QUERY LEARNING NEARLY PAULI SPARSE UNITARIES IN DIAMOND DISTANCE

Anonymous authors
Paper under double-blind review

ABSTRACT

We study the problem of learning nearly (s, ε)-sparse unitaries, meaning that the Pauli spectrum is concentrated on at most s components with at most ε residual mass in Pauli ℓ_1-norm. This class generalizes well-studied families, including sparse unitaries, quantum k-juntas, 2^k-Pauli dimensional channels, and compositions of depth $O(\log \log n)$ circuits with near-Clifford circuits. Given only forward query access to an unknown nearly sparse unitary U, our goal is to efficiently (both in time and query complexity) construct a quantum mapping that is close in diamond distance to U. We design a learning algorithm achieving this guarantee using $\tilde{O}(s^3/\varepsilon^2)$ (non-controlled) forward queries to U, and running time polynomial in relevant parameters. A key contribution is an efficient quantum algorithm that, given query access to an arbitrary unknown unitary U, estimates all Pauli coefficients (up to a shared global phase) whose magnitude exceeds a given threshold θ. We also study the broader class of unitaries with bounded Pauli ℓ_1-norm. For that class, we prove an exponential query lower bound $\Omega(2^{n/2})$. We introduce a more relaxed accuracy metric, which is the diamond distance restricted to a set of input states. Then, we show that, under this metric, unitaries with Pauli ℓ_1-norm uniformly bounded by L_1 are learnable with $\tilde{O}(L_1^4/\varepsilon^8)$ queries.

1 INTRODUCTION

Understanding the foundations of learning unknown quantum processes is a central problem with direct implications for the characterization and verification of quantum systems. This problem underlies a broad range of tasks, including quantum process tomography, Hamiltonian learning, and the verification and benchmarking of quantum devices (Caro, 2024; Arunachalam et al., 2025; Eisert et al., 2020). Without structural assumptions, however, learning quantum processes is intractable. In particular, learning an unknown n-qubit unitary with error ε in diamond distance requires $\Omega(4^n/\varepsilon)$ queries (Haah et al., 2023). This raises a fundamental question: *what natural structural assumptions and accuracy metrics enable efficient learning of quantum processes, and what are the resulting sample complexity bounds?*

A growing body of work has addressed this question by studying unitaries with locality or sparsity structures in a known basis; most prominently, the Pauli basis (Anshu et al., 2020; Arunachalam et al., 2025; Abbas et al., 2025; Grewal and Liang, 2025). Pauli sparsity leads to a simple description in a physically meaningful operator basis, and parallels classical sparse Boolean functions (Kushilevitz and Mansour, 1993). Formally, an n-qubit operator U can be expressed in the Pauli basis as $U = \sum_{\mathbf{s}} \alpha_{\mathbf{s}} \sigma^{\mathbf{s}}$, where $\sigma^{\mathbf{s}}$ are n-qubit Pauli operators indexed by $\mathbf{s} \in \{0, 1, 2, 3\}^n$ and $\alpha_{\mathbf{s}}$ are the corresponding Pauli coefficients. An operator is s-sparse if it has at most s non-zero Pauli coefficients. In this work, we study the more general class of nearly (s, ε)-sparse unitaries, such as the short-time evolution of the 1D Transverse Field Ising Model, whose Pauli spectrum is concentrated on a set $\mathcal{S}$ of size at most s dominant components, while allowing for a small residual mass ε, i.e., $\sum_{\mathbf{s} \notin \mathcal{S}} |\alpha_{\mathbf{s}}| \leq \varepsilon$.

Nearly sparse unitaries generalize several families, including exactly s-sparse, quantum k-juntas, and compositions of depth $O(\log \log n)$ circuits with near-Clifford circuits, which have been studied in prior work (Huang et al., 2024; Abbas et al., 2025). Particularly, quantum k-juntas, as a special case of 4^k-sparse unitaries, are learnable with $O(4^k/\varepsilon)$ queries (Chen et al., 2023; Bao and Yao, 2023; Montanaro and Osborne, 2010; Grewal and Liang, 2025; Heidari and Szpankowski, 2023).

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

More generally, unitaries whose Pauli support lies within a Pauli subgroup of size 2^k can be learned with optimal query complexity $\Theta(2^k/\varepsilon)$ (Grewal and Liang, 2025). While Grewal and Liang (2025) also give an $O(s^2/\varepsilon^2)$-query algorithm for exactly s-sparse unitaries, learning general nearly sparse unitaries, which allow an ℓ_1-small residual tail, remained open. This work addresses this gap.

1.1 Summary of the main results

Our first contribution is a primitive for recovering the dominant Pauli components of an unknown unitary using only forward queries. This leads to the following result, which is the main technical ingredient underlying our learning results, abbreviated from Theorem 1 in Section 3:

Result 1. *Given a threshold θ, there is an algorithm that accesses an unknown unitary U only through forward queries and estimates all Pauli coefficients of U with magnitude at least θ, up to a global phase, with additive error ε using $\tilde{O}(1/(\varepsilon^2\theta^2))$ queries.*

Unlike prior learning results (Arunachalam et al., 2025; Abbas et al., 2025; Montanaro and Osborne, 2010; Grewal and Liang, 2025; Caro, 2024), our coefficient-estimation primitive applies to arbitrary black-box unitaries without assuming sparsity, subgroup support, Boolean-valued spectra, or access to short-time evolution operators. The estimates $\hat{\alpha}_{\mathbf{s}}$ are close to $e^{i\phi}\alpha_{\mathbf{s}}$ for some unknown $\phi \in [0, 2\pi)$, and this phase ambiguity is unavoidable without access to controlled-U (Tang and Wright, 2025). A notable feature of our method is that it recovers relative phases among Pauli coefficients without requiring controlled-U, unlike Hadamard-test-based approaches (Arunachalam et al., 2024; Montanaro and Osborne, 2010).

Our approach builds upon the estimation algorithm of (Arunachalam et al., 2025) originally developed for Hamiltonian learning via access to short-time evolution operators of the form $U = e^{-iHt}$. We employ Bell sampling on the Choi state of U to efficiently locate the dominant Pauli components. However, the coefficient estimation techniques used in prior Hamiltonian learning works (Arunachalam et al., 2025; Abbas et al., 2025) do not directly extend to arbitrary unitaries, as they rely on linear approximations of the form $U \approx I - itH$, which require small evolution times. For a general unitary, or even the time evolution of a Hamiltonian at large times, such an approximation need not hold.

A further challenge is that existing shadow tomography methods for estimating Pauli observables, such as (Chen et al., 2024; King et al., 2025), are not directly applicable, since the Pauli coefficients of a unitary operator are generally complex-valued rather than real. We overcome these challenges by first estimating the magnitude of a large Pauli coefficient, and then using this estimation to combine Bell sampling with classical shadow tomography based on random Clifford measurements. This enables accurate estimation of both the magnitude and relative phase of the dominant Pauli coefficients of U. Our second contribution uses this coefficient-learning algorithm to learn nearly sparse unitaries in diamond distance. We have the following main results, abbreviated from Theorem 2 in Section 4:

Result 2. *There exists an algorithm that learns nearly (s, ε)-sparse unitaries in diamond distance using $\tilde{O}(s^3/\varepsilon^2)$ queries. Moreover, there exists a polynomial-time quantum algorithm that learns and implements nearly sparse unitaries using the same number of queries, and with additive error $O(\sqrt{s\varepsilon})$ in diamond distance. We additionally establish two refined bounds under practical structural assumptions: (i) if the target unitary has a dominant Pauli coefficient satisfying $|\alpha_{\mathbf{t}}| \geq 1/100$, the query complexity reduces to $\tilde{O}(s^2/\varepsilon^2)$; (ii) if the target unitary has bounded Pauli ℓ_1-norm, the implementation error improves to $O(\sqrt{\varepsilon})$.*

Learning unitaries under the diamond norm poses several challenges. Because the diamond distance depends on the Pauli ℓ_1-norm, one must accurately control the contribution of many coefficients, not just the largest ones. For exactly s-sparse unitaries, it suffices to estimate the s nonzero coefficients to ℓ_∞ accuracy. In contrast, for nearly sparse unitaries, one must contend with estimating a large number of small magnitude coefficients. Moreover, truncating to the largest Pauli coefficients does not, in general, yield a valid unitary operator, and any approximation must also admit an efficient implementation in polynomial time. We overcome these obstacles by leveraging the Linear Combination of Unitaries (LCU) framework (Berry et al., 2014), which allows us to construct an efficiently implementable approximation of the target unitary directly from the estimated Pauli coefficients. Our LCU-based implementation introduces an additional $O(\sqrt{s\varepsilon})$ error.

Table 1: Comparison of query complexities for various learning problems.

Work	Target Object	Query Complexity
(Abbas et al., 2025)	s-sparse Hamiltonian (H)	$\tilde{O}(s^2/(\varepsilon^{5/2}\|H\|_{\text{op}}^{3/2}))$ variable time queries to e^{-iHt}
(Arunachalam et al., 2025)	s-sparse Hamiltonian (H)	$\tilde{O}(s^2/\varepsilon^4)$ fixed time queries to e^{-itH}
(Grewal and Liang, 2025)	unitary with 2^k-dimensional Pauli subspace	$O(2^k/\varepsilon)$ queries to U
(Grewal and Liang, 2025)	s-sparse unitary U	$O(s^2/\varepsilon^2)$ queries to U
Theorem 2	nearly (s,ε)-sparse unitary	$\tilde{O}(s^3/\varepsilon^2)$ queries to U
Corollary 1	nearly (s,ε)-sparse unitary with $\|\alpha_{\mathbf{t}}\| > c^{\dagger}$	$\tilde{O}(s^2/\varepsilon^2)$ queries to U
Theorem 3	unitary with $\|U\|_{1,P} \le L_1$ under restricted diamond distance	$\tilde{O}(L_1^4/\varepsilon^8)$ queries to U

† For any constant $c > 1/100$.

Our third contribution concerns a broader class of unitaries with bounded Pauli ℓ_1-norm , i.e., $\sum_{\mathbf{s}} |\alpha_{\mathbf{s}}| \le L_1$. For this class, we prove that learning in diamond distance is intractable in the worst case. This implies that the diamond distance is too strong for learning this class. Motivated by this barrier, we introduce a restricted diamond distance that measures distinguishability only over a specified family of input states. Since worst-case diamond distance may be unnecessarily strong when the learned unitary is applied only to a restricted family of states, we introduce a restricted diamond distance. This is analogous to the time-concentrated distance used by Abbas et al. (2025) for Hamiltonian learning.

In particular, we focus on the uniformly restricted diamond distance. Unlike standard worst-case metrics, this measure restricts the optimization to states with a maximally mixed marginal, which physically corresponds to the average-case performance of the channel over all pure input states. This leads to the following results (abbreviated from Theorem 3 in Section 5, and Theorem 4 in Section 6):

Result 3. *Any polynomial-time quantum algorithm that learns unitaries with Pauli ℓ_1-norm uniformly bounded by L_1 under diamond distance needs $\Omega(2^{n/2})$ queries, but under the uniformly restricted diamond distance there exists a polynomial-time quantum algorithm with $\tilde{O}(L_1^4/\varepsilon^8)$ queries.*

Finally, in Table 1, we summarize our results in position within the broader literature on operator learning. Existing Hamiltonian learning algorithms (Abbas et al., 2025; Arunachalam et al., 2025) achieve favorable sample complexity by exploiting access to short-time evolutions, but rely on the ability to query the underlying generator. In contrast, our work considers fixed black-box unitaries with an ℓ_1-small Pauli tail and requires no time-control access; unlike the exact-sparse result of Grewal and Liang (2025), our guarantee applies to nearly sparse unitaries.

