The Learning Stabilizers with Noise Problem

Alexander Poremba
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA
Department of Computer Science and Department of Physics, Boston University, MA, USA

Yihui Quek
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA
School of Computer and Communication Sciences & School of Basic Sciences,
École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Peter Shor
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

Random classical codes have good error correcting properties, and yet they are notoriously hard to decode in practice. Despite many decades of extensive study, the fastest known algorithms still run in exponential time. The *Learning Parity with Noise* (LPN) problem, which can be seen as the task of decoding a random linear code in the presence of noise, has thus emerged as a prominent hardness assumption with numerous applications in both cryptography and learning theory.

Is there a natural quantum analog of the LPN problem? In this work, we introduce the *Learning Stabilizers with Noise* (LSN) problem, the task of decoding a random stabilizer code in the presence of local depolarizing noise. We give both polynomial-time and exponential-time quantum algorithms for solving LSN in various depolarizing noise regimes, ranging from extremely low noise, to low constant noise rates, and even higher noise rates up to a threshold. Next, we provide concrete evidence that LSN is hard. First, we show that LSN includes LPN as a special case, which suggests that it is at least as hard as its classical counterpart. Second, we prove worst-case to average-case reductions for variants of LSN. We then ask: what is the computational complexity of solving LSN? Because the task features quantum inputs, its complexity cannot be characterized by traditional complexity classes. Instead, we show that the LSN problem lies in a recently introduced (distributional and oracle) *unitary synthesis class*. Finally, we identify several applications of our LSN assumption, ranging from the construction of quantum bit commitment schemes to the computational limitations of learning from quantum data.

2012 ACM Subject Classification Theory of computation → Problems, reductions and completeness

Keywords and phrases Random quantum stabilizer codes, average-case hardness

Digital Object Identifier 10.4230/LIPIcs.ITCS.2026.108

Related Version *Full Version*: https://arxiv.org/abs/2410.18953

Funding Supported by the Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Systems Accelerator, under Grant number DOE DE-SC0012704, and Co-design Center for Quantum Advantage (C2QA) under contract number DE-SC0012704.

Acknowledgements The authors would like to thank Alexandru Gheorghiu, Anand Natarajan, Aram Harrow, Henry Yuen, John Bostanci, Jonas Haferkamp, Jonas Helsen, Jonathan Conrad, Soonwon Choi, Thomas Vidick and Vinod Vaikuntanathan for useful discussions.

1 Introduction

Coding theory has offered many valuable insights into the theory of computation, ranging from structural insights in complexity theory [21, 4], to the design of cryptographic primitives [47, 49, 26, 43, 3] and even to lower bounds in the field of computational learning theory [12, 25]. The existence of asymptotically good error correcting codes, in particular, is a major

cornerstone in the field. Thanks to the probabilistic method, we know that a random linear code already attains the so-called Gilbert-Varshamov bound [28, 50] with high probability. This suggests that asymptotically good error correcting codes not only exist in theory, but are in fact also abundant. Despite their remarkable error correcting properties, however, random linear codes have been found to be notoriously hard to decode in practice.

Learning Parity with Noise

The observation that better codes seem harder to decode is captured by the *Learning Parity with Noise* (LPN) problem [10]. In a nutshell, this assumption says that it is computationally difficult to decode a random linear code under Bernoulli noise. In other words, given as input

$$(\mathbf{A} \sim \mathbb{Z}_2^{n\times k}, \mathbf{A}\cdot\mathbf{x} + \mathbf{e} \pmod 2)$$

it is hard to find the secret string $\mathbf{x}$ which is chosen uniformly at random in $\mathbb{Z}_2^k$, and where $\mathbf{e} \sim \mathsf{Ber}_p^{\otimes n}$ is a random Bernoulli error for some appropriate noise rate $p \in (0, 1/2)$. Here, $\mathbf{A} \in \mathbb{Z}_2^{n\times k}$ serves as a random *generator matrix* of a linear code, for $n = \mathsf{poly}(k)$.

In practice, LPN is believed to be hard for both classical and quantum algorithms running in time $\mathsf{poly}(k)$ in various noise regimes, and the best known algorithms still run in exponential time. In some parameter regimes, however, LPN is also known to admit subexponential-time algorithms; for block length of $n = 2^{O(k/\log k)}$ and constant noise rate p, the celebrated BKW algorithm [12] solves LPN in subexponential time $2^{O(k/\log k)}$. Lyubashevksy [41] later gave an algorithm that succeeds in time $2^{O(k/\log\log k)}$ even when $n = k^{1+\epsilon}$, for any $\epsilon > 0$.

The conjectured hardness of LPN has found applications in both cryptography [35, 3, 39, 5, 38, 20, 6, 14] and learning theory [10, 25]. The *Learning with Errors* (LWE) problem [46] – a more structured variant of LPN – has since become the basis of modern cryptography and has even led to highly advanced cryptographic primitives, such as fully homomorphic encryption [27, 15] and the classical verification of quantum computations [42]. In the context of learning theory, it was shown that an efficient algorithm for LPN would allow us to learn important function classes, such as 2-DNF formulas, juntas, and even more general functions with sparse Fourier spectrum [25].

Because the LPN problem is so prevalent in many areas of computer science, a significant effort has been devoted to finding evidence of its hardness. One of these pieces of evidence is a *worst-to-average-case reduction* [14, 55]. Recall that the LPN problem is an *average-case* problem: the computational task is to decode a *random* code, secret and error. Ref. [14] studied a related worst-case problem – the *nearest codeword problem* (NCP) – and showed that it reduces to LPN. This reduction, later improved by [55], showed that LPN is at least as hard as (a mildly hard variant of) NCP in the worst case. [14] also found the first non-trivial complexity upper bound on the hardness of the LPN problem; specifically, they showed that (a variant of) the LPN problem is contained in the complexity class $\mathsf{SearchBPP}^{\mathsf{SZK}}$, and is thus unlikely to be NP-hard.

The hardness of decoding random stabilizer codes

Just like random linear codes, random *quantum* stabilizer codes[1] also possess remarkable error correcting properties [48, 31]. They are ubiquitous in quantum information science; for example, random stabilizer codes appear in the context of quantum authentication

[1] The stabilizer formalism was first developed by Gottesman [30] and incorporates the majority of quantum error correcting codes we know today [2].

schemes and the verification of quantum computations [1], quantum cryptography [22], the theory of quantum communication [48, 52], and even in black-hole physics and quantum gravity [33, 54, 32]. Characterizing the hardness and complexity of decoding random stabilizer codes is therefore not only important from the perspective of quantum error correction, but could also shed a new light on the computational limitations of many natural quantum information processing tasks. In this work, we ask:

How hard is it to decode a random quantum stabilizer code?

Surprisingly, this subject has seen little theoretical treatment. While prior work has shown that decoding quantum stabilizer codes is *worst-case* hard [36, 37] – via reduction from a purely classical decoding problem – the average-case complexity of decoding stabilizer codes on *typical* instances of stabilizer decoding was left as an open problem [37]. We re-formulate this as a question that bears on all of the areas mentioned above:

Can we find a natural quantum analog of the Learning Parity with Noise problem? In particular, what would its hardness imply for quantum information science as a whole?

Given the success of constructing cryptographic primitives from the hardness of LPN in a classical world, could such a quantum analog of LPN offer a new hardness assumption which would allow us to construct cryptographic protocols in a quantum world? Finally – and perhaps, even more interestingly – this assumption may turn out to be even harder to break than its classical counterpart of LPN.

2 Overview

We now give an overview of our contributions in this work, summarized in Table 1.

2.1 Learning Stabilizers with Noise

In this work, we introduce a natural quantum analog of LPN – the *Learning Stabilizers with Noise* (LSN) problem. In studying the LSN problem, we thoroughly characterize the hardness and complexity of decoding random stabilizer codes in different noise regimes. Similar to the LPN problem, which has found numerous applications in both cryptography and learning theory, we believe that our LSN assumption has the potential to occupy a similar role in quantum information more broadly. Before we introduce our LSN task formally, we first revisit LPN and draw a connection to quantum error correction.

A quantum analog of LPN?

Let $n, k \in \mathbb{N}$ be integers with $n = \mathsf{poly}(k)$, and let $p \in (0, 1/2)$ be a parameter. Recall that an instance of the LPN problem consists of a generator matrix for a random linear code, together with a noisy codeword for a random string; specifically, we consider samples of the form

$$(\mathbf{A} \sim \mathbb{Z}_2^{n\times k}, \mathbf{A}\cdot\mathbf{x} + \mathbf{e} \pmod 2)$$

where $\mathbf{A} \in \mathbb{Z}_2^{n\times k}$ is a random generator matrix, where $\mathbf{A}\cdot\mathbf{x} + \mathbf{e} \pmod 2$ is a noisy codeword which encodes uniformly random string $\mathbf{x} \sim \mathbb{Z}_2^k$, and where $\mathbf{e} \sim \mathsf{Ber}_p^{\otimes n}$ is a random Bernoulli error. Without loss of generality[2], we assume that the matrix $\mathbf{A}$ has full column-rank, i.e.,

[2] This happens with overwhelming probability for $\mathbf{A} \sim \mathbb{Z}_2^{n\times k}$ provided that $n \gg k$.

