Single-copy learning of stabilizers with correlated diagonal noise

Scott Schubr

October 23, 2026

Abstract

We study exact identification of an unknown qubit stabilizer from a state diagonal in its stabilizer eigenbasis. A positive lower bound on every target-signed stabilizer expectation unifies global white noise, strict-majority diagonal mixtures, and additional nonmajority correlated mixtures. For a supplied gap $\mu_0 \geq 2^{-(\log_2 n)^\gamma}$, fixed $\gamma < 1$, recovery uses $O(n^{1+\varepsilon})$ original copies at fixed confidence, polynomial ideal quantum work, and $2^{O(n/\log\log n)}$ classical time. The new step is a uniformly spread successful component in a learned, verified Pauli direction, giving contraction of surviving error odds without independent syndrome errors. We also give polynomial-time recovery when the background is an unknown target-containing stabilizer code projector and a relative diagonal-visibility promise holds. Dense CSS targets with sufficiently small background-code rank are included. Known noisy-parity solvers, polynomial subspace lifts and high-fidelity learning are explicit ingredients. A bounded-weight local-Pauli corollary supplies another polynomial regime.

1 Models and main results

Let $S \leq \mathbb{F}_2^{2n}$ be the unsigned Lagrangian of an unknown stabilizer $|s_0\rangle$, with common eigenbasis $\{|s_z\rangle : z \in \mathbb{F}_2^n\}$. The iid input is

$$\rho = \sum_z p_z |s_z\rangle\langle s_z|. \tag{1}$$

For $v \in S$ use the target-signed Hermitian Pauli G_v, and identify S with the syndrome character group only for analysis. Our positive-spectrum promise is

$$f(v) = \mathrm{Tr}(\rho G_v) = \sum_z p_z (-1)^{z\cdot v} \geq \mu_0 > 0 \quad (v \neq 0). \tag{2}$$

Only a rational floor μ_0 is supplied, not S, its signs, or the distribution. Operations/readout are ideal. Each unknown copy is processed separately, with arbitrary within-copy Clifford gates and classical adaptation, but no coherent coupling of different input copies. Original-copy counts include all rejected preparations. Except in exponential tails, logarithms have base two.

Theorem 1.1 (Positive-spectrum learning)**.** *Fix $\varepsilon > 0$, $0 < \gamma < 1$ and $\delta \in (0, 1/3)$. Under (1)–(2), if $\mu_0 \geq 2^{-(\log_2 n)^\gamma}$, exact signed-target-tableau recovery with success at least $1 - \delta$ uses $O_{\varepsilon,\gamma,\delta}(n^{1+\varepsilon})$ original copies, polynomial total ideal quantum operations, and classical time $2^{O_{\varepsilon,\gamma,\delta}(n/\log\log n)}$. This is asymptotic, with bounded finite-dimensional fallback branches. Input-encoding costs are additional.*

For global white noise $c|s_0\rangle\langle s_0| + (1-c)I/2^n$, the nonidentity expectations equal c. For arbitrary diagonal noise with $p_0 \geq (1+\mu_0)/2$, they are at least $2p_0 - 1 \geq \mu_0$. Thus one theorem contains

both cases. It strictly extends them: on three qubits, mixing the target with weight $1/5$ and the normalized projector of a rank-one target-containing stabilizer code with weight $4/5$ gives eigenvalues $2/5, 1/5, 1/5, 1/5, 0, 0, 0, 0$ and nonzero target-signed expectations 1 or $1/5$.

For a second regime let $C \leq S$ be an unknown signed target-containing stabilizer code of rank r, and let Π_C be its projector. Consider

$$\rho = c|s_0\rangle\langle s_0| + (1-c)\Pi_C/2^{n-r}, \qquad c \geq c_0 > 0. \tag{3}$$

Write $\mathcal{Z} = \{(0, z) : z \in \mathbb{F}_2^n\}$. Neither c, C nor r is required as input.