2 Preliminaries and Problem Formulation

Learning formulation We consider learning an unknown n-qubit unitary U through *coherent* queries. The learner may prepare an entangled state $|\phi\rangle$ on $n+m$ qubits, apply U to a chosen n-qubit subsystem, and jointly measure the resulting state $(U \otimes I)|\phi\rangle$.

Definition 1 (Learning unitaries)**.** *Let $\mathcal{C}$ be a concept class of n-qubit unitaries. We say that $\mathcal{C}$ is learnable with query complexity T and error ε with respect to a metric $d(\cdot,\cdot)$ if there exists a learning protocol that makes T queries to an unknown $U \in \mathcal{C}$ and outputs a hypothesis $\hat{U}$ such that, with probability at least $2/3$, $d(U,\hat{U}) \le \varepsilon$. The query complexity is the minimum such T. The learning is efficient if the protocol runs in time polynomial in n, T, and the other relevant parameters.*

Pauli decomposition The Pauli operators together with the identity $\{\mathbb{I}, X, Y, Z\}$ are denoted as $\{\sigma^0, \sigma^1, \sigma^2, \sigma^3\}$.

Fact 1. *Every linear operator A on n qubits admits a unique Pauli decomposition $A = \sum_{\mathbf{s}\in\{0,1,2,3\}^n} \alpha_{\mathbf{s}}\, \sigma^{\mathbf{s}}$, where the coefficients are given by $\alpha_{\mathbf{s}} = \frac{1}{2^n}\operatorname{Tr}(A\sigma^{\mathbf{s}})$.*

Algorithm 1 Shadow tomography

Input: m copies of a q qubit state $|\psi\rangle$ and observables $\mathcal{O} = \{O_i : i \in [k]\}$
Output: Estimates $\{\hat{o}_i\}_{i=1}^k$ of their expectation values
  function CST($\mathcal{O}, m, |\psi\rangle$)
    for $j \in [m]$ **do**
      Sample a uniformly random Clifford unitary V_j
      Apply V_j to the jth copy of $|\psi\rangle$
      Measure in the computational basis and record $\mathbf{b}_j \in \{0,1\}^n$
    for $O_i \in \mathcal{O}$ **do**
      Compute $\hat{o}_{i,j} \coloneqq (2^q + 1)\langle \mathbf{b}_j | V_j O_i V_j^\dagger | \mathbf{b}_j \rangle - \mathrm{Tr}(O_i)$ for every $j \in [m]$
      Apply the median-of-means estimator to $\{\hat{o}_{i,j}\}_{j=1}^m$ and record the result as $\hat{o}_i$
    return $\{\hat{o}_i\}_{i=1}^k$

Definition 2 (Nearly sparse unitaries)**.** *An n-qubit unitary U is called nearly (s, ε)-sparse in the Pauli basis if there is a set $\mathcal{S}$ of size at most s, such that $\sum_{\mathbf{s} \notin \mathcal{S}} |\alpha_\mathbf{s}| \leq \varepsilon$.*

Bell sampling and Choi-Jamiołkowski isomorphism The Choi-Jamiołkowski isomorphism associates quantum channels with Choi states. In particular, the Choi state of an n-qubit unitary U is the pure $2n$-qubit state

$$|J(U)\rangle \coloneqq (U \otimes \mathbb{I}_{2^n})(\frac{1}{2^{n/2}} \sum_{i=0}^{2^n-1} |i\rangle |i\rangle).$$

The (generalized) Bell states are $|\phi_\mathbf{s}\rangle = (\sigma^\mathbf{s} \otimes \mathbb{I}) |EPR\rangle^{\otimes n}, \mathbf{s} \in \{0,1,2,3\}^n$, where $|EPR\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. The family $\{|\phi_\mathbf{s}\rangle\}_{\mathbf{s} \in \{0,1,2,3\}^n}$ forms an orthonormal basis of the $2n$-qubit bipartite system. Measurement in this basis is called n-qubit Bell sampling (Hangleiter and Gullans, 2024).

Shadow tomography We use shadow tomography based on random Clifford measurements (Huang et al., 2020) to estimate the expectation values of a set of observables $\{O_j : j \in [k]\}$. If $\mathrm{tr}\left(O_j^2\right) \leq B$, then error ε is achieved using $O(B \log(k)/\varepsilon^2)$ copies of the state; see Algorithm 1.

Norms The diamond distance between mappings (e.g., quantum channels) $\mathcal{E}$ and $\mathcal{F}$ is defined as $d_\diamond(\mathcal{E}, \mathcal{F}) \coloneqq \max_\rho \|(\mathbb{I} \otimes (\mathcal{E} - \mathcal{F}))(\rho)\|_1$, where the norm is the trace norm.

Lastly, the Pauli-induced p-norm for an operator A with Pauli coefficients $\{\alpha_\mathbf{s}\}$ is given by $\|A\|_{p,P} \coloneqq (\sum_\mathbf{s} |\alpha_\mathbf{s}|^p)^{1/p}$. See Appendix A for relevant operator norms.

3 Finding and Estimating Large Pauli Coefficients

In this section, we present an algorithm for identifying and estimating the large Pauli coefficients of an unknown unitary U. The algorithm consists of two subroutines: (1) identifying the indices of the large coefficients and (2) estimating their values.

To identify these indices, we perform Bell sampling on the Choi state of U. Measuring $|J(U)\rangle$ in the Bell basis produces an outcome $\mathbf{s} \in \{0,1,2,3\}^n$ with probability $|\alpha_\mathbf{s}|^2$. Repeatedly preparing and measuring $|J(U)\rangle$ therefore produces samples from the distribution $\{|\alpha_\mathbf{s}|^2\}_{\mathbf{s} \in \{0,1,2,3\}^n}$. Specifically, if we take enough samples, we ensure that with high probability, every index whose Pauli coefficient has magnitude above a threshold θ is included in our resulting set $\mathcal{X}$. Algorithm 2 summarizes this procedure, and Lemma 1 establishes the guarantee; its proof appears in Appendix C.

Lemma 1. *Let $S_\theta \coloneqq \{\mathbf{s} : |\alpha_\mathbf{s}| \geq \theta\}$, Algorithm 2 outputs a set $\mathcal{X}$ such that, with probability at least $1 - \delta$, $S_\theta \subseteq \mathcal{X}$. Moreover, $|\mathcal{X}| = O\left(\frac{\log(1/(\theta\delta))}{\theta^2}\right)$.*

Magnitude estimation Given $\mathcal{X}$, next step is to estimate the corresponding Pauli coefficients. Since these coefficients are generally complex, the estimation process requires special care to extract

Algorithm 2 Finding large Pauli coefficients

Input: Query access to unitary U, threshold $\theta \in (0,1)$, and failure probability $\delta \in (0,1)$
Output: Index set $\mathcal{X}$

Set $m_1 \leftarrow O\left(\frac{\log(1/(\theta\delta))}{\theta^2}\right)$ and initialize $\mathcal{X} \leftarrow \emptyset$
Prepare m_1 copies of the Choi state $|J(U)\rangle = (U \otimes \mathbb{I}_{2^n})\,|EPR\rangle^{\otimes n}$
for $j \in [m_1]$ **do**
  Measure the jth copy in the Bell basis and obtain $\mathbf{s} \in \{0,1,2,3\}^n$
  $\mathcal{X} \leftarrow \mathcal{X} \cup \{\mathbf{s}\}$
return $\mathcal{X}$

Algorithm 3 Estimating Pauli coefficients

Input: Query access to a unitary U, index set $\mathcal{X}$, accuracy $\varepsilon \in (0,1)$, and failure probability $\delta \in (0,1)$
Output: Estimates $\{\hat{\alpha}_{\mathbf{s}} : \mathbf{s} \in \mathcal{X}\}$

Set $m_2 \leftarrow O\left(\frac{\log(|\mathcal{X}|/\delta)}{\varepsilon^2}\right)$
Prepare two batches of m_2 copies of the Choi state $|J(U)\rangle$
$\{\widetilde{M}_{\mathbf{s}}\}_{\mathbf{s}\in\mathcal{X}} \leftarrow \mathrm{CST}(\{M_{\mathbf{s}} : \mathbf{s} \in \mathcal{X}\}, m_2, |J(U)\rangle)$
$\mathbf{t} \leftarrow \arg\max_{\mathbf{s}\in\mathcal{X}} \widetilde{M}_{\mathbf{s}}$ and $\hat{\alpha}_{\mathbf{t}} \leftarrow \sqrt{\max\{\widetilde{M}_{\mathbf{t}}, 0\}}$
$\{\widetilde{R}_{\mathbf{s}}, \widetilde{I}_{\mathbf{s}}\}_{\mathbf{s}\in\mathcal{X}} \leftarrow \mathrm{CST}(\{R_{\mathbf{t},\mathbf{s}}, I_{\mathbf{t},\mathbf{s}} : \mathbf{s} \in \mathcal{X}\}, m_2, |J(U)\rangle)$
for $\mathbf{s} \in \mathcal{X} \setminus \{\mathbf{t}\}$ **do**
  $\hat{\alpha}_{\mathbf{s}} \leftarrow \dfrac{\widetilde{R}_{\mathbf{s}} + i\widetilde{I}_{\mathbf{s}}}{\hat{\alpha}_{\mathbf{t}}}$
return $\{\hat{\alpha}_{\mathbf{s}} : \mathbf{s} \in \mathcal{X}\}$

their relative phases. We first estimate $|\alpha_{\mathbf{s}}|^2$ for every $\mathbf{s} \in \mathcal{X}$ to accuracy ε and let $\mathbf{t}$ index the largest estimate. We set $\hat{\alpha}_{\mathbf{t}}$ to the estimated value of $|\alpha_{\mathbf{t}}|$, that is we ignore the phase of $\alpha_{\mathbf{t}}$. These magnitudes are obtained from the observables

$$M_{\mathbf{s}} \coloneqq |\phi_{\mathbf{s}}\rangle\langle\phi_{\mathbf{s}}|\,, \quad \mathbf{s} \in \mathcal{X},$$

since $\langle J(U)|M_{\mathbf{s}}|J(U)\rangle = |\alpha_{\mathbf{s}}|^2$. q We then estimate the remaining coefficients relative to $\alpha_{\mathbf{t}}$. Since we do not know the phase of $\alpha_{\mathbf{t}}$, we can only estimate other coefficients up to the same global phase. For each $\mathbf{s} \in \mathcal{X} \setminus \{\mathbf{t}\}$, define

$$R_{\mathbf{t},\mathbf{s}} \coloneqq \frac{1}{2}(|\phi_{\mathbf{t}}\rangle\langle\phi_{\mathbf{s}}| + |\phi_{\mathbf{s}}\rangle\langle\phi_{\mathbf{t}}|), \qquad I_{\mathbf{t},\mathbf{s}} \coloneqq \frac{1}{2}(-i\,|\phi_{\mathbf{t}}\rangle\langle\phi_{\mathbf{s}}| + i\,|\phi_{\mathbf{s}}\rangle\langle\phi_{\mathbf{t}}|)$$

Their expectation values satisfy

$$\langle J(U)|R_{\mathbf{t},\mathbf{s}}|J(U)\rangle = \mathrm{Re}(\alpha_{\mathbf{t}}^*\alpha_{\mathbf{s}}), \qquad \langle J(U)|I_{\mathbf{t},\mathbf{s}}|J(U)\rangle = \mathrm{Im}(\alpha_{\mathbf{t}}^*\alpha_{\mathbf{s}}),$$

and therefore determine every $\alpha_{\mathbf{s}}$ up to the global phase fixed by $\alpha_{\mathbf{t}}$. To reduce the sample complexity of calculating the expectation value, we use shadow tomography with Clifford measurements. The resulting sample complexity depends only logarithmically with $|\mathcal{X}|$ with polynomial time complexity. This is summarized in Algorithm 3.

