Mixed-state escape bounds and proper agnostic learning of stabilizer states in odd-prime dimensions

Scott Schubr

Corrected publication version — October 18, 2026

Abstract

Let p be an odd prime and let ρ be an arbitrary mixed state on n systems of dimension p. Write F for its maximum overlap with a pure stabilizer state. We prove a mixed-input escape estimate for known weighted Weyl measurements: for every proper linear subspace T of the label Lagrangian S of any maximizing stabilizer, a sampleable distribution assigns mass at least $3(p-1)F^6/(4p^3)$ to $S\setminus T$ when $p \equiv 3 \pmod 4$, and at least $(p-1)F^4/(2p^2)$ when $p \equiv 1 \pmod 4$. The draws use six and four input copies, respectively. A purity-sensitive single-system inequality supplies the mixed-input step. With stabilizer bootstrapping, coherent conditioning and bounded rejection sampling, the estimate yields a proper additive agnostic learner. For each fixed p and overlap promise $F \geq \tau$, the resources are $\exp(O_p(\log^2(2/\tau)))\,\mathrm{poly}_p(n, \varepsilon^{-1}, \log(2/\delta))$. We give explicit budgets and growing-p dependence in an exact Clifford and coherent-measurement model. The measurement primitives, arithmetic split and learning framework are attributed to prior work; the contribution concerns arbitrary mixed inputs.

1 Introduction and results

An agnostic stabilizer learner seeks a classical description of a pure stabilizer whose overlap with an unknown input is nearly the largest possible. The input need not itself be a stabilizer or follow a specified noise model. Chen, Gong, Ye and Zhang [1, Corollary 6.3] established a proper mixed-qubit learner using stabilizer bootstrapping. Qi, Xu, Feng and Li [2] developed agnostic learning for pure inputs in odd-prime local dimension; their Section VI identifies extension to mixed inputs as a further direction.

We supply a mixed-input escape estimate and complete the corresponding conditioning and resource analysis. The main analytic ingredient is a one-system inequality retaining the purity of a mixed conditioned block. The general learner follows the bootstrapping architecture of Chen et al. and its qudit adaptation by Qi et al. The weighted/skewed measurements and their residue-class arithmetic are already present in Allcock, Doriguello, Ivanyos and Santha [3]; related tensor-Weyl measurements occur in Nöller, Tran, Gachechiladze and Kueng [5]. Hinsche, Eisert and Carrasco [4] explicitly anticipate a qudit agnostic extension using their measurement tools.

Related tasks require care in comparison. Holt and Subramanian [7, Theorem 9.27] learn the exact stabilizer group of a mixed qudit state, rather than its best pure stabilizer comparator. Bittel et al. [8, Theorems III.1 and A.1] give single-copy algorithms for mixed-qubit inputs with a known stabilizer-overlap floor greater than $\sqrt{2/3}$ and for exact pure stabilizer learning in prime dimension, respectively.

For a density operator on $(\mathbb{C}^p)^{\otimes n}$ define

$$F = F_{\mathrm{stab}}(\rho) = \max_{s \in \mathcal{S}_{p,n}} \langle s|\rho|s\rangle,$$

where $\mathcal{S}_{p,n}$ denotes pure stabilizer states up to global phase. This is squared fidelity with a pure comparator. All logarithms are natural.

Theorem 1.1 (Mixed-input escape)**.** *Let $p \geq 3$ be prime, $n \geq 1$, and ρ any density operator. There is a physically sampleable law B_ρ on $\mathbb{F}_p^{2n}$ such that, for the label Lagrangian S of any maximizing stabilizer and every proper linear subspace $T < S$,*

$$B_\rho(S \setminus T) \geq \kappa_p F^{r_p}, \qquad (r_p, \kappa_p) = \begin{cases} (6,\, 3(p-1)/(4p^3)), & p \equiv 3 \pmod 4, \\ (4,\, (p-1)/(2p^2)), & p \equiv 1 \pmod 4. \end{cases} \tag{1}$$

A draw uses r_p independent input copies and a stabilizer measurement with polynomial overhead in n, p. These are sufficient copy counts.

