of our Theorem 1 is that it holds for any code and for any positive error rate. Complementary results are known for restricted classes of codes, and also for when the error rate tends to 1/2. We will now expand on some of these results.

A code $E : \mathbb{F}_2^k \to \mathbb{F}_2^n$ is *t-wise independent* if, for any t-subset of coordinates $S \subseteq [n]$ and a uniformly random $X \in \mathbb{F}_2^k$, the restriction $E(X)_{|S}$ is uniformly distributed over $\mathbb{F}_2^S$. Many codes have this property; for instance, the dual code of a linear code of distance d is $(d-1)$-wise independent. Under the same noise model considered here, Lee and Viola [21], using earlier work of Viola [37], showed that $\mathrm{NC}^0[\oplus]$ circuits cannot distinguish a corrupted uniformly random codeword of an $\omega(1)$-wise independent linear code from a uniformly random element of $\mathbb{F}_2^n$. Note that this problem is formally easier than (list) decoding.

Their result does not cover the Hadamard code, however, as it is not even 3-wise independent. Indeed, the Hadamard code is also easy to distinguish, as it contains the sub-code $(x_1, x_2, x_1 + x_2)$. Since the parity of these three bits is always zero, the parity under noise is biased towards zero and therefore easily distinguished from the parity of a random string.

In the very-high-error regime where the error rate approaches the information-theoretic limit of 1/2 (which is relevant for hardness amplification), stronger results are also known. For instance, Sudan (see [36, Section 6.2]) showed that list decoding with error parameter $1/2 - \varepsilon$ requires probabilistic $\mathrm{AC}^0[\oplus]$ circuits to have size $\exp(\mathrm{poly}(1/\varepsilon))$. Below we state his result when restricted to the Hadamard code, which is done for concreteness and better clarity; as can be easily seen from its proof, one could instead consider any other ECC.

▶ **Theorem 6** (MAJORITY from list-Hadamard). *Let $\mathcal{C}$ be a probabilistic circuit that solves the list-Hadamard problem $\mathrm{LH}_n(\varepsilon)$ with probability at least $3/4$. There exists a (deterministic) oracle AC^0 circuit $\mathcal{D}$ of size $\mathrm{poly}(n, 1/\varepsilon)$ which, when given oracle access to $\mathcal{C}$ and the ability to fix its random bits, computes MAJORITY on $\Omega(1/\varepsilon)$ bits.*

[4] The same separation of complexity classes can also be obtained by combining other previously-known results; see Section 2.3 for a discussion.

This result can be readily deduced from Sudan's arguments exposed in [36, Section 6.2]. As a corollary, the circuit lower bound for MAJORITY due to Razborov [26] and Smolensky [30] gives the following (known) hardness result for list decoding the Hadamard code.

▶ **Corollary 7** (Hardness of list-Hadamard). *If $\varepsilon(n) \leq 1/(\log n)^{\omega(1)}$, then the list-Hadamard problem $\mathrm{LH}_n(\varepsilon)$ cannot be solved by a probabilistic $\mathrm{AC}^0[\oplus]$ circuit with probability $\Omega(1)$.*

Combining this corollary with our $\mathrm{QNC}^0[\oplus]$ circuits for list-Hadamard given in Corollary 3, we obtain the second separation of complexity classes stated in Theorem 5.

The existence of the quantum circuits of Theorem 2 and Corollary 3 also follows from the Goldreich-Levin algorithm and the surprising fact that MAJORITY can be computed by a $\mathrm{QNC}^0[\oplus]$ circuit [20, 32].[5] However, whereas the circuit based on the Golreich-Levin algorithm depends on the error parameter ε (which influences the size of the MAJORITY gates), our quantum circuit is constructed independently of ε. Moreover, a key enabling sub-routine in the Høyer-Špalek circuit for MAJORITY [20] is the powerful quantum fan-out gate (see below for further details). In our circuit for Corollary 3, we construct this gate explicitly using only classical parity gates and single- and two-qubit gates; these gates are native to many quantum architectures and as such, may give an easier way to implement quantum fan-out. In the opposite direction, one can use the ideas behind the proof of Theorem 6 to show that our quantum circuit from Corollary 3 also gives a $\mathrm{QNC}^0[\oplus]$ circuit for MAJORITY, albeit not exact (the details are given in the full version [7, Appendix A.2]).