The following lemma establishes the formal accuracy guarantees for Algorithm 3, bounding the estimation error of the recovered Pauli coefficients. The full proof is deferred to Appendix D.

Lemma 2. *Let $\mathcal{X} \subseteq \{0,1,2,3\}^n$, and run Algorithm 3 with shadow-estimation accuracy ε and failure probability δ. If $\varepsilon \le |\alpha_{\mathbf{t}}|^2/2$, Then, with probability at least $1-\delta$, there is a phase $\phi = \arg(\alpha_{\mathbf{t}})$ such that, for every $\mathbf{s} \in \mathcal{X}$,*

$$\left|\hat{\alpha}_{\mathbf{s}} - e^{-i\phi}\alpha_{\mathbf{s}}\right| \le 4\sqrt{2}\,\frac{\varepsilon}{|\alpha_{\mathbf{t}}|}.$$

Theorem 1 (Estimation of large Pauli coefficients)**.** *Let $\theta, \varepsilon, \delta \in (0,1)$ with $\varepsilon \le \theta$, and define $S_\theta := \{\mathbf{s} : |\alpha_{\mathbf{s}}| \ge \theta\}$. Suppose that S_θ is nonempty. Given forward query access to an n-qubit*

unitary U there is an algorithm that, with probability at least $1-\delta$, outputs an index set X and estimates $\{\hat{\alpha}_{\mathbf{s}} : \mathbf{s} \in X\}$ satisfying $S_\theta \subseteq \mathcal{X}$ and

$$\left|\widehat{\alpha}_s - e^{-i\phi}\alpha_s\right| \leq \varepsilon \qquad \forall \mathbf{s} \in \mathcal{X},$$

for a shared global phase $\phi \in [0, 2\pi)$. The algorithm uses $O\left(\frac{\log(1/(\theta\delta))}{\varepsilon^2\theta^2}\right)$ queries to U and runs in time polynomial in n, $1/\theta$, $1/\varepsilon$, and $\log(1/\delta)$.

Proof. We run Algorithm 2 with threshold θ and failure probability $\delta/2$. By Lemma 1, with probability at least $1-\delta/2$, the resulting set $\mathcal{X}$ satisfies $S_\theta \subseteq \mathcal{X}$, and $|\mathcal{X}| = O\left(\frac{\log(1/(\theta\delta))}{\theta^2}\right)$. Next, we run Algorithm 3 on $\mathcal{X}$, with failure probability $\delta/2$ and accuracy ε_s. By the classical-shadow guarantee, with probability at least $1-\delta/2$, $\left|\widetilde{M}_{\mathbf{s}} - |\alpha_{\mathbf{s}}|^2\right| \leq \varepsilon_s$. Because S_θ is nonempty and $S_\theta \subseteq \mathcal{X}$, there exists $r \in \mathcal{X}$ such that $|\alpha_r| \geq \theta$. By the definition of $\mathbf{t}$, $\widetilde{M}_{\mathbf{t}} \geq \widetilde{M}_r$. It follows that

$$|\alpha_{\mathbf{t}}|^2 \geq \widetilde{M}_{\mathbf{t}} - \varepsilon_s \geq \widetilde{M}_r - \varepsilon_s \geq |\alpha_r|^2 - 2\varepsilon_s \geq \theta^2 - 2\varepsilon_s.$$

Choosing $\varepsilon_s \leq \theta^2/4$ ensures $|\alpha_{\mathbf{t}}|^2/2 \geq \varepsilon_s$. Then, by Lemma 2, with probability $1-\delta/2$, there is a phase $\phi = \arg(\alpha_{\mathbf{t}})$ such that, for every $\mathbf{s} \in \mathcal{X}$, $\left|\hat{\alpha}_{\mathbf{s}} - e^{-i\phi}\alpha_{\mathbf{s}}\right| \leq 4\sqrt{2}\,\varepsilon_s/|\alpha_{\mathbf{t}}|$. Choosing $\varepsilon = 8\varepsilon_s/\theta$, ensures $\left|\hat{\alpha}_{\mathbf{s}} - e^{-i\phi}\alpha_{\mathbf{s}}\right| \leq \varepsilon$. Finally, the total query complexity is the sum

$$m_1 + m_2 = O\left(\frac{\log(1/(\theta\delta))}{\theta^2}\right) + O\left(\frac{\log(|\mathcal{X}|/\delta)}{\varepsilon_s^2}\right) = O\left(\frac{\log(1/\theta\delta)}{\varepsilon^2\theta^2}\right).$$

The running time is polynomial in all parameters. In particular, because M_s, $R_{t,s}$, and $I_{t,s}$ are defined using the generalized Bell states $|\psi_{\mathbf{s}}\rangle = (\sigma^{\mathbf{s}} \otimes I)\,|EPR\rangle^{\otimes n}$, then by the stabilizer's formulation and the Gottesman-Knill theorem the random Clifford shadow terms in Algorithm 1 are computable in polynomial time for such observables. □

Although the algorithm estimates the Pauli coefficients up to a global phase, this ambiguity does not affect the final learning guarantee.

4 Learning Unitary Operations

Our goal is to learn an unknown n-qubit unitary U such that the learned unitary $\hat{U}$ is close to U in diamond distance. Two key challenges need to be addressed: (1) connecting the diamond distance to the errors in the estimated Pauli coefficients of the unitary, obtained from Theorem 1 and (2) efficiently constructing an estimate $\hat{U}$ to U from the estimated Pauli coefficients in polynomial time. We first address the second challenge.

Efficient implementation via approximate LCU Suppose we have identified $\hat{U} = \sum_{\mathbf{s}\in\mathcal{X}} \hat{\alpha}_{\mathbf{s}}\sigma_{\mathbf{s}}$ as a sparse approximation to the target unitary U. Although $\hat{U}$ is a linear combination of Pauli operators, it need not be exactly unitary and therefore may not be directly implementable as a quantum circuit. But we can "promote" $\hat{U}$ to an implementable (near-)unitary using the LCU technique. Suppose that $\hat{U}$ is γ-close to U in Pauli ℓ_1-norm, i.e., $a := \sum_{\mathbf{s}\in\mathcal{X}} |\hat{\alpha}_{\mathbf{s}}|$ and $\gamma := \|U - \hat{U}\|_{1,P}$.

Observation 1 (Approximate LCU, c.f. Thm 2.5 of (Kothari, 2014))**.** *There exists a quantum algorithm that implements $\hat{U}$ as a circuit W with error $O(a\sqrt{\gamma})$ in operator norm. Furthermore, the algorithm runs in poly$(a, n, |\mathcal{X}|)$ time.*

With this Observation, we just need to control the parameters a and γ to ensure an efficient implementation of a unitary close to U.

Learning nearly sparse unitaries We now apply the coefficient-estimation procedure to the class of nearly sparse unitaries introduced in Definition 2. The main idea is to identify all coefficients above a suitable threshold, estimate them using a large reference coefficient, and discard coefficients whose estimated magnitudes are too small.

Theorem 2. *Let U be a nearly (s,ε)-sparse n-qubit unitary, where $0 < \varepsilon \leq 1/2$, and let $\delta \in (0,1)$. There exists an algorithm that, with probability at least $1-\delta$, outputs a sparse operator $\hat{U}$ that is $O(\varepsilon)$-close to U in diamond distance. This algorithm uses $\tilde{O}(s^3/\varepsilon^2)$ queries and $\mathrm{poly}(n, s, 1/\varepsilon, \log(1/\delta))$ time. Moreover, the algorithm can construct a unitary circuit W using $\mathrm{poly}(n, s, 1/\varepsilon, \log(1/\delta))$ gates such that $d_\diamond(\mathcal{W},\mathcal{U}) = O(\sqrt{s\varepsilon})$.*

Proof. By Definition 2, there exists a set S of size at most s such that $\sum_{\mathbf{s}\notin S}|\alpha_\mathbf{s}| \leq \varepsilon$. Let $\tilde{U} = \sum_{\mathbf{s}\in S}\alpha_\mathbf{s}\sigma_\mathbf{s}$ be the sparse truncation of U to S. Since $\|U-\tilde{U}\|_{2,P} \leq \|U-\tilde{U}\|_{1,P} \leq \varepsilon$, and U is unitary, so $\|U\|_{2,P} = 1$, we have

$$\|\tilde{U}\|_{2,P} \geq \|U\|_{2,P} - \|U-\tilde{U}\|_{2,P} \geq 1-\varepsilon.$$

Hence, there exists $\mathbf{r} \in S$ such that $|\alpha_\mathbf{r}| \geq \frac{1-\varepsilon}{\sqrt{s}} \geq \frac{\varepsilon}{s}$ where the last inequality follows from $\varepsilon \leq 1/2$. We first run Algorithm 2 with threshold $\theta = \varepsilon/s$ and failure probability $\delta/2$. By Lemma 1, with probability at least $1-\delta/2$, the resulting set $\mathcal{X}$ contains every $\mathbf{s}$ satisfying $|\alpha_\mathbf{s}| \geq \theta$, in particular, contains $\mathbf{r}$. This uses $m_1 = \tilde{O}(s^2/\varepsilon^2)$ queries. We next run Algorithm 3 on $\mathcal{X}$, with failure probability $\delta/2$ and shadow estimation error $\varepsilon_s = ((1-\varepsilon)\varepsilon)/(32s^{3/2})$. Let $\mathbf{t}$ be the maximum index selected by the algorithm. By shadow tomography assumption, $|\hat{\alpha}_\mathbf{t}|^2 \geq |\alpha_\mathbf{r}|^2 - 2\varepsilon_s \geq \frac{(1-\varepsilon)^2}{2s}$. In particular, $\varepsilon_s \leq |\alpha_\mathbf{t}|^2/2$, so Lemma 2 gives a phase ϕ such that, for every $\mathbf{s} \in \mathcal{X}$,

$$\left|\hat{\alpha}_\mathbf{s} - e^{-i\phi}\alpha_\mathbf{s}\right| \leq 4\sqrt{2}\frac{\varepsilon_s}{|\alpha_\mathbf{t}|} \leq \frac{\varepsilon}{4s} =: \varepsilon_p.$$

By a union bound, the preceding events hold simultaneously with probability at least $1-\delta$. This would take $m_2 = \tilde{O}\left(s/\varepsilon^2\right) + \tilde{O}\left(1/\varepsilon_s^2\right) = \tilde{O}\left(s^3/\varepsilon^2\right)$ queries. From here on we work with the phase-corrected estimates $e^{i\phi}\hat{\alpha}_s$ and write $\hat{\alpha}_s$ for these; this costs nothing since ϕ enters only as a global rotation and is absorbed later by $d_\diamond(W,U) \leq 2\|W-U\|_{\mathrm{op}}$ via Lemma 4. We now filter the estimated coefficients by defining $\mathcal{X}_{\varepsilon_p} = \{\mathbf{s} \in \mathcal{X} : |\hat{\alpha}_\mathbf{s}| \geq 2\varepsilon_p\}$ and $\hat{U} = \sum_{\mathbf{s}\in\mathcal{X}_{\varepsilon_p}}\hat{\alpha}_\mathbf{s}\sigma_\mathbf{s}$. if $\mathbf{s} \in \mathcal{X}_{\varepsilon_p}$, then $|\alpha_\mathbf{s}| \geq |\hat{\alpha}_\mathbf{s}| - \varepsilon_p \geq \varepsilon_p$. Consequently, $\left|\mathcal{X}_{\varepsilon_p} \setminus S\right|\varepsilon_p \leq \varepsilon$ by definition of nearly sparse, and so $\left|\mathcal{X}_{\varepsilon_p}\right| \leq s + \frac{\varepsilon}{\varepsilon_p} \leq 5s$. Thus, although $\mathcal{X}_{\varepsilon_p}$ need not be contained in S, the filtered operator remains $O(s)$-sparse. For an index $\mathbf{s} \in S \setminus \mathcal{X}_{\varepsilon_p}$, either $\mathbf{s} \notin \mathcal{X}$, in which case $|\alpha_\mathbf{s}| < \theta$, or $\mathbf{s} \in \mathcal{X}$ but was removed by the filtering step, in which case $|\alpha_\mathbf{s}| \leq |\hat{\alpha}_\mathbf{s}| + \varepsilon_p \leq 3\varepsilon_p$. Consequently,

$$\begin{aligned}\left\|U-\hat{U}\right\|_{1,P} &\leq |\mathcal{X}_{\varepsilon_p}|\varepsilon_p + s\max\{\theta, 3\varepsilon_p\} + \sum_{\mathbf{s}\notin S}|\alpha_\mathbf{s}| \\ &\leq s\varepsilon_p + 2\varepsilon + s\max\{\theta, 3\varepsilon_p\} = O(\varepsilon).\end{aligned}$$

Writing $\gamma := \|U-\hat{U}\|_{1,P}$, we have $\|\hat{U}\|_{\mathrm{op}} \leq 1+\gamma$. Thus, by Lemma 4, $d_\diamond(\widehat{\mathcal{U}},\mathcal{U}) \leq (2+\gamma)\gamma = O(\varepsilon)$ and then the total query complexity is $m_1 + m_2 = \tilde{O}(\frac{s^3}{\varepsilon^2})$.