Table 1 Comparison of LPN and LSN in terms of hardness and complexity.

	Learning Parity with Noise	Learning Stabilizers with Noise (this problem)
Worst-case hardness	✓ NP-complete [9] as a decisional syndrome decoding task. Variant of the *Nearest Codeword Problem* (NCP) [14]	As a classical syndrome decoding task: NP-complete [36, 40] or #P-complete [37] depending on the decoding problem
Average-case hardness	✓ [14, 55]	This paper
Worst-to-average-case reductions	✓ [14, 55]	This paper
Algorithms for average-case problem	✓ $2^{O(k/\log k)}$ time complexity [11] in constant-noise regime.	This paper

the columns of $\mathbf{A}$ generate all of $\mathbb{Z}_2^k$. We now make a simple observation; namely, we can interpret the LPN instance $\mathbf{A}\cdot\mathbf{x}+\mathbf{e} \pmod 2$ as a particular noisy quantum codeword on n qubits[3], since

$$\begin{aligned}|\mathbf{A}\cdot\mathbf{x}+\mathbf{e} \pmod 2\rangle &= X^{\mathbf{e}}\,|\mathbf{A}\cdot\mathbf{x} \pmod 2\rangle \\ &= X^{\mathbf{e}}U_{\mathbf{A}}\left(|0^{n-k}\rangle\otimes|\mathbf{x}\rangle\right)\,, \end{aligned} \tag{1}$$

where $X^{\mathbf{e}} = X^{e_1}\otimes\cdots\otimes X^{e_n}$ is a product of low-weight Pauli-X operators and where the unitary operator $U_{\mathbf{A}}$ is defined to be the matrix multiplication operation

$$U_{\mathbf{A}} : |0^{n-k}\rangle\otimes|\mathbf{x}\rangle \to |\mathbf{A}\cdot\mathbf{x} \pmod 2\rangle\,. \tag{2}$$

Because $\mathbf{A}$ has full column-rank, $U_{\mathbf{A}}$ corresponds to a linear reversible circuit which can be described solely in terms of CNOT gates [45]. Therefore, $U_{\mathbf{A}}$ is a Clifford operator and thus maps Pauli operators to Pauli operators via conjugation.

We may also observe that $U_{\mathbf{A}}$ is the encoding circuit for a stabilizer code. The stabilizer group associated with this code is precisely the group of k commuting Pauli operators under which $U_{\mathbf{A}}(|0^{n-k}\rangle\otimes|\mathbf{x}\rangle)$ remains invariant. These are easily seen to be the Pauli operators

$$U_{\mathbf{A}}Z_iU_{\mathbf{A}}^\dagger, \quad \text{for } i\in[n-k]\,,$$

where Z_i denotes a Pauli operator which acts as a Pauli-Z operator on the i-th qubit, and is equal to the identity everywhere else. In other words, the Clifford encoding unitary $U_{\mathbf{A}}$ (derived from an instance of an LPN problem) gives rise to the quantum stabilizer code given by $S_{\mathbf{A}} = \langle U_{\mathbf{A}}Z_1U_{\mathbf{A}}^\dagger,\ldots,U_{\mathbf{A}}Z_{n-k}U_{\mathbf{A}}^\dagger\rangle$. This shows that every instance of LPN can be mapped to an instance of decoding stabilizer codes.

We now generalize this idea significantly, ultimately leading us to define the *Learning Stabilizers with Noise* (LSN) problem – the natural quantum analog of the LPN problem:

- (Random stabilizer code:) Note that the encoding Clifford unitary in Equation (2) generates a specific stabilizer code of the form $S_{\mathbf{A}} = \langle U_{\mathbf{A}}Z_1U_{\mathbf{A}}^\dagger,\ldots,U_{\mathbf{A}}Z_{n-k}U_{\mathbf{A}}^\dagger\rangle$. We consider stabilizer subgroups of the Pauli group which are chosen uniformly at random from the set of all stabilizer subgroups with $n-k$ generators, denoted by $\mathsf{Stab}(n,k)$.

[3] Strictly speaking, we should think of $|\mathbf{A}\cdot\mathbf{x}+\mathbf{e} \pmod 2\rangle$ as encoding the row vector $\mathbf{x}^\intercal\mathbf{A}+\mathbf{e}^\intercal \pmod 2$.

In fact, as we show in this work, this is equivalent to choosing random stabilizer codes which are described by the set of generators $\langle CZ_1C^\dagger, \ldots, CZ_{n-k}C^\dagger\rangle$, where $C \sim \mathsf{Cliff}_n$ is a random n-qubit Clifford operator.[4]

- (Local depolarizing noise:) Recall that the noisy codeword $X^{\mathbf{e}}U_{\mathbf{A}}(|0^{n-k}\rangle \otimes |\mathbf{x}\rangle)$ in Equation (1) is only affected by low-weight bit-flip errors $X^{\mathbf{e}}$, where $\mathbf{e} \sim \mathsf{Ber}_p^{\otimes n}$ comes from a Bernoulli distribution. In quantum systems, however, noise may also come in the form of phase errors. This leads us to consider a quantum noise model in the form of local depolarizing noise $\mathcal{D}_p^{\otimes n}$. Similar to the Bernoulli distribution, local depolarizing noise also produces low-weight errors with high probability, which therefore naturally generalizes the classical noise model.

In other words, we consider the task of decoding a random quantum stabilizer code in the presence of local depolarizing noise. Because the codeword is a *stabilizer state*, we call this the *Learning Stabilizers with Noise* (LSN) problem – in analogy to the classical LPN problem. We now give a formal definition of the problem.

Learning Stabilizers with Noise

The *Learning Stabilizers with Noise* (LSN) problem is to find $x \in \{0,1\}^k$ given as input

$$\left(S \in \mathsf{Stab}(n,k),\ E\,|\overline{\psi_x}\rangle^S\right) \quad \text{with} \quad |\overline{\psi_x}\rangle^S := U_{\mathrm{Enc}}^S(|0^{n-k}\rangle \otimes |x\rangle)\,,$$

where $S \sim \mathsf{Stab}(n,k)$ is a uniformly random stabilizer subgroup, where $E \sim \mathcal{D}_p^{\otimes n}$ is a Pauli error from a local depolarizing channel, where $x \sim \{0,1\}^k$ is a random string, and where U_{Enc}^S is some canonical encoding circuit for the stabilizer code associated with S. As mentioned before, the encoding circuit is typically given in the form of a random n-qubit Clifford operator.

At first sight, it may not be clear why the LSN problem is even well-defined, since a unique solution to the quantum learning problem may not exist in certain parameter regimes. The intuition behind our argument for the existence of a unique solution is as follows. Suppose that $p \in (0, 1/2)$ is a sufficiently small constant. Then, the quantum Gilbert-Varshamov bound implies that a random stabilizer code is non-degenerate and has distance at least $d = 3np + 1$ with overwhelming probability, in which case for any pair of codewords with $x, y \in \{0,1\}^k$, we have

$$\langle\overline{\psi_x}|\, E_a^\dagger E_b\, |\overline{\psi_y}\rangle = 0$$

by the *Knill-Laflamme conditions* – provided that E_a, E_b have weight at most $|E_a|, |E_b| \leq \frac{3}{2}np$. Fortunately, a simple Chernoff bound analysis reveals that this is the case with overwhelming probability for the local depolarizing channel $\mathcal{D}_p^{\otimes n}$. Therefore, Pauli errors which originate from a local depolarizing noise channel take orthogonal codewords to orthogonal codewords, and hence there must exist a measurement that perfectly distinguishes between them. This observation is also at the core of our algorithms for the LSN problem, which we describe next.

[4] While this is considered *folklore*, we are not aware of a proof which has previously appeared in the literature. Therefore, we decided to rigorously prove this statement.

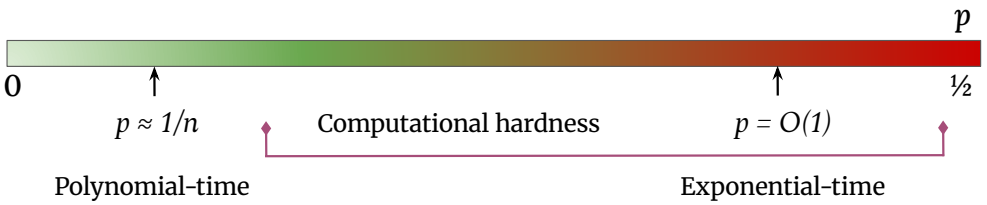

Figure 1 Algorithms for the LSN problem. In an extremely low noise regime with $p \approx 1/n$, the problem can be solved in polynomial-time. In a constant-noise regime with $p = O(1)$, the problem can be solved information-theoretically. Here, we give an exponential-time algorithm and conjecture that it is computationally hard.