Theorem 1.2 (Polynomial time with relative visibility)**.** *Fix rational constants $c_0 > 0$, $\beta \in (0, 1]$. Supply a polynomial-size list of known polynomial-size Clifford circuits, one of which satisfies*

$$\dim(US \cap \mathcal{Z}) - \dim(UC \cap \mathcal{Z}) \geq \beta n. \tag{4}$$

For (3), exact target recovery with failure at most δ uses original copies and time polynomial in n, the list/circuit sizes and $\log(1/\delta)$, for fixed c_0, β. The polynomial degree may depend on those constants. No low-weight-generator promise is needed.

Let D_X, D_Z be the X-only and Z-only target spaces and put $\kappa = n - \dim D_X - \dim D_Z$.

Corollary 1.3 (Dense CSS and CSS-defect targets)**.** *If $\kappa \leq an$ and $\dim C \leq bn$, for fixed nonnegative a, b with $b < (1-a)/2$, two frames $I, H^{\otimes n}$ suffice, with $\beta = (1-a)/2 - b$. In particular dense signed CSS targets and background rank bn, $b < 1/2$, are covered.*

Proof. One of $\dim D_X, \dim D_Z$ is at least $(1-a)n/2$. The corresponding frame makes it diagonal; the diagonal code space has dimension at most bn. □

1.1 Relationship to earlier work

The white-noise near-linear-copy theorem and its visible-subspace specialization were already public in [10]. This paper is a substantive same-family extension. Under white noise any correct seed improves the signal by a dimension formula. Under correlated errors, a correct seed may preserve every wrong eigenstate. Our increment proves a direction-wise success lower bound for a concrete *learned* finder and contracts the surviving wrong-to-target odds. Positive spectrum survives conditioning and unifies the models. For the polynomial regime we extend visible-subspace filtering to an unknown code background and use a relative, rather than absolute, diagonal dimension.

The star measurement is the partial-derivative primitive of [1], with an X/Y pivot for complex phases. Their Bell/LPN construction and Lyubashevsky's classical solver [2] supply quantitative ancestry. Full-support Clifford counts and efficient sampling are known [8, 7]. Bittel et al. [3] supply the final high-fidelity learner. General mixed-input stabilizer bootstrapping [11] and noise-robust Bell learning [9] use different inter-copy resources. Paid filtering and random parity hashing are established principles [11, 12]; our new obligation is their justified implementation with unknown-basis single-copy directions. The subspace lift is due to [4]. Section 6 is explicitly a standard shadow corollary [13], adjacent to graph and unknown-code learning [14, 15].

2 Conditioning and algorithmic inputs

Use Hermitian Paulis $W_{x,z} = i^{x \cdot z} X^x Z^z$, with integer overlap modulo four, retaining conjugation signs. Outside S all Pauli means vanish. Distinct pure stabilizers have squared overlap at most $1/2$.

Lemma 2.1 (Identifiability and closed conditioning)**.** *Under* (2)*,* $p_0 \geq \mu_0$ *and* $p_0 - p_z \geq \mu_0$ *for* $z \neq 0$*. Every other pure stabilizer has fidelity at most* $p_0 - \mu_0/2$*. A correct prefix* $D \leq S$ *accepts with* $q_D \geq p_0 \geq \mu_0$*, and its residual logical input again has positive-spectrum gap at least* μ_0*.*

Proof. Fourier inversion gives $p_0 = 2^{-n}\sum_v f(v) \geq \mu_0 + (1-\mu_0)2^{-n}$. Exactly half the characters detect nonzero z, whence $p_0 - p_z = 2^{-n}\sum_v (1-(-1)^{z\cdot v})f(v) \geq \mu_0$. For a different stabilizer t, put $r_z = |\langle t|s_z\rangle|^2$. Since $\sum_z r_z = 1$, $r_0 \leq 1/2$, its fidelity is at most $p_0 - \mu_0(1-r_0)$.

