Single-Copy Stabilizer Testing

Marcel Hinsche
Freie Universität Berlin
Berlin, Germany
IBM Quantum, IBM Research Europe – Zurich
Rüschlikon, Switzerland
m.hinsche@fu-berlin.de

Jonas Helsen
QuSoft, Centrum Wiskunde & Informatica (CWI)
Amsterdam, Netherlands
jonas@cwi.nl

Abstract

We consider the problem of testing whether an unknown n-qubit quantum state is a stabilizer state, with only single-copy access. We give an algorithm solving this problem using $O(n)$ copies, and conversely prove that $\Omega(\sqrt{n})$ copies are required for any algorithm. The main observation behind our algorithm is that when repeatedly measuring in a randomly chosen stabilizer basis, stabilizer states are the most likely among the set of all pure states to exhibit linear dependencies in measurement outcomes. Our algorithm is designed to probe deviations from this extremal behavior. For the lower bound, we first reduce stabilizer testing to the task of distinguishing random stabilizer states from the maximally mixed state. We then argue that, without loss of generality, it is sufficient to consider measurement strategies that a) lie in the commutant of the tensor action of the Clifford group and b) satisfy a Positive Partial Transpose (PPT) condition. By leveraging these constraints, together with novel results on the partial transposes of the generators of the Clifford commutant, we derive the lower bound on the sample complexity.

CCS Concepts

• **Theory of computation** → **Quantum information theory**; **Machine learning theory**.

Keywords

quantum information, quantum property testing, stabilizer states, Clifford group, single-copy, quantum memory

ACM Reference Format:
Marcel Hinsche and Jonas Helsen. 2025. Single-Copy Stabilizer Testing. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing (STOC '25), June 23–27, 2025, Prague, Czechia.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3717823.3718169

1 Introduction

Understanding properties of unknown quantum systems is crucial for applications ranging from benchmarking quantum devices to exploring fundamental physics. One key property of quantum states is their non-stabilizerness which quantifies their deviation from the set of stabilizer states. This property is intricately linked to the ability to enhance Clifford operations, ultimately enabling universal quantum computation. In this work, we are concerned with characterizing this property. Specifically, we study *stabilizer testing*, which has been extensively studied recently within the field of quantum property testing [3, 4, 8, 15, 21, 23, 41, 43].

Definition 1.1 (Stabilizer testing). Given access to copies of an unknown n-qubit pure state $|\psi\rangle$, decide if $|\psi\rangle$ is a stabilizer states or at least ϵ-far from all stabilizer states.

Motivated by practical feasibility considerations, recent research has made significant progress in understanding the statistical overheads associated with restricted quantum learning and testing protocols [9–11, 13, 14, 25, 29, 32, 38, 40, 45]. In particular, one often considers restricting the quantum learning/testing algorithm to process only one copy of the unknown state at a time. We refer to such algorithms as *single-copy* algorithms. This in contrast to *multi-copy* algorithms that can jointly operate across multiple copies at once. In practice, this limitation may arise due to a lack of quantum memory to store multiple copies which is why in some works single-copy algorithms are referred to as learning algorithms without quantum memory [12, 29]. It may however also arise due to the difficulty of implementing joint operations across multiple copies. The restriction to single-copy algorithms has been shown to lead to large, often even exponential, increases in the number of copies necessary for certain learning or testing tasks when compared to general multi-copy algorithms.

In the context of *learning* stabilizer states, this separation has been tightly characterized (c.f. Table 1). However, for stabilizer testing, an understanding of the separation between single- and multi-copy algorithms is still missing. While Ref. [23] gave a 2-copy algorithm using in total 6 copies of the unknown state (in the case of qubits), thereby settling the multi-copy complexity, the single-copy complexity remained open.

1.1 Results

In this work, we address this gap: First, we give a single-copy stabilizer testing algorithm using $O(n)$ copies, based on the computational difference sampling primitive introduced in [20].

Informal Theorem 1.2 (Upper bound). *There exists a non-adaptive single-copy algorithm for stabilizer testing that uses $t = O(n)$ copies running in time $O(n^3)$.*

We note that sample complexity of this algorithm beats a brute-force approach using classical shadows [30]. This approach would consist of estimating the fidelity with all $2^{O(n^2)}$ many n-qubit stabilizer states and would require $t = O(n^2)$ many copies (as well as exponential time complexity).

STOC '25, Prague, Czechia

ACM ISBN 979-8-4007-1510-5/25/06
https://doi.org/10.1145/3717823.3718169

Table 1: Upper and lower bounds on the number of copies t required for learning and testing stabilizer states with single-copy and multi-copy access, respectively.

	Single-copy	Multi-copy
Learning	$\Theta(n^2)$ [1, 2, 20]	$\Theta(n)$ [42]
Testing	$\Omega(\sqrt{n}) \leq t \leq O(n)$ (*This work*)	$\Theta(1)$ [23]

Secondly, we provide a lower bound of $\Omega(\sqrt{n})$ copies on the sample complexity of any single-copy algorithm, using the representation theory of the Clifford group [23] and a Positive Partial Transpose (PPT) relaxation technique due to Harrow [25]:

Informal Theorem 1.3 (Lower bound). *Any single-copy algorithm for stabilizer testing requires at least $t = \Omega(\sqrt{n})$ copies.*

As far as we know, this is the first work showing such a fine-grained single-copy/multi-copy separation for a property testing problem, as opposed to the exponential separations usually seen in the literature. We remark that establishing this lower bound for stabilizer testing requires tools well beyond those used to show the lower bounds for stabilizer state learning. The latter bounds follow from the Holevo bound in the multi-copy case [42] or require computing only second moments in the single-copy case, see Theorem 5 in [2]. In contrast, we use $t = \Omega(\sqrt{n})$-th moments of the Clifford group.

Extension to Qudits and Mixed State Inputs. For simplicity, we focus on n-qubit ($d = 2$) pure states throughout the manuscript. However, our single-copy stabilizer testing algorithm extends naturally to the case where the local dimension d is prime, maintaining the same time- and sample complexity as stated in Informal Theorem 1.2. Furthermore, the algorithm also directly applies to the case of mixed state inputs ρ as considered recently in [31]. That is, when applied to mixed states, the algorithm tests if ρ is at least ϵ-far from all pure stabilizer states. We further comment on these extensions in the full version of the paper [28].

1.2 Technical Overview

Upper Bound. The single-copy stabilizer testing algorithm we present is very much inspired by the learning algorithms for stabilizer states from [1, 16, 20]. In particular, in our algorithm, we utilize the *computational difference sampling* primitive introduced by [20]. It involves two copies of the unknown state $|\psi\rangle$, sampling each in the computational basis, and adding the outcome bitstrings mod 2 during classical post-processing (see Section 2.4 for more details).

Letting $|\psi\rangle$ be the unknown state we want to test, our single-copy algorithm proceeds as follows:

(1) Sample a uniformly random Clifford C from the Clifford group $\mathrm{Cl}(n)$.
(2) Use $2K = O(n)$ (but $K > n$) copies of $|\psi\rangle$ to repeatedly measure $C|\psi\rangle$ via computational difference sampling obtaining samples $\mathbf{v}_1, \ldots, \mathbf{v}_K$.
(3) Taking the samples as rows of a matrix, compute the rank of the $K \times n$ binary matrix in order to check if the samples span $\mathbb{F}_2^n$ or only a proper subspace of $\mathbb{F}_2^n$. In the former case, output 1, in the latter case output 0.

We refer to the expectation value (over the choice of random Clifford) of this random process as the *average spanning probability*:

$$\overline{\mathbb{P}_K}(|\psi\rangle) := \mathop{\mathbf{E}}_{C\sim\mathrm{Cl}(n)} \left[\Pr_{\mathbf{v}_1,\ldots,\mathbf{v}_K \sim r_{C|\psi\rangle}} \mathrm{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_K\} = \mathbb{F}_2^n \right] \quad (1)$$

Here, $r_{C|\psi\rangle}$ denotes the computational difference sampling distribution corresponding to the state $C|\psi\rangle$. By repeating the above 3-step process $O(1/\epsilon^2)$ many times, our algorithm estimates $\overline{\mathbb{P}_K}(|\psi\rangle)$ up to additive precision ϵ.

As we demonstrate, the average spanning probability $\overline{\mathbb{P}_K}(|\psi\rangle)$ is extremal for stabilizer states, taking a value of roughly $\overline{\mathbb{P}_K}(|S\rangle) \approx 0.42$ (for any $n \geq 10$) and increases beyond this value for non-stabilizer states. Our key technical contribution is to show that for a state that is ϵ-far from all stabilizer states, this increase is sufficiently large, namely of the order $\Omega(\epsilon)$. Hence, this increase can be detected sample-efficiently by estimating $\overline{\mathbb{P}_K}(|\psi\rangle)$ and comparing to the value for stabilizer states which can be efficiently computed exactly analytically. This gives rise to a single-copy stabilizer testing algorithm.

Understanding $\overline{\mathbb{P}_K}$ for stabilizer states is easy because their output distributions in the computational basis are well-known to be uniform over affine subspaces of $\mathbb{F}_2^n$ (c.f. Eq. (16)). The key challenge is to analyze and bound the average spanning probability for non-stabilizer states $|\psi\rangle$ because their output distributions in the computational basis are no longer uniform over a subspace. To overcome this challenge, we make use of a key relation between the computational difference sampling distribution r_ψ and the so-called *characteristic distribution* p_ψ. The characteristic distribution p_ψ associated to a pure n-qubit state is the distribution over $\mathbb{F}_2^{2n}$ given by

$$p_\psi(\mathbf{x}) = \frac{1}{2^n} |\langle\psi| P_{\mathbf{x}} |\psi\rangle|^2 , \quad (2)$$

where we identify Pauli operators in $\{I, X, Y, Z\}^n$ with bitstrings $\mathbf{x} = (\mathbf{a}, \mathbf{b})$ via $P_{\mathbf{x}} = i^{\mathbf{a}\cdot\mathbf{b}} X^{\mathbf{a}} Z^{\mathbf{b}}$. The relation was established in [20] and posits that the computational difference sampling distribution r_ψ can be written in terms of p_ψ as follows:

$$r_\psi(\mathbf{a}) = \sum_{\mathbf{b}\in\mathbb{F}_2^n} p_\psi(\mathbf{a}, \mathbf{b}) . \quad (3)$$

Using this relation, we are able to rewrite $\overline{\mathbb{P}_K}(|\psi\rangle)$ in terms of p_ψ. Lastly, to connect back to stabilizer testing, we use the fact that the weight of p_ψ on any commuting subgroup of $\{I, X, Y, Z\}^n$ relates to the *stabilizer fidelity* of $|\psi\rangle$, defined as $\max_{|S\rangle\in\mathrm{Stab}(n)} |\langle S|\psi\rangle|^2$, where $\mathrm{Stab}(n)$ denotes the set of all n-qubit stabilizer states. In particular, for any such stabilizer group corresponding to an isotropic subspace $M \subset \mathbb{F}_2^{2n}$, it holds that [21, 23],

$$\max_{|S\rangle\in\mathrm{Stab}(n)} |\langle S|\psi\rangle|^2 \geq \sum_{\mathbf{x}\in M} p_\psi(\mathbf{x}) . \quad (4)$$

Lower Bound. For the lower bound, we first establish a reduction between stabilizer testing and the task of distinguishing uniformly random stabilizer states from the maximally mixed state. This reduction is analogous to the reduction between purity testing and the task of distinguishing Haar random states from the maximally

mixed state, which was previously considered in [12, 14, 25]. The difference is that the unitary group $U(n)$ is replaced with the Clifford group $\mathrm{Cl}(n)$. Based on this analogy, we adapt the proof strategy due to Harrow [25] to the Clifford group. We now outline the main ideas behind this proof strategy in more detail.