Algorithms for Learning Stabilizers with Noise

Previously, we discussed why random stabilizer codes give rise to a single-shot decoding problem which exhibits unique solutions with high probability. This suggests that LSN can be solved information-theoretically. Can we also find efficient algorithms for solving the LSN problem? Not surprisingly, the answer depends on the specific noise regime of the error distribution (see Figure 1). We give both polynomial-time and exponential-time quantum algorithms for solving the LSN problem in various noise-regimes:

- **Extremely low-noise** regime with parameter $p \approx \frac{1}{n}$. In this parameter regime, we show that a simple projection onto the stabilizer codespace (see Algorithm 1) suffices to solve the LSN problem in time $O(n^3)$ with constant success probability.
- **Low constant-noise** regime for a small constant $p \in (0, 1/2)$. In this regime, we show that the *Pretty Good Measurement* (PGM) [8, 44] (with only a single sample) succeeds with high probability. Concretely, we view the LSN problem as a state discrimination task where the goal is to identify 1-out-of-2^k (almost) orthogonal states. The success of the PGM is inextricably linked to the structure of our problem: although the distance between two arbitrary orthogonal states contracts tremendously – in fact, exponentially in system size (see, e.g. Proposition IV.7 in [34]) – under a layer of local depolarizing noise, a random stabilizer code encodes orthogonal states into *orthogonal* subspaces that are "protected" from such destructive contraction even under noise. This means that the information of the LSN secret is preserved under noise, and this is the intuition behind our approach in Algorithm 2.
- **Higher constant-noise** regime. In this regime, decoding is still possible at the cost of more samples. We derive the scaling of the sample complexity with noise, up to a certain noise threshold.

Connection to syndrome decoding

Recall that an instance of the LSN problem by definition features *quantum inputs*, which naturally led us to consider *quantum* algorithms as a means of solving the problem. However, despite the fact that LSN is a quantum problem, it reduces to a natural *classical* problem – *stabilizer syndrome decoding*. One way of solving the problem is to just measure the stabilizer syndromes of the noisy quantum codeword and to tackle the classical decoding problem instead. However, there are good reasons to believe that the classical (average-case) syndrome

decoding problem is actually harder: while LSN trivially reduces to stabilizer syndrome decoding, the reverse direction is highly unclear; in particular, it is not at all obvious why an algorithm for the LSN problem would also imply an algorithm for the classical stabilizer syndrome decoding problem.

Worst-case to average-case reductions

Recall that the LSN problem is stated as an *average-case* problem, where the success probability of an algorithm is measured on average over the random choice of stabilizer $S \in \mathsf{Stab}(n,k)$, secret $x \in \mathbb{Z}_2^k$ and error $E \sim \mathcal{D}_p^{\otimes n}$. While the quantum Gilbert-Varshamov bound does in fact guarantee that an average-case instance of LSN can be solved information-theoretically, our results indicate that the problem becomes computationally intractable for large k – even in constant-noise regimes. This raises the question of whether we can find concrete evidence for the average-case hardness of the LSN problem, beyond the fact that it subsumes the classical LPN problem.

Recently, a number of works showed that there is in fact evidence of *worst-case* hardness for LPN; specifically, by studying a related worst-case problem – the *nearest codeword problem* (NCP) [14, 55]. Using the *sample amplification* technique [41], Brakerski, Lyubashevsky, Vaikuntanathan and Wichs [14] gave a worst-case to average-case reduction from NCP to LPN. Here, an instance to the former problem consists of $(\mathbf{C} \in \mathbb{Z}_2^{m\times k}, \mathbf{t} = \mathbf{C}\cdot\mathbf{s} + \mathbf{w} \pmod 2)$ for some $\mathbf{s} \in \mathbb{Z}_2^k$, with the promise that the generator matrix $\mathbf{C}$ is balanced[5] and that the Hamming weight of the error $\mathbf{w} \in \mathbb{Z}_2^m$ is known. On a high level, the reduction in [14] proceeds in two steps:

- (Re-randomization of the secret) A random string $\mathbf{u} \sim \mathbb{Z}_2^k$ is chosen, and the worst-case instance $(\mathbf{C}, \mathbf{t})$ gets mapped via an additive shift to

  $$(\mathbf{C},\, \mathbf{t} + \mathbf{C}\cdot\mathbf{u} \pmod 2) = (\mathbf{C}, \mathbf{C}\cdot(\mathbf{s}+\mathbf{u}) + \mathbf{w} \pmod 2).$$

  Note that, whereas the initial secret $\mathbf{s} \in \mathbb{Z}_2^k$ was fixed, the new secret $\mathbf{s}+\mathbf{u}$ is now distributed according to the uniform distribution over $\mathbb{Z}_2^k$.
- (Re-randomization of the code and error) A random $\mathbf{R} \in \mathbb{Z}_2^{n\times m}$ is sampled from a *smoothing distribution* $\mathcal{R}_w^{n\times m}$, and the previous sample gets mapped to

  $$(\mathbf{R}\cdot\mathbf{C}, \mathbf{R}\cdot(\mathbf{C}\cdot(\mathbf{s}+\mathbf{u}) + \mathbf{w} \pmod 2)) = (\mathbf{R}\cdot\mathbf{C}, \mathbf{R}\cdot\mathbf{C}\cdot(\mathbf{s}+\mathbf{u}) + \mathbf{R}\cdot\mathbf{w} \pmod 2))$$

[14] show that the resulting sample is statistically close to an (average-case) LPN sample – provided that the *smoothing distribution* $\mathcal{R}_w^{n\times m}$ is chosen appropriately.

By taking a similar approach, we develop a worst-case to average-case reduction for the LSN problem. Here, the starting point is a worst-case stabilizer decoding instance $(S, E\,|\overline{\psi}_x\rangle^S)$ for some stabilizer $S \in \mathrm{Stab}(n,k)$, Pauli error E of bounded weight (at most $\log^c(n)$ for any $c > 0$), and secret $x \in \{0,1\}^k$. First, we observe that, in order to re-randomize the secret x, we need to act on the encoded data $|\overline{\psi}_x\rangle^S$ itself. Hence, it suffices to choose a random string $u \sim \{0,1\}^k$ and to apply the *logical* Pauli operator $\overline{X}^u$ associated with S to the noisy codeword itself, resulting in the desired transformation $E\,|\overline{\psi}_{x\oplus u}\rangle^S$.

To re-randomize the code and the error, we first observe that the shifted codeword $|\overline{\psi}_{x\oplus u}\rangle^S$ can be written as

$$|\overline{\psi}_{x\oplus u}\rangle^S = U_{\mathrm{Enc}}^S(|0^{n-k}\rangle \otimes |x\oplus u\rangle)$$

[5] Roughly speaking, this means that the minimum and maximum distance of the linear code generated by $\mathbf{C} \in \mathbb{Z}_2^{n\times m}$ is neither too small nor too large.

for some (not necessarily random) encoding Clifford $U^S_{\mathrm{Enc}} \in \mathsf{Cliff}_n$. Because Cliff_n forms a finite group, this suggests that one could simply sample a uniformly random $C \sim \mathsf{Cliff}_n$ and consider the state $CE\left|\overline{\psi}_{x\oplus u}\right\rangle^S = (CEC^\dagger)C\left|\overline{\psi}_{x\oplus u}\right\rangle^S$, where $C\left|\overline{\psi}_{x\oplus u}\right\rangle^S$ now comes from a random stabilizer code for a uniformly random encoding Clifford $C \cdot U^S_{\mathrm{Enc}}$. While this does seem to result in a re-randomized stabilizer code, the aforementioned transformation could potentially *blow up* the weight of the Pauli error E. In fact, it is well-known that random Cliffords are *Pauli-mixing* [19, 1]: they take any non-identity Pauli operator and map it to a uniformly random non-identity Pauli operator via conjugation. This seems to suggest that any naive worst-case to average-case reduction for LSN is doomed to fail: the weight of the re-randomized Pauli error now appears to follow a Binomial distribution with parameter 3/4, thereby inducing a high-noise distribution which is statistically far from any reasonable noise channel (e.g., such as a local depolarizing channel).

In an attempt to overcome this barrier, we develop a re-randomization strategy which is much more *gentle* on the error (i.e., it does not cause it to blow up), and yet still ensures that the code gets *somewhat* re-randomized; however, this occurs at an expense: our random self-reduction only applies to a very restrictive case of the LSN problem. Our strategy is to follow the random Pauli operator with a *twirl* – another random unitary consisting of a random permutation operator, followed by a layer of random single-qubit Cliffords. Similar ensembles of unitaries been used in the *randomized compiling and benchmarking* literature in order to tailor noise with arbitrary coherence and spatial correlations into a symmetric Pauli channel [51, 23].[6] Under this twirl, a worst-case error of weight w is transformed into a uniformly random error of weight w; in particular, the weight of the error remains invariant. However, that the distributions of the error and the code are now *correlated*, which requires a much more refined analysis.