The prefix projector is $P_D = |D|^{-1}\sum_{d\in D} G_d$ and fixes the target. Hence $q_D \geq p_0$. For $v \in S\backslash D$, its conditional signed mean is

$$\frac{\mathrm{Tr}(\rho P_D G_v)}{q_D} = \frac{|D|^{-1}\sum_{d\in D} f(v+d)}{q_D} \geq \mu_0. \tag{5}$$

Every summand has a nonzero label and $q_D \leq 1$. A known Clifford decoder maps the prefix to coordinate Z constraints. Measure only those coordinates and discard them; surviving eigenstates give the residual diagonal basis. □

Proposition 2.2 (Classical parity input [2])**.** *Fix* $a > 0$, $0 < \theta < 1$*. Ordinary LPN in dimension* d*, with uniform features and independent error bits of correctness advantage* $\eta > 2^{-(\log d)^\theta}$*, is solvable from* $O(d^{1+a})$ *examples in time* $2^{O_{a,\theta}(d/\log\log d)}$ *with success tending to one. We use success at least* $9/10$ *above a fixed threshold.*

Here error probability is $1/2 - \eta$, so correctness advantage is half the correlation bias. The sample and weak-bias exponents are independent fixed parameters. The source theorem and its BKW preprocessing yield budgets from a lower advantage bound. Every call is capped, including off-promise calls. A rate grid below covers formulations supplying the exact rate; a uniform secret mask covers random-secret formulations. This is an external classical theorem.

Proposition 2.3 (High-fidelity endpoint [3])**.** *At fixed known stabilizer fidelity floor* $f_0 > \sqrt{2/3}$ *and additive accuracy* $\xi > 0$*, a mixed* m*-qubit input can be learned within* ξ *of its optimal stabilizer overlap with failure* ζ*, using* $O(m + \log(1/\zeta))$ *single-copy measurements and polynomial processing and gates.*

We use $f_0 = .9$, $\xi = .1$. At target weight $a \geq .9$, any other stabilizer has fidelity at most $a/2 + (1-a) \leq .55$. A successful additive-.1 output is therefore exactly the target.

3 A learned random component in a verified direction

Work after a correct prefix on m logical qubits. A uniformly random Clifford U gives a full-support target Lagrangian, transverse to $\mathcal{Z}$, with probability

$$p_m = \frac{2^{m(m+1)/2}}{\prod_{j=1}^m (2^j+1)} = \prod_{j=1}^m (1+2^{-j})^{-1} > e^{-1} > 1/4. \tag{6}$$

Indeed transverse Lagrangians are graphs of symmetric binary matrices. This standard count is also recorded in [8].

Lemma 3.1 (Exact diagonal star law)**.** *In a full-support frame there is a unique unsigned target direction with* X *label* e_1*, of form* $P_1 \otimes Z^{a_*}$*,* $P = X$ *or* Y*. Measuring that pivot axis and* Z *elsewhere gives*

$$\mathbb{P}(x,y) = 2^{-m}\big[1 + \lambda_U(-1)^{y + a_*\cdot x + b}\big], \qquad \lambda_U \geq \mu_0, \tag{7}$$

where b is its target sign bit. Thus x is uniform and label noise is feature-independent, with correctness advantage $\lambda_U/2$.

Proof. The X-projection is bijective. All common eigenstates have flat computational amplitudes and differ by linear phases. In each eigenstate the features are uniform and the generator equation fixes the pivot parity, shifted by its syndrome character. Averaging these laws yields (7); the bias is the corresponding conditional target-signed expectation, bounded by Lemma 2.1. No independence between syndrome coordinates was used. □

For each frame take separate datasets of $N_m = O((m-1)^{1+a})$ logical copies on both pivot axes. Try both intercept shifts and an advantage grid on $[\mu_0/2, 1/2]$, with endpoints and mesh at most $1/(100N_m)$, using at most $100N_m + 2$ points. Uniform masks r change a shifted label to $y + t + r \cdot x$ and are undone on proposed secrets. Datasets are reused only classically across grid calls. For the nearest grid advantage η',

$$\mathrm{TV}(\mathcal{D}_{\eta}^{N_m}, \mathcal{D}_{\eta'}^{N_m}) \le N_m|\eta - \eta'| \le 1/100. \tag{8}$$

The correct-axis/intercept call therefore succeeds with probability at least .89 on every favorable frame. At the noiseless endpoint use linear elimination. Malformed outputs are discarded; all calls are capped. Lift each complete candidate with identity on the prefix and validate on fresh original copies.

