Expander qLDPC Codes against Long-range Correlated Errors in Memory

Yash Deepak Kashtikar,[1,*] Pranay Mathur,[1,*] Sudharsan Senthil,[1,*] and Avhishek Chatterjee[1,†]

[1] *Department of Electrical Engineering, Indian Institute of Technology Madras, India*

(Dated: October 7, 2025)

Fault-tolerance using constant space-overhead against long-range correlated errors is an important practical question. In the pioneering works [1–3], fault-tolerance using poly-logarithmic overhead against long-range correlation modeled by pairwise joint Hamiltonian was proven when the total correlation of an error at a qubit location with errors at other locations was $O(1)$, i.e., the total correlation at a location did not scale with the number of qubits. This condition, under spatial symmetry, can simply be stated as the correlation between locations decaying faster than $\frac{1}{\text{dist}^{\text{dim}}}$. However, the pairwise Hamiltonian model remained intractable for constant overhead codes. Recently, [4] introduced and analyzed the generalized hidden Markov random field (MRF) model, which provably captures all stationary distributions, including long-range correlations [5]. It resulted in a noise threshold in the case of long-range correlation, for memory corrected by the linear-distance Tanner codes [6] for super-polynomial time. In this paper, we prove a similar result for square-root distance qLDPC codes and provide an explicit expression for the noise threshold in terms of the code rate, for up to $o(\sqrt{\#\text{qubits}})$ scaling of the total correlation of error at a location with errors at other locations.

Quantum low density parity check codes (qLDPC) codes [7–10] are known to have a polynomial distance and constant space overhead. This is a significant improvement over the poly-logarithmic space overhead required for surface codes, and may lead to practically useful quantum computers using sub-million physical qubits. Expander qLDPC codes with sublinear distance [7] were shown to provide fault-tolerance against i.i.d. and local stochastic errors [8, 9, 11, 12]. Furthermore, sublinear-distance qLDPC codes have shown encouraging performance in moderately sized practical quantum memories.[13].

Despite significant advances in theory and practice towards achieving fault-tolerance, some criticisms remain. Two main issues raised by critics are: (i) whether syndromes can be extracted from constant rate codes using (topologically) local gate operations, and (ii) whether the constant overhead fault-tolerance schemes can correct long-range correlated errors. Recently, some interesting work has been done to address the first issue, where promising local and semi-local syndrome extraction methods have been proposed [14, 15]. In this work, we try to address the second issue, that of fault-tolerance against long-range correlated errors.

Fault-tolerance against the local stochastic error model is an important step towards studying correlation. However, since this model has an exponential correlation decay [1], the problem of fault-tolerance against long-range correlation remains open. The seminal series of papers [1–3] showed fault-tolerance against long-range correlated errors using poly-logarithmic space overhead.

Though a well-chosen system and bath joint Hamiltonian can model any correlation, it is often analytically intractable. In [1–3], pairwise joint Hamiltonian models were introduced and it was shown that if for any qubit its total correlation with all other qubits is $O(1)$, fault-tolerance can be achieved. Under spatial symmetry, this condition is satisfied if the correlation between two locations decays faster than $\frac{1}{\text{dist}^{\text{dim}}}$.

However, studying constant overhead codes such as qLDPC codes, using the pairwise joint Hamiltonian model remained intractable. Recently, in [4], the generalized hidden Markov random field model was proposed, which could model all stationary distributions including the ones where any qubit's total correlation with other locations scale as $(\#\text{qubits})^{\gamma}$, for $0 \leq \gamma \leq 1$. This model was proven to capture correlation structures that are not captured by pairwise joint Hamiltonian models [4]. Also, any quantum memory error corrected using linear distance Tanner codes (from [6, 16]) was shown to have a super-polynomial retention time when the total correlation of a qubit location with others grows no faster than $\sqrt{\#\text{qubits}}$.

In this paper, we study fault-tolerance of the square-root distance expander qLDPC codes [7–10] in long-range correlated errors modeled as a generalized hidden Markov random field (MRF) [4] and show a positive noise threshold (of $\Theta(1)$) when the total correlation of a qubit grows no faster than $\sqrt{\#\text{qubits}}$. To our knowledge, this is the first provable noise threshold result for sublinear distance constant overhead codes against long-range correlations where the total correlation at a location increases with the number of qubits.

Fault-tolerant memory

The memory contains a state of n qubits that encodes k logical qubits using a stabilizer code. The evolution of this memory is modeled as a sequence of periodic phases, each of which contains a *rest phase* and an *error correction phase*. The n-qubit state can decohere and ac-

* Student authors are listed in alphabetical order.