Complexity of Learning Stabilizers with Noise

What is the computational complexity of solving the LSN problem? Notice that the description of the learning task features quantum inputs, which means that its complexity cannot be characterized by traditional complexity classes such as BQP or QMA which deal with classical-input decision problems. Instead, we show that (a variant of) the LSN problem lies in the complexity class called $\mathsf{avgUnitaryBQP}^{\mathsf{NP}}$, a (distributional and oracle) unitary synthesis class which was recently formalized by Bostanci et al. [13]. At a high level, this follows from the quantum Gilbert-Varshamov bound which ensures that a random LSN instance carries a unique stabilizer syndrome with overwhelming probability. Once the syndrome is computed, it can then be decoded with the help of an NP oracle. In other words, any polynomial-time quantum machine which access to the oracle can *synthesize* the appropriate Pauli correction for the noisy quantum codeword.

In the full version of the paper, we give a brief review of unitary complexity and also prove our aforementioned complexity upper bound on the hardness of the LSN problem.

[6] To our knowledge, however, our work is the first to identify the twirl as a useful tool in the context of a worst-case to average-case reduction.

2.2 Applications

Learning from quantum data

Just as LPN has been fundamental to lower bounds in classical learning theory, we expect that LSN will be a useful tool for proving lower bounds in quantum learning theory. In the full version of the paper, we identify one such learning setting: learning from quantum data [16, 18, 24, 17], a generalization of Probably Approximately Correct (PAC) learning to the quantum setting. Here, the goal is to learn a map $\rho : \mathcal{X} \to L(\mathcal{H}_d)$ from classical to quantum data – for example, a Hamiltonian can be construed as a map from temperatures to Gibbs states, or time-evolved states.

In one special case of this task known as *learning state preparation processes*, the learner is allowed to observe input-output pairs for an unknown map, but the inputs are sampled from a distribution and are not identical. Refs. [18, 24] gave sample-efficient algorithms for learning in this setting, thus showing that it is information theoretically possible. But we show that these algorithms can never be computationally efficient – assuming the hardness of our LSN assumption. While the lack of computational efficiency was implicit for the fully general learning setting due to a result of [7], that result does not apply when there are physically natural restrictions on the concept class, such as learning processes that involve quantum noise. Our LSN hardness assumption fills this gap.

Constructing quantum bit commitment schemes

In the full version of the paper, we give a cryptographic application of LSN and show how to construct a statistically hiding and computationally binding quantum commitment scheme [53]. This is a fundamental cryptographic primitive that allows two parties (called a *sender* and *receiver*) to engage in a two-phase quantum communication protocol: in the first phase (the "commit phase"), the sender sends a commitment (i.e., a quantum register) to a bit b to the receiver; the *hiding* property of a bit commitment scheme ensures that the receiver cannot decide the value of b from the commitment alone. In the second phase (the "reveal phase"), the sender sends another quantum register to the receiver that allows the receiver to compute the value of b; the *binding* property of commitments ensures that the sender can only reveal the correct value of b, i.e. if the sender sent a reveal register that was meant to convince the receiver it had committed to a different value of b, the receiver would detect this.

3 Random stabilizer codes

A random $[[n, k]]$ quantum stabilizer code is a uniformly random choice of abelian subgroup $S \in \mathsf{Stab}(n, k) \leq \mathcal{P}_n$ with $n - k$ generators. Note, here the Pauli signs are important and will determine what subspace is stabilized by S. Because the members of S must commute, choosing $n - k$ elements uniformly from $\mathcal{P}_n$ will not always give a valid stabilizer code.

We now give a rigorous proof of the fact that a random element of the Clifford group Cliff_n, acting on any initial choice of S, generates a uniformly random S, and hence a uniformly random $[[n, k]]$ stabilizer code.

▶ **Theorem 1.** *Let $n \in \mathbb{N}$ be an integer and let $S = \langle \mathbf{g}_1, \ldots, \mathbf{g}_{n-k} \rangle$ be any stabilizer with generators $\mathbf{g}_1, \ldots, \mathbf{g}_{n-k} \in \mathcal{P}_n$. Then, the conjugated stabilizer code*

$$USU^\dagger = \langle U\mathbf{g}_1U^\dagger, \ldots, U\mathbf{g}_{n-k}U^\dagger \rangle, \quad \text{for } U \sim \mathsf{Cliff}_n,$$

yields a uniformly stabilizer in the set $\mathsf{Stab}(n, k)$.

Proof. First, we show that the Clifford group Cliff_n acts transitively on the set of stabilizers $\mathsf{Stab}(n,k)$. Let $S = \langle \mathbf{g}_1, \ldots, \mathbf{g}_{n-k} \rangle$ be an arbitrary stabilizer with $n-k$ generators. From [31], we know that there exists a Clifford operator $C \in \mathsf{Cliff}_n$ and a Pauli $P \in \bar{\mathcal{P}}_n$ such that the composition of the two operations maps S to the canonical stabilizer Z. In particular, we can let $V = PC$ such that

$$VSV^\dagger = \langle V\mathbf{g}_1V^\dagger, \ldots, V\mathbf{g}_{n-k}V^\dagger \rangle = \langle Z_1, \ldots, Z_{n-k} \rangle.$$

Likewise, from [31], we also know that once we have the canonical all-Z stabilizer $Z = \langle Z_1, \ldots, Z_{n-k} \rangle$, we can obtain any other stabilizer $S' = \langle \mathbf{g}'_1, \ldots, \mathbf{g}'_{n-k} \rangle$ via some other composition of operators $W = DQ$, where $D \in \mathsf{Cliff}_n$ and $Q \in \bar{\mathcal{P}}_n$, i.e.,

$$S' = \langle \mathbf{g}'_1, \ldots, \mathbf{g}'_{n-k} \rangle = \langle WZ_1W^\dagger, \ldots, WZ_{n-k}W^\dagger \rangle.$$

Therefore, for any pair of distinct stabilizers $S, S' \in \mathsf{Stab}(n,k)$ there exists an operator WV that maps S to S'. By using the fact that Cliffords are the normalizer of the Pauli group, WV can be realized as a single Pauli operation followed by a single Clifford operation.

Finally, we show that the probability that a random Clifford applied to an arbitrary stabilizer $S = \langle \mathbf{g}_1, \ldots, \mathbf{g}_{n-k} \rangle$ yields any pair of distinct stabilizers S_1, S_2 with exactly the same probability. From before, there exists a Clifford $C \in \mathsf{Cliff}_n$ such that

$$\begin{aligned}
\Pr_{U \sim \mathsf{Cliff}_n}[USU^\dagger = S_1] &= \Pr_{U \sim \mathsf{Cliff}_n}[(CU)S(CU)^\dagger = CS_1C^\dagger] \\
&= \Pr_{U \sim \mathsf{Cliff}_n}[(CU)S(CU)^\dagger = S_2] \\
&= \Pr_{U \sim \mathsf{Cliff}_n}[USU^\dagger = S_2].
\end{aligned}$$

The last line follows from the fact that Cliff_n is a group, and hence the uniform distribution over Cliff_n is Clifford invariant. ◀

4 The Learning Stabilizers with Noise problem

In this section, we formally define the *Learning Stabilizers with Noise* (LSN) problem as the natural quantum analog of the LPN problem. We begin with a set of definitions for the problem (and its variants), and then show that the problem is well-defined (i.e., it admits a unique solution) for appropriate choices of parameters.

Definition

We now provide a formal definition of our learning task.

▶ **Definition 2** (Learning Stabilizers with Noise problem). *Let $k \in \mathbb{N}$ be the security parameter and let $n = \mathsf{poly}(k)$ be an integer. Let $p \in (0, 1/2)$ be a parameter. The Learning Stabilizers with Noise ($\mathsf{LSN}_{n,k,\mathcal{D}_p^{\otimes n}}$) problem is to find $x \in \{0,1\}^k$ given as input a sample*

$$\left(S \in \mathsf{Stab}(n,k), E\,|\overline{\psi_x}\rangle^S\right),$$

where $S \sim \mathsf{Stab}(n,k)$ is a uniformly random stabilizer (specified in terms of a classical description of S), $E \sim \mathcal{D}_p^{\otimes n}$ is a Pauli error with $E \in \bar{\mathcal{P}}_n$, $x \sim \{0,1\}^k$ is a random string , and $|\overline{\psi_x}\rangle^S \in C(S)$ is the codeword

$$|\overline{\psi_x}\rangle^S := U^S_{\mathrm{Enc}}(|0^{n-k}\rangle \otimes |x\rangle)$$

for some canonical encoding circuit U_{Enc}^S for the stabilizer code associated with S. We say that a quantum algorithm solves the $\mathsf{LSN}_{n,k,\mathcal{D}_p^{\otimes n}}$ problem if it runs in time $\mathsf{poly}(k)$ and succeeds with probability at least $1/\mathsf{poly}(k)$ over the choice of S, E and x, and its internal randomness.

Let us first state a few remarks.

▶ Remark 3 (Density matrix formulation). In Definition 2, the input to the learning algorithm is stated in the form $(S \in \mathsf{Stab}(n,k), E\,|\overline{\psi_x}\rangle^S)$, where $S \sim \mathsf{Stab}(n,k)$, $E \sim \mathcal{D}_p^{\otimes n}$ and $x \sim \{0,1\}^k$ is a random string. The pure state $E\,|\overline{\psi_x}\rangle^S$, however, should rather be understood as a density matrix of the form

$$\rho_x^S = \mathcal{D}_p^{\otimes n}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S) = \sum_{E\in\bar{\mathcal{P}}_n} \Pr_{E\sim\mathcal{D}_p^{\otimes n}}[E]\cdot E\,|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S\,E^\dagger\,,$$

where we think of $\mathcal{D}_p^{\otimes n}$ as a product of local depolarizing channels with parameter p.