Theorem 1.2 (Proper mixed-state learner)**.** *Let $p \geq 3$ be prime, $n \geq 1$, $0 < \tau \leq 1$, and $0 < \varepsilon, \delta < 1$. Given independent copies of any density operator ρ with $F_{\mathrm{stab}}(\rho) \geq \tau$, an algorithm returns a pure stabilizer tableau s satisfying*

$$\langle s|\rho|s\rangle \geq F_{\mathrm{stab}}(\rho) - \varepsilon$$

with probability at least $1 - \delta$. In the operation model below, setting $L = \log(2/\tau)$ gives the conservative resource bound

$$\exp\bigl(O(p^2 L(L + \log p))\bigr)\, \mathrm{poly}(n, p, \varepsilon^{-1}, \log(2/\delta)). \tag{2}$$

The constant in $O(\cdot)$ is absolute. For every fixed p, this is $\exp(O_p(L^2))\, \mathrm{poly}_p(n, \varepsilon^{-1}, \log(2/\delta))$. The explicit copy budget is given in Section 7.

Operation model. We allow independent input copies, exact prime-dimensional Clifford operations (including Fourier and SUM), coherent Weyl eigenspace measurements, and finite-field arithmetic, with implementation overhead polynomial in n, p. Classical estimates use finite precision. The access resources are the independent input copies and exact operations specified here. The resource bound retains its displayed dependence on p, including the p-dependent quasipolynomial exponent.

Relationship to the qutrit case. At $p = 3$, Theorem 1.1 gives $F^6/18$. This manuscript consolidates the earlier private qutrit draft as the $p = 3$ case of the same result family. The all-prime statement is a generalization, rather than a second independent contribution containing that earlier theorem.

2 Weyl conventions and a positive overlap law

Set $D = p^n$, $V = \mathbb{F}_p^{2n}$ and $\omega = e^{2\pi i/p}$. For $u = (a, b)$ use

$$W_u = \omega^{-a\cdot b/2} Z^a X^b, \qquad [u, v] = a \cdot b' - b \cdot a', \qquad W_u W_v = \omega^{[u,v]/2} W_{u+v}.$$

Here $X|j\rangle = |j+1\rangle$ and $Z|j\rangle = \omega^j |j\rangle$; all exponents are in $\mathbb{F}_p$. Label subspaces are linear. A pure stabilizer is specified by a maximal isotropic subspace, or Lagrangian, together with compatible eigenvalues. We use the usual odd-prime stabilizer closure and Clifford synthesis facts [10, 2].

Write $\chi_\rho(u) = \mathrm{Tr}(\rho W_u)$ and define

$$A_\rho(x) = D^{-1} \mathrm{Tr}(\rho W_x \rho W_x^\dagger), \qquad \widehat{A}_\rho(u) = \sum_x A_\rho(x) \omega^{[u,x]}. \tag{3}$$

Positivity of the trace product gives $A_\rho \geq 0$. Weyl twirling, $\sum_x W_x \rho W_x^\dagger = DI$, gives $\sum_x A_\rho(x) = 1$. Weyl orthogonality and conjugation give

$$\widehat{A}_\rho(u) = |\chi_\rho(u)|^2. \tag{4}$$

For example, expanding $\rho = D^{-1}\sum_v \chi_\rho(-v)W_v$ in (3) yields $A_\rho(x) = D^{-2}\sum_v |\chi_\rho(v)|^2\omega^{[x,v]}$, which is the inversion of (4). The auxiliary law is normalized for mixed inputs. In contrast, the characteristic weight $p_\rho(u) = D^{-1}|\chi_\rho(u)|^2$ has total mass $\mathrm{Tr}(\rho^2)$. They agree for pure inputs, but must not be interchanged for mixed inputs. We do not assume a physical two-copy sampler for A_ρ. Positive Fourier-dual overlap distributions for mixed qubits already appear in Iyer and Liang [6, Lemma 2.12, Remark 2.13 and Facts 2.14–2.15]. Equation (4) is the standard Weyl-overlap identity in our odd-prime conventions. General Fourier-difference sampling is also treated by Holt and Subramanian [7, Lemma 8.4].