Lastly, applying LCU to implement $\hat{U}$ incurs additional error $O(a\sqrt{\gamma})$ in operator norm per Observation 1. In our case, $\gamma = 2\varepsilon$ and $a = \sum_{\mathbf{s}\in\mathcal{X}_{\varepsilon_p}}|\hat{\alpha}_\mathbf{s}| \leq O(\sqrt{s})\left\|\hat{U}\right\|_{2,p} = O(\sqrt{s})$, as $\hat{U}$ is $O(s)$-sparse. Thus, the overall error is bounded by $O(\sqrt{s\varepsilon})$. Because both W and U are exactly unitary, we can now safely apply Lemma 4, thus $d_\diamond(\mathcal{W},\mathcal{U}) \leq 2\|W-U\|_{op} = O(\sqrt{s\varepsilon})$. All steps run in $\mathrm{poly}(n, s, 1/\varepsilon, \log(1/\delta))$ time. □

Corollary 1 (Improved complexity for unitaries with dominant Pauli coefficient)**.** *If in addition to being nearly (s,ε)-sparse in Theorem 2, the target unitary U has a dominant Pauli coefficient satisfying $|\alpha_\mathbf{t}| \geq c$ for some constant $c > 1/100$. Then the query complexity reduces to $\tilde{O}(\frac{s^2}{\varepsilon^2})$.*

Corollary 2 (Improved complexity for bounded-norm unitaries)**.** *If, in addition to being nearly (s,ε)-sparse in Theorem 2, the ℓ_1 norm is bounded by a constant, then there exists a polynomial time algorithm that learns U with $O(\sqrt{\varepsilon})$ error in diamond distance using $\tilde{O}(\frac{s^3}{\varepsilon^2})$. If additionally, $|\alpha_\mathbf{t}| \geq c$ for some constant $c > 1/100$, then the algorithm needs $\tilde{O}(\frac{s^2}{\varepsilon^2})$ queries.*

We note that improving the $\tilde{O}(s^3/\varepsilon^2)$ query complexity appears nontrivial within our approach. The main bottleneck is that diamond-distance learning requires controlling the total Pauli ℓ_1 error, so each significant coefficient must be estimated to precision on the order of ε/s. Moreover, our estimation procedure is based on recovering coefficient ratios, and without any assumptions beyond sparsity the largest coefficient can be as small as $\Theta(1/\sqrt{s})$, which tightens the required per-coefficient accuracy to roughly $\varepsilon/s^{3/2}$. Finally, estimating a Pauli coefficient to additive error δ using shadow tomography leads to the $\tilde{O}(s^3/\varepsilon^2)$ scaling.

5 Learning unitaries with bounded ℓ_1 norm

In this section, we show that the learnability holds under a more relaxed notion of closeness in diamond norm, which we will define later. Our lower bound results in Section 6 show that bounding the Pauli ℓ_1-norm alone does not suffice to ensure efficient learnability. In many practical settings, one is not concerned with the action of a quantum channel on *all* possible input states, but rather with its behavior on a specific subset of states of interest. This naturally leads to a relaxed distinguishability measure that evaluates closeness only on such restricted inputs.

Definition 3 (Restricted diamond distance)**.** *Let $\mathcal{U}$ and $\hat{\mathcal{U}}$ be the quantum channels induced by unitary U and linear mapping $\hat{U}$ acting on a system A, and let $\mathcal{A}$ be a specified set of states on A. The* restricted diamond distance *between $\mathcal{U}$ and $\hat{\mathcal{U}}$ with respect to $\mathcal{A}$ is defined as*

$$d_{\diamond,\mathcal{A}}(\mathcal{U},\hat{\mathcal{U}}) := \sup_{R} \sup_{\rho_{RA}:\,\mathrm{Tr}_R(\rho_{RA})\in\mathcal{A}} \left\|(\mathbb{I}_R\otimes\mathcal{U})(\rho_{RA}) - (\mathbb{I}_R\otimes\hat{\mathcal{U}})(\rho_{RA})\right\|_1, \tag{1}$$

where the supremum is taken over all finite-dimensional reference systems R.

Under this definition, the distinguishability of two unitaries is evaluated only on those joint input states whose marginal on the system of interest lies in $\mathcal{A}$. Thus, $d_{\diamond,\mathcal{A}}$ quantifies the worst-case deviation between $\mathcal{U}$ and $\hat{\mathcal{U}}$ when restricted to the relevant class of inputs. At one extreme, when $\mathcal{A}$ consists of all states on A, the restricted diamond distance reduces to the standard diamond distance.

Connection to ℓ_2-norm For unitary channels, the diamond distance admits dimension-dependent upper and lower bounds in terms of the Hilbert–Schmidt inner product (equivalently, the entanglement infidelity). We now establish a sharper characterization for the restricted diamond distance, which avoids this dimensional dependence and directly relates the distance to expectation values of the overlap operator with respect to admissible input states.

We study the connection of uniformly restricted diamond distance—which captures average-case performance—directly via the Pauli ℓ_2-norm (see Lemma 5 in Appendix E). Furthermore, we establish in Lemma 6 (Appendix E) that ℓ_1-bounded approximations trivially imply the existence of sparse approximations, allowing us to present our main result on learning unitaries with bounded Pauli ℓ_1-norm under the restricted diamond distance, which is proved in Appendix F.

Theorem 3. *The query complexity of learning n-qubit unitaries, with Pauli ℓ_1-norm bounded by L_1, and with $O(\varepsilon)$ error in the restricted diamond distance $d_{\diamond,\{\mathbb{I}/N\}}$, is $\tilde{O}\left(L_1^4/\varepsilon^8\right)$.*

6 Lower bounds

This section establishes lower bounds for learning sparse and nearly sparse unitaries, as well as for unitaries with bounded Pauli ℓ_1-norm. We first provide a lower bound for learning exactly sparse unitaries, and then extend it to nearly sparse unitaries.

Corollary 3 (c.f. (Grewal and Liang, 2025))**.** *Any algorithm that learns the class of n-qubit unitary U with at most s non-zero Pauli coefficients to error ε in diamond distance and with success probability at least $2/3$ must make at least $\Omega(\frac{s}{\varepsilon})$ queries to U.*

Our next result shows that the bounded Pauli ℓ_1 norm alone does not suffice to ensure efficient learnability, which is proved in Appendix G.

Theorem 4. *Any algorithm that learns the class of n-qubit unitary U with Pauli ℓ_1 norm at most $L = O(1)$ to error $\varepsilon = 1/10$ in diamond distance and with success probability at least $2/3$ must make at least $\Omega(2^{n/2})$ queries to U.*

7 SPECIAL CASES AND EXAMPLES

This section presents some special unitaries that are nearly sparse or have bounded ℓ_1 norm. The technical proofs of the claims are available in Appendix H.

Lemma 3 (Propagation of Pauli ℓ_1-norm Hamiltonians)**.** *Let H be a Hamiltonian with bounded Pauli ℓ_1-norm, i.e., $\|H\|_{1,P} \leq L_H$. Then Pauli ℓ_1-norm of the unitary evolution $U = e^{iH}$ is bounded by e^{L_H}. Consequently, U can be learned using Theorem 3 with sample complexity dependent on e^{L_H}.*

The relationship between classical Boolean functions and diagonal Hamiltonians provides a rich source of examples for bounded-norm unitaries. Given a real-valued function $f : \{0,1\}^n \to \mathbb{R}$ define the Hamiltonian $H_f = \sum_x f(x) |x\rangle\langle x|$, and its corresponding unitary evolution $U_f = e^{iH_f}$.

Classes such as DNFs and decision trees have bounded L1 norm (Blum et al., 1994; Khardon, 1994); therefore, their corresponding unitaries and Hamiltonians have bounded ℓ_1 norm. These structures appear naturally in combinatorial optimization problems, such as the Quantum Approximate Optimization Algorithm (QAOA) (Farhi et al., 2014). For instance, the normalized Max-Cut Hamiltonian for a graph $G = (V, E)$ is given by $H_C = 1/(2|E|) \sum_{(i,j)\in E(G)} (Z_i Z_j - I)$. Here, the Pauli ℓ_1-norm is $O(1)$. Consequently, Lemma 3 implies that the unitary e^{iH_C} can be learned efficiently by our algorithm.

Remark 1. *Identify each Z-string with a bitmask $z \in \{0,1\}^n$, where $z_j = 1$ indicates a Z acting on qubit j, and $\chi_z(x) := (-1)^{z\cdot x}$ be the characters. Then, the Z-Pauli coefficients of U_f are the Boolean Fourier coefficients of the function $g(x) := e^{if(x)}$:*

$$\alpha_z = \hat{g}_z := 2^{-n} \sum_{x\in\{0,1\}^n} e^{if(x)} \chi_z(x),$$

Example 1. *Let $U = \prod_{j=1}^m e^{-i\theta_j P_j}$, where P_j are Pauli strings, and set $\Lambda = \sum_j |\tan\theta_j|$. Truncating the expansion to products containing at most r rotations gives a sparse operator with $s \leq \sum_{k=0}^r \binom{m}{k}$,$\|U - U_{\leq r}\|_{1,P} \leq e^{\Lambda} \frac{\Lambda^{r+1}}{(r+1)!}$, so U is nearly $\left(s, e^{\Lambda}\Lambda^{r+1}/(r+1)!\right)$-sparse.*

Proposition 1 (Multi-controlled phase unitary)**.** *For any k-qubit system ($k \leq n$), define the unitary that applies a phase ϕ to the $|1^k\rangle$ state as below. Then, $\|U_{k,\phi}\|_{1,P} \leq 3$.*

$$U_{k,\phi} := I + (e^{i\phi} - 1) |1^k\rangle\langle 1^k|.$$

Proposition 2. *For any n-qubit Grover diffusion operator D, $\|D\|_{1,P} < 3$.*

8 CONCLUSION AND FUTURE WORK

This work studies the problem of learning unitaries that are nearly sparse in the Pauli basis. We provide an upper bound on the query complexity and establish an efficient algorithm for learning this class of unitaries in diamond distance. We also propose an algorithm for recovering large Pauli coefficients. We then extend our results to unitaries with bounded Pauli ℓ_1-norm and introduce a relaxed distance measure, called the restricted diamond distance, for which we provide upper and lower bounds on the query complexity of learning this class of unitaries. Lastly, we present several examples of unitaries with small Pauli ℓ_1-norm, including Hamiltonians corresponding to Boolean functions with bounded ℓ_1-norm.