▶ Remark 4 (Clifford representation). Recall that the input of the learner in Definition 2 consists of a random stabilizer $S \in \mathsf{Stab}(n,k)$. In Theorem 1, we showed that random stabilizer codes can be equivalently described by uniformly random Clifford encoding circuits. Therefore, we can alternatively think of an $\mathsf{LSN}_{n,k,\mathcal{D}_p^{\otimes n}}$ instance as

$$\big(C \in \mathsf{Cliff}_n,\ E\,C(|0^{n-k}\rangle\otimes|x\rangle)\big)$$

where C is a random Clifford and the first argument refers to a classical description of C.

General variants

Recall that the input in Definition 2 consists of a sample

$$\big(S, E\,|\overline{\psi_x}\rangle^S\big) \ \sim\ \mathsf{LSN}_{n,k,\mathcal{D}_p^{\otimes n}}$$

where $S \sim \mathsf{Stab}(n,k)$ is a (classical description of) uniformly random stabilizer, where $E \sim \mathcal{D}_p^{\otimes n}$ is a Pauli error $E \in \bar{\mathcal{P}}_n$, where $x \sim \{0,1\}^k$ is a random string. Occasionally, we also consider more general variants of the problem, denoted by $\mathsf{LSN}_{n,k,\mathcal{N},\mathcal{S},\mathcal{I}}$, that feature samples

$$\big(S, E\,|\overline{\psi_x}\rangle^S\big) \ \sim\ \mathsf{LSN}_{n,k,\mathcal{N},\mathcal{S},\mathcal{I}}$$

which are captured by the following set of distributions (which depend on n and k):

- $\mathcal{N}$ is a *general noise distribution* with support over n-qubit Pauli errors $E \in \bar{\mathcal{P}}_n$.
- $\mathcal{S}$ is a *general distribution over Stabilizer codes*, either with support over the set of stabilizers $S \in \mathsf{Stab}(n,k)$, or over Clifford encoding circuits in $C \in \mathsf{Cliff}_n$.
- $\mathcal{I}$ is a *general distribution over input strings* of the form $x \in \{0,1\}^k$.

Depending on the choice of distributions $\mathcal{N}$, $\mathcal{S}$ and $\mathcal{I}$, the learning task may either become easier or harder than the canonical learning problem in Definition 2.

5 Quantum Algorithms for Learning Stabilizers with Noise

In this section, we give both polynomial-time and exponential-time quantum algorithms for solving LSN in various depolarizing noise regimes.

Algorithm 1 Projection onto the Stabilizer Codespace.

Input: Instance $(S \in \mathsf{Stab}(n,k), \mathcal{D}_p^{\otimes n}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S)) \sim \mathsf{LSN}_{n,k,\mathcal{D}_p^{\otimes n}}$.
Output: A string $x' \in \{0,1\}^k$.
1 Let $\rho_x^S \leftarrow \mathcal{D}_p^{\otimes n}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S)$ denote the ensemble instance.
2 Find an encoding Clifford U_{Enc}^S associated with $S \in \mathsf{Stab}(n,k)$.
3 Compute $(U_{\mathrm{Enc}}^S)^\dagger \rho_x^S U_{\mathrm{Enc}}^S$ and measure in the computational basis.
4 Output the string x' corresponding to the last k bits of the measurement outcome.

5.1 Single-Shot Decoding for Extremely Low Noise Rates

In this section, we first consider an extremely low noise regime in which the LSN problem becomes computationally tractable. Specifically, we consider the $\mathsf{LSN}_{n,k,\mathcal{D}_p^{\otimes n}}$ problem with parameter $p = 1/n$ and $n = \mathsf{poly}(k)$. We show that a simple projection onto the stabilizer code space suffices to solve $\mathsf{LSN}_{n,k,\mathcal{D}_p^{\otimes n}}$ in time $O(n^3)$ with good success probability.

▶ **Theorem 5** (Single-Shot Decoding for Extremely Low Noise Rates). *Let $n, k \in \mathbb{N}$ be integers with $n \geq k$ and $n = \mathsf{poly}(k)$. Let $p = \frac{1}{n}$ be a noise parameter. Then, Algorithm 1 runs in time $O(n^3)$ and solves the $\mathsf{LSN}_{n,k,\mathcal{D}_p^{\otimes n}}$ problem with probability at least 0.25.*

Proof. Suppose we are given as input an instance of the $\mathsf{LSN}_{n,k,\mathcal{D}}$ problem, i.e.,

$$\left(S \in \mathsf{Stab}(n,k), \mathcal{D}_p^{\otimes n}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S)\right) \sim \mathsf{LSN}_{n,k,\mathcal{D}}$$

where $S \sim \mathsf{Stab}(n,k)$ describes a random stabilizer code and $x \sim \{0,1\}^k$ is a random element. Let $\rho_x^S \leftarrow \mathcal{D}_p^{\otimes n}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S)$ denote the ensemble instance. Then, the probability that Algorithm 1 succeeds is trivially bounded below by the probability that the depolarizing noise channel is error-less, and thus

$$\mathop{\mathbb{E}}_{x\sim\{0,1\}^k} \mathrm{Tr}\left[(\mathbb{I}_{n-k} \otimes |x\rangle\langle x|)(U_{\mathrm{Enc}}^S)^\dagger \rho_x^S U_{\mathrm{Enc}}^S\right] \geq \left(1 - \frac{3p}{4}\right)^n \geq 1 - n \cdot \frac{3p}{4} \geq 0.25\,,$$

where the second inequality follows from Bernoulli's inequality, and the last inequality follows from our assumption that $p = 1/n$. ◀

5.2 Single-Shot Decoding for Low Constant Noise Rates

Recall that LSN is the following decoding problem: given as input an instance

$$\left(S \in \mathsf{Stab}(n,k), E\,|\overline{\psi_x}\rangle^S\right),$$

where $S \sim \mathsf{Stab}(n,k)$ is a random stabilizer code and $E \sim \mathcal{D}_p^{\otimes n}$ is a random n-qubit Pauli error $E \in \bar{\mathcal{P}}_n$, and where $x \sim [2^k]$ is a random element. In other words, we have a noisy state discrimination task $\{\rho_x\}_{x\in[2^k]}$ where ρ_x is the mixed state

$$\rho_x = \mathcal{D}_p^{\otimes n}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|) = \sum_{E\in\bar{\mathcal{P}}_n} \Pr_{E\sim\mathcal{D}_p^{\otimes n}}[E] \cdot E\,|\overline{\psi_x}\rangle\langle\overline{\psi_x}|\,E^\dagger. \tag{3}$$

Our algorithm for solving LSN is to implement a *Pretty Good Measurement* (PGM) [8, 44], but with a twist that enables us to bound its success probability: we will implement an approximation of the PGM that works for *truncated depolarizing noise*, noise from which

we have culled the highest weight (and thus most destructive) errors. The reason this works is that the PGM is actually the optimal measurement for orthogonal states and discriminates them perfectly. Using the Gilbert-Varshamov bound, we are able to harness the fact that under truncated depolarizing noise, encoded orthogonal states remain approximately orthogonal, and thus the PGM still works well for the task.[7]

Pretty Good Measurement

We recall the following result by Montanaro:

▶ **Lemma 6** ([44]). *Let $\{\rho_x\}$ be an ensemble of quantum states and let $\mathbf{\Lambda}_x = \mathbf{\Sigma}^{-\frac{1}{2}}\rho_x\mathbf{\Sigma}^{-\frac{1}{2}}$ with $\mathbf{\Sigma} = \sum_x \rho_x$, where inverses of $\mathbf{\Sigma}$ are taken with respect to its support. Then, the worst-case error of the pretty good measurement ensemble $\{\mathbf{\Lambda}_x\}$ is at most*

$$p_{err} = \max_x \left(1 - \mathrm{Tr}\left\{\mathbf{\Lambda}_x\rho_x\right\}\right) \leq \sum_{x\neq y} \sqrt{\mathsf{F}(\rho_x,\rho_y)}$$

where $\mathsf{F}(\rho_x,\rho_y)$ denotes the fidelity between ρ_x and ρ_y. Moreover, the PGM is optimal if the states in $\{\rho_x\}$ are pair-wise orthogonal, as then $\mathsf{F}(\rho_x,\rho_y) = 0$.

We're going to use the following block-encoding based algorithm in [29] for implementing the PGM. Let κ_ρ denote the reciprocal of the smallest eigenvalue of a density matrix ρ. We use the following theorem.