3 A mixed local inequality and the all-system lift

Lemma 3.1 (Local mixed estimate)**.** *Let τ be a density matrix on $\mathbb{C}^p$ such that $|0\rangle$ maximizes overlap with single-system stabilizers, and put $f = \tau_{00}$. Then*

$$\sum_j \tau_{jj}^2 - \frac{1}{p}\,\mathrm{Tr}(\tau^2) \geq c_p f^2, \qquad c_p = \frac{p-1}{4p}. \tag{5}$$

Proof. The $p+1$ stabilizer mutually unbiased bases correspond to the lines in $\mathbb{F}_p^2$ [9, 2]. If their outcome probabilities are z_v, Weyl orthogonality gives

$$\sum_v z_v = p+1, \qquad \sum_v z_v^2 = 1 + P, \qquad P = \mathrm{Tr}(\tau^2).$$

Since every $z_v \leq f$, we have $f \geq 1/p$ and $P \leq \min(1, (p+1)f - 1)$. Cauchy–Schwarz on the other $p-1$ computational outcomes gives $\sum_j \tau_{jj}^2 \geq f^2 + (1-f)^2/(p-1)$. Let g denote the left side of (5) and put

$$h_0 = \frac{2}{p+1}, \qquad K_0 = \frac{(3p-1)(p+1)}{4p(p-1)}.$$

Subtraction and factorization yield

$$g - c_p f^2 \geq K_0(f-h_0)\left(f - \frac{2(2p-1)}{3p-1}\right) \geq 0, \qquad 1/p \leq f \leq h_0,$$
$$g - c_p f^2 \geq K_0(f-h_0)\left(f - \frac{2}{3p-1}\right) \geq 0, \qquad h_0 \leq f \leq 1.$$

In the first line use $P \leq (p+1)f - 1$; its other root exceeds one. In the second use $P \leq 1$; its other root is at most h_0. This includes both endpoints. □

Lemma 3.2 (Hyperplane masses)**.** *Let S belong to a global optimizing stabilizer of overlap F. For every hyperplane $H < S$,*

$$A_\rho(H) \geq F^2/p, \qquad A_\rho(S \setminus H) \geq c_p F^2. \tag{6}$$

Proof. Clifford covariance and a computational linear change put the optimizer at $|0^n\rangle$, with $S = \mathrm{span}(Z_1, \ldots, Z_n)$ and $H = \mathrm{span}(Z_2, \ldots, Z_n)$. For $y \in \mathbb{F}_p^{n-1}$ let σ_y be the positive principal block with entries $(\sigma_y)_{jk} = \rho_{(j,y),(k,y)}$. Character sums give

$$A_\rho(H) = p^{-1} \sum_y \mathrm{Tr}(\sigma_y^2), \tag{7}$$

$$A_\rho(S \setminus H) = \sum_y \left(\sum_j (\sigma_y)_{jj}^2 - p^{-1} \mathrm{Tr}(\sigma_y^2) \right). \tag{8}$$

Every summand in (8) is nonnegative: positivity implies $\mathrm{Tr}(\sigma^2) \le (\mathrm{Tr}\,\sigma)^2 \le p \sum_j \sigma_{jj}^2$. The target block has $(\sigma_0)_{00} = F$, giving the first bound. Set $t = \mathrm{Tr}\,\sigma_0 > 0$. The normalized block $\tau = \sigma_0/t$ has $f = F/t$ and $|0\rangle$ is its stabilizer optimizer: otherwise $|v\rangle \otimes |0^{n-1}\rangle$ would improve the original global optimum. Lemma 3.1 makes the target summand at least $t^2 c_p (F/t)^2 = c_p F^2$. For $n = 1$ there is a single block and $H = \{0\}$, so the same proof applies. □

4 Known weighted measurements and mixed-input escape

We describe the measurement explicitly to fix signs and mixed-input normalization. It is an instance of the known weighted/skewed stabilizer measurements [3, 5]. Choose nonzero $t_1, \ldots, t_k \in \mathbb{F}_p$ with $\sum_i t_i^2 = 0$ and put

$$R_u = \bigotimes_{i=1}^{k} W_{t_i u}, \qquad P_y = D^{-2} \sum_{u \in V} \omega^{[y,u]} R_u.$$

The cancellation makes $R_u R_v = R_{u+v}$ and $R_u^\dagger = R_{-u}$. Character orthogonality shows that the P_y are orthogonal Hermitian projectors summing to I, with rank D^{k-2}. Indeed, all nonzero R_u have trace zero. If Y is the outcome on $\rho^{\otimes k}$, then

$$\mathbb{E}\omega^{[u,Y]} = \prod_i \chi_\rho(t_i u).$$

For independent outcomes Y, Y', define B_ρ to be the law of $Y - Y'$. Consequently

$$\widehat{B}_\rho(u) = \prod_i |\chi_\rho(t_i u)|^2, \qquad B_\rho = A_{t_1} * \cdots * A_{t_k}, \quad A_t(x) = A_\rho(t^{-1}x). \tag{9}$$

The scalars are nonzero, so every A_t is normalized. For each physical position the two generating Paulis on the k registers commute; standard symplectic Clifford synthesis implements their joint measurement. The construction factors across the n positions. It retains the joint eigenspaces and does not separately measure the tensor factors.