Lemma 3.2 (Direction minorization)**.** *Analyze an ideal version using exact membership/sign validation. Conditional on every correct history, one attempt returns each nonzero logical target direction v with probability at least*

$$\frac{\alpha}{2^m - 1}, \qquad \alpha = 1/8. \tag{9}$$

The remaining probability may give other genuine directions or no output; the actual output distribution need not be uniform.

Proof. On the favorable event G, let $u(U)$ be the unique graph direction with X label e_1, and $v(U) = U^{-1}u(U)$. Every linear automorphism g of S extends in symplectic coordinates to an automorphism g, g^{-T} of the full label space, hence to a Clifford V preserving S. Right multiplication $U \mapsto UV$ preserves the uniform frame law and G, but $v(UV) = V^{-1}v(U)$. The transitive action on $S \setminus 0$ makes $v(U)$ uniform conditional on G. This does not assume invariance of the diagonal error distribution.

Every genuine candidate on a favorable frame has the same X label e_1, so it is the same unique direction. An earlier accepted candidate cannot divert the successful correct call to another genuine direction. Its pointwise .89 success bound gives for every v a probability at least $p_m(.89)/(2^m - 1) \ge \alpha/(2^m - 1)$. Bad-frame genuine outputs add nonnegative mass. Couple actual finite validation to this process until its first estimation error, rather than conditioning the frame law on success. □

Lemma 3.3 (Correlated-error odds contraction)**.** *Let $Z_D = (q_D - p_0)/p_0$ be surviving wrong-to-target odds. A correct attempt/filter, leaving the prefix unchanged on failure, satisfies*

$$\mathbb{E}[Z_{\mathrm{next}} \mid \textit{history}] \le (15/16)Z_D. \tag{10}$$

After $T = \lceil 16 \ln(36/(\mu_0\delta)) \rceil$ attempts, target weight is at least .9 except with probability $\delta/4$.

Proof. Each surviving wrong eigenstate has nonzero residual syndrome w. Of the $2^m - 1$ nonzero directions exactly 2^{m-1} detect it. Minorization removes it with probability at least $1/16$. Correct additional constraints never increase surviving odds. Summing nonnegative wrong masses proves (10), without independent errors or independent finder outcomes. Initially $Z_0 \le 1/\mu_0$. Thus $\mathbb{E}Z_T \le e^{-T/16}/\mu_0$ and $\mathbb{P}(Z_T > 1/9) \le 9e^{-T/16}/\mu_0 \le \delta/4$. Residual target weight equals $1/(1+Z_T)$. If no logical qubits remain, set $Z = 0$ and absorb. At most T constraints are added. □

4 Finite caps and the original-copy ledger

Set $a = \min(\varepsilon/4, 1/4)$ and $\theta = (1+\gamma)/2$. For sufficiently large n, all feature dimensions obey $d = n - O_\delta(1 + (\log n)^\gamma)$ and

$$\mu_0/2 \ge 2^{-(\log n)^\gamma - 1} > 2^{-(\log d)^\theta}. \tag{11}$$

Let $N = \max N_m = O(n^{1+a})$, $J = 100N + 2$ and $B = 4TJ$ bound candidate slots. Validate each separately with

$$V = \left\lceil 32\mu_0^{-2} \ln(32\max(1, B)/\delta) \right\rceil \tag{12}$$

fresh original copies. Hoeffding bounds an error exceeding $\mu_0/4$ by $2e^{-V\mu_0^2/32}$. Union-bound over slots, conditioning on past data before each fresh dataset, to obtain simultaneous accuracy except with probability $\delta/16$. The threshold $|\widehat{f}| \ge \mu_0/2$ then separates off-S zero means from genuine magnitudes at least μ_0 and recovers target signs. Logical lifts have original means obeying (2); validation needs no prefix postselection.

The endpoint requests $Q_H = O(n + \ln(1/\delta))$ conditional copies with failure budget $\delta/4$. Set $Q = 2TN + Q_H$, and cap each request at

$$L = \left\lceil \mu_0^{-1} \ln(16\max(1, Q)/\delta) \right\rceil \tag{13}$$

fresh originals, each measured through the complete prefix once. Acceptance is at least $p_0 \ge \mu_0$, so any preparation cap fails with total probability at most $\delta/16$. Accepted copies have the exact conditional iid law. Couple capped/uncapped processes until their first charged error; retain all caps even on incorrect branches.