▶ **Theorem 7** ([29]). *The PGM measurement channel for $\{\rho_x\}_{x\in[2^k]}$ can be implemented with error ϵ (in terms of diamond distance) in time*

$$\widetilde{\mathcal{O}}\left(\sqrt{2^k\kappa_{\bar{\rho}}}N_\rho(\kappa_{\bar{\rho}} + \min(\kappa_\rho, 2^k\cdot\kappa_{\bar{\rho}}/\epsilon^2)\right),$$

where $\bar{\rho} = 2^{-k}\sum_{x\in[2^k]}\rho_x$ and where N_ρ denotes the size of the quantum circuit needed to implement a purification of ρ_x.

In our case, since a purification of ρ_x is

$$\sum_{E\in\bar{\mathcal{P}}_n} \sqrt{\Pr_{E_a\sim\mathcal{D}_p^{\otimes n}}[E_a]}\,|a\rangle\otimes E_a U_{\mathrm{Enc}}^S(|0^{n-k}\rangle\otimes|x\rangle), \tag{4}$$

we may prepare this purification simply by applying U_{Enc}^S on a coherent superposition. The number of gates required to implement an n-qubit Clifford is $N_\rho = O(n^2)$. We show the following theorem.

▶ **Theorem 8** (Single-Shot Decoding for Low Constant Noise Rates). *Let $n,k\in\mathbb{N}$ and $\epsilon\in(0,1)$. Let $\mathcal{D}_p^{\otimes n}$ be the n-qubit depolarizing channel, for some $p\in(0,1/2)$ such that*

$$H(3p) + 3\log_2(3)p + k/n < .99 - \frac{\log(1/\epsilon)}{n}. \tag{5}$$

Then, there exists a quantum algorithm which runs in time

$$\widetilde{\mathcal{O}}\left(n^2\sqrt{2^k\kappa_{\bar{\rho}}}(\kappa_{\bar{\rho}} + \min(\kappa_\rho, 2^k\cdot\kappa_{\bar{\rho}}/\epsilon^2)\right).$$

and solves the $\mathsf{LSN}_{n,k,\mathcal{D}_p^{\otimes n}}$ problem with probability at least $1 - O(\epsilon)$.

[7] This can be understood in another way: the purpose of a good error-correcting code is to encode quantum information in subspaces that do not contract much under quantum channels (i.e. two initial orthogonal states remain approximately orthogonal under the action of quantum channels) – a random stabilizer code is an example of such a code.

Proof. Suppose we are given as input an instance of the $\mathsf{LSN}_{n,k,\mathcal{D}_p^{\otimes n}}$ problem, i.e.,

$$\left(S \in \mathsf{Stab}(n,k), \mathcal{D}_p^{\otimes n}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S)\right) \sim \mathsf{LSN}_{n,k,\mathcal{D}_p^{\otimes n}}$$

where $S \sim \mathsf{Stab}(n,k)$ describes a random stabilizer code and $x \sim \{0,1\}^k$ is a random element. We now show that running Algorithm 2 on input $(S, \mathcal{D}_p^{\otimes n}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S))$ and parameter $\epsilon \in (0,1)$ yields x with the desired success probability.

Algorithm 2, in fact, implements the PGM with respect to a slightly different ensemble as compared to the true ensemble of problem instances. That is to say, it will suffice to implement the pretty good measurement with respect to the ensemble $\{\tilde{\rho}_x^S\}_{x\in\{0,1\}^k}$ where

$$\tilde{\rho}_x^S := \tilde{\mathcal{D}}^{(\frac{3np}{2})}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S) = \sum_{E\in\bar{\mathcal{P}}_n} \tilde{\mathbf{p}}^{(\frac{3n}{2})}(E)\cdot E\,|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S\,E^\dagger, \tag{6}$$

where $\tilde{\mathcal{D}}^{(\frac{3np}{2})}$ is the truncated depolarizing noise channel to be defined shortly; while we remind readers that the true ensemble of problem instances is

$$\rho_x^S := \tilde{\mathcal{D}}^{(\frac{3np}{2})}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S) = \sum_{E\in\bar{\mathcal{P}}_n} \tilde{\mathbf{p}}^{(n)}(E)\cdot E\,|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S\,E^\dagger. \tag{7}$$

We now define the *truncated depolarizing noise channel* as the channel which acts on an input state ρ as

$$\tilde{\mathcal{D}}^{(w)}(\rho) := \sum_{E\in\mathcal{P}_n:|E|\le w} \tilde{\mathbf{p}}^{(w)}(E) E\rho E^\dagger \tag{8}$$

where $p \in (0, 3/4)$ denotes the depolarizing noise parameter, and the truncation lies in the fact that we restrict the support to Paulis with bounded weight. We define the truncated probability distribution via

$$\tilde{\mathbf{p}}^{(w)}(E) := \frac{1}{N}\left(\frac{p}{3}\right)^{|E|}(1-p)^{n-|E|} \tag{9}$$

and N is a normalization factor, i.e. $N = \sum_{E\in\bar{\mathcal{P}}_n:|E|\le w} \tilde{\mathbf{p}}^{(w)}(E)$, that ensures that $\tilde{\mathbf{p}}^{(w)}$ is a probability distribution over $\bar{\mathcal{P}}_n$. The distribution corresponding to the n-qubit local depolarizing noise channel corresponds to $w = n$, i.e. $\tilde{\mathbf{p}}^{(n)}(E) = \Pr_{E\sim\mathcal{D}_p^{\otimes n}}[E]$ and the truncated channel corresponds to acting only with the Paulis with weight at most w, with the same relative probabilities as in n-qubit local depolarizing noise. Because the weights of the Pauli channels in the decomposition of $\mathcal{D}_p^{\otimes n}$ are distributed as a binomial $w \sim \mathsf{Binom}(n,p)$, one can show that for $w = 3np/2$, the total variation distance between the probability distribution over the n-qubit Pauli channels induced by $\mathcal{D}_p^{\otimes n}$ and $\tilde{\mathcal{D}}^{(w)}$ is

$$\delta_{\mathsf{TV}}\left(\tilde{\mathbf{p}}^{(3np/2)}, \tilde{\mathbf{p}}^{(n)}\right) \le \Pr_{E\sim\mathcal{D}_p^{\otimes n}}\left[|E| \ge \frac{3}{2}np\right] \le \exp\left(-\frac{np}{12}\right), \tag{10}$$

using a Chernoff bound.

Algorithm 2 implements the PGM measurement channel with (diamond distance) error $\epsilon \in (0,1)$ with the stated time complexity. Let us first analyze the error probability of the (ideal) pretty good measurement $\{\tilde{\mathbf{\Lambda}}_x^S\}_x$:

Algorithm 2 Pretty Good Measurement for LSN.

Input: Instance $(S \in \mathsf{Stab}(n,k), \mathcal{D}_p^{\otimes n}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S)) \sim \mathsf{LSN}_{n,k,\mathcal{D}_p^{\otimes n}}$ and $\epsilon \in (0,1)$.

Output: A string $x' \in \{0,1\}^k$.

1 Let $\rho_x^S \leftarrow \mathcal{D}_p^{\otimes n}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S)$ denote the ensemble instance.

2 Use the algorithm in Theorem 7 with precision ϵ to measure ρ_x^S via the POVM

$$\{\tilde{\mathbf{\Lambda}}_x^S\}_{x\in\{0,1\}^k} \quad \text{with} \quad \tilde{\mathbf{\Lambda}}_x^S = \mathbf{\Sigma}^{-\frac{1}{2}} \tilde{\rho}_x^S \mathbf{\Sigma}^{-\frac{1}{2}}$$

where $\mathbf{\Sigma} = \sum_x \tilde{\rho}_x^S$ and where inverses of $\mathbf{\Sigma}$ are taken with respect to its support, and where the state $\tilde{\rho}_x^S$ is defined in Equation (6).

3 Output the measurement outcome $x' \in \{0,1\}^k$.

$$\begin{aligned}
\max_x \left(1 - \mathrm{tr}(\widetilde{\mathbf{\Lambda}}_x^S \rho_x^S)\right) &= \max_x \left(1 - \mathrm{tr}\left(\widetilde{\mathbf{\Lambda}}_x \left(\rho_x^S - \tilde{\rho}_x^S\right)\right) - \mathrm{tr}\left(\widetilde{\mathbf{\Lambda}}_x \tilde{\rho}_x^S\right)\right) && (11)\\
&\leq \max_x \left(1 + \delta_{\mathsf{TD}}\left(\rho_x^S, \tilde{\rho}_x^S\right) - \mathrm{tr}\left(\widetilde{\mathbf{\Lambda}}_x \tilde{\rho}_x^S\right)\right) && (12)\\
&\leq \delta_{\mathsf{TV}}\left(\tilde{\mathbf{p}}^{(n)}, \tilde{\mathbf{p}}^{(w)}\right) + \max_x \left(1 - \mathrm{tr}\left(\widetilde{\mathbf{\Lambda}}_x \tilde{\rho}_x^S\right)\right) && (13)\\
&= e^{-\frac{np}{12}} + \max_x \left(1 - \mathrm{tr}\left(\widetilde{\mathbf{\Lambda}}_x \tilde{\rho}_x^S\right)\right) && (14)\\
&\leq e^{-\frac{np}{12}} + \sum_{x\neq y} \sqrt{\mathsf{F}(\tilde{\mathcal{D}}^{(\frac{3np}{2})}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S), \tilde{\mathcal{D}}^{(\frac{3np}{2})}(|\overline{\psi_y}\rangle\langle\overline{\psi_y}|^S))}. && (15)
\end{aligned}$$

The second inequality comes from the convexity of trace distance and Equations (6) and (7). The second last equality comes from Eq. 10. The last inequality comes from Lemma 6.