Consider a single-copy algorithm with access to t copies of either a random stabilizer state $|S\rangle$ or the maximally mixed state. The first main insight is that since the states $\mathbf{E}_{|S\rangle\sim\mathrm{Stab}(n)}\left[|S\rangle\langle S|^{\otimes t}\right]$ and $I^{\otimes t}/2^{nt}$, which the algorithm is supposed to distinguish, both commute with all $C^{\otimes t}$ where $C \in \mathrm{Cl}(n)$, we can restrict our attention to measurements in the commutant of the t-fold tensor action of the Clifford group, defined as

$$\{A \in \mathcal{L}((\mathbb{C}^2)^{\otimes n})^{\otimes t}) \mid [A, C^{\otimes t}] = 0, \forall C \in \mathrm{Cl}(n)\}\,. \tag{5}$$

This commutant was fully characterized in the seminal work [23]. In particular, it is spanned by operators $R(T)$ which are associated with certain subspaces $T \subset \mathbb{F}_2^{2t}$. Similar to the permutation operators spanning the commutant of the tensor action $U^{\otimes t}$ of the unitary group, these operators $R(T)$ are also approximately orthogonal with respect to the Hilbert-Schmidt inner product in the regime where $t \ll n$.

The second main insight is that any distinguishing two-outcome POVM $\{M_0, M_1\}$ corresponding to a single-copy measurement strategy is also *Positive Partial Transpose* (PPT), i.e., it satisfies

$$0 \preceq M_i^{\Gamma_S} \preceq I \quad \forall S \subseteq [t]\,. \tag{6}$$

Here, $^{\Gamma_S}$ denotes taking the partial transpose with respect to the subset of copies indexed by some $S \subseteq \{1, \ldots, t\}$. Motivated by this insight, we study the partial transposes of the unitary operators $R(O)$ which form an important subgroup of the commutant (see Section 4.2 in [23]). In particular, we are able to characterize the singular values of the partial transposes $R(O)^{\Gamma_S}$.

Finally, using our insights about the partial transposes, we are able to leverage both the PPT constraint in Eq. (6) and the approximate orthogonality of the operators $R(T)$ to derive the sample complexity lower bound.

1.3 Related Work

Learning Stabilizer States and Their Generalizations. Our work is closely related to a line of work exploring the learnability in a tomographic sense of stabilizer states and their generalizations. This line started with [1, 42] giving single- and multi-copy algorithms for learning stabilizer states, respectively. Since then, these works were generalized to learning states with large stabilizer dimension [16, 20, 35, 37], meaning states stabilized by a non-maximal abelian subgroup of Pauli operators, and to higher degree binary phase states [2], where degree 2 corresponds to stabilizer states. In all of these works, the underlying assumption is that the unknown state $|\psi\rangle$ belongs to the restricted class of states that is to be learned. This assumption sets these works apart from a testing scenario where the very goal is to test the validity of such an assumption.

Agnostic Tomography of Stabilizer States. Recent works have shifted towards agnostic learning of states (also called agnostic tomography) where $|\psi\rangle$ can be arbitrary and the goal is to learn the best approximation from a given class of states. This agnostic learning framework naturally also gives rise to testing algorithms. For instance, [21] addressed agnostic tomography of stabilizer states, [19] focused on stabilizer product states, and [15] extended the approach to stabilizer states, states with large stabilizer dimensions, and product states.

Tolerant Stabilizer Testing. Another exciting direction is *tolerant* stabilizer testing, where the goal is to determine if an unknown state $|\psi\rangle$ is approximately a stabilizer state, rather than exactly one or far from it. Recent work [3, 4, 21, 41], building on [23], has shown that by repeatedly running the algorithm from [23], one can achieve tolerant testing across a broad range of parameters.

Measuring Non-Stabilizerness. Characterizing the amount of non-stabilizerness in a quantum state is also an active topic of research in a more physics-oriented literature. In particular, [26, 27] showed how to use 2-copy Bell measurements to quantify non-stabilizerness in quantum states. However, to the best of our knowledge, there is to date no efficient single-copy protocol for measuring non-stabilizerness. While it has been realized that the so-called *stabilizer entropies* [36] can also be measured via single-copy protocols [36, 44], the proposed protocols feature an exponential sample complexity.

1.4 Open Problems

As a result of this work, we identified multiple interesting open problems:

Matching Upper and Lower bounds. Currently, our $\Omega(\sqrt{n})$ lower bound and our $O(n)$ upper bound on the sample complexity are not quite matching. The lower bound can potentially be tightened by a sharper analysis of the orthogonality of the operators $R(T)$ spanning the commutant as well as an even more refined understanding of their partial transposes.

Lower Bound via the Tree Representation Framework. As mentioned before, [12] obtains a lower bound for purity testing by considering an analogous distinguishing task to the one we consider. However, they bound the sample complexity of this distinguishing task via the so-called tree representation framework instead of the PPT relaxation. We discuss obstacles to this approach in our case in the full version of our paper [28]. It remains an open question, if our lower bound can also be obtained by means of the tree representation framework.

Tolerant Single-Copy Stabilizer Testing, Efficiently Measuring Non-Stabilizerness. Can our single-copy stabilizer testing algorithm be made tolerant? A related open question is: Are there efficient single-copy algorithms for measuring some measure of non-stabilizerness?

Multi-Copy Access but Restricted Memory. Ref. [14] considers an intermediate scenario between single- and multi-copy access. Therein, multi-copy access is allowed, however the overall quantum memory is only $k < 2n$ qubits, so one cannot operate across two full copies of the unknown state $|\psi\rangle$. In this setting, they establish a phase transition for the sample complexity of purity testing. Due to the structural similarities between purity testing and stabilizer testing, it is interesting to try to generalize their result to stabilizer testing.

Full Version. Due to space constraints, most proofs have been omitted. These can be found in the full version of this manuscript [28].

2 Preliminaries

We begin by setting some notation. We often denote pure states by $\psi = |\psi\rangle\langle\psi|$. For a positive integer n, we define $[n] := \{1, \ldots, n\}$. For a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_K\}$, we denote their span by $\langle \mathbf{v}_1, \ldots, \mathbf{v}_K \rangle$. We denote by $\mathbb{F}_2$ the finite field of 2 elements and by $\mathbb{F}_2^n$ the n-dimensional vector space over this field. We denote by $\mathbf{0}_n \in \mathbb{F}_2^n$ the zero-vector in this vector space. For p a distribution over a set S, we denote drawing a sample $x \in S$ according to p by $x \sim p$. For p a distribution over $\mathbb{F}_2^n$ and $H \subseteq \mathbb{F}_2^n$ a subset, we denote the weight of p on the set H by $p(H) := \sum_{\mathbf{x}\in H} p(\mathbf{x})$.

We will make extensive use of the following standard fact of probability theory on linear spaces.

Lemma 2.1. *Let p be a distribution over $\mathbb{F}_2^n$ and let H be a subspace of $\mathbb{F}_2^n$ and let $p(H)$ be the probability weight of p on H. Then, for i.i.d. $\mathbf{x}_1, \ldots, \mathbf{x}_K \sim p$,*

$$\Pr_{\mathbf{x}_1,\ldots,\mathbf{x}_K\sim p} \left(\langle \mathbf{x}_1, \ldots, \mathbf{x}_K\rangle \subseteq H\right) = \left(p(H)\right)^K . \tag{7}$$

2.1 Vector Spaces Over $\mathbb{F}_2$

In this work, we will deal repeatedly with the vector spaces $\mathbb{F}_2^n$ and $\mathbb{F}_2^{2n}$ and two different inner products on them:

Definition 2.2 (Standard inner product). For $\mathbf{a}, \mathbf{b} \in \mathbb{F}_2^n$, we define their *standard inner product* as

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + \cdots + a_n b_n \tag{8}$$

where operations are performed over $\mathbb{F}_2$.

Note that the standard inner product is also defined on $\mathbb{F}_2^{2n}$. For instance, for $\mathbf{x} = (\mathbf{a}, \mathbf{b}), \mathbf{y} = (\mathbf{v}, \mathbf{w}) \in \mathbb{F}_2^{2n}$ we have $\mathbf{x} \cdot \mathbf{y} = \mathbf{a} \cdot \mathbf{v} + \mathbf{b} \cdot \mathbf{w}$.

Definition 2.3 (Orthogonal complement). Let $H \subseteq \mathbb{F}_2^n$ be a subspace. The *orthogonal complement* of H, denoted by $H^\perp$, is defined by

$$H^\perp := \left\{a \in \mathbb{F}_2^n : \mathbf{a} \cdot \mathbf{b} = 0 ,\ \forall\, \mathbf{b} \in H\right\} . \tag{9}$$

Fact 2.4. *Let H be a subspace of $\mathbb{F}_2^n$. Then:*

- $H^\perp$ is a subspace.
- $\left(H^\perp\right)^\perp = H$.
- $\dim(H) + \dim\left(H^\perp\right) = n$

For elements of $\mathbb{F}_2^{2n}$, we additionally also introduce the symplectic inner product.

Definition 2.5 (Symplectic inner product). For $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^{2n}$, we define their *symplectic inner product* as

$$[\mathbf{x}, \mathbf{y}] = x_1 y_{n+1} + x_2 y_{n+2} + \cdots + x_{2n} y_n \tag{10}$$

where operations are performed over $\mathbb{F}_2$.

A subspace $T \subset \mathbb{F}_2^{2n}$ is said to be *isotropic* when for all $\mathbf{x}, \mathbf{y} \in T$, $[\mathbf{x}, \mathbf{y}] = 0$. A *Lagrangian* subspace $M \subset \mathbb{F}_2^{2n}$ is an isotropic subspace of maximal dimension, namely of dimension $\dim(M) = n$.

2.2 Phaseless Pauli Operators and the Characteristic Distribution

In this section we recall some well-known facts about the Pauli group. The single-qubit Pauli matrices are denoted by $\{I, X, Y, Z\}$. The n-qubit Pauli group $\mathcal{P}_n$ is the set $\{\pm 1, \pm i\} \times \{I, X, Y, Z\}^{\otimes n}$. The Clifford group is the normalizer of the Pauli group. We denote the n-qubit Clifford group by $\mathrm{Cl}(n)$.