Our approach has two primary limitations. First, our implementation relies on the LCU framework to construct a circuit from the estimated Pauli coefficients, which introduces an unavoidable additive error of $O(\sqrt{s\varepsilon})$ in diamond distance. Reducing or eliminating this overhead is a fundamental open problem in quantum simulation that falls outside the scope of this work. Second, our framework assumes active query access to U, whereas passive learning from input-output samples is arguably more practical. Extending our results to the passive (PAC) learning setting is a natural direction for future work. Other future directions include improving the query complexity of learning nearly sparse unitaries in diamond distance using different techniques, such as Heisenberg scaling and improved methods for implementing the approximated unitaries.

AI USE STATEMENT

In this work, we used generative AI tools to assist with mathematical reasoning and manuscript preparation. In particular, a generative AI tool suggested Equation (4) that led us to improve one of our theoretical bounds. The authors independently derived and verified this identity, checked the conditions under which it applies, and reworked and verified the complete resulting argument and all downstream claims. Generative AI tools were also used to help draft and rephrase portions of the manuscript and to improve its clarity, grammar, and presentation.

We did not use generative AI tools to generate or analyze data, design or conduct experiments, implement the proposed methods, or interpret empirical results; these tasks are not applicable to this theoretical work. Generative AI tools were not used as sources of scientific evidence or bibliographic authority. All mathematical statements, proofs, citations, and related-work claims were reviewed and verified by the authors against the relevant arguments and primary sources.

We have reviewed all AI-assisted work and take responsibility for the final content of this paper, including all text, mathematical claims, proofs, and other material produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

In this work, we have provided all the needed theorems and proofs. Proof of main theorems are in the main body, and the remaining are in the appendix. This paper is entirely theoretical, and does not contain any experiments, datasets, or computational resources to reproduce the results.

REFERENCES

A. Abbas, N. Cerrato, F. E. Gutiérrez, D. Grinko, F. A. Mele, and P. Sinha. Nearly optimal algorithms to learn sparse quantum hamiltonians in physically motivated distances. *arXiv preprint arXiv:2509.09813*, 2025.

H. Aizenstein, A. Blum, R. Khardon, E. Kushilevitz, L. Pitt, and D. Roth. On learning read-k-satisfy-j DNF. *SIAM Journal on Computing*, 27(6):1515–1530, 1998.

A. Angrisani. Learning unitaries with quantum statistical queries. *Quantum*, 9:1817, 2025.

A. Anshu, S. Arunachalam, T. Kuwahara, and M. Soleimanifar. Sample-efficient learning of quantum many-body systems. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 685–691. IEEE, 2020.

S. Arunachalam and R. de Wolf. A survey of quantum learning theory. *arXiv:1701.06806*, 2017.

S. Arunachalam and R. De Wolf. Optimal quantum sample complexity of learning algorithms. *J. Mach. Learn. Res.*, 19(1):2879–2878, Jan. 2018. ISSN 1532-4435.

S. Arunachalam, A. Dutt, F. E. Gutiérrez, and C. Palazuelos. Learning low-degree quantum objects. In *51st International Colloquium on Automata, Languages, and Programming (ICALP 2024)*, volume 297 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 13:1–13:19. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024. doi: 10.4230/LIPIcs.ICALP.2024.13.

S. Arunachalam, A. Dutt, and F. Escudero Gutiérrez. Testing and learning structured quantum hamiltonians. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 1263–1270, 2025.

Z. Bao and P. Yao. On testing and learning quantum junta channels. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 1064–1094. PMLR, 2023.

K. Beer, D. Bondarenko, T. Farrelly, T. J. Osborne, R. Salzmann, D. Scheiermann, and R. Wolf. Training deep quantum neural networks. *Nature communications*, 11(1):808, 2020.

A. Belovs. Quantum algorithms for learning symmetric juntas via the adversary bound. *computational complexity*, 24(2):255–293, 2015.

D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma. Exponential improvement in precision for simulating sparse hamiltonians. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 283–292, 2014.

A. Blum, M. Furst, J. Jackson, M. Kearns, Y. Mansour, and S. Rudich. Weakly learning dnf and characterizing statistical query learning using fourier analysis. In *Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*, pages 253–262, 1994.

N. H. Bshouty. Exact learning boolean function via the monotone theory. *Information and Computation*, 123(1):146–153, 1995.

N. H. Bshouty, J. C. Jackson, and C. Tamon. More efficient PAC-learning of DNF with membership queries under the uniform distribution. *Journal of Computer and System Sciences*, 68(1):205–234, 2004.

M. C. Caro. Learning quantum processes and hamiltonians via the pauli transfer matrix. *ACM Transactions on Quantum Computing*, 5(2):1–53, 2024.

S. Chen, W. Gong, and Q. Ye. Optimal tradeoffs for estimating pauli observables. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1086–1105. IEEE, 2024.

T. Chen, S. Nadimpalli, and H. Yuen. Testing and learning quantum juntas nearly optimally. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1163–1185. SIAM, 2023.

J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi. Quantum certification and benchmarking. *Nature Reviews Physics*, 2(7):382–390, 2020.

E. Farhi, J. Goldstone, and S. Gutmann. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014.

V. Feldman. Attribute-efficient and non-adaptive learning of parities and DNF expressions. *Journal of Machine Learning Research*, 8:1431–1460, 2007.

V. Feldman. Learning DNF expressions from Fourier spectrum. In S. Mannor, N. Srebro, and R. C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 17.1–17.19, 2012.

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*, pages 25–32. ACM Press, 1989.

S. Grewal and D. Liang. Efficient learning of structured quantum circuits via pauli dimensionality and sparsity. *arXiv preprint arXiv:2510.00168*, 2025.

G. Gutoski and N. Johnston. Process tomography for unitary quantum channels. *Journal of Mathematical Physics*, 55(3), 2014.

J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu. Sample-optimal tomography of quantum states. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 913–925, 2016.

J. Haah, R. Kothari, R. O'Donnell, and E. Tang. Query-optimal estimation of unitary channels in diamond distance. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 363–390. IEEE, 2023.

D. Hangleiter and M. J. Gullans. Bell sampling from quantum circuits. *Physical Review Letters*, 133(2):020601, 2024.

M. Heidari and R. Khardon. Learning DNF through generalized Fourier representations. In *Proceedings of the Conference on Learning Theory*, 2025. Full paper available as arXiv preprint 2506.01075.

M. Heidari and W. Szpankowski. Learning k-qubit quantum operators via pauli decomposition. In *International Conference on Artificial Intelligence and Statistics*, pages 490–504. PMLR, 2023.

M. Heidari and W. Szpankowski. New bounds on quantum sample complexity of measurement classes. In *2024 IEEE International Symposium on Information Theory (ISIT)*, pages 1515–1520. IEEE, July 2024. doi: 10.1109/isit57864.2024.10619538.

M. Heidari, A. Padakandla, and W. Szpankowski. A theoretical framework for learning from quantum data. In *2021 IEEE International Symposium on Information Theory (ISIT)*. IEEE, jul 2021. doi: 10.1109/isit45174.2021.9517721.

L. Hellerstein, D. Kletenik, L. Sellie, and R. A. Servedio. Tight bounds on proper equivalence query learning of DNF. In *The 25th Annual Conference on Learning Theory*, pages 31.1–31.18, 2012.

H.-Y. Huang, R. Kueng, and J. Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, 2020.

H.-Y. Huang, Y. Liu, M. Broughton, I. Kim, A. Anshu, Z. Landau, and J. R. McClean. Learning shallow quantum circuits. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1343–1351, 2024.

J. C. Jackson. An efficient membership-query algorithm for learning DNF with respect to the uniform distribution. *Journal of Computer and System Sciences*, 55(3):414–440, dec 1997. doi: 10.1006/jcss.1997.1533.

A. T. Kalai, A. Samorodnitsky, and S.-H. Teng. Learning and smoothed analysis. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 395–404. IEEE, Oct. 2009.

R. Khardon. On using the fourier transform to learn disjoint dnf. *Information Processing Letters*, 49(5):219–222, 1994.

B. T. Kiani, S. Lloyd, and R. Maity. Learning unitaries by gradient descent. *arXiv preprint arXiv:2001.11897*, 2020.

R. King, D. Gosset, R. Kothari, and R. Babbush. Triply efficient shadow tomography. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 914–946. SIAM, 2025.

R. Kothari. Efficient algorithms in quantum query complexity. 2014.

E. Kushilevitz. A simple algorithm for learning O(log n)-term DNF. In *Proceedings of the Ninth Annual Conference on Computational Learning Theory*, pages 266–269, 1996.

E. Kushilevitz and Y. Mansour. Learning decision trees using the Fourier spectrum. *SIAM Journal on Computing*, 22(6):1331–1348, Dec. 1993.

R. Levy, D. Luo, and B. K. Clark. Classical shadows for quantum process tomography on nearterm quantum computers. *Physical Review Research*, 6(1):013029, 2024.

A. Montanaro and T. J. Osborne. Quantum boolean functions. *arXiv:0810.2435*, Oct. 2010.

E. Tang and J. Wright. Are controlled unitaries helpful? *arXiv preprint arXiv:2508.00055*, 2025.

L. G. Valiant. A theory of the learnable. *Communications of the ACM*, 27(11):1134–1142, Nov. 1984.

H. Zhao, L. Lewis, I. Kannan, Y. Quek, H.-Y. Huang, and M. C. Caro. Learning quantum states and unitaries of bounded gate complexity. *PRX Quantum*, 5(4):040306, Oct. 2024. ISSN 2691-3399. doi: 10.1103/prxquantum.5.040306.

A OTHER RELEVANT DISTANCE METRICS AND NORMS

For any linear operator A, the trace norm is defined as $\|A\|_1 := \mathrm{Tr}\left(\sqrt{A^\dagger A}\right)$, the Frobenius norm is $\|A\|_2 := \sqrt{\mathrm{Tr}(A^\dagger A)}$, and the operator norm is $\|A\|_{op} := \max_{|\psi\rangle:\||\psi\rangle\|_2=1} \|A|\psi\rangle\|_2$. The Frobenius and Pauli ℓ_2-norms satisfy

$$\|A\|_2 = \sqrt{N}\|A\|_{2,P}, \tag{2}$$

where $N = 2^n$ is the Hilbert space dimension.

Definition 4 (diamond distance)**.** *For a pair of quantum mappings $\mathcal{E}$ and $\mathcal{F}$, the diamond distance is defined as*

$$d_\diamond(\mathcal{E},\mathcal{F}) := \max_\rho \|(\mathbb{I}\otimes(\mathcal{E}-\mathcal{F}))(\rho)\|_1,$$

where the maximization is over all density operators ρ acting on a space of dimension N^2.

Definition 5 (Phase-aligned operator distance)**.** *For operators U, V, the* phase-aligned operator norm distance *is defined as*

$$d_{optphase}(U,V) := \min_{\theta\in[0,2\pi)} \left\|U - e^{i\theta}V\right\|_{\mathrm{op}}. \tag{3}$$

Next, we establish an inequality connecting the above two distances.