† Contact author: avhishek@ee.iitm.ac.in

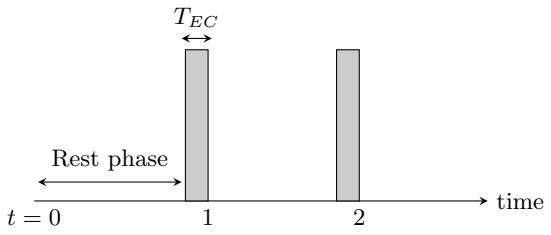

FIG. 1. Illustration of the quantum memory model. The labeled discrete times correspond to the end of a periodic phase. Each of these phases consists of a rest phase and an error correction phase. Qubits can decohere during the rest phase, it is hence followed by a T_{EC}-long error-correction phase. The error-correction phase includes syndrome extraction as well as correction.

cumulate errors during the rest phase. Since an arbitrary error can be decomposed as a linear combination of Pauli flips [17, 18], following the existing literature on fault-tolerance using qLDPC codes [9, 19, 20], we consider only Pauli errors.

The error correction phase has three main sub-phases. The first sub-phase is the extraction of syndrome bits by employing ancillary qubits. Next, based on the syndrome bits, a decoding algorithm estimates the locations and types of Pauli flips (X or Z) that have occurred. Finally, appropriate Pauli operators are applied to those locations to correct the errors. Depending on the choice of code and decoding algorithm, and the distribution of Pauli errors, all qubit errors may or may not be corrected in an error correction phase.

We index the periodic phases by discrete time-steps $t = 1, 2, \ldots$ and use the notation $\{L_{i,t}\}$ to denote the locations at which Pauli errors occur during the t^{th} rest phase. An error at location $i \in [n]$ at time-step t is indicated by setting $L_{i,t} = 1$, otherwise $L_{i,t}$ is set to 0. We use the notation $E_t \subset [n]$ to denote $\{i : L_{i,t} = 1\}$. We call E_t^e the effective error before the error correction phase t, which includes E_t and the uncorrected errors from previous phases that have been carried forward.

Long-range Correlated Qubit Errors

Long-range correlations between errors can be spatial and temporal. In phase t, there are correlations between $\{L_{i,t} : i \in [n]\}$, and there are correlations also between phases, i.e., between $\{L_{i,t} : i \in [n]\}$ and $\{L_{i,t'} : i \in [n]\}$ or, in other words, between E_t and $E_{t'}$.

For almost all systems in equilibrium with the environment, errors are stationary in time and space, i.e., $\mathbb{P}(L_{i,t} = 1)$ is the same for all t and i and we denote it by $\bar{p}$. It is known from the probability literature [5] that a hidden Markov random field model can approximate any stationary distribution with arbitrarily high precision. Motivated by this fact and the analytical tractability of hidden random fields, a generalized hidden Markov random field model was proposed in [4] to model long-range correlated errors. This error model was proven to be broader than the well known pairwise Hamiltonian model for long-range correlation [1–3], and to include long-range correlations not captured by the said model. In this work, we study fault-tolerance using qLDPC under this generalized hidden Markov random field model.

In the system, $\{L_{i,t} : i \in [n]\}$ are the manifest errors, which are caused by underlying (hidden) physical scenarios in the vicinity of the qubit locations. The physical scenario in the vicinity of a qubit is dictated by the local temperature, local circuit disturbances, local effects of external fields and radiations [21, 22], etc. The hidden parameters at time t are denoted by $\{J_{i,t} : i \in [n]\}$. Each $J_{i,t}$ is a vector of potentially multiple local physical parameters. We assume $J_{i,t}$ to be from a countable set, possibly infinite.

The probability of having an error at location i, $\mathbb{P}[L_{i,t} = 1]$, is potentially affected by multiple of these hidden parameters, and these hidden parameters are also highly correlated with each other. We use the following generalized hidden Markov random field model from [4] to capture long-range correlation.

$$\mathbb{P}(\{L_{i,t} : i \in [n]\}) = \mathbb{P}(\{J_{k,t} : k \in [n]\}) \times \prod_i \mathbb{P}(L_{i,t}|\{J_{k,t} : k \in [n]\}), \text{ and}$$

$$\mathbb{P}(\{J_{k,t} : k \in [n]\}) = \mathbb{P}(J_{1,t}) \prod_{k=1}^{n-1} \mathbb{P}(J_{k+1,t}|J_{k,t}).$$