In this work, we will deal almost exclusively with phaseless Pauli operators (also often referred to as Weyl operators, c.f. [20, 23]), i.e., elements of the set $\{I, X, Y, Z\}^{\otimes n}$. Hence, whenever we refer to a Pauli operator, we refer to its phaseless version unless explicitly stated otherwise. Since $\{I, X, Y, Z\}^{\otimes n}$ is one-to-one with $\mathbb{F}_2^{2n}$, we can label elements of this set by bitstrings of length $2n$ as follows. Let $\mathbf{x} = (\mathbf{a}, \mathbf{b}) = (a_1, a_2, \ldots, a_n, b_1, b_2 \ldots, b_n) \in \mathbb{F}_2^{2n}$. We then define

$$P_{\mathbf{x}} = i^{\mathbf{a}\cdot\mathbf{b}} \left(X^{a_1} Z^{b_1}\right) \otimes \cdots \otimes \left(X^{a_n} Z^{b_n}\right) = i^{\mathbf{a}\cdot\mathbf{b}} X^{\mathbf{a}} Z^{\mathbf{b}} . \tag{11}$$

Here, as an exception, the inner product $\mathbf{a} \cdot \mathbf{b}$ on the phase in front is understood as being an integer rather than mod 2. Throughout this work, we will often identify bitstrings in $\mathbf{x} \in \mathbb{F}_2^{2n}$ with their corresponding (phaseless) Pauli operator $P_{\mathbf{x}}$. For instance, we will identify $\mathcal{Z} := \{I, Z\}^{\otimes n}$ with $\mathbf{0}_n \times \mathbb{F}_2^n$.

The commutation relations between Pauli operators are captured by the symplectic inner product of their corresponding bitstrings. In particular, for $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^{2n}$, the corresponding Pauli operators $P_{\mathbf{x}}, P_{\mathbf{y}}$ commute if $[\mathbf{x}, \mathbf{y}] = 0$ and anticommute if $[\mathbf{x}, \mathbf{y}] = 1$.

Any n-qubit state ρ can be expanded in the Pauli basis as

$$\rho = \frac{1}{2^n} \sum_{\mathbf{x}\in\mathbb{F}_2^n} \mathrm{tr}\left(\rho P_{\mathbf{x}}\right) P_{\mathbf{x}} . \tag{12}$$

The squared coefficients of this expansion form a probability distribution known as the *characteristic distribution*:

Definition 2.6 (Characteristic distribution [20, 23]). Let $|\psi\rangle$ be an n-qubit pure state, then its characteristic distribution p_ψ is defined via

$$p_\psi(\mathbf{x}) = \frac{1}{2^n} \mathrm{tr}\left(\psi P_{\mathbf{x}}\right)^2 . \tag{13}$$

Note that this distribution is automatically normalized because $\frac{1}{2^n}\sum_{\mathbf{x}} \mathrm{tr}\left(\psi P_{\mathbf{x}}\right)^2 = \mathrm{tr}\left(\psi^2\right) = 1$.

2.3 Single-Copy Algorithms

This work characterizes the complexity of stabilizer testing under the restriction to single-copy algorithms. Here, we provide a formal definition of such algorithms and note that similar definitions can be found in [12, 14] under the term *learning algorithms without quantum memory*. In general, a quantum learning algorithm accesses multiple identical copies of an unknown state ρ and uses quantum measurements and classical post-processing to predict its properties. Throughout the paper, we take the total number of copies of the unknown state to be t. Single-copy algorithms form a restricted class of quantum learning algorithms defined as follows:

Definition 2.7 (Single-copy algorithm). Let ρ be an n-qubit quantum state. A *single-copy* learning/testing algorithm using in total t copies of ρ operates in t sequential rounds. In each round, the algorithm measures a fresh copy of ρ using a POVM in $\mathbb{C}^{2^n \times 2^n}$ which could be chosen adaptively, i.e., it could depend on the measurement

results of previous rounds. After all t rounds, the algorithm predicts the desired properties of ρ based on the measurement results of all t rounds.

In contrast, general quantum learning algorithms are allowed to process all t copies simultaneously by applying a global POVM in $\mathbb{C}^{2^{nt}\times 2^{nt}}$.

2.4 Computational Difference Sampling

The core measurement routine of the testing algorithm introduced in this paper (Theorem 1.2) relies on *computational difference sampling*. This primitive was introduced in [20] and refers to measuring two copies of a state in the computational basis and adding the obtained outcomes via bit-wise addition mod 2 in classical post-processing. More formally, we have the following definition:

Definition 2.8 (Computational difference sampling). Computational difference sampling a quantum state $|\psi\rangle$ corresponds to the following quantum measurement. (1) Measure $|\psi\rangle$ in the computational basis to get outcome $\mathbf{a} \in \mathbb{F}_2^n$. (2) Measure an additional copy of $|\psi\rangle$ in the computational basis to get another outcome $\mathbf{b} \in \mathbb{F}_2^n$. (3) Output $\mathbf{a}+\mathbf{b} \in \mathbb{F}_2^n$ (addition mod 2). Let $r_\psi(\mathbf{a})$ denote the probability of sampling $\mathbf{a} \in \mathbb{F}_2^n$ through this process.

Importantly, computational difference sampling only uses single-copy measurements. In this work, we will repeatedly make use of the following key correspondence between subspace weights of the computational difference sampling distribution r_ψ and the characteristic distribution p_ψ:

Lemma 2.9 (Subspace weight correspondence between r_ψ and p_ψ). *Let $|\psi\rangle$ be a pure n-qubit quantum state. Given a subspace $H \subseteq \mathbb{F}_2^n$, consider its orthogonal complement $H^\perp$ (with respect to the standard inner product on $\mathbb{F}_2^n$). Then,*

$$r_\psi(H) = |H|\, p_\psi\left(\mathbf{0}_n \times H^\perp\right), \tag{14}$$

where $\mathbf{0}_n \times H^\perp \subseteq \mathcal{Z}$.

Importantly, Lemma 2.9 says that the weight of r_ψ on any subspace only depends on the weight of p_ψ on a corresponding subspace of Paulis in $\mathcal{Z}$. We note that Lemma 2.9 is already implicit in Corollary 8.5. in [20]. However, their statement is presented with respect to the symplectic complement since they work throughout with the symplectic inner product on $\mathbb{F}_2^{2n}$ instead of the standard inner product on $\mathbb{F}_2^n$. Here, we find it more convenient to state this relationship in terms of the orthogonal complement.

2.5 Stabilizer States and Isotropic Subspaces

A pure n-qubit state is a *stabilizer state* if there exists an Abelian group $S \subset \mathcal{P}_n$ of 2^n Pauli operators $P \in \mathcal{P}_n$ (with phase of +1 or −1) such that

$$S = \{P \in \mathcal{P}_n : P|\psi\rangle = |\psi\rangle\}. \tag{15}$$

This abelian group is the *stabilizer group* of the stabilizer state and determines it uniquely. Henceforth, we denote stabilizer states by $|S\rangle$ and denote the set of all pure n-qubit stabilizer states by $\mathrm{Stab}(n)$. By considering the phaseless (or unsigned) versions of the Pauli operators forming a stabilizer group, every stabilizer group can be associated to a Lagrangian subspace M. That is, Lagrangian subspaces are in a one-to-one correspondence with unsigned stabilizer groups. There are 2^n many different stabilizer states corresponding to the same Lagrangian subspace M, with each stabilizer state corresponding to a different choice of signs of the Pauli operators making up the unsigned stabilizer group. For more background on the connection between symplectic geometry and the stabilizer formalism, see [22, 23].

Every n-qubit pure stabilizer state can (up to a global phase) be written in the form

$$\frac{1}{|A|}\sum_{x\in A}(-1)^{q(x)}\, i^{l(x)}\, |x\rangle, \tag{16}$$

where A is an affine subspace of $\mathbb{F}_2^n$ and $l, q : \mathbb{F}_2^n \to \mathbb{F}_2$ are linear and quadratic (respectively) polynomials over $\mathbb{F}_2$ [18, 46]. This implies that when measured in the computational basis, the output distributions of stabilizer states are uniform over affine subspaces. Furthermore, computational difference sampling is convenient precisely because it removes any affine shift and hence leads to distributions that are uniform over linear subspaces of $\mathbb{F}_2^n$.

2.6 Stabilizer Fidelity and Stabilizer Testing

We denote the set of all pure n-qubit stabilizer states by $\mathrm{Stab}(n)$. A rich topic in quantum information is concerned with determining the non-stabilizerness of quantum states, also commonly referred to as magic. For more background on a range of magic measures, see e.g. [39].

Here, we will specifically focus on the so-called stabilizer fidelity:

Definition 2.10 (Stabilizer fidelity, [6]). Let $|\psi\rangle$ be a pure n-qubit quantum state. The *stabilizer fidelity* of $|\psi\rangle$ is defined as

$$F_{\mathrm{Stab}}(|\psi\rangle) := \max_{|S\rangle\in\mathrm{Stab}(n)} |\langle S|\psi\rangle|^2. \tag{17}$$

The stabilizer (in-)fidelity serves as a magic measure for pure states. Arguably, a simpler problem than trying to determine or estimate the non-stabilizerness of an unknown quantum state $|\psi\rangle$, is simply to decide if $|\psi\rangle$ is a stabilizer state or has at least a bit of magic without necessarily quantifying how much magic. This task is known as *stabilizer testing*, and it has received attention in the past in the context of quantum property testing [3, 8, 20, 23]. Formally, the task is the following:

Definition 2.11 (Stabilizer testing). Given $\epsilon > 0$ and access to copies of an unknown n-qubit pure state $|\psi\rangle$, decide if $|\psi\rangle$ is

(1) either a stabilizer state, i.e., $F_{\mathrm{Stab}}(|\psi\rangle) = 1$,
(2) or at least ϵ far from all stabilizer states, i.e., $F_{\mathrm{Stab}}(|\psi\rangle) \le 1-\epsilon$.

It is promised that $|\psi\rangle$ satisfies one of the two cases.

In this work, we propose and analyze a single-copy algorithm for stabilizer testing. Our analysis relies on the fact that the stabilizer fidelity can be related to the characteristic distribution. To state this relation, we start by defining the following set:

Definition 2.12 (M_ψ). Let $|\psi\rangle$ be a pure n-qubit quantum state. Then define M_ψ as the following subset of $\mathbb{F}_2^{2n}$ corresponding to $|\psi\rangle$.

$$M_\psi := \{\mathbf{y} \in \mathbb{F}_2^{2n} : \mathrm{tr}\left(\psi P_{\mathbf{y}}\right)^2 > \tfrac{1}{2}\}. \tag{18}$$

We have the following fact for M_ψ.

Fact 2.13 ([20, 23]). *For all $|\psi\rangle$, the set M_ψ is isotropic. That is, for all $\mathbf{x}, \mathbf{y} \in M_\psi$, $[\mathbf{x}, \mathbf{y}] = 0$. Furthermore, $|M_\psi| \leq 2^n$.*

Note that, in general, the set M_ψ is not necessarily a subspace, in particular, it need not be closed under addition. However, for a stabilizer state $|S\rangle$, its corresponding set M_S corresponds precisely to the Lagrangian subspace associated with the stabilizer state. We now state a lower bound on the stabilizer fidelity $F_{\text{Stab}}(|\psi\rangle)$ of a state $|\psi\rangle$ in terms of its characteristic distribution p_ψ and the corresponding set M_ψ:

Fact 2.14 (Proof of Theorem 3.3 in [23], Corollary 7.4. in [21]). *Let $|\psi\rangle$ be an n-qubit pure state and M be a Lagrangian subspace of $\mathbb{F}_2^{2n}$, then*

$$F_{\text{Stab}}(|\psi\rangle) \geq p_\psi(M) . \tag{19}$$

In particular, it follows that

$$F_{\text{Stab}}(|\psi\rangle) \geq p_\psi\left(M_\psi\right), \tag{20}$$

since M_ψ can always be completed to a Lagrangian subspace.