Finally, quantum list-decoding of classical error correcting codes was also studied in [40]. The model considered there consists of a faulty quantum circuit that implements the encoding, which differs from our setting.

2.4 Future directions

We conjecture that the correct rate of decay in Theorem 1 is exponential in the message length, as suggested by the results we obtain in the high-characteristic setting (exposed in the full version [7, Section 7]). This raises several intriguing questions related to notions of rank for tensors and polynomial maps [6].

Our results leave open the problem of decoding more general classes of error correcting codes by shallow quantum circuits, or by efficient quantum algorithms. A particular class of interest consists of low-degree Reed-Muller codes, which generalize the Hadamard code.

A related question is if the problem of distinguishing random corrupted codewords of some code from uniformly random strings gives similar separations of the quantum and classical complexity classes considered here. For example, Lee and Viola [21] proved $\mathrm{NC}^0[\oplus]$-hardness of distinguishing $\omega(1)$-wise independent codes. Is there a $\mathrm{QNC}^0[\oplus]$ distinguisher for such a code?

3 Techniques

To establish our main results, we use techniques from two different areas. Broadly speaking, Theorem 1 builds on ideas from higher-order Fourier analysis [33, 19], while Theorem 2 (unsurprisingly) uses ideas from quantum computing [24].

[5] The above-mentioned classical hardness of MAJORITY thus also implies a separation between $\mathrm{AC}^0[\oplus]$ and $\mathrm{QNC}^0[\oplus]$, showing that despite its simplicity, the latter class of quantum circuits is remarkably powerful.

3.1 Polynomial equidistribution

The proof of Theorem 1 uses the basic observation that any function $\mathbb{F}_2^n \to \mathbb{F}_2^k$ that is computable by an $\mathrm{NC}^0[\oplus]$ circuit can be given by a collection of k constant-degree polynomials over $\mathbb{F}_2$ in n variables. Indeed, any gate with fan-in d implements a function $\mathbb{F}_2^d \to \mathbb{F}_2$ and any such function can be represented by a d-variable polynomial of total degree at most d. Degree is multiplicative under composition and composition occurs only between different layers of the circuit. Since the parities amount to addition in $\mathbb{F}_2$ and NC^0 circuits have constant depth, the total degree of the output is bounded.

We will therefore study the distribution of polynomial maps under biased input distributions. We will do so in a slightly more general setting over arbitrary finite fields of prime order.[6] For a prime p, let $\mathbb{F}_p$ denote the finite field with p elements. For $\rho \in [0, 1]$, an $\mathbb{F}_p$-valued random variable Z is *ρ-biased* if with probability ρ it equals 0 and with probability $1 - \rho$ it is uniformly distributed over $\mathbb{F}_p$. Note that this corresponds to the noise $\mathcal{N}_\rho(0)$ added by the symmetric channel when the alphabet is $\mathbb{F}_p$.

A mapping $\phi : \mathbb{F}_p^n \to \mathbb{F}_p^k$ is a *polynomial map* if there exist polynomials $f_1, \ldots, f_k \in \mathbb{F}_p[x_1, \ldots, x_n]$ such that $\phi = (f_1, \ldots, f_k)$. The degree of ϕ is the maximal degree among the f_i. To prove Theorem 1, it thus suffices to prove the following result.

▶ **Theorem 8** (Impossibility of decoding by polynomial maps). *For any $d \in \mathbb{N}$ and $\rho, \varepsilon \in (0, 1)$ there exists an integer $k_0 = k_0(p, d, \rho, \varepsilon)$ such that the following holds. Let $k \geq k_0$ and n be integers, $\phi : \mathbb{F}_p^n \to \mathbb{F}_p^k$ be a polynomial map of degree at most d and $E : \mathbb{F}_p^k \to \mathbb{F}_p^n$ be an arbitrary function. Then*

$$\Pr_{x \in \mathbb{F}_p^k, Z \sim \mathcal{N}_\rho(0)} \left[\phi\big(E(x) + Z\big) = x\right] \leq \varepsilon.$$