To finish the proof, it suffices to bound the pair-wise fidelities in Equation (15). Here, we exploit the special structure of the encoded states. We appeal to the quantum Gilbert-Varshamov bound, which states that random stabilizer codes are non-degenerate (i.e. have good distance) with high probability. This means that errors of weight at most $\frac{3}{2}np$ acting on orthogonal states keep them orthogonal. Concretely, this implies that

$$\begin{aligned}
&\Pr_{S\sim\mathsf{Stab}(n,k)}\left[E_a^\dagger E_b \notin N(S), \forall\, |E_a|, |E_b| \leq \frac{3}{2}np\right] && (16)\\
&= \Pr_{S\sim\mathsf{Stab}(n,k)}\left[\langle\overline{\psi_x}|^S E_a^\dagger E_b |\overline{\psi_y}\rangle^S = 0 \,\forall x \neq y, \forall |E_a|, |E_b| \leq \frac{3}{2}np\right] && (17)\\
&\geq 1 - 3/2np \cdot 2^{nH(3/2p)} \cdot 3^{3/2np} \cdot 2^{-n+k}. && (18)
\end{aligned}$$

The first equality follows from the observations mentioned above. Moreover, it follows from Uhlmann's theorem that

$$\begin{aligned}
&\mathsf{F}(\tilde{\mathcal{D}}^{(\frac{3np}{2})}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S), \tilde{\mathcal{D}}^{(\frac{3np}{2})}(|\overline{\psi_y}\rangle\langle\overline{\psi_y}|^S)) && (19)\\
&= \max_U |\langle\phi^{\rho_x}| (U \otimes I) |\phi^{\rho_y}\rangle|^2 && (20)\\
&= \max_U \left| \sum_{E_a, E_b \in \bar{\mathcal{P}}_n} \sqrt{\tilde{\mathbf{p}}^{(\frac{3np}{2})}(E_a)\tilde{\mathbf{p}}^{(\frac{3np}{2})}(E_b)}\, \langle a| U |b\rangle \cdot \langle\overline{\psi_x}| E_a^\dagger E_b |\overline{\psi_y}\rangle \right|^2 && (21)
\end{aligned}$$

where, for $x \in \{0,1\}^k$, we defined the purification

$$|\phi^{\rho_x}\rangle = \sum_{E_a \in \bar{\mathcal{P}}_n} \sqrt{\tilde{\mathbf{p}}^{(\frac{3np}{2})}(E_a)}\ |a\rangle \otimes E_a\, |\overline{\psi_x}\rangle^S .$$

Note that the superposition ranges over Pauli errors of weight at most $\frac{3}{2}np$ as we are using the truncated distribution over Paulis. Therefore, conditioned on the event that the stabilizer code S is non-degenerate and has distance at least $d = 3np+1$, the Knill-Laflamme error correction conditions imply that $\langle\overline{\psi_x}|\, E_a^\dagger E_b\, |\overline{\psi_y}\rangle = 0$, for any pair of codewords with $x, y \in \{0,1\}^k$, and thus

$$\mathsf{F}(\tilde{\mathcal{D}}^{(\frac{3np}{2})}(|\overline{\psi_x}\rangle\langle\overline{\psi_x}|^S), \tilde{\mathcal{D}}^{(\frac{3np}{2})}(|\overline{\psi_y}\rangle\langle\overline{\psi_y}|^S)) = 0. \tag{22}$$

Further conditioning on the event that the implementation of the PGM succeeded, the probability of error due to the PGM mis-identifying the state is $e^{-np/12}$.

We can now put everything together to compute the final success probability of Algorithm 2, union bounding over all the three error sources:

1. S is degenerate.
2. Diamond-distance approximation error of implementing the PGM channel.
3. PGM measurement mis-identifies x.

We get that Algorithm 2 successfully outputs x with probability at least $1-\delta$, where $\delta \leq 3np \cdot 2^{nH(3p)} \cdot 3^{3np} \cdot 2^{-n+k} + \epsilon + e^{-np/12}$.

This algorithm succeeds with constant probability for any noise rate p such that the term

$$3np \cdot 2^{nH(3p)} \cdot 3^{3np} \cdot 2^{-n+k} = \exp[\log(3np) + n(H(3p) + 3\log(3)p + k/n - 1)] \tag{23}$$

does not blow up. Noting that $k/n = O(1)$ in the definition of LSN, we can then check that as long as p is a constant that satisfies

$$H(3p) + 3\log_2(3)p + k/n < .99 - \frac{\log(1/\epsilon)}{n}, \tag{24}$$

this term vanishes exponentially in n and the total probability of error is $O(\epsilon)$. ◀

References

1 Dorit Aharonov, Michael Ben-Or, Elad Eban, and Urmila Mahadev. Interactive proofs for quantum computations, 2017. arXiv:1704.04487.

2 Victor V. Albert and Philippe Faist, editors. *The Error Correction Zoo*. Online website, 2024. URL: https://errorcorrectionzoo.org/.

3 Michael Alekhnovich. More on average case vs approximation complexity. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '03, page 298, USA, 2003. IEEE Computer Society.

4 Anurag Anshu, Nikolas P. Breuckmann, and Chinmay Nirkhe. Nlts hamiltonians from good quantum codes. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, pages 1090–1096, New York, NY, USA, 2023. Association for Computing Machinery. doi:10.1145/3564246.3585114.

5 Benny Applebaum, David Cash, Chris Peikert, and Amit Sahai. Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In *Proceedings of the 29th Annual International Cryptology Conference on Advances in Cryptology*, CRYPTO '09, pages 595–618, Berlin, Heidelberg, 2009. Springer-Verlag. doi:10.1007/978-3-642-03356-8_35.

6 Benny Applebaum, Naama Haramaty-Krasne, Yuval Ishai, Eyal Kushilevitz, and Vinod Vaikuntanathan. Low-Complexity Cryptographic Hash Functions. In Christos H. Papadimitriou, editor, *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*, volume 67 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 7:1–7:31, Dagstuhl, Germany, 2017. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ITCS.2017.7.

7 Srinivasan Arunachalam, Alex Bredariol Grilo, and Aarthi Sundaram. Quantum hardness of learning shallow classical circuits. *SIAM J. Comput.*, 50(3):972–1013, 2021. doi:10.1137/20M1344202.

8 H. Barnum and E. Knill. Reversing quantum dynamics with near-optimal quantum and classical fidelity, 2000. arXiv:quant-ph/0004088.

9 E. Berlekamp, R. McEliece, and H. van Tilborg. On the inherent intractability of certain coding problems (corresp.). *IEEE Transactions on Information Theory*, 24(3):384–386, 1978. doi:10.1109/TIT.1978.1055873.

10 Avrim Blum, Merrick Furst, Michael Kearns, and Richard J. Lipton. Cryptographic primitives based on hard learning problems. In Douglas R. Stinson, editor, *Advances in Cryptology — CRYPTO' 93*, pages 278–291, Berlin, Heidelberg, 1994. Springer Berlin Heidelberg.

11 Avrim Blum, Adam Kalai, and Hal Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model, 2000. arXiv:cs/0010022.

12 Avrim Blum, Adam Kalai, and Hal Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. *J. ACM*, 50(4):506–519, July 2003. doi:10.1145/792538.792543.

13 John Bostanci, Yuval Efron, Tony Metger, Alexander Poremba, Luowen Qian, and Henry Yuen. Unitary complexity and the uhlmann transformation problem, 2023. doi:10.48550/arXiv.2306.13073.

14 Zvika Brakerski, Vadim Lyubashevsky, Vinod Vaikuntanathan, and Daniel Wichs. Worst-case hardness for LPN and cryptographic hashing via code smoothing. Cryptology ePrint Archive, Paper 2018/279, 2018. URL: https://eprint.iacr.org/2018/279.

15 Zvika Brakerski and Vinod Vaikuntanathan. Efficient fully homomorphic encryption from (standard) lwe. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 97–106, 2011. doi:10.1109/FOCS.2011.12.

16 Matthias C. Caro. Binary classification with classical instances and quantum labels. *Quantum Machine Intelligence*, 3(1), May 2021. doi:10.1007/s42484-021-00043-z.

17 Matthias C. Caro, Tom Gur, Cambyse Rouzé, Daniel Stilck França, and Sathyawageeswar Subramanian. Information-theoretic generalization bounds for learning from quantum data. In Shipra Agrawal and Aaron Roth, editors, *Proceedings of Thirty Seventh Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pages 775–839. PMLR, 30 June–03 July 2024. URL: https://proceedings.mlr.press/v247/caro24a.html.

18 Kai-Min Chung and Han-Hsuan Lin. Sample efficient algorithms for learning quantum channels in pac model and the approximate state discrimination problem, 2021. arXiv:1810.10938.