Lemma 4 (Generalized Distance Bound)**.** *Let U and V be arbitrary linear operators acting on n qubits. Let $\mathcal{U}(\rho) = U\rho U^\dagger$ and $\mathcal{V}(\rho) = V\rho V^\dagger$. Then the diamond distance is bounded by:*

$$d_\diamond(\mathcal{U},\mathcal{V}) \le (\|U\|_{op} + \|V\|_{op}) d_{optphase}(U,V)$$

Moreover,

$$d_{optphase}(U,V) \le \min_{\phi\in[0,2\pi]} \|U - e^{i\phi}V\|_{1,P}$$

Proof. Let ρ_{RA} be an arbitrary state on the joint system. Let ϕ be an arbitrary phase and define $E = U - e^{i\phi}V$. We can rewrite the difference acting on ρ_{RA} by adding and subtracting $(\mathbb{I}\otimes U)\rho_{RA}(\mathbb{I}\otimes e^{-i\phi}V^\dagger)$. Therefore, as $\mathcal{U}$ and $\mathcal{V}$ are independent of the global phases of U and V we have that

$$\begin{aligned}(\mathbb{I}\otimes\mathcal{U})(\rho_{RA}) - (\mathbb{I}\otimes\mathcal{V})(\rho_{RA}) &= (\mathbb{I}\otimes U)\rho_{RA}(\mathbb{I}\otimes U^\dagger) - (\mathbb{I}\otimes e^{i\phi}V)\rho_{RA}(\mathbb{I}\otimes e^{-i\phi}V^\dagger)\\ &= (\mathbb{I}\otimes U)\rho_{RA}(\mathbb{I}\otimes E^\dagger) + (\mathbb{I}\otimes E)\rho_{RA}(\mathbb{I}\otimes e^{-i\phi}V^\dagger).\end{aligned}$$

Applying the triangle inequality together with the inequalities (a) $\|XYZ\|_1 \le \|X\|_{op}\|Y\|_1\|Z\|_{op}$; (b) $\|\rho_{RA}\|_1 = 1$; and (c) $\|A\otimes B\|_{op} = \|A\|_{op}\|B\|_{op}$:

$$\begin{aligned}\|(\mathbb{I}\otimes\mathcal{U} - \mathbb{I}\otimes\mathcal{V})(\rho_{RA})\|_1 &\le \|U\|_{op}\cdot 1\cdot\|E^\dagger\|_{op} + \|E\|_{op}\cdot 1\cdot\|V^\dagger\|_{op}\\ &= \|U\|_{op}\|E\|_{op} + \|E\|_{op}\|V\|_{op}\end{aligned}$$

Taking the supremum over all states ρ_{RA} and the minimum over all phases ϕ establishes the first inequality. For the second part of the lemma, let the Pauli decomposition $U = \sum_{\mathbf{s}} \alpha_{\mathbf{s}}\sigma^{\mathbf{s}}$ and $V = \sum_{\mathbf{s}} \beta_{\mathbf{s}}\sigma^{\mathbf{s}}$. Then, for any ϕ, we can write

$$\begin{aligned}d_{optphase}(U,V) \le \left\|U - e^{i\phi}V\right\|_{op} &= \left\|\sum_{\mathbf{s}}(\alpha_{\mathbf{s}} - e^{i\phi}\beta_{\mathbf{s}})\sigma_{\mathbf{s}}\right\|_{op}\\ &\le \sum_{\mathbf{s}} \left|\alpha_{\mathbf{s}} - e^{i\phi}\beta_{\mathbf{s}}\right|\|\sigma_{\mathbf{s}}\|_{op} = \left\|U - e^{i\phi}V\right\|_{1,P},\end{aligned}$$

where we used the fact that each Pauli operator has operator norm equal to one. □

B Related Work

While learning a general unitary requires $O(d^2)$ queries (Gutoski and Johnston, 2014; Haah et al., 2016), recent work by Haah et al. (2023) refined these bounds specifically for the diamond distance metric, establishing the query optimality for distinguishing unitary channels. Because of this barrier, many recent works have employed heuristic methods, including gradient descent (Kiani et al., 2020), classical shadows (Levy et al., 2024), and quantum neural networks (Beer et al., 2020).

While these methods are useful, they often lack guarantees of efficiency. Thus, a valid question is: *which classes of unitaries are efficiently learnable?* Zhao et al. studied how to efficiently learn unitaries of bounded gate complexity (Zhao et al., 2024). Similarly, testing and learning quantum k-juntas has been a topic of interest due to their real-world application (Chen et al., 2023; Belovs, 2015).

The problem of learning sparse Hamiltonians or unitaries in the Pauli basis is closely related to learning sparse Boolean functions. The celebrated KM algorithm (Kushilevitz and Mansour, 1993) or Goldreich-Levin (Goldreich and Levin, 1989) have been shown to be effective for learning several classes of Boolean functions including Disjunctive Normal Form (DNF) expressions, juntas, and decision trees (Jackson, 1997; Valiant, 1984; Bshouty, 1995; Kushilevitz, 1996; Aizenstein et al., 1998; Hellerstein et al., 2012). The approach was elaborated and extended by several authors (Bshouty et al., 2004; Feldman, 2007; Kalai et al., 2009; Feldman, 2012; Heidari and Khardon, 2025).

Several studies have attempted to develop a quantum variant of these algorithms, including Montanaro (Montanaro and Osborne, 2010) and Angrisani et al. (Angrisani, 2025). Crucially, Angrisani's method yields a sample complexity that depends on the system dimension n, typically to locate the support of the operators.

There is another line of work on learning sparse Hamiltonians. Traditional methods often rely on preparing either an eigenstate or the thermal (Gibbs) state (Anshu et al., 2020). Recently, Arunachalam et al. used Bell sampling and shadow tomography to introduce a protocol for learning Pauli coefficients of a sparse Hamiltonian with $O(s^2/\varepsilon^4)$ queries (Arunachalam et al., 2025). Abbas et al. also implemented a method using the isolation technique to learn the coefficients using $\tilde{O}(s\log(1/\varepsilon))$ experiments, achieving Heisenberg-limited scaling (Abbas et al., 2025).

PAC learning of quantum processes has also been studied in recent literature (Heidari and Szpankowski, 2024; Heidari et al., 2021; Arunachalam and De Wolf, 2018; Arunachalam and de Wolf, 2017). These works focus on learning quantum channels under the PAC framework, where the learner aims to approximate an unknown quantum channel based on input-output examples drawn from a specific distribution. Their results provide sample complexity bounds for learning quantum channels with respect to various distance measures, such as the diamond norm and trace norm. However, these works do not specifically address the efficient learnability of sparse unitaries in the Pauli basis, which is the focus of our study.

In the context of unitary learning, the learnability depends heavily on the available access model and the structural assumptions. If the unitary is generated by a sparse Hamiltonian and one has access to the time-evolution operator at variable times, Abbas et al. (2025) provides a Hamiltonian learning algorithm with optimal linear dependence on sparsity. In the standard black-box setting (fixed U), Arunachalam et al. (2025) show that unitaries generated by sparse Hamiltonians can be learned up to Pauli ℓ_∞-norm. Additionally, unitaries whose Pauli spectrum is supported on a small subgroup are learnable with complexity $\tilde{O}(2^k/\epsilon)$.

Our work targets a broader class: general unitaries that are sparse (or nearly sparse) in the Pauli basis, without assuming a underlying sparse Hamiltonian structure or time-control access. While Abbas et al. (2025) achieve better complexity, they require stronger access (variable time evolution). We introduce an efficient learning algorithm $\tilde{O}(s^3/\epsilon^2)$ using only standard access to the fixed unitary U.

C Proof of Lemma 1

Proof. Suppose S_θ is not empty. This problem is related to the non-uniform variant of the famous Coupon Collection problem. The probability of observing any specific target index $\mathbf{s} \in S_\theta$ in a

single shot is

$$p(\mathbf{s}) = |\alpha_{\mathbf{s}}|^2 \geq \theta^2.$$

After m_1 independent draws, the probability that a fixed $\mathbf{s} \in S_\theta$ is never observed is

$$\begin{aligned}\Pr[\mathbf{s} \notin \mathcal{X}] &= (1 - p(\mathbf{s}))^{m_1} \\ &\leq e^{-p(\mathbf{s})m_1} \\ &\leq e^{-\theta^2 m_1}.\end{aligned}$$

By the union bound, the probability of missing an index in S_θ after m_1 samples is

$$Pr[S_\theta \nsubseteq \mathcal{X}] \leq |S_\theta| e^{-\theta^2 m_1}.$$

Hence, to have this bounded by δ, we need

$$m_1 \geq \frac{\log(|S_\theta|/\delta)}{\theta^2}.$$

By Parseval's identity, $\sum_{\mathbf{s}} |\alpha_{\mathbf{s}}|^2 = 1$. Thus,

$$1 \geq \sum_{\mathbf{s} \in S_\theta} |\alpha_{\mathbf{s}}|^2 \geq \theta^2 |S_\theta|,$$

and hence $|S_\theta| \leq 1/\theta^2$. Therefore,

$$\Pr[S_\theta \nsubseteq \mathcal{X}] \leq \frac{1}{\theta^2} e^{-\theta^2 m_1}.$$

Choosing

$$m_1 = \left\lceil \frac{\log\big(1/(\theta^2\delta)\big)}{\theta^2} \right\rceil$$

makes this probability at most δ.

Finally, since $\mathcal{X}$ contains at most one distinct index per sample,

$$|\mathcal{X}| \leq m_1 \leq \frac{\log\big(1/(\theta^2\delta)\big)}{\theta^2} + 1.$$

Thus,

$$|\mathcal{X}| = O\left(\frac{\log(1/(\theta\delta))}{\theta^2}\right).$$

□

D Proof of Lemma 2

Proof. Let the true values of the traces be $R_{\mathbf{s}} = \mathrm{Tr}[R_{\mathbf{t},\mathbf{s}} |J(U)\rangle]$ and $I_{\mathbf{s}} = \mathrm{Tr}[I_{\mathbf{t},\mathbf{s}} |J(U)\rangle]$. By the classical-shadow guarantee, with probability at least $1 - \delta$, simultaneously for every $\mathbf{s} \in \mathcal{X}$,

$$\left|\widetilde{M_{\mathbf{s}}} - |\alpha_{\mathbf{s}}|^2\right| \leq \varepsilon,$$

and

$$|\widetilde{R}_{\mathbf{s}} - R_{\mathbf{s}}| \leq \varepsilon, \qquad |\widetilde{I}_{\mathbf{s}} - I_{\mathbf{s}}| \leq \varepsilon,$$

where

$$R_{\mathbf{s}} + iI_{\mathbf{s}} = \alpha_{\mathbf{s}} \alpha_{\mathbf{t}}^*.$$

Write

$$\hat{\alpha}_{\mathbf{t}} = \sqrt{\widetilde{M_{\mathbf{t}}}} \quad \text{and} \quad Z_{\mathbf{s}} := \alpha_{\mathbf{s}} \alpha_{\mathbf{t}}^*, \qquad \widetilde{Z}_{\mathbf{s}} := \widetilde{R}_{\mathbf{s}} + i\widetilde{I}_{\mathbf{s}}.$$

The magnitude estimate satisfies

$$|\hat{\alpha}_{\mathbf{t}} - |\alpha_{\mathbf{t}}|| = \frac{|\hat{\alpha}_{\mathbf{t}}^2 - |\alpha_{\mathbf{t}}|^2|}{\hat{\alpha}_{\mathbf{t}} + |\alpha_{\mathbf{t}}|} \tag{4}$$
$$= \frac{|\widetilde{M}_{\mathbf{t}} - |\alpha_{\mathbf{t}}|^2|}{\hat{\alpha}_{\mathbf{t}} + |\alpha_{\mathbf{t}}|}$$
$$\leq \frac{\varepsilon}{|\alpha_{\mathbf{t}}|}.$$

Consequently,

$$\hat{\alpha}_{\mathbf{t}} \geq \alpha_{\mathbf{t}} - \frac{\varepsilon}{\alpha_{\mathbf{t}}} > 0.$$

We need to compare the magnitude of $\alpha_{\mathbf{s}}$ with that of the selected reference coefficient $\alpha_{\mathbf{t}}$. Since $\mathbf{t}$ maximizes $\widetilde{M}_{\mathbf{s}}$,

$$|\alpha_{\mathbf{s}}|^2 \leq \widetilde{M}_{\mathbf{s}} + \varepsilon$$
$$\leq \widetilde{M}_{\mathbf{t}} + \varepsilon$$
$$\leq |\alpha_{\mathbf{t}}|^2 + 2\varepsilon.$$