2.7 The Action of (Random) Cliffords on $\mathbb{F}_2^{2n}$

Clifford circuits map Pauli operators to Pauli operators under conjugation. In particular, when applying a Clifford circuit C to a phaseless Pauli operator $P_{\mathbf{x}}$, the resulting Pauli is not necessarily a phaseless one. However, when identifying Pauli operators with their phaseless versions, we can define the Clifford action on $\mathbb{F}_2^{2n}$: We have that $C(\mathbf{x}) = \mathbf{y}$ if

$$CP_{\mathbf{x}}C^\dagger = \pm P_{\mathbf{y}} . \tag{21}$$

Similarly, for a subset $S \subseteq \mathbb{F}_2^{2n}$, we write $C(S)$ to denote $C(S) = \{C(\mathbf{x}) : \mathbf{x} \in S\}$. In fact, the action of the Clifford group on $\mathbb{F}_2^{2n}$ is precisely that of the symplectic group $\text{Sp}(2n, \mathbb{F}_2)$. Hence, in particular, the Clifford action preserves the symplectic inner product, $[C(\mathbf{x}), C(\mathbf{y})] = [\mathbf{x}, \mathbf{y}]$.

When we apply a Clifford unitary to a state $|\psi\rangle$, we shuffle the individual probabilities of the characteristic distribution p_ψ around. In particular, the mapping is precisely described by the Clifford action on $\mathbb{F}_2^{2n}$ since the characteristic distribution p_ψ only depends on the phaseless Pauli operators. Formally, we have:

Fact 2.15. *Let $|\psi\rangle$ be an n-qubit pure state, let $C \in \text{Cl}(n)$ be a Clifford unitary, then for all $\mathbf{x} \in \mathbb{F}_2^{2n}$*

$$p_{C|\psi\rangle}(\mathbf{x}) = p_\psi\left(C^\dagger(\mathbf{x})\right) . \tag{22}$$

Random Cliffords. In this paper, we are mainly interested in applying uniformly random Clifford unitaries to pure states. In particular, a core part of the analysis of our algorithm relies on the action of random Cliffords, and hence random symplectic transformations, on isotropic or even Lagrangian subspaces of $\mathbb{F}_2^{2n}$. As demonstrated in several works [5, 7, 33], random Clifford unitaries can be sampled and compiled into Clifford circuits efficiently:

Fact 2.16. *There is a classical algorithm that samples a uniformly random element C of the n-qubit Clifford group $\text{Cl}(n)$ and outputs a Clifford circuit implementation of C in time $O\left(n^2\right)$.*

A uniformly random Clifford C corresponds to a uniformly random symplectic transformation. When such a transformation acts on an isotropic subspace T of $\mathbb{F}_2^{2n}$, it results in a uniformly random isotropic subspace of the same dimension. In particular, a random Clifford C acting on a Lagrangian subspace M results in a uniformly random Lagrangian subspace $C(M)$. In this work, the Lagrangian subspace $\mathcal{Z} = \mathbf{0}_n \times \mathbb{F}_2^n$ plays a distinguished role because of its relation to computational difference sampling. In particular, let us now answer the following question: What is the probability that a uniformly random Lagrangian subspace $C(M)$ has a k-dimensional intersection with the fixed Lagrangian subspace $\mathcal{Z}$? This is captured by the following definition:

Definition 2.17 (Probability of k-dimensional intersection). Let M be a Lagrangian subspace of $\mathbb{F}_2^{2n}$ and let $0 \leq k \leq n$. Then, we define

$$Q(n,k) := \Pr_{C\sim\text{Cl}(n)}\left(\dim\left(C(M) \cap \mathcal{Z}\right) = k\right). \tag{23}$$

This probability can be expressed as follows,

$$Q(n,k) = \frac{|\text{Cl}(n,k)|}{|\text{Cl}(n)|}, \tag{24}$$

where we introduced

$$\text{Cl}(n,k) := \{C \in \text{Cl}(n) : \dim\left(C(M) \cap \mathcal{Z}\right) = k\} . \tag{25}$$

To characterize $Q(n,k)$, we can count the relevant sets of Lagrangian subspaces. First, we start with the total number of them:

Fact 2.18. *The total number of Lagrangian subspaces of $\mathbb{F}_2^{2n}$ is given by*

$$\mathcal{T}(n) = \prod_{i=1}^{n}\left(2^i + 1\right) = 2^{\frac{1}{2}n(n+1)} \prod_{i=1}^{n}\left(1 + \frac{1}{2^i}\right). \tag{26}$$

Next, the number of Lagrangian subspaces whose intersection with $\mathcal{Z}$ is k-dimensional was obtained already in Corollary 2 in Ref. [34]:

Lemma 2.19 (Corollary 2 in [34]). *Let M be a fixed Lagrangian subspace of $\mathbb{F}_2^{2n}$. The number of Lagrangian subspaces N whose intersection with M is k-dimensional is given by*

$$\kappa(n,k) := \binom{n}{k}_2 \cdot 2^{\frac{1}{2}(n-k)(n-k+1)}, \tag{27}$$

where $\binom{n}{k}_2$ is the Gaussian binomial coefficient given by (for $k \leq n$)

$$\binom{n}{k}_2 = \prod_{i=0}^{k-1}\frac{2^{n-i}-1}{2^{k-i}-1} . \tag{28}$$

In particular, for $k = 0, 1, 2$, we have

$$\kappa(n,0) = 2^{\frac{1}{2}n(n+1)}, \tag{29}$$

$$\kappa(n,1) = \left(2^n - 1\right) 2^{\frac{1}{2}(n-1)n}, \tag{30}$$

$$\kappa(n,2) = \frac{\left(2^n-1\right)\left(2^{n-1}-1\right)}{3} 2^{\frac{1}{2}(n-1)(n-2)} . \tag{31}$$

Hence, we can characterize $Q(n,k)$ as follows:

Corollary 2.20. *Let $0 \leq k \leq n$. Then,*

$$Q(n,k) = \frac{\kappa(n,k)}{\mathcal{T}(n)} . \tag{32}$$

In particular, for $k = 0, 1, 2$, we have

$$Q(n,0) = \prod_{i=1}^{n} \left(\frac{1}{1+2^{-i}} \right), \tag{33}$$

$$Q(n,1) = \frac{2^n - 1}{2^n} Q(n,0), \tag{34}$$

$$Q(n,2) = \frac{1}{3} \frac{(2^n-1)(2^n-2)}{2^{2n}} Q(n,0). \tag{35}$$

Note that for fixed n, $Q(n,k)$ forms a probability distribution over $k \in \{0,\dots,n\}$ and is normalized as $\sum_{k=0}^{n} Q(n,k) = 1$. We remark two important aspects about the distribution $Q(n,k)$: First, the distribution converges quickly for increasing n, so that the values of $Q(n,k)$ do not change significantly above $n = 10$. Secondly, the distribution decays very quickly and, regardless of the value of n, the primary contributions come from $k = 0, 1, 2, 3$. The largest contribution is always $Q(n,0)$ whose limiting value is given by

$$Q(n,0) \overset{n\to\infty}{\to} 0.41942244\ldots \tag{36}$$

It approaches this value monotonously from above. For large enough n, say $n \geq 7$, it is useful to keep the following approximate values in mind $Q(n,0) \approx 0.42$ and $Q(n,1) \approx Q(n,0)$ and $Q(n,2) \approx \frac{Q(n,1)}{3}$.

2.8 Commutant of Clifford Tensor Powers

A key ingredient for deriving our lower bound is the commutant of the t-fold tensor power action of the Clifford group $\mathrm{Cl}(n)$, i.e., the linear space of operators on $(\mathbb{C}^2)^{\otimes n})^{\otimes t}$ that commute with $C^{\otimes t}$ for all $C \in \mathrm{Cl}(n)$. Formally, we define it as follows:

Definition 2.21 (Commutant of t-th Clifford tensor power action). We define $\mathrm{Comm}(\mathrm{Cl}(n), t)$ as follows

$$\mathrm{Comm}(\mathrm{Cl}(n), t) := \{A \in \mathcal{L}((\mathbb{C}^2)^{\otimes n})^{\otimes t}) \mid [A, C^{\otimes t}] = 0, \forall C \in \mathrm{Cl}(n)\}. \tag{37}$$

The seminal work [23] fully characterized this commutant in terms of so-called stochastic Lagrangian subspaces defined as follows:

Definition 2.22 (Stochastic Lagrangian subspaces). The set $\Sigma_{t,t}$ denotes the set of all subspaces $T \subseteq \mathbb{F}_2^{2t}$ with the following properties:

(1) $\mathbf{x}\cdot\mathbf{x} = \mathbf{y}\cdot\mathbf{y} \mod 4$ for all $(\mathbf{x},\mathbf{y}) \in T$,
(2) $\dim(T) = t$,
(3) $\mathbf{1}_{2t} = (1,\dots,1) \in T$.

We refer to elements in $\Sigma_{t,t}$ as *stochastic Lagrangian subspaces*.

In particular, the key result of [23] is that $\mathrm{Comm}(\mathrm{Cl}(n), t)$ is spanned by operators $R(T)$ associated with the stochastic Lagrangian subspaces $T \in \Sigma_{t,t}$.

Theorem 2.23 (Theorem 4.3 in [23]). *If $n \geq t-1$, then $\mathrm{Comm}(\mathrm{Cl}(n), t)$ is spanned by the linearly independent operators $R(T) := r(T)^{\otimes n}$, where $T \in \Sigma_{t,t}$ and*

$$r(T) := \sum_{(\mathbf{x},\mathbf{y})\in T} |x\rangle\langle y| . \tag{38}$$

Alongside this central characterization, [23] proved several results about the operators $R(T)$:

Fact 2.24 (Eq. (4.10) in [23]). *For all $T \in \Sigma_{t,t}$ and all stabilizer states $|S\rangle$, it holds that*

$$\langle S|^{\otimes t} R(T) |S\rangle^{\otimes t} = 1 . \tag{39}$$

Fact 2.25 (Traces of $R(T)$, Remark 5.1 in [23]). *Let $T \in \Sigma_{t,t}$ and let $\Delta = \{(\mathbf{x},\mathbf{x}) \mid \mathbf{x} \in \mathbb{F}_2^t\}$ be the diagonal subspace. Let $l = t - \dim(T \cap \Delta)$, then*

$$\mathrm{tr}\, R(T) = (\mathrm{tr}\, r(T))^n = 2^{n(t-l)} . \tag{40}$$

Furthermore,

$$\sum_{T\in\Sigma_{t,t}} \mathrm{tr}\, R(T) = 2^{nt} (-2^{-n};2)_{t-1} , \tag{41}$$

where the q-Pochhammer symbol $(-2^{-n};2)_{t-1}$ is given by $(-2^{-n};2)_{t-1} = \prod_{k=0}^{t-2}(1+2^{-n+k})$.

We note that $\mathrm{tr}(R(T)) = 2^{nt}$, i.e. $l = 0$, only for the identity element $T = e$.

Fact 2.26 (Cardinality of $\Sigma_{t,t}$, Theorem 4.10 in [23]).

$$\left|\Sigma_{t,t}\right| = \prod_{k=0}^{t-2} \left(2^k + 1\right) \leq 2^{\frac{1}{2}(t^2+5t)} . \tag{42}$$

Ref. [23] also characterized the commutant further by uncovering an important group structure within. To this end, need the following definition.