For each $i \in [n]$, we define $g_i(\{J_{k,t} : k \in [n]\}) := \mathbb{P}(L_{i,t}|\{J_{k,t} : k \in [n]\})$ such that $\mathbf{E}[g_i(\{J_{k,t} : k \in [n]\}) = \bar{p}$. We assume that each g_i varies smoothly with changes in $\{J_{i,t}\}$. Formally, we assume g_i to be c_n-Lipschitz with respect to the Hamming distance, with $c_n = O(n^{-0.5-\epsilon_g})$ for $0 < \epsilon_g \leq 0.5$. In [4, Prop. 1], an information theoretic converse was proven for a class of non-smooth $\{g_i\}$: despite the total correlation being $O(1)$ at any location, the memory retention time is $O(n^2)$. Hence, the assumption regarding the smoothness of $\{g_i\}$ seems unavoidable for achieving fault-tolerance over a long duration (high polynomial or super-polynomial).

Furthermore, the assumption on $\{g_i\}$ is a generalization of a condition that would likely arise in quantum chips when they scale. A major reason for the success of classical chips was that the area of the chips did not scale with the number of bits. This slow scaling of the area is also desired for quantum chips for them to be useful in practice. Such scaling would essentially require a linear number of qubits to be within a short physical distance, in turn causing the probability of error of a given qubit to be affected by the local environments of a linear number of qubits. However, g_i being a measure of probability, can be 1 at most. The individual impact of each of these $\Theta(n)$ neighboring locations would therefore have to be $\Theta(n^{-1})$. This scenario is equivalent to g_i being c_n-Lipschitz with $c_n = O(n^{-1})$. Thus, the scenario

we consider, i.e., $c_n = O(n^{-0.5-\epsilon_g})$ for $0 < \epsilon_g \leq 0.5$, is a more general, relaxed condition and would also be true advanced technologies in the future with improved insulation between qubits on a chip.

Metrics for Long-range Dependence

For most physical systems, the physical parameters are stationary in a stochastic sense. Hence, it is natural to assume that $\{J_{i,t}\}$ are stationary in time, which in turn would imply that $\{L_{i,t}\}$ are stationary in time. However, we do not assume any temporal correlation decay or temporal Markovian property for the hidden and manifest variables, $\{J_{i,t}\}$ and $\{L_{i,t}\}$.

Next, we discuss the notion of correlation in the study of long-range correlation. In the pioneering series of works on long-range correlation [1–3], the model involved pairwise terms H_{ij}, for qubit locations i and j, in the system and bath joint Hamiltonian. The condition under which a positive fault-tolerance threshold was proved in [3] is: for each $i \in [n]$, $\sum_j ||H_{ij}|| = O(1)$.

For characterizing how the interactions or correlations between errors should decay with distance to have fault-tolerance ,this condition was simplified to: $||H_{ij}||$ decaying faster than $|i-j|^{-\dim}$. However, $\sum_j ||H_{ij}|| = O(1)$ is a broader condition, since it can be true despite the decay of $||H_{ij}||$ being slower than $|i-j|^{-\dim}$.

To quantify the dependence between two random variables Y_1 and Y_2, a well-known metric in probability theory, computer science, and statistical physics [23] is the total variation distance between their joint distribution and the product of their marginal distributions. Here, we denote that by $\mathrm{Dep}(Y_1, Y_2)$. When this metric is close to 0, i.e., Y_1 and Y_2 have low dependence, and Y_1, Y_2 take values in a finite set, their covariance, which we denote by $\mathrm{cov}(Y_i, Y_j)$, is a Lipschitz continuous function with respect to this metric and hence can be bounded by a constant times the metric. Thus, when $\{\mathrm{Dep}(Y_i, Y_j)\}$ are close to 0, the total accumulated covariance at i, $C(Y_i) := \sum_j \mathrm{cov}(Y_i, Y_j)$ is equivalent to $\mathrm{Dep}(Y_i) = \sum_j \mathrm{Dep}(Y_i, Y_j)$, the total dependence of Y_i on $[n]$, in the order sense, i.e., $C(Y_i) = \Theta(\mathrm{Dep}(Y_i))$.

The total number of errors $|E|$, also given by $\sum_i L_i$, has the mean $\bar{p} \cdot n = \Theta(n)$. Its variance, given by $\sum_i C(L_i)$, is $O(n)$, when $C(L_i) = O(1)$. It is implicit that fault-tolerance is not possible without having all but a vanishing fraction of $\{\mathrm{Dep}(L_i, L_j)\}$ and $\{\mathrm{cov}(L_i, L_j)\}$ close to 0, otherwise the variance of $|E|$ would be $\Theta(n^2)$ and error correction would be impossible in an information theoretic sense.