Therefore,

$$\frac{|\alpha_{\mathbf{s}}|}{|\alpha_{\mathbf{t}}|} \leq \sqrt{1 + \frac{2\varepsilon}{|\alpha_{\mathbf{t}}|^2}}.$$

Let $\phi = \arg(\alpha_{\mathbf{t}})$, so that $\alpha_{\mathbf{t}} = e^{i\phi}|\alpha_{\mathbf{t}}|$. Algorithm 3 outputs

$$\hat{\alpha}_{\mathbf{s}} = \frac{\widetilde{Z}_{\mathbf{s}}}{\hat{\alpha}_{\mathbf{t}}},$$

whereas

$$e^{-i\phi}\alpha_{\mathbf{s}} = \frac{\alpha_{\mathbf{s}}\alpha_{\mathbf{t}}^*}{|\alpha_{\mathbf{t}}|} = \frac{Z_{\mathbf{s}}}{|\alpha_{\mathbf{t}}|}.$$

we can write the error as a difference of fractions, to separate the error contributions, we add and subtract the term $\frac{Z_{\mathbf{s}}}{\hat{\alpha}_{\mathbf{t}}}$:

$$\hat{\alpha}_{\mathbf{s}} - e^{-i\phi}\alpha_{\mathbf{s}} = \frac{\widetilde{Z}_{\mathbf{s}}}{\hat{\alpha}_{\mathbf{t}}} - \frac{Z_{\mathbf{s}}}{\hat{\alpha}_{\mathbf{t}}} + \frac{Z_{\mathbf{s}}}{\hat{\alpha}_{\mathbf{t}}} - \frac{Z_{\mathbf{s}}}{|\alpha_{\mathbf{t}}|}$$
$$= \frac{\widetilde{Z}_{\mathbf{s}} - Z_{\mathbf{s}}}{\hat{\alpha}_{\mathbf{t}}} + Z_{\mathbf{s}}\left(\frac{1}{\hat{\alpha}_{\mathbf{t}}} - \frac{1}{|\alpha_{\mathbf{t}}|}\right).$$

Hence

$$\left|\hat{\alpha}_{\mathbf{s}} - e^{-i\phi}\alpha_{\mathbf{s}}\right| \leq \frac{|\widetilde{Z}_{\mathbf{s}} - Z_{\mathbf{s}}|}{\hat{\alpha}_{\mathbf{t}}} + |Z_s|\left|\frac{1}{\hat{\alpha}_{\mathbf{t}}} - \frac{1}{|\alpha_{\mathbf{t}}|}\right|$$
$$= \frac{|\widetilde{Z}_{\mathbf{s}} - Z_{\mathbf{s}}|}{\hat{\alpha}_{\mathbf{t}}} + \frac{|\alpha_{\mathbf{s}}||\hat{\alpha}_{\mathbf{t}} - |\alpha_{\mathbf{t}}||}{\hat{\alpha}_{\mathbf{t}}}.$$

Since the real and imaginary parts are each estimated within ε,

$$|\widetilde{Z}_{\mathbf{s}} - Z_{\mathbf{s}}| \leq \sqrt{2}\,\varepsilon.$$

Using the bounds above gives

$$\left|\hat{\alpha}_{\mathbf{s}} - e^{-i\phi}\alpha_{\mathbf{s}}\right| \leq \frac{\sqrt{2}\,\varepsilon + |\alpha_{\mathbf{s}}|\varepsilon/|\alpha_{\mathbf{t}}|}{|\alpha_{\mathbf{t}}| - \varepsilon/|\alpha_{\mathbf{t}}|}$$
$$\leq \frac{\varepsilon}{|\alpha_{\mathbf{t}}| - \varepsilon/|\alpha_{\mathbf{t}}|}\left(\sqrt{2} + \sqrt{1 + \frac{2\varepsilon}{|\alpha_{\mathbf{t}}|^2}}\right).$$

If $\varepsilon \le |\alpha_{\mathbf{t}}|^2/2$, then

$$|\alpha_{\mathbf{t}}| - \frac{\varepsilon}{|\alpha_{\mathbf{t}}|} \ge \frac{|\alpha_{\mathbf{t}}|}{2}, \qquad \sqrt{1 + \frac{2\varepsilon}{|\alpha_{\mathbf{t}}|^2}} \le \sqrt{2},$$

and therefore

$$\left|\hat{\alpha}_{\mathbf{s}} - e^{-i\phi}\alpha_{\mathbf{s}}\right| \le 4\sqrt{2}\,\frac{\varepsilon}{|\alpha_{\mathbf{t}}|}.$$

This holds simultaneously for all $\mathbf{s} \in \mathcal{X}$, completing the proof. □

E Connection to ℓ_2-norm

Here, we consider the specific instance where $\mathcal{A} = \{\mathbb{I}/N\}$, the maximally mixed state. This case is of particular importance as it corresponds to the performance of the channel averaged uniformly over all pure input states.

Lemma 5. *Let $\mathcal{U}$ be the unitary channel for unitary U, and $\hat{\mathcal{U}}$ be the channel for operator $\hat{U}$. If $\left\|U - \hat{U}\right\|_{2,P} = \nu$, then*

$$d_{\diamond,\{\mathbb{I}/N\}}(\mathcal{U},\hat{\mathcal{U}}) \le 2\nu + \nu^2.$$

Proof. Let

$$|\phi\rangle_{R_1 A} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |i\rangle_{R_1} |i\rangle_A$$

be the canonical purification of I_A/N. We first show that it suffices to evaluate the restricted distance using $|\phi\rangle\langle\phi|$.

Let ρ_{RA} be any state satisfying $\mathrm{Tr}_R(\rho_{RA}) = I_A/N$, and let $|\psi\rangle_{REA}$ be a purification of ρ_{RA}. Since $|\psi\rangle_{REA}$ and $|\varphi\rangle_{R_1 A}$ are both purifications of I_A/N, there exists an isometry $W : R_1 \to R \otimes E$ such that

$$|\psi\rangle_{REA} = (W \otimes I_A)|\phi\rangle_{R_1 A}.$$

Thus, defining the quantum channel

$$\Lambda(X) = \mathrm{Tr}_E(WXW^\dagger),$$

we have

$$\rho_{RA} = (\Lambda \otimes I_A)(|\phi\rangle\langle\phi|).$$

Because Λ acts only on the reference system, it commutes with both mappings acting on A. Hence,

$$\begin{aligned}&(I_R \otimes \mathcal{U})(\rho_{RA}) - (I_R \otimes \hat{\mathcal{U}})(\rho_{RA}) \\ &\quad = (\Lambda \otimes I_A)\left[(I_{R_1} \otimes \mathcal{U})(|\phi\rangle\langle\phi|) - (I_{R_1} \otimes \hat{\mathcal{U}})(|\phi\rangle\langle\phi|)\right].\end{aligned}$$

The operator inside the brackets is Hermitian. Therefore, by trace-norm contractivity under quantum channels,

$$\left\|(I_R \otimes \mathcal{U})(\rho_{RA}) - (I_R \otimes \hat{\mathcal{U}})(\rho_{RA})\right\|_1 \le \left\|(I_{R_1} \otimes \mathcal{U})(|\phi\rangle\langle\phi|) - (I_{R_1} \otimes \hat{\mathcal{U}})(|\phi\rangle\langle\phi|)\right\|_1.$$

Also, $|\phi\rangle\langle\phi|$ is itself an valid state. Therefore,

$$d_{\diamond,\{I/N\}}(\mathcal{U},\hat{\mathcal{U}}) = \left\|(I_{R_1} \otimes \mathcal{U})(|\phi\rangle\langle\phi|) - (I_{R_1} \otimes \hat{\mathcal{U}})(|\phi\rangle\langle\phi|)\right\|_1.$$

Therefore, it suffices to evaluate the distance using $|\phi\rangle\langle\phi|$. Let $|\phi_{\mathcal{U}}\rangle = (\mathbb{I}\otimes U)|\phi\rangle$ and define $|\phi_{\hat{\mathcal{U}}}\rangle$ analogously. Letting $|\delta\rangle = |\phi_{\mathcal{U}}\rangle - |\phi_{\hat{\mathcal{U}}}\rangle$, we evaluate its squared norm:

$$\begin{aligned}
\||\delta\rangle\|_2^2 &= \langle\phi|\left(\mathbb{I}\otimes(U-\hat{U})^\dagger(U-\hat{U})\right)|\phi\rangle \\
&= \frac{1}{N}\sum_{i,j=0}^{N-1}\left(\langle i|_R\langle i|_A\right)\left(\mathbb{I}_R\otimes(U-\hat{U})^\dagger(U-\hat{U})\right)\left(|j\rangle_R|j\rangle_A\right) \\
&= \frac{1}{N}\sum_{i=0}^{N-1}\langle i|_A(U-\hat{U})^\dagger(U-\hat{U})|i\rangle_A \\
&= \frac{1}{N}\operatorname{Tr}\left((U-\hat{U})^\dagger(U-\hat{U})\right) \\
&= \left\|U-\hat{U}\right\|_{2,P}^2 = \nu^2.
\end{aligned}$$

Thus, $\||\delta\rangle\|_2 = \nu$. Finally, expanding the density operators, we bound the trace distance:

$$\begin{aligned}
\left\||\phi_{\mathcal{U}}\rangle\langle\phi_{\mathcal{U}}| - |\phi_{\hat{\mathcal{U}}}\rangle\langle\phi_{\hat{\mathcal{U}}}|\right\|_1 &= \||\phi_{\mathcal{U}}\rangle\langle\phi_{\mathcal{U}}| - (|\phi_{\mathcal{U}}\rangle - |\delta\rangle)(\langle\phi_{\mathcal{U}}| - \langle\delta|)\|_1 \\
&= \||\phi_{\mathcal{U}}\rangle\langle\delta| + |\delta\rangle\langle\phi_{\mathcal{U}}| - |\delta\rangle\langle\delta|\|_1 \\
&\leq 2\||\phi_{\mathcal{U}}\rangle\langle\delta|\|_1 + \||\delta\rangle\langle\delta|\|_1 \\
&= 2\nu + \nu^2.
\end{aligned}$$

The final equality relies on the fact that $\||\psi_1\rangle\langle\psi_2|\|_1 = \||\psi_1\rangle\|_2\||\psi_2\rangle\|_2$, and that $|\phi_{\mathcal{U}}\rangle$ is perfectly normalized since U is unitary. □

We next have the following result that shows if U is approximated by a ℓ_1-bounded operator, then U is also approximateable by a sparse operator.