Definition 2.27 (Stochastic orthogonal group). The stochastic orthogonal group, denoted O_t, is defined as the group of $t \times t$ binary matrices O such that

$$O\mathbf{x}\cdot O\mathbf{x} = \mathbf{x}\cdot\mathbf{x} \mod 4 \qquad \forall \mathbf{x} \in \mathbb{F}_2^t . \tag{43}$$

Note that, for any $O \in O_t$, the subspace $T_O = \{(O\mathbf{x},\mathbf{x}) \mid \mathbf{x} \in \mathbb{F}_2^t\}$ is a stochastic Lagrangian subspace. That is, $T_O \in \Sigma_{t,t}$ for all $O \in O_t$. In the following, we will thus view O_t as a subset of $\Sigma_{t,t}$, i.e., $O_t \subset \Sigma_{t,t}$. We will denote the identity element in O_t by e, it corresponds to the diagonal subspace $\Delta = \{(\mathbf{x},\mathbf{x}) \mid \mathbf{x} \in \mathbb{F}_2^t\}$. Notice also that the symmetric group on t elements, denoted $\mathcal{S}_t$, can be viewed as a subgroup of O_t by considering its matrix representation on $\mathbb{F}_2^t$.

While for $O \in O_t$, the corresponding operators $R(O)$ are unitary, this is not the case for the operators $R(T)$ for $T \in \Sigma_{t,t} \setminus O_t$. Here, we record a bound on the trace-norm of these operators which we will require later.

Fact 2.28 (c.f. Lemma 1 in [24]). *Let $T \in \Sigma_{t,t} \setminus O_t$, then*

$$\|R(T)\|_1 \leq 2^{n(t-1)} . \tag{44}$$

Lastly, similar to [25], we want to quantify the orthogonality of the operators $R(T)$ spanning the commutant $\mathrm{Comm}(\mathrm{Cl}(n), t)$. To this end, we define their corresponding Gram matrix as follows:

Definition 2.29 (Gram matrix G corresponding to $\Sigma_{t,t}$). We define the Gram matrix corresponding to $\{R(T)\}_{T\in\Sigma_{t,t}}$ as the $|\Sigma_{t,t}| \times |\Sigma_{t,t}|$-matrix with entries given by

$$G^{(n,t)}_{T,T'} := \mathrm{tr}\left(R(T)^\dagger R(T')\right) \quad \text{for } T, T' \in \Sigma_{t,t} . \tag{45}$$

For convenience, we will often drop the superscript (n,t) on $G^{(n,t)}$. Next, we state a straightforward bound on the off-diagonal row-sums of the Gram matrix:

Fact 2.30 (Bound on the off-diagonal row-sum of the Gram matrix). *Let $T \in \Sigma_{t,t}$, then*

$$\sum_{T' \in \Sigma_{t,t} \setminus \{T\}} G_{T,T'} \leq |\Sigma_{t,t}| \cdot 2^{n(t-1)} \leq 2^{n(t-1)+\frac{1}{2}(t^2+5t)}. \tag{46}$$

3 Single-Copy Algorithm for Stabilizer Testing

In this section, we will prove the following theorem:

Theorem 3.1 (Upper bound for single-copy testing). *Let $n \geq 3, \epsilon > 3 \cdot 2^{-n}$ and let $|\psi\rangle$ be an n-qubit pure state. There exists a single-copy algorithm for stabilizer testing an unknown state $|\psi\rangle$ that uses $t = O\left(n/\epsilon^2\right)$ copies of $|\psi\rangle$, runs in time $O(n^3)$, and succeeds with high probability.*

This section is organized as follows: In Section 3.1, we introduce the main quantity of interest $\overline{\mathbb{P}_K}(|\psi\rangle)$, which we call *average spanning probability*, alongside a single-copy algorithm for estimating $\overline{\mathbb{P}_K}(|\psi\rangle)$ to additive precision ϵ. The average spanning probability $\overline{\mathbb{P}_K}(|\psi\rangle)$ captures the probability that $K = O(n)$ samples drawn randomly according to the computational difference sampling distribution $r_{C|\psi\rangle}$ of $C|\psi\rangle$ span the full space $\mathbb{F}_2^n$ for C drawn uniformly randomly from the Clifford group. The movation behind this quantity is that the stabilizer states are extremal with respect to this quantity. In Section 3.2, we derive an exact expression for this extremal value of the average spanning probability for stabilizer states. In Section 3.3, we discuss how the relation between r_ψ and p_ψ from Lemma 2.9 allows us to extend these calculations to non-stabilizer states. Lastly, in Section 3.4, we show that for a state $|\psi\rangle$ that is ϵ-far from all stabilizer states, the value of $\overline{\mathbb{P}_K}(|\psi\rangle)$ deviates by at least $\Omega(\epsilon)$ from the extremal value attained by the stabilizer states. This immediately implies the main result Theorem 3.1, because we have an algorithm to estimate $\overline{\mathbb{P}_K}(|\psi\rangle)$ and can efficiently compute the exact extremal stabilizer value.

3.1 Average Spanning Probability

Definition 3.2 (Spanning probability). Let $|\psi\rangle$ be an n-qubit pure state and let r_ψ be its computational difference sampling distribution. Let $\mathbf{v}_1, \ldots, \mathbf{v}_K \sim r_\psi$ and consider the event that $\mathbf{v}_1, \ldots, \mathbf{v}_K \in \mathbb{F}_2^n$ span the full space $\mathbb{F}_2^n$, in short $\langle \mathbf{v}_1, \ldots, \mathbf{v}_K \rangle = \mathbb{F}_2^n$. The *$K$-spanning probability* of $|\psi\rangle$, denoted $\mathbb{P}_K(|\psi\rangle)$ is defined as the probability of this event, i.e.,

$$\mathbb{P}_K(|\psi\rangle) := \Pr_{\mathbf{v}_1,\ldots,\mathbf{v}_K \sim r_\psi} \left(\langle \mathbf{v}_1, \ldots, \mathbf{v}_K \rangle = \mathbb{F}_2^n\right). \tag{47}$$

Note that $\mathbb{P}_K(|\psi\rangle)$ depends on the (positive integer) parameter K and we have that $\mathbb{P}_K(|\psi\rangle) = 0$ for $K < n$. Throughout this work, we will choose K such that $K \geq n$ but also $K = O(n)$. For this canonical choice, we will sometimes drop the K and refer to $\mathbb{P}_K(|\psi\rangle)$ simply as the spanning probability of $|\psi\rangle$.

Definition 3.3 (Average spanning probability). Let $|\psi\rangle$ be an n-qubit pure state. The *average spanning probability* $\overline{\mathbb{P}_K}(|\psi\rangle)$ of $|\psi\rangle$ is defined as

$$\overline{\mathbb{P}_K}(|\psi\rangle) := \mathop{\mathbf{E}}_{C \sim \mathrm{Cl}(n)} \left[\mathbb{P}_K(C|\psi\rangle)\right]. \tag{48}$$

The motivation for defining the average spanning probability is that it precisely captures the expectation value of the following simple random process consuming single copies of an n-qubit pure state $|\psi\rangle$.

(1) Draw a uniformly random Clifford $C \sim \mathrm{Cl}(n)$.
(2) Sample $\mathbf{v}_1, \ldots, \mathbf{v}_K \sim r_{C|\psi\rangle}$ by performing computational difference sampling on $C|\psi\rangle$.
(3) If $\langle \mathbf{v}_1, \ldots, \mathbf{v}_K \rangle = \mathbb{F}_2^n$, output 1, else output 0.

Note that the third step involves computing the rank of the $K \times n$ binary matrix M constructed by taking the samples $\mathbf{v}_1, \ldots, \mathbf{v}_K$ as the rows. We have that $\langle \mathbf{v}_1, \ldots, \mathbf{v}_K \rangle = \mathbb{F}_2^n$ is equivalent to rank $(M) = n$. By repeating the above 3-step process and averaging the outcomes, one can estimate $\overline{\mathbb{P}_K}(|\psi\rangle)$. Hence, we record the following lemma:

Lemma 3.4 (Estimating $\overline{\mathbb{P}_K}(|\psi\rangle)$ with single copies). *Let $|\psi\rangle$ be an n-qubit pure state and let $K \geq n$. Then, there exists a single-copy algorithm that, with probability $1 - \delta$, produces an estimate $\hat{r}$ such that*

$$\left|\hat{r} - \overline{\mathbb{P}_K}(\psi)\right| \leq \epsilon. \tag{49}$$

The algorithm consumes $O\left(\frac{K \log(1/\delta)}{\epsilon^2}\right)$ copies of $|\psi\rangle$ and runs in time $O\left(Kn^2\right)$.

As we will demonstrate through the course of this section, taking $K = O(n)$ and estimating $\overline{\mathbb{P}_K}(|\psi\rangle)$ to additive precision ϵ is sufficient for stabilizer testing. In particular, the time and sample complexities stated in Theorem 3.1 follow directly from Lemma 3.4.

3.2 The Stabilizer Value of $\overline{\mathbb{P}_K}(|\psi\rangle)$

Because of the average over the Clifford group $\mathrm{Cl}(n)$ in Definition 3.3, $\overline{\mathbb{P}_K}(|\psi\rangle)$ takes the same value for all stabilizer states $|S\rangle \in \mathrm{Stab}(n)$. We call this value the *stabilizer value*, and define it formally as follows:

Definition 3.5 (Stabilizer value). Let $|S\rangle \in \mathrm{Stab}(n)$ be an n-qubit pure stabilizer state. We define the stabilizer value of the average spanning probability as

$$\overline{\mathbb{P}_K}(\mathrm{Stab}(n)) := \overline{\mathbb{P}_K}(|S\rangle). \tag{50}$$

Next, we derive an expression for the stabilizer value in terms of n, K which can be efficiently computed exactly. To this end, we first characterize the K-spanning probability for stabilizer states.

Lemma 3.6 (Spanning probability for stabilizer states). *Let $K \geq n$ and let $|S\rangle$ be a stabilizer state. Furthermore, let M_S be the Lagrangian subspace corresponding to the stabilizer group of $|S\rangle$. Then, the K-spanning probability is given by*

$$\mathbb{P}_K(|S\rangle) = \begin{cases} \prod_{j=0}^{n-1} \left(1 - 2^{j-K}\right) & \dim(M_S \cap \mathcal{Z}) = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{51}$$

With Lemma 3.6 established, we can now obtain the stabilizer value for any combination of K and n. Recall, that for $K < n$, the spanning probability is always trivially zero, hence the following theorem focuses on the case $K \geq n$. For convenience, we point the reader to the definition of $Q(n, k)$ given in Definition 2.17.