In the pairwise joint Hamiltonian model, $||H_{ij}||$ captures the dependence between error at i and j, and thus is equivalent to $\mathrm{Dep}(L_i, L_j)$. So, the quantity $\sum_j ||H_{ij}||$ is the total influence or correlation of all qubits on the qubit at location i and is equivalent to the total dependence of qubit i on $[n]$, $\mathrm{Dep}(L_i) = \sum_j \mathrm{Dep}(L_i, L_j)$, and the total accumulated covariance at i, $C(L_i) = \sum_j \mathrm{cov}(L_i, L_j)$.

Thus, in this sense, the condition derived in [1–3] for the dependence or correlation structure to achieve fault-tolerance can also be seen in three equivalent ways: (i) $\mathrm{var}(|E|)$ is $O(n)$ i.e., $\mathrm{var}(\sum_i L_i)$ is $O(n)$, (ii) $\mathrm{Dep}(L_i)$ is $O(1)$, and (iii) $C(L_i)$ is $O(1)$.

In the generalized hidden MRF model [4], the metric used to capture the total influence or correlation with all qubits on qubit i, denoted by $V(J_i)$, is defined as

$$\sum_j \sup_{a,b} \frac{1}{2} \sum_h |\mathbb{P}(J_j = h | J_i = a) - \mathbb{P}(J_j = h | J_i = b)|.$$

Each term in the summation corresponding to j is equivalent to $\mathrm{Dep}(J_i, J_j)$, the total variation distance between the joint and the product of the marginals of J_i and J_j. Thus, $V(J_i)$ is equivalent to $\mathrm{Dep}(J_i) = \sum_j \mathrm{Dep}(J_i, J_j)$ in the order sense.

When $\{g_i\}$ are smooth, fault-tolerance using linear distance Tanner codes is achieved when the total correlation at qubit i with all other qubits, $V(J_i)$, is $o(n^{0.5})$ [4]. This is a significant improvement over the previous condition on variance and total correlation in the pairwise Hamiltonian model, $C(L_i) = O(n)$, if $V(J_i)$ is equivalent to $\mathrm{Dep}(L_i)$ and $C(L_i)$ in the order sense.

By the law of total covariance and the distribution of $\mathbb{P}(\{L_{i,t} : i \in [n]\})$ defined above, the covariance between L_i and L_j is the same as the covariance between $g_i(\{J_k : k \in [n]\})$ and $g_j(\{J_k : k \in [n]\})$. Consider the case where $g_i(\{J_k : k \in [n]\}) = \sum_k c_{i,k} J_k$ for $c_{i,k} \geq 0$, $J_k \geq 0$ and J_k belong to a finite alphabet. We need to have $\sum_k c_{i,k} = O(1)$ since g_i are probabilities. Then, the covariance of L_i and L_j can be written as

$$\mathrm{cov}(L_i, L_j) = \sum_{k,k'} \mathrm{cov}(J_k, J_{k'}) c_{i,k} c_{j,k'}.$$

Thus,

$$\begin{aligned} C(L_i) &= \sum_j \sum_{k,k'} \mathrm{cov}(J_k, J_{k'}) c_{i,k} c_{j,k'} \\ &= \Theta(1) \cdot \sum_{k,k'} \mathrm{cov}(J_k, J_{k'}) c_{i,k} \\ &= \Theta(1) \cdot \sum_k C(J_k) c_{i,k} = \Theta(C(J_i)), \end{aligned}$$

where the last step is due to spatial symmetry and the fact that $\sum_k c_{i,k} = O(1)$. Since $C(J_i)$ is equivalent to $\mathrm{Dep}(J_i)$, which is equivalent to $V(J_i)$, we see that $C(L_i)$ and $V(L_i)$ are equivalent to $V(J_i)$ in the order sense.

Hence, the condition under which fault-tolerance is proved for linear distance Tanner codes in [4] is indeed equivalent to $C(L_i)$ and $V(L_i)$ being $o(n^{0.5})$ and thus is an order-wise improvement over $C(L_i)$ and $V(L_i)$ being $O(1)$ in [2, 3, 24].

Since $V(J_i)$ is equivalent to $V(L_i)$ and $C(L_i)$, we use $V(J_i)$ as our metric for the total correlation of qubit i with all qubits, as in [4], and prove fault-tolerance for $V(J_i) = o(n^{\epsilon_g})$. As discussed above, in the likely case of large quantum chips, ϵ_g is 0.5.