Proof of Theorem 1.1. If $p \equiv 3 \pmod 4$, the sets of squares and of minus one minus squares both have size $(p+1)/2$, so they intersect. Choose a, b with $a^2 + b^2 = -1$. Both are nonzero since -1 is a nonsquare. Use $k = 3$ and $(t_1, t_2, t_3) = (1, a, b)$. Scalar multiplication preserves H and $S \setminus H$, so each factor in (9) has the same masses there as A_ρ. The three disjoint events with exactly one factor in $S \setminus H$ and two in H all sum into $S \setminus H$. Thus Lemma 3.2 gives

$$B_\rho(S \setminus H) \ge 3 A_\rho(H)^2 A_\rho(S \setminus H) \ge \frac{3(p-1)}{4p^3} F^6.$$

If $p \equiv 1 \pmod 4$, choose $a^2 = -1$ and use $k = 2$ and $(1, a)$. The two disjoint events with one factor in each set give

$$B_\rho(S \setminus H) \geq 2A_\rho(H)A_\rho(S \setminus H) \geq \frac{p-1}{2p^2}F^4.$$

Enlarging any proper $T < S$ to a hyperplane H proves the claim. A draw uses two k-copy measurements. The residue-class split is known measurement arithmetic [3]. □

Remark 4.1. For pure input $A_\rho = p_\rho$, so the convolution agrees with the pure-input marginal used in Qi et al. [2, Eq. (27)], after its skew change of variables. Zero-weight registers are omitted, never inverted. In particular the two-register branch handles $p = 5$, where $a^2 + b^2 = -1$ has no solution with both a, b nonzero. The pure-input algorithm of Qi et al. uses its own four-copy primitives.

5 Coherent conditioning and correlation thresholds

Lemma 5.1 (Preservation of a global optimizer)**.** *Let s maximize stabilizer overlap F and let P be a consistent stabilizer-code projector fixing s. Then $w = \operatorname{Tr}(P\rho) \geq F$, and $\rho_P = P\rho P/w$ has stabilizer optimum exactly F/w, still attained by s.*

Proof. Stabilizer measurement closure [10] gives, for any pure stabilizer t, either $P|t\rangle = 0$ or $P|t\rangle = \sqrt{c}\, e^{i\theta}|t'\rangle$ for another stabilizer t' and $0 < c \leq 1$. Thus $\langle t|P\rho P|t\rangle = c\langle t'|\rho|t'\rangle \leq F$, whereas s attains F. Since P contains $|s\rangle\langle s|$ as a subprojection, $w \geq F$. Division by w proves the claim, including tied optimizers. □

For any state σ set $c_\sigma(x) = |\operatorname{Tr}(\sigma W_x)|^2$ and define rational constants

$$\eta = \frac{1}{4p^2}, \qquad q = 1 - \frac{\eta}{4} = 1 - \frac{1}{16p^2}. \tag{10}$$

Lemma 5.2 (High correlations commute; low outcomes amplify)**.** *Labels with $c_\sigma(x) > 1 - 3\eta$ commute pairwise. If σ is supported on a stabilizer code with label span C, every label with nonzero expectation commutes with C. If $c_\sigma(x) \leq 1 - \eta$, every eigenspace outcome of W_x has probability at most q.*

Proof. For a nonzero label let z_j be its spectral probabilities and put $\Delta = 1 - \cos(2\pi/p) = 2\sin^2(\pi/p) \geq 8/p^2$. Then

$$1 - c_\sigma(x) = \sum_{j,l} z_j z_l (1 - \cos(2\pi(j-l)/p)) \geq \Delta(1 - \sum_j z_j^2).$$