Exact Clifford sampling [7] uses polynomial work; cap integer rejection with total failure at most $\delta/16$. Odds-tail and endpoint failures are each $\delta/4$; validation, preparation and sampling are each $\delta/16$. Their sum is $11\delta/16 < \delta$. On the complementary event the endpoint gives the exact residual target, lifted through its decoder and prefix. A deterministic original-copy bound is

$$QL + BV. \tag{14}$$

For fixed $\varepsilon, \gamma, \delta$, T, μ_0^{-1} and logarithmic factors are $n^{o(1)}$. Hence this is $n^{1+a+o(1)} = O(n^{1+\varepsilon})$. A polynomial number of parity calls preserves $2^{O(n/\log\log n)}$ classical time. Quantum gates per original are polynomial. Finitely many dimensions below the common threshold use exhaustive Pauli estimation to $\mu_0/4$ and signed independent-generator recovery. The promised floor is positive there, so fallback changes only parameter-dependent constants. For $n = 1$ use direct Pauli estimation. This proves Theorem 1.1. Confidence is fixed in its near-linear statement, and constants are not uniform as $\gamma \uparrow 1$.

5 Polynomial time for an unknown code background

For (3), target-signed means are 1 on C, c on $S \setminus C$, and zero off S. In a supplied frame put $D = US \cap \mathcal{Z}$, $E = UC \cap \mathcal{Z}$, and identify their Z labels with linear spaces. The target computational

law is $\mathrm{Unif}(z+A)$, $A = D^{\perp}$, while the code-background law is $\mathrm{Unif}(z+B)$, $B = E^{\perp}$. Since signs agree, one affine shift z suffices; $A \le B$. Two fresh separate measurements give difference law

$$c^2\,\mathrm{Unif}(A) + (1-c^2)\,\mathrm{Unif}(B). \tag{15}$$

Any convolution involving the ambient component is uniform on B. The relative promise becomes $\dim B - \dim A \ge \beta n$.

If $c \le .9$, the ambient component weight is at least .19. Draw

$$M_B = \lceil (2/.19)(n\ln 2 + \ln(1/\zeta)) \rceil \tag{16}$$

differences and take their span $\widehat{B}$. Each proper hyperplane of B is escaped with probability at least .19/2, so a union bound over fewer than 2^n hyperplanes gives $\mathbb{P}(\widehat{B} \ne B) \le \zeta$. Use fresh differences for all subsequent work. Abort a trial if any lies outside $\widehat{B}$; otherwise map it to coordinates in that space. This makes incorrect ambient guesses total bounded branches.

On the good event identify B with $\mathbb{F}_2^b$. Set $w_0 = \min(c_0^2, .19)$. For $c \in [c_0, .9]$, both weights in (15) are at least w_0, and $\dim A \le b - \beta n \le (1-\beta)b$. Lemma 5.2 below recovers A with success at least .9 and polynomial resources for fixed c_0, β. Ambient-span recovery is an established part of the finite-field mixture framework [4].

Compute $D = A^{\perp}$ in the full physical frame, lift a basis, and validate every candidate and sign on fresh originals. Incorrect recoveries have no discovery guarantee but retain finite caps. For a correct full D, its filter accepts with

$$q_D = c + (1-c)2^{-g}, \qquad c_D = \frac{c}{c + (1-c)2^{-g}}, \qquad g = \dim D - \dim(C \cap D). \tag{17}$$

Within the background code, each of the g independent new constraints halves the support dimension. The conditional background is again a normalized target-containing code projector. In the promised frame $C \cap D = E$, so $g \ge \beta n$, and $c_D \ge .9$ for sufficiently large n. The endpoint therefore recovers the target with its declared success probability.

Repeat per frame, using fresh data, and also run the high-fidelity endpoint directly on original inputs. If $c > .9$, this direct branch succeeds. If $c \le .9$, the ambient/lift procedure succeeds on a promised frame. No estimate of c is needed. For every endpoint call enforce its resource cap, lift its full signed tableau, reject malformed, noncommuting or dependent outputs, and verify all n signed generators on fresh originals. Outside validation error, every accepted full tableau is exactly the target. This final check is essential for bad frames or insufficient prefixes.