Studying the distribution of polynomial maps in many variables over a finite field falls within the purview of additive combinatorics. In the "unbiased" situation where Z is uniformly distributed there are powerful tools from higher-order Fourier analysis that can be used to study the distribution of $\phi(Z)$. In particular, Green and Tao [14] proved that if ϕ is "regular" (random-like), then $\phi(Z)$ is approximately uniformly distributed over $\mathbb{F}_p^k$. This implies that the probability of the event $\{\phi(E(x) + Z) = x\}$ considered is small for every x. A "regularity-type" lemma proved in [14] shows that one can "force" ϕ to be regular by restricting it to a partition defined by sufficiently many polynomial equations of degree less than the degree of ϕ. However, these techniques cause the size of the polynomial map ϕ considered to blow up considerably, and are only effective if k is an extremely slowly growing function of n.

In order to deal with this issue, and to adapt these results to the case where Z is no longer uniform but biased, we employ a dichotomy often used in additive combinatorics that studies the "pseudorandom" case of regular maps separately from the "structured" case of maps that carry a certain algebraic structure. This is done by defining and studying a new notion of rank for (high-dimensional) polynomial maps, which we call the *analytic rank*,[7] and which measures how well-equidistributed the values taken by the considered map are.

[6] The restriction to prime order is done for notational reasons and for ease of exposition. Our arguments can be readily adapted to the case of non-prime finite fields.

[7] A very similar notion of rank was defined for multilinear forms by Gowers and Wolf [13], who coined the term analytic rank. We use the same name to highlight the similarity between our two notions, which are relevant for distinct types of mathematical objects.

In the pseudorandom case, a key tool we use is a new random restriction result for high-rank polynomial maps proved in a companion paper [6]. We use this to show that the distribution of values taken by a high-rank polynomial map will be close to uniform even under a biased input distribution. This implies that the event considered in the theorem has very low probability for any fixed x, in which case we can conclude by averaging.

In the structured case we deal instead with polynomial maps of low rank, whose values are in a sense poorly distributed. Results from higher-order Fourier analysis then imply that they can be determined by "few" lower-degree polynomial maps (plus a few extra polynomials); by a simple Fourier-analytic argument we can reduce the analysis of a low-rank polynomial map to those lower-degree maps which specify it, making it amenable to an inductive argument.

3.2 Building the quantum circuit

The quantum circuit of Theorem 2 is inspired by a distributed version of the Bernstein-Vazirani algorithm [3]. Given a corrupted Hadamard codeword $H(x)$, this single-query quantum algorithm returns x with probability $\Omega(\varepsilon^2)$. The distributed version describes an entangled strategy for a particular non-local game [8] consisting of n players who, when given unique coordinates of $H(x)$, must each return an element of $\mathbb{F}_2^k$. They win if and only if the sum of their answers equals x. It turns out that by sharing an n-partite GHZ state of local dimension 2^k, they can simulate the Bernstein-Vazirani algorithm and achieve the same success probability. We then turn this entangled strategy into a quantum circuit that only uses single and two-qubit gates and classical parity gates. For this we use two constant-depth sub-routines, one for preparing GHZ states and another for the quantum *fan-out gate* [25], which implements the map $|x\rangle|y_1\rangle\dots|y_n\rangle \mapsto |x\rangle|y_1 \oplus x\rangle\dots|y_n \oplus x\rangle$.

To generate the GHZ state, we use a poor man's cat state [2], which is a GHZ state with some of its qubits flipped. We correct this poor man's cat state to a GHZ state by flipping qubits based on parity computations. The input for these parity computations follows from the procedure that generates the poor man's cat state.

To implement the quantum fan-out gate, we use ideas from distributed quantum computing. These ideas use GHZ states and classical parity gates together with single and two-qubit gates. With the quantum fan-out gate, we also obtain the quantum parity gate, by conjugating the quantum fan-out gate with Hadamard gates.

Part of the circuit is applying phase-flips, conditional on the bits of the corrupted codeword. To do this, we use quantum fan-out gates in a circuit, exponential in size in k to correctly apply the phase-flips [32].

The depth of the list-decoding circuit is constant, whereas the circuit size is $O(n^2 \log n)$. We also show how to reduce this complexity to $O(n \log n \log\log n)$, while increasing the depth by only a small constant number. We do this by preparing a state on $\lceil \log(k+1) \rceil$ qubits. Evaluating an OR on this newly prepared state yields the same result as evaluating an OR on the original k qubits [20]. Applying the same exponential size circuit as before on this newly prepared state indeed gives the reduced circuit size.