Hence $c_\sigma(x) > 1 - 3\eta$ implies $\max_j z_j \geq \sum_j z_j^2 > 1 - 3\eta/\Delta \geq 29/32 > 7/8$. For spectral projectors E, G of noncommuting Weyls, character orthogonality gives $EGE = E/p$ and $GEG = G/p$. The two-projection norm bound is $\|E + G\| \leq 1 + p^{-1/2} < 7/4$. Two such high noncommuting labels would therefore contradict $\operatorname{Tr}(\sigma(E + G)) > 7/4$. The zero label causes no difficulty. Definite eigenvalues for code Weyls force all expectations of noncommuting labels to vanish by conjugation. Finally the reverse triangle inequality gives $z_b \leq (1 + |\operatorname{Tr}(\sigma W_x)|)/2$ for each outcome. If $c_\sigma(x) \leq 1 - \eta$, then $z_b \leq (1 + \sqrt{1-\eta})/2 \leq 1 - \eta/4 = q$. □

On a prefix of code projections fixing s, Lemma 5.1 keeps it globally optimal. A correct additional low-correlation projection multiplies its overlap by at least $1/q$. The acceptance of the entire current code on an original copy is at least the original F; successive preparation costs need not be multiplied.

6 Robust capture of high correlations

Write $r = r_p$ and $\kappa = \kappa_p$ from (1), and put

$$a = \kappa\tau^r/2, \quad t_0 = \eta/4, \quad m = \left\lceil \frac{4n^2 \log p + \log 8}{a} \right\rceil, \quad h = \left\lceil 128 t_0^{-2} \log(32m) \right\rceil. \tag{11}$$

For a code-supported σ, take m independent B_σ samples. For each sampled label estimate c_σ using h fresh copies of σ; keep labels whose estimates exceed $1 - 2\eta$. Include the current code span C and extend the retained span to a full Lagrangian H. Abort on observed nonisotropy. The procedure does not test the mass bound below.

Lemma 6.1 (Capture)**.** *In the uncapped conditional-copy model this procedure gives, with probability at least* $3/4$*, a full Lagrangian* $H \supset C$ *with*

$$B_\sigma(G \setminus H) \le a, \qquad G = \{x : c_\sigma(x) > 1 - \eta\}. \tag{12}$$

The estimate and retention steps require finite classical precision only.

Proof. Measure W_x spectrally and average its unit-modulus eigenvalues to estimate $\chi_\sigma(x)$. Hoeffding for real and imaginary parts at tolerance $t_0/8$ gives failure at most $4\exp(-ht_0^2/128) \le 1/(8m)$. Both true and empirical complex means have modulus at most one, so the corresponding squared-modulus error is less than $t_0/2$. Evaluate the empirical squared modulus with deterministic error at most $t_0/2$, using rational enclosures of roots of unity. Precision $O(\log p)$ beyond counter arithmetic suffices because $t_0^{-1} = 16p^2$. A union bound gives total error at most t_0 for every sampled label except with probability $1/8$.

On this event all sampled G labels are retained; every retained label has $c_\sigma(x) > 1 - 9\eta/4 > 1 - 3\eta$. Lemma 5.2 makes their span and C isotropic. There are at most p^{4n^2} subspaces of V, by the overcount using the column spans of $2n$ by $2n$ matrices. For a fixed H violating (12), the probability that all m samples miss $G \setminus H$ is at most e^{-am}. A union bound over subspaces is at most $1/8$. This handles the data dependence of the output H. The two failure probabilities sum to at most $1/4$. □

7 A capped learner with explicit resources

We now complete the bootstrap analysis using the framework of [1, 2]. Set

$$J = \left\lceil \frac{\log(1/\tau)}{-\log q} \right\rceil + 1, \qquad b_0 = \frac{3a}{4p}, \qquad s_0 = b_0^J, \tag{13}$$

$$Q = J(rm + mh + r + 1), \qquad K = \left\lceil \tau^{-1} \log(4Q/s_0) \right\rceil, \tag{14}$$

$$R = \left\lceil \frac{2}{s_0} \log(2/\delta) \right\rceil, \qquad N = R(J+1), \qquad v = \left\lceil \frac{2}{\varepsilon^2} \log(4N/\delta) \right\rceil. \tag{15}$$

All parameters can be chosen as conservative computable integer upper bounds; the displayed ceilings describe convenient budgets, rather than a requirement for exact logarithm comparisons. Rational accuracy and confidence inputs suffice.