For explicit confidence accounting let F be the frame-list length, $R = \lceil \ln(8/\delta)/\ln 2 \rceil$, and $H = FR + 1$ bound endpoint calls. With ambient error .1 and lift error .1, an attempt succeeds with probability at least $.8 > 1/2$ on the promised frame. Missing all good attempts costs at most $\delta/8$. Give each endpoint failure $\delta/(8H)$ and let $B_* = 2nH$ bound prefix and final validation slots. Use $\lceil 32 c_0^{-2} \ln(16\max(1, B_*)/\delta) \rceil$ originals per slot. If Q_* bounds all conditional-copy requests, cap each full-prefix preparation at $\lceil c_0^{-1} \ln(8\max(1, Q_*)/\delta) \rceil$ originals. Endpoint, validation and preparation each fail with total probability at most $\delta/8$. Other branches retain polynomial caps. The finitely many dimensions below the lift/amplification thresholds use exhaustive Pauli estimation. All thresholds depend only on fixed c_0, β. This proves Theorem 1.2.

5.1 The fixed-positive-weight lift

The following is the needed specialization of the polynomial-lift mechanism of [4], already used for white noise in [10]. We give the full proof because the printed numerical weight floor in [4] cannot silently be dropped. Write $\binom{d}{\le \ell} = \sum_{j=0}^{\min(d,\ell)} \binom{d}{j}$, and let $\phi_\ell(x)$ list squarefree monomials of degree at most ℓ, including the constant.

Lemma 5.1 (Known extremal rank ingredient)**.** *If $T \subseteq \mathbb{F}_2^n$ has at least 2^r distinct points, then* $\dim \operatorname{span} \phi_\ell(T) \geq \binom{r}{\leq \ell}$.

This is the extremal polynomial-rank result used as Lemmas 18–19 of [4], attributed there to [5] and equivalently to [6]. Also $\dim \operatorname{span} \phi_\ell(A) = \binom{d}{\leq \ell}$ for a d-dimensional linear A: an invertible linear coordinate change preserves degree, and restriction to A gives the multilinear polynomial space on d variables.

Lemma 5.2 (Recovery at every fixed positive mixture weight)**.** *Fix rational $w_0 \in (0, 1/2]$ and $\beta \in (0, 1]$. Given independent samples from*

$$w \operatorname{Unif}(A) + (1-w) \operatorname{Unif}(\mathbb{F}_2^n), \qquad w_0 \leq w \leq 1 - w_0, \quad \dim A \leq (1-\beta)n, \tag{18}$$

there is a polynomial-time and polynomial-sample algorithm recovering A with constant probability at least $9/10$, for fixed w_0, β. The algorithm does not know w or $\dim A$.

Proof. Choose a fixed integer $\ell \geq 2$ with $(1-\beta/2)^\ell \leq w_0/100$, and put $N = \binom{n}{\leq \ell}$, $\varrho = 49w_0/100$, $\eta = \varrho/(4\ell)$. For sufficiently large n, require

$$n \geq 2\ell/\beta, \qquad n \geq 8\ell^2/\varrho, \tag{19}$$

and

$$e^{-w_0 N/8} + N 2^{-\eta n} + 2^n e^{-4w_0 N/25} \leq 1/10. \tag{20}$$

These are effective conditions and hold eventually for fixed parameters.

Draw N samples x_i and mark precisely the indices for which $\phi_\ell(x_i)$ lies in the span of all other sampled columns. Return the span of the marked original vectors. All marks can be computed by one nullspace calculation: an index is marked exactly when some kernel vector has a nonzero coordinate there. Thus the procedure is bounded and polynomial in N, n, even on off-promise data.

Let $d = \dim A$ and $D_A = \binom{d}{\leq \ell}$. By (19),

$$\frac{D_A}{N} \leq \frac{\binom{d+\ell}{\ell}}{\binom{n}{\ell}} \leq \left(\frac{d+\ell}{n}\right)^\ell \leq (1-\beta/2)^\ell \leq w_0/100. \tag{21}$$

Label mixture components only for analysis. The number N_1 of inlier draws from A is at least $w_0 N/2$, except with probability $e^{-w_0 N/8}$.