Theorem 3.7 (Stabilizer value in terms of n, K). *Let $K \geq n$, then*

$$\overline{\mathbb{P}_K}(\mathrm{Stab}(n)) = Q(n, 0) \times \prod_{j=0}^{n-1} \left(1 - 2^{j-K}\right). \tag{52}$$

Proof. Let $|S\rangle$ be any n-qubit pure stabilizer state and let M_S be the Lagrangian subspace corresponding to the stabilizer group of $|S\rangle$. We partition the Clifford group as $\mathrm{Cl}(n) = \bigcup_{k=0}^{n} \mathrm{Cl}(n,k)$ where

$$\mathrm{Cl}(n,k) := \{C \in \mathrm{Cl}(n) : \dim(C(M_S) \cap \mathcal{Z}) = k\} . \quad (53)$$

Then,

$$\mathop{\mathbf{E}}_{C\sim\mathrm{Cl}(n)} [\mathbb{P}_K(C|S\rangle)] = \frac{1}{|\mathrm{Cl}(n)|} \sum_{k=0}^{n} \sum_{C\in\mathrm{Cl}(n,k)} \mathbb{P}_K(C|S\rangle) , \quad (54)$$

$$= \underbrace{\frac{|\mathrm{Cl}(n,0)|}{|\mathrm{Cl}(n)|}}_{=Q(n,0)} \times \prod_{j=0}^{n-1} \left(1 - 2^{j-K}\right) . \quad (55)$$

□

3.3 Going Beyond Stabilizer States

In the previous subsection, we found that the K-spanning probability $\mathbb{P}_K(|\psi\rangle)$ for a state $|\psi\rangle$ is easy to treat in the case of stabilizer states $|S\rangle$ because the distribution $r_{|S\rangle}$ is either uniform over $\mathbb{F}_2^n$ or confined to some proper subspace of $\mathbb{F}_2^n$. However, how can we obtain $\mathbb{P}_K(|\psi\rangle)$ for general pure states $|\psi\rangle$ where r_ψ does not take such a simple form?

Instead of calculating $\mathbb{P}_K(|\psi\rangle)$ exactly, in this section, we outline our approach for obtaining bounds on $\mathbb{P}_K(|\psi\rangle)$. Firstly, note that we can rewrite the K-spanning probability as follows,

$$\mathbb{P}_K(|\psi\rangle) = 1 - \Pr_{\mathbf{v}_1,\ldots,\mathbf{v}_K\sim r_\psi} \left(\bigcup_{L\in\mathcal{L}\setminus\{\mathbb{F}_2^n\}} (\langle \mathbf{v}_1,\ldots,\mathbf{v}_K\rangle = L) \right), \quad (56)$$

where the union goes over all subspaces $L \in \mathcal{L} \setminus \{\mathbb{F}_2^n\}$ and $\mathcal{L}$ is the set of all subspaces of $\mathbb{F}_2^n$. Note that we can restrict the union in Eq. (56) to range only over the $(n-1)$-dimensional subspaces of $\mathbb{F}_2^n$ because these contain all other smaller subspaces in $\mathcal{L}$. So, let us define $\mathcal{L}^{n-1}$ to be the set of all $2^n - 1$ many $(n-1)$-dimensional subspaces of $\mathbb{F}_2^n$. Then, we can rewrite $\mathbb{P}_K(|\psi\rangle)$ as

$$\mathbb{P}_K(|\psi\rangle) = 1 - \Pr_{\mathbf{v}_1,\ldots,\mathbf{v}_K\sim r_\psi} \left(\bigcup_{L\in\mathcal{L}^{n-1}} (\langle \mathbf{v}_1,\ldots,\mathbf{v}_K\rangle \subseteq L) \right). \quad (57)$$

Using a union bound and the general relation between r_ψ and p_ψ from Lemma 2.9, we find the following lower bound to the spanning probability:

Lemma 3.8 (Lower bound on the spanning probability from union bound). *Let $|\psi\rangle$ be an n-qubit pure state. Then, the spanning probability $\mathbb{P}_K(|\psi\rangle)$ can be lower bounded as follows,*

$$\mathbb{P}_K(|\psi\rangle) \geq 1 - \sum_{\mathbf{y}\in\mathcal{Z}\setminus\{I^{\otimes n}\}} \left[\frac{1 + \mathrm{tr}\left(\psi P_{\mathbf{y}}\right)^2}{2} \right]^K . \quad (58)$$

In the next section, we will make use of Lemma 3.8 in order to lower bound the average spanning probability for arbitrary states.

3.4 Bound in Terms of Stabilizer Fidelity

Below, we will prove our main technical theorem establishing a lower bound on $\overline{\mathbb{P}_K}(|\psi\rangle)$ in terms of p_ψ. However, let us briefly remark that the following direct approach fails: Apply the union bound from Lemma 3.8 and take the average over all Cliffords to obtain

$$\overline{\mathbb{P}_K}(|\psi\rangle) = \mathop{\mathbf{E}}_{C\sim\mathrm{Cl}(n)} [\mathbb{P}_K(C|\psi\rangle)] ,$$

$$\geq 1 - \sum_{\mathbf{y}\in\mathcal{Z}\setminus\{I^{\otimes n}\}} \mathop{\mathbf{E}}_{C\sim\mathrm{Cl}(n)} \left[\frac{1 + \mathrm{tr}\left(|\psi\rangle\langle\psi| C^\dagger P_{\mathbf{y}} C\right)^2}{2} \right]^K ,$$

$$= 1 - \frac{2^n - 1}{4^n - 1} \sum_{\mathbf{y}\in\mathbb{F}_2^{2n}\setminus\{\mathbf{0}_{2n}\}} \left[\frac{1 + \mathrm{tr}\left(|\psi\rangle\langle\psi| P_{\mathbf{y}}\right)^2}{2} \right]^K .$$

Here, in the second line, we used that for a random Clifford C, the rotated Pauli $C^\dagger P_{\mathbf{y}} C$ is just a uniformly random Pauli operator.

To see that this bound is not useful, we can e.g. evaluate it for any stabilizer state which, after some calculation, yields an exponentially small lower bound. This bound would not at all capture the stabilizer value which we know is the correct bound from Section 3.2 for any stabilizer state. The underlying issue is that the union bound from Lemma 3.8 is far too loose to be useful when applied like this.

Instead, to circumvent this issue, our approach is to again first partition the Clifford group similar to Eq. (53). Crucially, after conditioning on C belonging to $\mathrm{Cl}(n,k)$, the union bound actually gives rise to sharp bounds.

Theorem 3.9 (Lower bound on $\overline{\mathbb{P}_K}(|\psi\rangle)$ in terms of p_ψ). *Let $n \geq 3, K \geq 5n$ and let $|\psi\rangle$ be an n-qubit pure state. Furthermore, let $M_\psi = \{\mathbf{y} \in \mathbb{F}_2^{2n} : \mathrm{tr}\left(\psi P_{\mathbf{y}}\right)^2 > \frac{1}{2}\}$ and let M be any Lagrangian subspace of $\mathbb{F}_2^{2n}$ such that $M_\psi \subseteq M$. Note that such an M always exists. Then,*

$$\overline{\mathbb{P}_K}(\psi) - \overline{\mathbb{P}_K}(\mathrm{Stab}(n)) \geq Q(n,1)\left[1 - p_\psi(M)\right] - 2^{-n}. \quad (59)$$

From Fact 2.14, we have that for any Lagrangian subspace M, the stabilizer fidelity is lower bounded via $F_{\mathrm{Stab}}(\psi) \geq p_\psi(M)$. Hence, we immediately obtain the following corollary:

Corollary 3.10 (Difference from stabilizer value). *Let $n \geq 3, K \geq 5n$ and let $|\psi\rangle$ be an n-qubit pure state with stabilizer fidelity $F_{\mathrm{Stab}}(|\psi\rangle)$, then*

$$\overline{\mathbb{P}_K}(|\psi\rangle) - \overline{\mathbb{P}_K}(\mathrm{Stab}(n)) \geq Q(n,1)\left[1 - F_{\mathrm{Stab}}(|\psi\rangle)\right] - 2^{-n}. \quad (60)$$

Recall from Corollary 2.20 that $Q(n,1)$ is a large constant. Hence, Corollary 3.10 implies that for state $|\psi\rangle$ with stabilizer infidelity of at least ϵ, $\overline{\mathbb{P}_K}(|\psi\rangle)$ differs from the stabilizer value by $\Omega(\epsilon)$. This difference will hence be detected upon estimating $\overline{\mathbb{P}_K}(|\psi\rangle)$ to precision ϵ. Combining this insight with Lemma 3.4 yields Theorem 3.1.

4 Lower Bound for Single-Copy Stabilizer Testing

In this section, we will prove the following theorem:

Theorem 4.1 (Lower bound for single-copy stabilizer testing). *Any single-copy algorithm for stabilizer testing to accuracy $0 < \epsilon < 1 - n^2/2^n$ requires at least $t = \Omega(\sqrt{n})$ copies.*

This section is organized as follows: Firstly, in Section 4.1, we will argue that the sample complexity lower bound stated in Theorem 4.1 can be derived from the many-versus-one distinguishing task of distinguishing uniformly random stabilizer states from the maximally mixed state. This task is analogous to the reduction of purity testing to the task of distinguishing Haar random states from the maximally mixed state, which was previously considered in [12, 14, 25]. The main difference is that the unitary group is replaced with the Clifford group. Based on this observation, in Section 4.2, we adapt the proof strategy of [25] to the Clifford group. In particular, we use the fact that the operators spanning the commutant $\mathrm{Comm}(\mathrm{Cl}(n), t)$ are approximately orthogonal with respect to the Hilbert-Schmidt inner product in the regime where $t \ll n$ and combine this with new insights about partial transposes of the operators $R(O)$ for $O \in O_t$ to derive lower bounds on distinguishing uniformly random states from the maximally mixed state. In Section 4.3, we collect our results about partial transposes of $R(O)$ for $O \in O_t$. This subsection is modular and might be of independent interest.

4.1 Reduction to Distinguishing Random Stabilizer States From the Maximally Mixed State

Our starting point is Le Cam's two-point method, that is, we consider distinguishing tasks between ensembles of states:

Definition 4.2 (t-copy distinguishing task). Let μ and ν be two ensembles of n-qubit quantum states. We consider the following two events to happen with equal probability of 1/2:

- The unknown state ρ is sampled according to μ.
- The unknown state ρ is sampled according to ν.

Given t copies of ρ, the goal is to give a quantum algorithm that correctly distinguishes between these two events with probability $\geq 2/3$.

Throughout, we fix the number of copies to be t so that the algorithm has access to $\mathbf{E}[\rho^{\otimes t}]$ where the expectation is taken with respect to one of the ensembles μ, ν. For the purpose of this work, we consider the following three ensembles of states and corresponding states that the algorithm has access to:

(H) Haar random n-qubit states, $\rho_H := \mathbf{E}_{|\psi\rangle\sim\mu_H}\left[|\psi\rangle\langle\psi|^{\otimes t}\right]$
(S) uniformly random n-qubit stabilizer states, $\rho_S := \mathbf{E}_{|S\rangle\sim\mathrm{Stab}(n)}\left[\,|S\rangle\langle S|^{\otimes t}\,\right]$
(I) the maximally mixed n-qubit state, $\rho_I := I^{\otimes t}/2^{nt}$.