One run. Maintain a consistent code projector P, initially I, and its label span C, initially zero. For at most J stages:

1. Prepare independent copies of $\sigma = \rho_P$ as needed by capped rejection sampling below. Construct H by Section 6; abort this run on preparation failure or observed nonisotropy.
2. Measure one fresh *original* ρ copy in the stabilizer basis of H and record the resulting pure stabilizer candidate.
3. Draw a fresh independent $x \sim B_\sigma$. Abort if $x \in H$ or x does not commute with C. Otherwise choose an eigenvalue uniformly from the p possibilities, and extend the code by that eigenvalue of W_x.
4. If the code reaches rank n, record its uniquely specified pure stabilizer and stop. Preserve all candidates when any run aborts.

The algorithm never tests whether a sample lies in the unknown target's label space, has low correlation, or satisfies the hidden capture guarantee. Since $C \subset H$, every accepted outside-H extension increases rank and is a nonzero consistent code.

Capped conditional copies. For each request for ρ_P, try at most K fresh original copies. Coherently measure the commuting code eigenvalues and retain only the desired joint outcome. Clifford conjugation to coordinate Z measurements followed by undoing implements the code projector while retaining coherence within it. Separate measurements of the physical tensor factors would not implement this projector. If all trials fail, abort the run. A zero-weight or tiny-weight code still has bounded cost. There are at most Q conditional-copy requests in a run: rm for capture draws, mh for estimates, and r for the progress draw per stage; the additional one in Q is slack.

Proof of Theorem 1.2. Fix one original global optimizer s with Lagrangian S. Analyze an auxiliary favorable witness using uncapped conditional copies only while it is active. It is absorbed as a success at the first recording of s. Otherwise it continues only on a target-fixing low-correlation jump, and is killed at the first other continuation or abort. Thus a favorable prefix consists only of target-fixing low-correlation jumps; no ideal conditional state need be defined beyond a killed or absorbed witness. Lemma 5.1 gives $F_{\mathrm{stab}}(\sigma) \geq \tau$ and preserves s as a global optimizer on every active prefix.

With probability at least $3/4$ the capture step is good. When its $H = S$, the fresh original basis measurement records s with probability $F \geq \tau$. When $H \neq S$, escape with the proper subspace $S \cap H$ gives $B_\sigma(S \setminus H) \geq \kappa\tau^r = 2a$. Removing $G \setminus H$ loses at most a. Thus a target label of low correlation outside H occurs with probability at least a. It commutes with $C \subset S$, and guessing its correct eigenvalue costs $1/p$. Partitioning over good H, the conditional probability of either recording s or making a favorable jump is at least

$$\frac{3}{4}\min(\tau, a/p) = b_0.$$

The equality follows from $r \geq 4$, $\kappa \leq 1$ and $0 < \tau \leq 1$. Each favorable jump multiplies the conditional optimum by at least $1/q$. By the choice of J, J such nonabsorbing jumps would force $\tau q^{-J} > 1$. Early rank-n favorable termination already records s. Induction with recording as an absorbing success therefore gives ideal witness probability at least $s_0 = b_0^J$.

On this stopped favorable witness, the code weight on an original copy is always at least $F \geq \tau$. Each preparation cap loses probability at most $e^{-\tau K}$. Coupling only up to the recording time and

using at most Q requests loses at most

$$Qe^{-\tau K} \le s_0/4.$$

Thus an actual capped run records s with probability at least $s_0/2$. This argument imposes no acceptance lower bound on wrong prefixes, which are still capped. It also avoids assuming free postselection on successful ideal trajectories after the target has been recorded.

Run independently R times. Equation (15) makes the probability of never listing s at most $e^{-Rs_0/2} \le \delta/2$. At most N candidates occur. On v fresh original copies per candidate, measure its rank-one stabilizer projector and estimate its overlap. Conditional on the completed list, Hoeffding and a union bound give simultaneous error at most $\varepsilon/2$ except with probability $\delta/2$. Return a candidate with largest estimate; a listed optimizer then gives true overlap at least $F - \varepsilon$. Return any fixed stabilizer on an empty list.