19 Richard Cleve, Debbie Leung, Li Liu, and Chunhao Wang. Near-linear constructions of exact unitary 2-designs. *Quantum Info. Comput.*, 16(9–10):721–756, July 2016. doi:10.26421/QIC16.9-10-1.

20 Bernardo David, Rafael Dowsley, and Anderson C. A. Nascimento. Universally composable oblivious transfer based on a variant of LPN. In Dimitris Gritzalis, Aggelos Kiayias, and Ioannis G. Askoxylakis, editors, *Cryptology and Network Security - 13th International Conference, CANS 2014, Heraklion, Crete, Greece, October 22-24, 2014. Proceedings*, volume 8813 of *Lecture Notes in Computer Science*, pages 143–158. Springer, 2014. doi:10.1007/978-3-319-12280-9_10.

21 Irit Dinur. The pcp theorem by gap amplification. In *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '06, pages 241–250, New York, NY, USA, 2006. Association for Computing Machinery. doi:10.1145/1132516.1132553.

22 Yfke Dulek and Florian Speelman. Quantum ciphertext authentication and key recycling with the trap code, 2018. arXiv:1804.02237.

23 Joseph Emerson, Marcus Silva, Osama Moussa, Colm Ryan, Martin Laforest, Jonathan Baugh, David G. Cory, and Raymond Laflamme. Symmetrized characterization of noisy quantum processes. *Science*, 317(5846):1893–1896, September 2007. doi:10.1126/science.1145699.

24 Marco Fanizza, Yihui Quek, and Matteo Rosati. Learning quantum processes without input control. *PRX Quantum*, 5:020367, June 2024. doi:10.1103/PRXQuantum.5.020367.

25 Vitaly Feldman, Parikshit Gopalan, Subhash Khot, and Ashok Kumar Ponnuswami. New results for learning noisy parities and halfspaces. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '06, pages 563–574, USA, 2006. IEEE Computer Society. doi:10.1109/FOCS.2006.51.

26 Jean-Bernard Fischer and Jacques Stern. An efficient pseudo-random generator provably as secure as syndrome decoding. In Ueli M. Maurer, editor, *Advances in Cryptology - EUROCRYPT '96, International Conference on the Theory and Application of Cryptographic Techniques, Saragossa, Spain, May 12-16, 1996, Proceeding*, volume 1070 of *Lecture Notes in Computer Science*, pages 245–255. Springer, 1996. doi:10.1007/3-540-68339-9_22.

27 Craig Gentry. *A fully homomorphic encryption scheme*. PhD thesis, Stanford University, Stanford, CA, USA, 2009. AAI3382729.

28 Edgar N. Gilbert. A comparison of signalling alphabets. *Bell System Technical Journal*, 31:504–522, 1952. URL: https://api.semanticscholar.org/CorpusID:120421020.

29 András Gilyén, Seth Lloyd, Iman Marvian, Yihui Quek, and Mark M. Wilde. Quantum algorithm for petz recovery channels and pretty good measurements. *Phys. Rev. Lett.*, 128:220502, June 2022. doi:10.1103/PhysRevLett.128.220502.

30 Daniel Gottesman. Stabilizer codes and quantum error correction, 1997. arXiv:quant-ph/9705052.

31 Daniel Gottesman. Surviving as a quantum computer in a classical world, 2024-03 2024. URL: https://www.cs.umd.edu/class/spring2024/cmsc858G/QECCbook-2024-ch1-15.pdf.

32 Daniel Harlow and Patrick Hayden. Quantum computation vs. firewalls. *Journal of High Energy Physics*, 2013(6), June 2013. doi:10.1007/jhep06(2013)085.

33 Patrick Hayden and John Preskill. Black holes as mirrors: quantum information in random subsystems. *Journal of High Energy Physics*, 2007(09):120, September 2007. doi:10.1088/1126-6708/2007/09/120.

34 Christoph Hirche, Cambyse Rouzé, and Daniel Stilck França. Quantum differential privacy: An information theory perspective, 2023. arXiv:2202.10717.

35 Nicholas J. Hopper and Manuel Blum. Secure human identification protocols. In *Proceedings of the 7th International Conference on the Theory and Application of Cryptology and Information Security: Advances in Cryptology*, ASIACRYPT '01, pages 52–66, Berlin, Heidelberg, 2001. Springer-Verlag. doi:10.1007/3-540-45682-1_4.

36 Min-Hsiu Hsieh and Fran çois Le Gall. Np-hardness of decoding quantum error-correction codes. *Phys. Rev. A*, 83:052331, May 2011. doi:10.1103/PhysRevA.83.052331.

37 Pavithran Iyer and David Poulin. Hardness of decoding quantum stabilizer codes. *IEEE Trans. Inf. Theor.*, 61(9):5209–5223, September 2015. doi:10.1109/TIT.2015.2422294.

38 Abhishek Jain, Stephan Krenn, Krzysztof Pietrzak, and Aris Tentes. Commitments and efficient zero-knowledge proofs from learning parity with noise. In Xiaoyun Wang and Kazue Sako, editors, *Advances in Cryptology – ASIACRYPT 2012*, pages 663–680, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg. doi:10.1007/978-3-642-34961-4_40.

39 Ari Juels and Stephen A. Weis. Authenticating pervasive devices with human protocols. In *Proceedings of the 25th Annual International Conference on Advances in Cryptology*, CRYPTO'05, pages 293–308, Berlin, Heidelberg, 2005. Springer-Verlag. doi:10.1007/11535218_18.

40 Kao-Yueh Kuo and Chung-Chin Lu. On the hardness of decoding quantum stabilizer codes under the depolarizing channel. In *2012 International Symposium on Information Theory and its Applications*, pages 208–211, 2012. URL: https://ieeexplore.ieee.org/document/6400919/.

41 Vadim Lyubashevsky. The parity problem in the presence of noise, decoding random linear codes, and the subset sum problem. In *Proceedings of the 8th International Workshop on Approximation, Randomization and Combinatorial Optimization Problems, and Proceedings of the 9th International Conference on Randamization and Computation: Algorithms and Techniques*, APPROX'05/RANDOM'05, pages 378–389, Berlin, Heidelberg, 2005. Springer-Verlag. doi:10.1007/11538462_32.

42 Urmila Mahadev. Classical verification of quantum computations. *SIAM J. Comput.*, 51(4):1172–1229, 2022. doi:10.1137/20M1371828.

43 Robert J. McEliece. A public-key cryptosystem based on algebraic coding theory. DSN Progress Report 42-44, Jet Propulsion Laboratory, Pasadena, CA, 1978.

44 Ashley Montanaro. On the distinguishability of random quantum states. *Communications in Mathematical Physics*, 273(3):619–636, March 2007. doi:10.1007/s00220-007-0221-7.

45 Ketan N. Patel, Igor L. Markov, and John P. Hayes. Optimal synthesis of linear reversible circuits. *Quantum Info. Comput.*, 8(3):282–294, 2008. doi:10.26421/QIC8.3-4-4.

46 Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *Journal of the ACM (JACM)*, 56(6):1–40, 2009. doi:10.1145/1568318.1568324.

47 Adi Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, 1979. doi:10.1145/359168.359176.

48 Graeme Stewart Baird Smith. *Upper and lower bounds on quantum codes*. PhD thesis, California Institute of Technology, USA, 2006. AAI3235592.

49 Jacques Stern. A new identification scheme based on syndrome decoding. In Douglas R. Stinson, editor, *Advances in Cryptology - CRYPTO '93, 13th Annual International Cryptology Conference, Santa Barbara, California, USA, August 22-26, 1993, Proceedings*, volume 773 of *Lecture Notes in Computer Science*, pages 13–21. Springer, 1993. doi:10.1007/3-540-48329-2_2.

50 R. R. Varshamov. Estimation of number of signals in codes with correction of non-symmetric errors. *Avtomat. i Telemekh*, 25(11):1628–1629, 1964. URL: https://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=at&paperid=11783&option_lang=eng.

51 Joel J. Wallman and Joseph Emerson. Noise tailoring for scalable quantum computation via randomized compiling. *Physical Review A*, 94(5), November 2016. doi:10.1103/physreva.94.052325.

52 Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013.

53 Jun Yan. General properties of quantum bit commitments (extended abstract). In Shweta Agrawal and Dongdai Lin, editors, *Advances in Cryptology – ASIACRYPT 2022*, pages 628–657, Cham, 2022. Springer Nature Switzerland. doi:10.1007/978-3-031-22972-5_22.

54 Beni Yoshida and Alexei Kitaev. Efficient decoding for the hayden-preskill protocol, 2017. arXiv:1710.03363.

55 Yu Yu and Jiang Zhang. Smoothing out binary linear codes and worst-case sub-exponential hardness for lpn. In *Advances in Cryptology – CRYPTO 2021: 41st Annual International Cryptology Conference, CRYPTO 2021, Virtual Event, August 16–20, 2021, Proceedings, Part III*, pages 473–501, Berlin, Heidelberg, 2021. Springer-Verlag. doi:10.1007/978-3-030-84252-9_16.