References

1 Sanjeev Arora and Boaz Barak. *Computational complexity*. Cambridge University Press, Cambridge, 2009. A modern approach. doi:10.1017/CBO9780511804090.

2 Adam Bene Watts, Robin Kothari, Luke Schaeffer, and Avishay Tal. Exponential separation between shallow quantum circuits and unbounded fan-in shallow classical circuits. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 515–526. ACM, June 2019. doi:10.1145/3313276.3316404.

3 Ethan Bernstein and Umesh Vazirani. Quantum complexity theory. *SIAM Journal on Computing*, 26(5):1411–1473, 1997. doi:10.1137/S0097539796300921.

4 Sergey Bravyi, David Gosset, and Robert König. Quantum advantage with shallow circuits. *Science*, 362(6412):308–311, October 2018. doi:10.1126/science.aar3106.

5 Sergey Bravyi, David Gosset, Robert König, and Marco Tomamichel. Quantum advantage with noisy shallow circuits. *Nature Physics*, 16(10):1040–1045, 2020. Preliminary version in FOCS'19. doi:10.1038/s41567-020-0948-z.

6 Jop Briët and Davi Castro-Silva. Random restrictions of high-rank tensors and polynomial maps, 2022. arXiv:2212.13728. doi:10.48550/arXiv.2212.13728.

7 Jop Briët, Harry Buhrman, Davi Castro-Silva, and Niels M. P. Neumann. Noisy decoding by shallow circuits with parities: classical and quantum, 2023. arXiv:2302.02870.

8 Richard Cleve, Peter Høyer, Benjamin Toner, and John Watrous. Consequences and limits of nonlocal strategies. In *Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004.*, pages 236–249. IEEE, 2004. doi:10.1109/CCC.2004.1313847.

9 Matthew Coudron, Jalex Stark, and Thomas Vidick. Trading locality for time: certifiable randomness from low-depth circuits. *Communications in Mathematical Physics*, 382(1):49–86, 2021. doi:10.1007/s00220-021-03963-w.

10 P Elias. List decoding for noisy channels. In *IRE WESCON Convention Record, 1957*, volume 2, pages 94–104, 1957.

11 François Le Gall. Average-Case Quantum Advantage with Shallow Circuits. In Amir Shpilka, editor, *34th Computational Complexity Conference (CCC 2019)*, volume 137 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 21:1–21:20, Dagstuhl, Germany, 2019. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CCC.2019.21.

12 O. Goldreich and L. A. Levin. A hard-core predicate for all one-way functions. In *Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing*, STOC '89, pages 25–32, New York, NY, USA, 1989. Association for Computing Machinery. doi:10.1145/73007.73010.

13 W. T. Gowers and J. Wolf. Linear forms and higher-degree uniformity for functions on $\mathbb{F}_p^n$. *Geom. Funct. Anal.*, 21(1):36–69, 2011. doi:10.1007/s00039-010-0106-3.

14 Ben Green and Terence Tao. The distribution of polynomials over finite fields, with applications to the Gowers norms. *Contrib. Discrete Math.*, 4(2):1–36, 2009. doi:10.11575/cdm.v4i2.62086.

15 Frederic Green, Steven Homer, Cristopher Moore, and Christopher Pollett. Counting, fanout and the complexity of quantum *ACC*. *Quantum Info. Comput.*, 2(1):35–65, December 2002. doi:10.26421/QIC2.1-3.

16 Daniel Grier and Luke Schaeffer. Interactive shallow clifford circuits: Quantum advantage against NC^1 and beyond. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC'20)*, pages 875–888, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3357713.3384332.

17 Venkatesan Guruswami and Salil Vadhan. A lower bound on list size for list decoding. *IEEE Transactions on Information Theory*, 56(11):5681–5688, 2010. doi:10.1109/TIT.2010.2070170.

18 R. W. Hamming. Error detecting and error correcting codes. *The Bell System Technical Journal*, 29(2):147–160, 1950. doi:10.1002/j.1538-7305.1950.tb00463.x.

19 Hamed Hatami, Pooya Hatami, and Shachar Lovett. Higher-order Fourier analysis and applications. *Found. Trends Theor. Comput. Sci.*, 13(4):247–448, 2019. doi:10.1561/0400000064.