For each pair of ensembles, we can consider an associated distinguishing task. For instance, the pair (H, S) corresponds to distinguishing Haar random states from uniformly random stabilizer states. This task is also the natural starting point for proving a lower bound on the sample complexity of stabilizer testing since it can be reduced to stabilizer testing: With overwhelming probability, a Haar random state is far from all stabilizer states, so a stabilizer testing algorithm would likely reject it but accept a uniformly random stabilizer state. This observation is formalized via the following lemma:

Lemma 4.3. *Let $0 < \epsilon < 1 - n^2/2^n$. Then, any algorithm for stabilizer testing to accuracy ϵ using t copies can solve the t-copy distinguishing task of Haar random states versus uniformly random stabilizer states with probability $1 - 2^{-O(n^2)}$.*

This reduction, i.e., applying a stabilizer testing algorithm to distinguish between the Haar random ensemble and uniformly random stabilizer ensemble, may fail with a small probability, as indicated in Lemma 4.3, namely when the Haar randomly sampled state happens to be ϵ-close to a stabilizer state. In this event, it is not guaranteed that a stabilizer testing algorithm correctly distinguishes the two ensembles. To prove Lemma 4.3, we hence bound the probability of this event as follows:

Fact 4.4. *Let $0 < \epsilon < 1 - n^2/2^n$. Then, for a Haar random n-qubit state $|\psi\rangle$,*

$$\Pr_{|\psi\rangle}\left[\max_{|S\rangle\in\mathrm{Stab}(n)} |\langle S|\psi\rangle|^2 \geq 1-\epsilon\right] \leq 2^{-O(n^2)}. \tag{61}$$

Next, we argue that, when considering single-copy algorithms, any sample complexity lower bound for distinguishing uniformly random stabilizer states versus the maximally mixed state (the pair (S, I)) leads to a lower bound for the pair (H, S). This essentially follows from a triangle inequality between the three pairs as we now explain: Consider a single-copy distinguishing algorithm for the pair (H, S) using t copies of the unknown state ρ guessing 0 or 1, corresponding to the case (H) and (S), respectively. We can model this distinguishing algorithm via a two-outcome POVM $\{M_0, M_1\}$ acting on ρ_H or ρ_S, respectively. Letting $M = M_0 - M_1$, the bias achieved by the algorithm is given by $|\mathrm{tr}(M(\rho_H - \rho_S))|$. In order to correctly distinguish between case (H) and (S) with probability $\geq 2/3$, the bias needs to be at least $2/3$. Our goal is to upper bound the bias $|\mathrm{tr}(M(\rho_H - \rho_S))|$ in terms of the number of copies t.

By the triangle inequality, we have

$$|\mathrm{tr}(M(\rho_H - \rho_S))| \leq |\mathrm{tr}(M(\rho_H - \rho_I))| + |\mathrm{tr}(M(\rho_S - \rho_I))|. \tag{62}$$

For the pair (H, I), it is shown in [25] that any single-copy algorithm using at most t copies of ρ can achieve a bias of at most $|\mathrm{tr}(M(\rho_H - \rho_I))| = O(t^2\, 2^{-n/2})$. For the regime of $t \leq n$, this contribution to the RHS is hence exponentially small in n. As we will show $|\mathrm{tr}(M(\rho_S - \rho_I))|$ will be the dominant contribution to the RHS. Hence, for our purposes, it suffices to focus on bounding the bias for the pair (S, I), corresponding to distinguishing a uniformly random stabilizer state from the maximally mixed state.

We note that [12] proved an even stronger lower bound of $t = \Omega(2^{n/2})$ on the number of copies for distinguishing Haar random states from the maximally mixed state (the pair (H, I)). This translates to a bound $|\mathrm{tr}(M(\rho_H - \rho_I))| = O(t\, 2^{-n/2})$ on the bias.

4.2 Lower Bound via the Commutant of the Clifford Tensor Action

In this section, we focus on bounding the bias $|\mathrm{tr}(M(\rho_S - \rho_I))|$ where $\rho_S = \mathbf{E}_{|S\rangle}\left[|S\rangle\langle S|^{\otimes t}\right]$ and $\rho_I = I^{\otimes t}/2^{nt}$. That is we focus on the task of distinguishing uniformly random stabilizer states from the maximally mixed state given t copies of the unknown state. Our proof strategy closely resembles that of Ref. [25]. The key observation that we will use is that any measurement POVM $\{M_0, M_1\}$ implementable by an LOCC protocol [17] (corresponding to a single-copy measurement strategy), is also *Positive Partial*

Transpose (PPT), i.e.

$$0 \preceq M_i^{\Gamma_S} \preceq I \quad \forall S \subseteq [t]\,. \tag{63}$$

This implies that the difference $M = M_0 - M_1$ satisfies $-I \preceq M^{\Gamma_S} \preceq I$. Here, the superscript $^{\Gamma_S}$ denotes taking a partial transpose with respect to a subset $S \subset [t]$. Furthermore, note that ρ_S, ρ_I both commute with all $C^{\otimes t}$ for all $C \in \mathrm{Cl}(n)$, i.e., $\rho_S, \rho_I \in \mathrm{Comm}(\mathrm{Cl}(n), t)$. The second key insight is that, hence, WLOG we can restrict ourselves to consider M such that $\left[M, C^{\otimes t}\right] = 0$ as well. To see this, consider

$$\mathrm{tr}\left(M\ (\rho_S - \rho_I)\right) = \mathop{\mathbf{E}}_C \left[\mathrm{tr}\left(MC^{\otimes t}\ (\rho_S - \rho_I)\, C^{\dagger\otimes t}\right)\right], \tag{64}$$

$$= \mathrm{tr}\left(\mathop{\mathbf{E}}_C \left[C^{\dagger\otimes t} M C^{\otimes t}\right]\ (\rho_S - \rho_I)\right), \tag{65}$$

where $\mathbf{E}_C \left[C^{\dagger\otimes t}(\cdot)C^{\otimes t}\right]$ projects onto $\mathrm{Comm}(\mathrm{Cl}(n), t)$. Hence, we will henceforth assume that $M \in \mathrm{Comm}(\mathrm{Cl}(n), t)$ and so M can be expanded as $M = \sum_{T\in\Sigma_{t,t}} m_T R(T)$, where $R(T)$ are the operators spanning the commutant of the Clifford group, as discussed in Section 2.8.

We will now begin by showing an upper bound on $|\mathrm{tr}(R(T)\,M)|$ for all $T \neq e$. To do so, we use an upper bound on $\|R(O)^{\Gamma_S}\|_1 \leq 2^{n(t-1)}$ established in Corollary 4.11 in the next subsection, namely Section 4.3. This is ultimately also how we make use of the PPT constraint on the measurement.

Lemma 4.5 (Upper bound on $|\mathrm{tr}(R(T)\,M)|$). *Let $M = \sum_{T\in\Sigma_{t,t}} m_T\, R(T)$ be such that it satisfies the condition*

$$-I \preceq M^{\Gamma_S} \preceq I \quad \forall S \subseteq [t]\,. \tag{66}$$

Then, for all $T \in \Sigma_{t,t}$ such that $T \neq e$,

$$|\mathrm{tr}(R(T)\,M)| \leq 2^{n(t-1)}\,. \tag{67}$$

Next, we establish a bound on the magnitudes $|m_T|$ for all $T \neq e$ when M is *traceless* and PPT:

Lemma 4.6 (Upper bound on $|m_T|$). *Let $M = \sum_{T\in\Sigma_{t,t}} m_T\, R(T)$ be such that $\mathrm{tr}(M) = 0$ and it satisfies the condition*

$$-I \preceq M^{\Gamma_S} \preceq I \quad \forall S \subseteq [t]\,. \tag{68}$$

Provided that $\frac{1}{2}(t^2+5t)+1 \leq n$ we have that $|m_T| \leq 2^{-n+1}$ for all $T \in \Sigma_{t,t} \setminus \{e\}$.

We are now in the position to prove our main result of this section, the bound on the bias $|\mathrm{tr}(M(\rho_S - \rho_I))|$.

Theorem 4.7 (Bound on bias). *Let $\{M_0, M_1\}$ be a PPT measurement, let $M = M_0 - M_1$ and $\frac{1}{2}(t^2+5t)+2 \leq n/2$, and let $\rho_S = \mathbf{E}_{|S\rangle}\left[|S\rangle\langle S|^{\otimes t}\right]$ and $\rho_I = I^{\otimes t}/2^{nt}$, then*

$$|\mathrm{tr}(M(\rho_S - \rho_I))| \leq 2^{-n/2}\,. \tag{69}$$

As we explained in Section 4.1, this bound on the bias immediately implies the following corollary:

Corollary 4.8 (Single-copy lower bound for random stabilizer vs. maximally mixed). *Any single-copy algorithm for distinguishing the maximally mixed state $I/2^n$ from random n-qubit stabilizer states, with probability at least 2/3, requires at least $t = \Omega(\sqrt{n})$ many copies of the unknown state.*

Furthermore, by virtue of Lemma 4.3 and the triangle inequality in Eq. (62) between the three ensembles (H, S, I), Theorem 4.7 also implies our main result, Theorem 4.1.

4.3 Partial Transposes of $R(O)$

In this section, we study partial transposes of $R(O)$ for $O \in O_t$. Let $S \subset [t]$, then S and its complement $\bar{S}$ in $[t]$ form a partition of $[t]$. The operators $R(O)$ act on t subsystems. We consider taking partial transposes with respect to a subset $S \subset [t]$ of subsystems and will denote this operation by $^{\Gamma_S}$. Concretely, for $S \subset [t]$, we denote by $R(O)^{\Gamma_S}$ its partial transpose with the respect to the subsystems indexed by S.

The next lemma is a generalization of [25, Lemma 4] from the permutation operators to the operators $R(O)$ (which include the permutation operators since $\mathcal{S}_t \subseteq O_t$).

Lemma 4.9 (Singular values of partial transposes of $R(O)$). *Choose a set $S \subset [t]$. For any $O \in O_t$, let $k = \dim\left(\ker O_{S,S}\right)$. Then $R(O)^{\Gamma_S}$ has $2^{n(t-2k)}$ non-zero singular values each equal to 2^{kn}. Consequently we have*

$$\left\|R(O)^{\Gamma_S}\right\|_1 = 2^{n(t-k)}\,. \tag{70}$$

Lemma 4.9 provides a characterization of the singular values of $R(O)^{\Gamma_S}$ and the trace-norm $\|R(O)^{\Gamma_S}\|_1$ in terms of $\dim\left(\ker O_{S,S}\right)$. For our purposes, we are interested in proving upper bounds on $\|R(O)^{\Gamma_S}\|_1$ for all $O \in O_t$ except the identity $O = e$. Hence, we will now show that for all $O \neq e$ one can always choose the set S on which the partial transpose is applied, such that $\dim\left(\ker O_{S,S}\right)$ is non-zero, i.e., $\ker O_{S,S}$ is non-trivial. To prove this, it is sufficient to show that O has at least a single zero principal minor, i.e., there exists $S \subset [t]$ such that $\det\left(O_{S,S}\right) = 0$. The following proposition implies this sufficient condition as a corollary:

Proposition 4.10 (All principal minors equal 1 implies conjugate to upper triangular). *For any binary matrix $A \in \mathbb{F}_2^{t\times t}$ such that $\det\left(A_{S,S}\right) = 1$ for all $S \subset [t]$, there exists $\pi \in \mathcal{S}_t$ such that $\pi A \pi^{-1}$ is upper triangular.*

Note that for orthogonal matrices O such that $O^T O = I$, O being upper triangular implies that O is the identity, i.e., $O = e$. Hence, we obtain the following corollary:

Corollary 4.11. *For all $O \in O_t \setminus \{e\}$, there exists $S \subset [t]$ such that $\|R(O)^{\Gamma_S}\|_1 \leq 2^{n(t-1)}$.*

Acknowledgments

We thank Daniel Liang, Matthis Lehmkühler, Jonas Haferkamp, Felipe Montealegre, and Robert Huang for helpful discussions. MH acknowledges funding by the BMBF (MUNIQC-Atoms). JH acknowledges funding from the Dutch Research Council (NWO) through Veni No.VI.Veni.222.331 and the Quantum Software Consortium (NWO Zwaartekracht Grant No.024.003.037).