The deterministic total original-copy budget is

$$RQK + RJ + Nv. \tag{16}$$

The terms cover capped conditional preparation, original basis samples and final validation. Tableau arithmetic, isotropic completion, coherent code measurement and the weighted measurements have polynomial overhead in n, p and these parameter counts. With $L = \log(2/\tau)$,

$$\begin{aligned} J &= O(p^2 L), & \kappa^{-1} &= O(p^2), & r &\le 6, \\ m &= O(\kappa^{-1}\tau^{-r}(n^2 \log p + 1)), & h &= O(p^4 \log(32m)), & & \\ \log(1/s_0) &= O(p^2 L(L + \log p)). & & & & \end{aligned}$$

The remaining factors in (16) give (2), including $\tau = 1$, where $J = 1$. The classical retention thresholds are rational, and the rounding budget in Lemma 6.1 avoids exact comparison with transcendental numbers. □

Corollary 7.1 (Unknown optimum)**.** *For every fixed odd prime p, arbitrary mixed input admits a proper additive-ε learner with resources* $\exp(O_p(\log^2(2/\varepsilon)))\,\mathrm{poly}_p(n, \log(2/\delta))$.

Proof. Set $\tau = \varepsilon$. If $F < \varepsilon$, every stabilizer already has overlap at least $0 > F - \varepsilon$; otherwise apply Theorem 1.2. Polynomial inverse-accuracy factors are absorbed in the exponential. □

8 Scope, attribution and verification

The local proof uses standard complete-MUB second moments. The weighted sampling mechanism, its arithmetic split, and the bootstrap architecture are prior ingredients. The proposed increment is the mixed-state local estimate, its operational all-prime escape consequence, and a complete mixed-input learner analysis with bounded preparation costs.

The output guarantee is an approximately optimal *pure stabilizer comparator*. Multi-copy lower bounds for density-matrix reconstruction or estimation of every Weyl amplitude [5] concern those respective guarantees. The learner analysis uses the stated exact operation model for its measurements.

Exact mixed-state stabilizer-group learning [7] is also a different objective: a full-rank depolarized stabilizer can have no nonidentity Weyl of unit-modulus expectation while retaining substantial overlap with a pure stabilizer. Conversely, the single-copy results of [8] provide important alternatives in their stated high-fidelity or exact-pure regimes.

The all-n theorem is proved analytically. The supporting package contains reproducible finite numerical checks for specified one- and two-system cases, including independent constructions of the measurement projectors. AI-assisted review was used.

Development disclosure. This paper was developed with AI assistance in derivation, programming, literature discovery and proof checking. The manuscript presents analytic proofs, and the supporting reproducibility programs provide floating-point diagnostics.

References

[1] S. Chen, W. Gong, Q. Ye and Z. Zhang, *Stabilizer bootstrapping: A recipe for efficient agnostic tomography and magic estimation*. arXiv:2408.06967v3, 5 December 2024.

[2] W. Qi, B. Xu, S. Feng and L. Li, *Agnostic learning of qudit stabilizer states*. arXiv:2607.15559v1, 17 July 2026.

[3] J. Allcock, J. F. Doriguello, G. Ivanyos and M. Santha, *Reconquering Bell sampling on qudits: stabilizer learning and testing, quantum pseudorandomness bounds, and more*. arXiv:2510.06848v3, 25 August 2026.

[4] M. Hinsche, J. Eisert and J. Carrasco, *The abelian state hidden subgroup problem: Learning stabilizer groups and beyond*. arXiv:2505.15770v3, 27 May 2026.

[5] J. Nöller, V. T. Tran, M. Gachechiladze and R. Kueng, *An infinite hierarchy of multi-copy quantum learning tasks*. arXiv:2510.08070v2, 1 September 2026.

[6] V. Iyer and D. Liang, *Tolerant Testing of Stabilizer States with Mixed State Inputs*. arXiv:2411.08765v2, 12 May 2025.

[7] I. Holt and S. Subramanian, *Learning quantum symmetries*. arXiv:2609.38085v2, 30 September 2026.

[8] L. Bittel, J. Eisert, W. Gong, A. A. Mele and L. Schatzki, *Adaptivity is all you need: Optimal stabilizer learning using just single-copy measurements*. arXiv:2610.02031v1, 1 October 2026.

[9] W. K. Wootters and B. D. Fields, *Optimal state-determination by mutually unbiased measurements*, Annals of Physics **191** (1989), 363–381. doi:10.1016/0003-4916(89)90322-9.

[10] E. Hostens, J. Dehaene and B. De Moor, *Stabilizer states and Clifford operations for systems of arbitrary dimensions, and modular arithmetic*, Physical Review A **71** (2005), 042315. doi:10.1103/PhysRevA.71.042315.