20 Peter Høyer and Robert Špalek. Quantum fan-out is powerful. *Theory of Computing*, 1(5):81–103, 2005. doi:10.4086/toc.2005.v001a005.

21 Chin Ho Lee and Emanuele Viola. Some limitations of the sum of small-bias distributions. *Theory of Computing*, 13(16):1–23, 2017. doi:10.4086/toc.2017.v013a016.

22 Cristopher Moore. Quantum circuits: Fanout, parity, and counting. *arXiv preprint*, 1999. arXiv:quant-ph/9903046.

23 Cristopher Moore and Martin Nilsson. Parallel quantum computation and quantum codes. *SIAM Journal on Computing*, 31(3):799–815, 2001. doi:10.1137/S0097539799355053.

24 Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, 2010. doi:10.1017/CBO9780511976667.

25 Paul Pham and Krysta M. Svore. A 2D nearest-neighbor quantum architecture for factoring in polylogarithmic depth. *Quantum Info. Comput.*, 13(11–12):937–962, 2013. doi:10.26421/QIC13.11-12-3.

26 A. A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mathematical Notes of the Academy of Sciences of the USSR*, 41(4):333–338, April 1987. doi:10.1007/bf01137685.

27 Atri Rudra and Steve Uurtamo. Two theorems on list decoding. In Maria Serna, Ronen Shaltiel, Klaus Jansen, and José Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 696–709, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

28 C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656, 1948. doi:10.1002/j.1538-7305.1948.tb01338.x.

29 P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal of Computing*, 26(5):1484–1509, 1997. Preliminary version in FOCS'94. doi:10.1137/S0097539795293172.

30 R. Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, pages 77–82, New York, NY, USA, 1987. Association for Computing Machinery. doi:10.1145/28395.28404.

31 M. Sudan, L. Trevisan, and S. Vadhan. Pseudorandom generators without the xor lemma. In *Proceedings. Fourteenth Annual IEEE Conference on Computational Complexity (Formerly: Structure in Complexity Theory Conference) (Cat.No.99CB36317)*, pages 4–, 1999. doi:10.1109/CCC.1999.766253.

32 Yasuhiro Takahashi and Seiichiro Tani. Collapse of the hierarchy of constant-depth exact quantum circuits. In *2013 IEEE Conference on Computational Complexity*, pages 168–178, 2013. doi:10.1109/CCC.2013.25.

33 Terence Tao. *Higher order Fourier analysis*, volume 142 of *Graduate Studies in Mathematics*. American Mathematical Society, 2012. doi:10.1090/gsm/142.

34 Luca Trevisan. Some applications of coding theory in computational complexity. In *Complexity of computations and proofs*, volume 13 of *Quad. Mat.*, pages 347–424. Dept. Math., Seconda Univ. Napoli, Caserta, 2004.

35 Luca Trevisan and Salil Vadhan. Pseudorandomness and average-case complexity via uniform reductions. *Computational Complexity*, 16(4):331–364, 2007. doi:10.1007/s00037-007-0233-x.

36 Emanuele Viola. *The Complexity of Hardness Amplification and Derandomization.* PhD thesis, Harvard University, USA, 2006. AAI3217914. doi:10.5555/1195277.

37 Emanuele Viola. The sum of d small-bias generators fools polynomials of degree d. *Comput. Complexity*, 18(2):209–217, 2009. Preliminary version in CCC'08. doi:10.1007/s00037-009-0273-5.

38 Adam Bene Watts and Natalie Parham. Unconditional quantum advantage for sampling with shallow circuits, 2023. arXiv:2301.00995.

39 John M Wozencraft. List decoding. *Quarterly Progress Report*, 48:90–95, 1958.

40 Tomoyuki Yamakami. Quantum list decoding of classical block codes of polynomially small rate from quantumly corrupted codewords. *Baltic J. Modern Computing*, 4(4):753–788, 2016. doi:0.22364/bjmc.2016.4.4.12.

41 Andrew Chi-Chin Yao. Probabilistic computations: Toward a unified measure of complexity. In *18th Annual Symposium on Foundations of Computer Science (sfcs 1977)*, pages 222–227, 1977. doi:10.1109/SFCS.1977.24.