Lemma 6. *Suppose that U is approximated by $\tilde{U}$ such that*

$$\left\|U-\tilde{U}\right\|_{2,P}^2 \leq \frac{\varepsilon}{4} \qquad \textit{and} \qquad \left\|\tilde{U}\right\|_{1,P} = L_1 > 0.$$

Then U can be approximated by an $\frac{8L_1^2}{\varepsilon}$-sparse operator $\hat{U}$ generated from the Pauli coefficients of U larger than $\frac{\varepsilon}{2L_1}$, such that

$$\left\|U-\hat{U}\right\|_{2,P}^2 \leq 2\varepsilon.$$

Proof. We first show the existence of a sparse V that approximates $\tilde{U}$ and hence U. Write the Pauli expansion $\tilde{U} = \sum_{\mathbf{s}}\tilde{\alpha}_{\mathbf{s}}\sigma^{\mathbf{s}}$. Let $S' = \left\{\mathbf{s} : |\tilde{\alpha}_{\mathbf{s}}| \geq \frac{\varepsilon}{4L_1}\right\}$, and define $V = \sum_{\mathbf{s}\in S'}\tilde{\alpha}_{\mathbf{s}}\sigma^{\mathbf{s}}$. Then

$$\|V\|_{1,P} \leq \left\|\tilde{U}\right\|_{1,P} = L_1$$

and

$$|S'| \leq \frac{4L_1^2}{\varepsilon}.$$

Moreover,

$$\begin{aligned}
\left\|\tilde{U}-V\right\|_{2,P}^2 &= \sum_{\mathbf{s}\notin S'}|\tilde{\alpha}_{\mathbf{s}}|^2 \\
&\leq \max_{\mathbf{s}\notin S'}|\tilde{\alpha}_{\mathbf{s}}|\sum_{\mathbf{s}\notin S'}|\tilde{\alpha}_{\mathbf{s}}| \\
&\leq \frac{\varepsilon}{4L_1}L_1 \\
&= \frac{\varepsilon}{4}.
\end{aligned}$$

Using $\|A+B\|_{2,P}^2 \le 2\left(\|A\|_{2,P}^2 + \|B\|_{2,P}^2\right)$, we obtain

$$\|U-V\|_{2,P}^2 \le 2\left(\left\|U-\tilde{U}\right\|_{2,P}^2 + \left\|\tilde{U}-V\right\|_{2,P}^2\right) \le \varepsilon.$$

Now define $S = \left\{\mathbf{s} : |\alpha_{\mathbf{s}}| \ge \frac{\varepsilon}{2L_1}\right\}$ and $\hat{U} = \sum_{\mathbf{s}\in S} \alpha_{\mathbf{s}}\sigma^{\mathbf{s}}$. From the above construction,

$$\varepsilon \ge \|U-V\|_{2,P}^2 = \sum_{\mathbf{s}\notin S'} |\alpha_{\mathbf{s}}|^2 + \sum_{\mathbf{s}\in S'} |\alpha_{\mathbf{s}} - \tilde{\alpha}_{\mathbf{s}}|^2.$$

In particular,

$$\sum_{\mathbf{s}\notin S'} |\alpha_{\mathbf{s}}|^2 \le \varepsilon. \tag{5}$$

The error of approximating U by $\hat{U}$ is bounded as follows:

$$\begin{aligned}
\left\|U-\hat{U}\right\|_{2,P}^2 &= \sum_{\mathbf{s}\notin S} |\alpha_{\mathbf{s}}|^2 \\
&= \sum_{\mathbf{s}\notin S\cup S'} |\alpha_{\mathbf{s}}|^2 + \sum_{\mathbf{s}\in S'\setminus S} |\alpha_{\mathbf{s}}|^2 \\
&\le \varepsilon + |S'|\frac{\varepsilon^2}{4L_1^2} \\
&\le 2\varepsilon.
\end{aligned}$$

Finally, we bound the sparsity of $\hat{U}$. Since

$$|S\cap S'| \le |S'| \le \frac{4L_1^2}{\varepsilon},$$

it remains to bound $|S\setminus S'|$. Every $\mathbf{s}\in S\setminus S'$ satisfies

$$|\alpha_{\mathbf{s}}|^2 \ge \frac{\varepsilon^2}{4L_1^2}.$$

Therefore, by (5),

$$|S\setminus S'|\frac{\varepsilon^2}{4L_1^2} \le \sum_{\mathbf{s}\in S\setminus S'} |\alpha_{\mathbf{s}}|^2 \le \varepsilon,$$

which implies

$$|S\setminus S'| \le \frac{4L_1^2}{\varepsilon}.$$

Consequently,

$$\begin{aligned}
|S| &= |S\cap S'| + |S\setminus S'| \\
&\le 2\cdot\frac{4L_1^2}{\varepsilon} \\
&= \frac{8L_1^2}{\varepsilon}.
\end{aligned}$$

□

F Proof of Theorem 3

Proof. Set

$$\theta = \frac{\varepsilon^2}{4L_1}, \qquad \varepsilon_p = \frac{\theta}{2} = \frac{\varepsilon^2}{8L_1}.$$

Apply Theorem 1 with threshold θ, accuracy ε_p, and constant failure probability. It outputs a set $\mathcal{X}$ containing all $\mathbf{s}$ such that $|\alpha_{\mathbf{s}}| \geq \theta$, together with estimates satisfying

$$|\hat{\alpha}_{\mathbf{s}} - e^{-i\phi}\alpha_{\mathbf{s}}| \leq \varepsilon_p \qquad \text{for every } \mathbf{s} \in \mathcal{X},$$

for a shared phase ϕ. Since this phase does not change the induced channel, we absorb it into U. Define

$$\hat{\mathcal{S}} := \{\mathbf{s} \in \mathcal{X} : |\hat{\alpha}_{\mathbf{s}}| \geq \theta\}, \qquad \hat{U} := \sum_{\mathbf{s}\in\hat{\mathcal{S}}} \hat{\alpha}_{\mathbf{s}}\sigma^{\mathbf{s}}.$$

If $\mathbf{s} \notin \hat{\mathcal{S}}$, then $|\alpha_{\mathbf{s}}| < \theta + \varepsilon_p = 3\theta/2$. Moreover, every $\mathbf{s} \in \hat{\mathcal{S}}$ satisfies $|\alpha_{\mathbf{s}}| \geq \theta - \varepsilon_p = \theta/2$, and hence $|\hat{\mathcal{S}}| \leq 2L_1/\theta$. Therefore,

$$\begin{aligned}\left\|U - \hat{U}\right\|_{2,P}^2 &= \sum_{\mathbf{s}\notin\hat{\mathcal{S}}} |\alpha_{\mathbf{s}}|^2 + \sum_{\mathbf{s}\in\hat{\mathcal{S}}} |\alpha_{\mathbf{s}} - \hat{\alpha}_{\mathbf{s}}|^2 \\ &\leq \frac{3\theta}{2}\|U\|_{1,P} + |\hat{\mathcal{S}}|\varepsilon_p^2 \\ &\leq \frac{3\theta L_1}{2} + \frac{2L_1}{\theta}\varepsilon_p^2 = \frac{\varepsilon^2}{2}.\end{aligned}$$

Thus, $\left\|U - \hat{U}\right\|_{2,P} = O(\varepsilon)$. By Lemma 5,

$$d_{\diamond,\{\mathbb{I}/N\}}(\mathcal{U},\hat{\mathcal{U}}) \leq 2\left\|U - \hat{U}\right\|_{2,P} + \left\|U - \hat{U}\right\|_{2,P}^2 = O(\varepsilon).$$

Finally, Theorem 1 gives query complexity

$$O\left(\frac{\log(1/\theta)}{\varepsilon_p^2\theta^2}\right) = \widetilde{O}\left(\frac{L_1^4}{\varepsilon^8}\right),$$

which completes the proof. □

G Proof of Theorem 4

Proof. We prove the claim by considering the Oracle family of unitaries. Consider the *phase oracle* (or Grover oracle) family

$$U_x = I - 2\,|x\rangle\langle x|\,, \qquad x \in \{0,1\}^n.$$

The Pauli ℓ_1 norm of each U_x is bounded by a constant independent of n. To see this, note that each U_x admits an expansion using only Z-type Pauli operators:

$$\alpha_I = 1 - \frac{2}{N},$$

and for each non-identity Pauli Z_s,

$$\alpha_{Z_s} = -\frac{2}{N}(-1)^{s\cdot x},$$

where $s \cdot x$ denotes the mod-2 inner product.

Therefore,

$$\begin{aligned}\|U_x\|_{1,P} &= \left|1 - \frac{2}{N}\right| + (N-1)\frac{2}{N} \\ &= 3 - \frac{4}{N} < 3.\end{aligned}$$

Thus, the family $\{U_x\}_{x\in\{0,1\}^n}$ has uniformly bounded Pauli ℓ_1 norm with $L < 3$, independent of the number of qubits n.

For $x \neq y$, the unitary channels induced by U_x and U_y are perfectly distinguishable: their diamond distance satisfies

$$\|U_x - U_y\|_\diamond = 2.$$

Consequently, any learner that outputs an estimate $\hat{U}$ such that

$$\left\|U_x - \hat{U}\right\|_\diamond < 1$$

must, in effect, identify the marked string x exactly. However, determining x given black-box access to U_x is exactly the unstructured search problem and is known to require $\Omega(2^{n/2})$ queries. This lower bound is tight, as Grover's algorithm achieves it and is optimal by the BBBV theorem. □

H Technical proofs of Section 7

H.1 Proof of Lemma 3

Proof. We expand the unitary U using its Taylor series definition:

$$U = e^{iH} = \sum_{k=0}^{\infty} \frac{i^k}{k!} H^k.$$

Using the triangle inequality for the Pauli ℓ_1-norm, we have:

$$\|U\|_{1,P} = \left\| \sum_{k=0}^{\infty} \frac{i^k}{k!} H^k \right\|_{1,P} \leq \sum_{k=0}^{\infty} \frac{1}{k!} \left\|H^k\right\|_{1,P}.$$

We utilize the submultiplicativity of the Pauli ℓ_1-norm. For any operators A and B, $\|AB\|_{1,P} \leq \|A\|_{1,P}\|B\|_{1,P}$. Applying this inductively to H^k, we obtain:

$$\left\|H^k\right\|_{1,P} \leq \|H\|_{1,P}^k \leq L_H^k.$$

Substituting this back into the summation:

$$\|U\|_{1,P} \leq \sum_{k=0}^{\infty} \frac{L_H^k}{k!} = e^{L_H}.$$

Since $\|U\|_{1,P}$ is bounded by $L_U = e^{L_H}$, we can invoke Theorem 3. To learn U up to restricted diamond distance ε, we require a sparse approximation with error parameter ε satisfying the condition from Lemma 6:

$$\varepsilon^2 \leq \frac{\varepsilon}{4L_U^2} = \frac{\varepsilon}{4e^{2L_H}}.$$

Thus, Algorithm 3 can estimate U by targeting coefficient precision $O(\varepsilon^3)$, ensuring the learned unitary $\hat{U}$ satisfies $\left\|U - \hat{U}\right\|_{2,P} \leq \sqrt{\varepsilon}$. □

H.2 Proof of Proposition 1

Proof. Using $|1\rangle\langle 1| = \frac{1}{2}(I - Z)$, we obtain

$$|1^k\rangle\langle 1^k| = \tfrac{1}{2^k} \sum_{S \subseteq [k]} (-1)^{|S|} Z_S,$$

where $Z_S := \prod_{j \in S} Z_j$.

Therefore, the Pauli coefficients of $U_{k,\phi}$ are:

$$\alpha_{Z_S} = \begin{cases} \dfrac{e^{i\phi} - 1}{2^k}(-1)^{|S|}, & S \neq \emptyset, \\ 1 + \dfrac{e^{i\phi} - 1}{2^k}, & S = \emptyset. \end{cases}$$

Thus,

$$\|U_{k,\phi}\|_{1,P} = \left|1 + \frac{e^{i\phi} - 1}{2^k}\right| + (2^k - 1)\frac{|e^{i\phi} - 1|}{2^k}$$
$$\leq 1 + |e^{i\phi} - 1| \leq 3.$$

□

H.3 Proof of Proposition 2

Proof. The proof is similar to the previous example. The Grover diffusion operator (up to a global sign) is $D := 2\,|+\rangle\langle+|^{\otimes n} - I$. Since $|+\rangle\langle+| = \frac{1}{2}(I + X)$, we have $|+\rangle\langle+|^{\otimes n} = \frac{1}{2^n}\sum_{S\subseteq[n]} X_S$, where $X_S := \prod_{j\in S} X_j$. Thus,

$$D = -I + \frac{2}{2^n}\sum_{S\subseteq[n]} X_S.$$

The Pauli coefficients are:

$$\alpha_I = -1 + 2^{1-n}, \qquad \alpha_{X_S} = 2^{1-n} \quad (S \neq \emptyset).$$

Hence,

$$\|D\|_{1,P} = \left|-1 + 2^{1-n}\right| + (2^n - 1)2^{1-n}$$
$$= 3 - 2^{2-n} < 3.$$

□