Condition on all component labels satisfying that count and on every sample other than a full-space draw x_i. The span H_i of the other lifted columns together with the entire $\phi_\ell(A)$ has dimension

$$\dim H_i \leq N - 1 - N_1 + D_A \leq (1-\varrho)N.$$

It is independent of the uniform x_i. Set $r = \lfloor (1-\eta)n \rfloor$. Equation (19) gives

$$\begin{aligned} \frac{\binom{r}{\leq \ell}}{N} &\geq \frac{\binom{r}{\ell}}{\binom{n+\ell}{\ell}} \geq \left(\frac{r-\ell+1}{n+\ell}\right)^\ell \\ &\geq (1 - \varrho/(2\ell))^\ell \geq 1 - \varrho/2 > 1 - \varrho. \end{aligned}$$

By Lemma 5.1, fewer than 2^r original points can have their lift in H_i. A full-space draw therefore lies there with probability at most $2^{-\eta n}$. Union-bound over N indices. Except with probability $N 2^{-\eta n}$, no outlier column is marked. This also accounts for an outlier that accidentally lands in A.

Unmarked inlier columns are coloops of the lifted matrix and hence linearly independent among themselves. There are at most D_A of them, so the marked set contains at least $.98N_1$ inliers and no outliers. For any proper hyperplane of A, Hoeffding bounds the chance it contains at least $.9N_1$ uniform inliers by $e^{-8N_1/25}$. There are fewer than $2^d \le 2^n$ such hyperplanes. Except with probability $2^n e^{-4w_0 N/25}$, every subset containing $.9N_1$ inliers spans A. For $d = 0$ spanning is automatic. Equation (20) proves the recovery guarantee. Below the finite size threshold one can learn the finite classical distribution and enumerate subspaces; for fixed w_0, β this changes only constants. In the quantum application we instead use the finite-dimensional Pauli fallback below. □

The degree is $O(1 + \beta^{-1} \log(1/w_0))$. In the quantum application finite-dimensional fallback is Pauli estimation.

6 A standard bounded-weight polynomial regime

This separate sufficient condition is a local-shadow consequence of the credited low-weight-generator learning methods [13, 14, 15].

Proposition 6.1 (Sign-consistent low-weight generators)**.** *Suppose* (1) *has nonnegative target-signed expectations on all target stabilizer directions of weight at most w, and admits independent target generators of weight at most w whose expectations are at least a supplied $\mu > 0$. Put $M = \sum_{j=1}^{\min(w,n)} 3^j \binom{n}{j}$. Exact target recovery with failure at most δ uses*

$$N = \left\lceil 40\, 3^w \mu^{-2} \ln(2M/\delta) \right\rceil \tag{22}$$

random local-Pauli measurements and $O(NMw + Mn^2 + n^3)$ classical operations. For fixed w this is polynomial in $n, 1/\mu, \log(1/\delta)$.

Proof. Measure each qubit independently in a uniform X, Y, Z basis. For a candidate Pauli P of weight $k \le w$, set X_P to 3^k times its outcome product if all its axes match, and to zero otherwise. Then $\mathbb{E}X_P = \mathrm{Tr}(\rho P)$, $\mathbb{E}X_P^2 = 3^k \le 3^w$ and $|X_P| \le 3^w$. Bernstein gives simultaneous error at most $\mu/4$ with (22), by union bound. Keep candidates with $|\widehat{f}_P| \ge 3\mu/4$ and their estimated signs. Off-S zero means are rejected. Retained members have magnitude at least $\mu/2$, so nonnegativity makes their signs correct. All promised generators survive. Gaussian elimination gives n independent signed generators. Correlations between estimates do not affect the union bound. □

For an unknown graph of maximum degree D, or its unknown local Clifford image, independent local depolarization $\mathcal{D}_\lambda(\sigma) = \lambda\sigma + (1-\lambda)I/2$, $\lambda > 0$, gives $f(v) = \lambda^{\mathrm{wt}(G_v)} > 0$. Take $w = D + 1$, $\mu = \lambda^{D+1}$. This allows a fixed local error rate even when global target fidelity is exponentially small. Arbitrarily correlated diagonal distributions meeting the same sign/generator conditions are also included. A mixture of product depolarizing channels with a common positive contraction lower bound is one example. This illustrates a standard tractable subclass, rather than a new solution to local-noise graph learning. The uniform spectrum floor in Theorem 1.1 need not hold here.