References

[1] Scott Aaronson and Daniel Gottesman. 2008. Identifying Stabilizer States. (Aug. 2008). doi:08080052

[2] Srinivasan Arunachalam, Sergey Bravyi, Arkopal Dutt, and Theodore J. Yoder. 2023. Optimal Algorithms for Learning Quantum Phase States. In *18th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC*

2023). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 3:1–3:24. doi:10.4230/LIPIcs.TQC.2023.3

[3] Srinivasan Arunachalam and Arkopal Dutt. 2024. Polynomial-Time Tolerant Testing Stabilizer States. arXiv:2408.06289

[4] Zongbo Bao, Philippe van Dordrecht, and Jonas Helsen. 2024. Tolerant Testing of Stabilizer States with a Polynomial Gap via a Generalized Uncertainty Relation. arXiv:2410.21811

[5] Ewout Van Den Berg. 2021. A Simple Method for Sampling Random Clifford Operators. In *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*. IEEE Computer Society, 54–59. doi:10.1109/QCE52317.2021.00021

[6] Sergey Bravyi, Dan Browne, Padraic Calpin, Earl Campbell, David Gosset, and Mark Howard. 2019. Simulation of Quantum Circuits by Low-Rank Stabilizer Decompositions. *Quantum* 3 (Sept. 2019), 181. doi:10.22331/q-2019-09-02-181

[7] Sergey Bravyi and Dmitri Maslov. 2021. Hadamard-Free Circuits Expose the Structure of the Clifford Group. *IEEE Transactions on Information Theory* 67, 7 (July 2021), 4546–4563. doi:10.1109/TIT.2021.3081415

[8] Kaifeng Bu, Weichen Gu, and Arthur Jaffe. 2023. Stabilizer Testing and Magic Entropy. arXiv:2306.09292

[9] Sebastien Bubeck, Sitan Chen, and Jerry Li. 2020. Entanglement Is Necessary for Optimal Quantum Property Testing. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*. 692–703. doi:10.1109/FOCS46700.2020.00070

[10] Matthias C. Caro. 2024. Learning Quantum Processes and Hamiltonians via the Pauli Transfer Matrix. *ACM Transactions on Quantum Computing* 5, 2 (June 2024), 14:1–14:53. doi:10.1145/3670418

[11] Sitan Chen, Jordan Cotler, Hsin-Yuan Huang, and Jerry Li. 2021. A Hierarchy for Replica Quantum Advantage. arXiv:2111.05874

[12] Sitan Chen, Jordan Cotler, Hsin-Yuan Huang, and Jerry Li. 2022. Exponential Separations Between Learning With and Without Quantum Memory. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*. 574–585. doi:10.1109/FOCS52979.2021.00063

[13] Sitan Chen and Weiyuan Gong. 2023. Efficient Pauli Channel Estimation with Logarithmic Quantum Memory. arXiv:2309.14326

[14] Sitan Chen, Weiyuan Gong, and Qi Ye. 2024. Optimal Tradeoffs for Estimating Pauli Observables. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*. 1086–1105. doi:10.1109/FOCS61266.2024.00072

[15] Sitan Chen, Weiyuan Gong, Qi Ye, and Zhihan Zhang. 2024. Stabilizer Bootstrapping: A Recipe for Efficient Agnostic Tomography and Magic Estimation. arXiv:2408.06967

[16] Nai-Hui Chia, Ching-Yi Lai, and Han-Hsuan Lin. 2024. Efficient Learning of t-Doped Stabilizer States with Single-Copy Measurements. *Quantum* 8 (Feb. 2024), 1250. doi:10.22331/q-2024-02-12-1250

[17] Eric Chitambar, Debbie Leung, Laura Mančinska, Maris Ozols, and Andreas Winter. 2014. Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask). *Communications in Mathematical Physics* 328, 1 (May 2014), 303–326. doi:10.1007/s00220-014-1953-9

[18] Jeroen Dehaene and Bart De Moor. 2003. Clifford Group, Stabilizer States, and Linear and Quadratic Operations over GF(2). *Physical Review A* 68, 4 (Oct. 2003), 042318. doi:10.1103/PhysRevA.68.042318

[19] Sabee Grewal, Vishnu Iyer, William Kretschmer, and Daniel Liang. 2024. Agnostic Tomography of Stabilizer Product States. arXiv:2404.03813

[20] Sabee Grewal, Vishnu Iyer, William Kretschmer, and Daniel Liang. 2024. Efficient Learning of Quantum States Prepared With Few Non-Clifford Gates. arXiv:2305.13409

[21] Sabee Grewal, Vishnu Iyer, William Kretschmer, and Daniel Liang. 2024. Improved Stabilizer Estimation via Bell Difference Sampling. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing (STOC 2024)*. Association for Computing Machinery, New York, NY, USA, 1352–1363. doi:10.1145/3618260.3649738

[22] D. Gross. 2006. Hudson's Theorem for Finite-Dimensional Quantum Systems. *J. Math. Phys.* 47, 12 (Dec. 2006), 122107. doi:10.1063/1.2393152

[23] David Gross, Sepehr Nezami, and Michael Walter. 2021. Schur-Weyl Duality for the Clifford Group with Applications: Property Testing, a Robust Hudson Theorem, and de Finetti Representations. *Communications in Mathematical Physics* 385, 3 (Aug. 2021), 1325–1393. doi:10.1007/s00220-021-04118-7 arXiv:1712.08628

[24] J. Haferkamp, F. Montealegre-Mora, M. Heinrich, J. Eisert, D. Gross, and I. Roth. 2023. Efficient Unitary Designs with a System-Size Independent Number of Non-Clifford Gates. *Communications in Mathematical Physics* 397, 3 (Feb. 2023), 995–1041. doi:10.1007/s00220-022-04507-6

[25] Aram W. Harrow. 2023. Approximate Orthogonality of Permutation Operators, with Application to Quantum Information. *Letters in Mathematical Physics* 114, 1 (Dec. 2023), 1. doi:10.1007/s11005-023-01744-1 arXiv:2309.00715

[26] Tobias Haug and M.S. Kim. 2023. Scalable Measures of Magic Resource for Quantum Computers. *PRX Quantum* 4, 1 (Jan. 2023), 010301. doi:10.1103/PRXQuantum.4.010301

[27] Tobias Haug, Soovin Lee, and M. S. Kim. 2024. Efficient Quantum Algorithms for Stabilizer Entropies. *Physical Review Letters* 132, 24 (June 2024), 240602. doi:10.1103/PhysRevLett.132.240602

[28] Marcel Hinsche and Jonas Helsen. 2024. Single-Copy Stabilizer Testing. arXiv:2410.07986

[29] Hsin-Yuan Huang, Michael Broughton, Jordan Cotler, Sitan Chen, Jerry Li, Masoud Mohseni, Hartmut Neven, Ryan Babbush, Richard Kueng, John Preskill, and Jarrod R. McClean. 2022. Quantum Advantage in Learning from Experiments. *Science* 376, 6598 (June 2022), 1182–1186. doi:10.1126/science.abn7293

[30] Hsin-Yuan Huang, Richard Kueng, and John Preskill. 2020. Predicting Many Properties of a Quantum System from Very Few Measurements. *Nature Physics* 16, 10 (Oct. 2020), 1050–1057. doi:10.1038/s41567-020-0932-7

[31] Vishnu Iyer and Daniel Liang. 2024. Tolerant Testing of Stabilizer States with Mixed State Inputs. doi:10.48550/arXiv.2411.08765 arXiv:2411.08765

[32] Robbie King, David Gosset, Robin Kothari, and Ryan Babbush. 2025. Triply Efficient Shadow Tomography. *PRX Quantum* 6, 1 (Feb. 2025), 010336. doi:10.1103/PRXQuantum.6.010336

[33] Robert Koenig and John A. Smolin. 2014. How to Efficiently Select an Arbitrary Clifford Group Element. *J. Math. Phys.* 55, 12 (Dec. 2014), 122202. doi:10.1063/1.4903507

[34] Richard Kueng and David Gross. 2015. Qubit Stabilizer States Are Complex Projective 3-Designs. arXiv:1510.02767

[35] Ching-Yi Lai and Hao-Chung Cheng. 2022. Learning Quantum Circuits of Some T Gates. *IEEE Transactions on Information Theory* 68, 6 (June 2022), 3951–3964. doi:10.1109/TIT.2022.3151760

[36] Lorenzo Leone, Salvatore F. E. Oliviero, and Alioscia Hamma. 2022. Stabilizer Rényi Entropy. *Physical Review Letters* 128, 5 (Feb. 2022), 050402. doi:10.1103/PhysRevLett.128.050402

[37] Lorenzo Leone, Salvatore F. E. Oliviero, and Alioscia Hamma. 2024. Learning T-Doped Stabilizer States. *Quantum* 8 (May 2024), 1361. doi:10.22331/q-2024-05-27-1361

[38] Yuhan Liu and Jayadev Acharya. 2024. The Role of Randomness in Quantum State Certification with Unentangled Measurements. In *Proceedings of Thirty Seventh Conference on Learning Theory*. PMLR, 3523–3555.

[39] Zi-Wen Liu and Andreas Winter. 2022. Many-Body Quantum Magic. *PRX Quantum* 3, 2 (May 2022), 020333. doi:10.1103/PRXQuantum.3.020333

[40] Angus Lowe and Ashwin Nayak. 2022. Lower Bounds for Learning Quantum States with Single-Copy Measurements. arXiv:2207.14438

[41] Saeed Mehraban and Mehrdad Tahmasbi. 2024. Improved Bounds for Testing Low Stabilizer Complexity States. arXiv:2410.24202

[42] Ashley Montanaro. 2017. Learning Stabilizer States by Bell Sampling. arXiv:1707.04012

[43] Ashley Montanaro and Ronald de Wolf. 2016. *A Survey of Quantum Property Testing*. Number 7 in Graduate Surveys. Theory of Computing Library. doi:10.4086/toc.gs.2016.007

[44] Salvatore F. E. Oliviero, Lorenzo Leone, Alioscia Hamma, and Seth Lloyd. 2022. Measuring Magic on a Quantum Processor. *npj Quantum Information* 8, 1 (Dec. 2022), 148. doi:10.1038/s41534-022-00666-5 arXiv:2204.00015

[45] Aadil Oufkir. 2023. Sample-Optimal Quantum Process Tomography with Non-Adaptive Incoherent Measurements. In *2023 IEEE International Symposium on Information Theory (ISIT)*. 1919–1924. doi:10.1109/ISIT54713.2023.10206538

[46] Maarten Van den Nest. 2009. Classical Simulation of Quantum Computation, the Gottesman-Knill Theorem, and Slightly Beyond. arXiv:0811.0898

Received 2024-10-30; accepted 2025-02-01