7 Scope and reproducibility

The main theorem requires a positive bound on all target-signed expectations. A strict majority is sufficient but unnecessary. General Pauli noise need not satisfy that promise. The polynomial code

theorem requires the supplied relative-visible frame list. Each theorem retains its stated regime and hypotheses. The main classical search remains subexponential. Its upper reduction maps the problem to LPN.

The public exact checker corroborates finite Fourier conditioning, target gaps, odds contraction, nested-support differences and relative code-filter laws through four binary coordinates. These are identity checks, not executions of the full learner or the external LPN solver. The all-dimension claims rest on the analytic proofs and cited algorithmic theorems. Generic polynomial-copy availability is already obtainable from global Clifford shadows and exhaustive stabilizer comparison [13]; our main result concerns the joint copy, classical-time and input-promise regime.

AI assistance

AI tools assisted with derivation, programming, literature discovery and proof checking.

References

[1] S. Arunachalam, S. Bravyi, A. Dutt and T. J. Yoder, *Optimal algorithms for learning quantum phase states*, TQC 2023, LIPIcs **266**, 3:1–3:24. doi:10.4230/LIPIcs.TQC.2023.3. Full version: arXiv:2208.07851v2.

[2] V. Lyubashevsky, *The parity problem in the presence of noise, decoding random linear codes, and the subset sum problem*, APPROX/RANDOM 2005, LNCS **3624**, 378–389. doi:10.1007/11538462_32. Author manuscript.

[3] 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 (2026).

[4] A. Chen, A. De and A. Vijayaraghavan, *Learning a mixture of two subspaces over finite fields*, ALT 2021, PMLR **132**, 481–504. Proceedings paper. Full version: arXiv:2010.02841v2.

[5] I. Ben-Eliezer, R. Hod and S. Lovett, *Random low-degree polynomials are hard to approximate*, Computational Complexity **21** (2012), 63–81.

[6] P. Keevash and B. Sudakov, *Set systems with restricted cross-intersections and the minimum rank of inclusion matrices*, SIAM Journal on Discrete Mathematics **18** (2005), 713–727.

[7] R. Koenig and J. A. Smolin, *How to efficiently select an arbitrary Clifford group element*, Journal of Mathematical Physics **55** (2014), 122202. doi:10.1063/1.4903507. arXiv:1406.2170.

[8] S. Arunachalam and L. Schatzki, *Optimal stabilizer testing and learning with limited quantum memory*, arXiv:2607.02444v2 (2026).

[9] L. Tang and T. Steckmann, *Learning noise-robust stabilizer structure via Bell sampling*, arXiv:2610.03307v2 (2026).

[10] S. Schubr, *Single-copy learning of noisy stabilizers by certified filtering*, Luminary Research Archive, edition 1 (2026). Public predecessor.

[11] S. Chen, W. Gong, Q. Ye and Z. Zhang, *Stabilizer bootstrapping: A recipe for efficient agnostic tomography and magic estimation*, arXiv:2408.06967 (2024), STOC 2025.

[12] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin and W. K. Wootters, *Mixed-state entanglement and quantum error correction*, Phys. Rev. A 54 (1996), 3824–3851. quant-ph/9604024.

[13] H.-Y. Huang, R. Kueng and J. Preskill, *Predicting many properties of a quantum system from very few measurements*, Nature Physics 16 (2020), 1050–1057. arXiv:2002.08953.

[14] Y. Ouyang and M. Tomamichel, *Learning quantum graph states with product measurements*, arXiv:2205.06432 (2022).

[15] H. Leitch, S. Tirukkovalluri and Y. Ouyang, *Learning unknown stabilizer codes using product measurements*, arXiv:2609.04997v2 (2026).