From a physical perspective, $\{g_i\}$ are continuous non-decreasing functions of $\sum_k c_{i,k} J_k$, like the sigmoid function. Since in fault-tolerance, we are interested in the case when probabilities of errors g_i and g_j are close to 0 (low noise), in that range g_i and g_j can be lower and upper bounded by two lines through the origin. Hence, for small error rate, the above order-wise equivalence shown between $C(J_i)$ (or $V(J_i)$) and $C(L_i)$ (or $V(L_i)$) for linear $\{g_i\}$ also holds for nonlinear $\{g_i\}$.

qLDPC for fault-tolerant memory

Quantum Low Density Parity Check(qLDPC) codes are a well-studied group of stabilizer codes. Like stabilizer codes [25], they are constructed by combining two classical codes, one to correct Pauli X flips and the other to correct Pauli Z flips. The corresponding parity check matrices H_X and H_Z are such that the Z code is the dual of the X code. For constructing qLDPC codes, we need the two classical codes to be LDPC codes. The low density parity check property of a classical code means that its parity check matrix is sparse, where sparsity refers to the weights of the rows and columns being constant, i.e., not scaling with the size of the codeword.

An effective way to construct classical low density parity check codes is to use expander codes [10, 11]. Expander codes are constructed using bipartite expander graphs that serve as Tanner graphs for the code. The left nodes are code bits and the right nodes are parity checks.

In this work, we use a particular class of quantum LDPC (qLDPC) codes, the quantum expander codes introduced in [10]. This code has a distance of $\Omega(\sqrt{n})$, as shown in [11, 12]. In [12], these codes were shown to have a small set flip decoding algorithm with locality property. It was also shown to have fault-tolerance against errors distributed according to the local stochastic noise model that has a fast (exponential) correlation decay [1].

In this paper, our main contribution is to show that the same codes can offer fault-tolerance against long-range correlated errors with significantly slower correlation decay (polynomial with degree 1) . Thus, this work gives an affirmative answer to whether constant overhead fault-tolerance is possible against long-range correlated errors.

Our proof techniques extend to other qLDPC codes that have a polynomial distance and share a similar locality property of the decoding algorithm. There are recent qLDPC codes with these properties [20, 26]. However, in this paper, for the sake of simplicity in presentation, we limit our discussion to the $\Omega(\sqrt{n})$-distance expander qLDPC codes that were studied in [10–12].

As argued in previous work on fault-tolerance [11, 12], it is enough to prove fault-tolerance using expander qLDPC codes assuming only the possibility of Pauli-X errors , as the case of Pauli-Z errors follows similarly. Hence, for simplicity in presentation, the variables $\{L_{i,t}\}$, E_t and E_t^e can be assumed to correspond only to Pauli-X errors.

As discussed above, the fault-tolerant memory model considered here is standard: a state is stored after encoding it using the aforementioned expander qLDPC codes, and a polynomial time decoding algorithm is run in each error-correction phase. We employ the small-set-flip decoding algorithm [11, Alg. 2].

Noise threshold against long-range correlated errors

In this section, we present our main result on the existence of a noise threshold for memory affected by long-range correlated errors. In that regard, we first introduce the essential parameters that describe the expander qLDPC code, the decoding algorithms, and an appropriate metric for spatial correlation among errors.

The expander qLDPC code used here is the symmetric version of the code in [11], where the chosen bipartite expander graph has symmetric expansion coefficients for the left and right sets of vertices, i.e., $\delta_A = \delta_B$ and $\gamma_A = \gamma_B$, as was the choice in [12]. However, the left and right degrees (d_A and d_B, or d_V and d_C, in [11] and [12] respectively) are different, and their ratio is denoted by r. The rate of the expander qLDPC code has a lower bound of $\frac{(1-r)^2}{1+r^2}$, which follows from [7, 10, 11].

As mentioned above, in each error correction phase, the linear-time small set flip decoding algorithm (Alg. 2 in [11]) is run to periodically correct the errors that accumulate over time. The number of adversarial errors that can be corrected by the original small set flip decoding algorithm has a lower bound of $c'\sqrt{n}$, where $c' = \frac{r(1-8\delta_A)}{4+2r(1-8\delta_A)}\gamma_A \frac{r^2}{\sqrt{1+r^2}}$, by Proposition 4.16 in [12].

Definition 1. *Adjacency graph $\mathcal{G}$. [11].*

The set of qubits is represented as a graph $\mathcal{G} = (\{1, \ldots, n\}, \mathcal{E})$ called an adjacency graph with vertices representing the qubits of the code and edges representing the presence of a stabilizer generator acting on both qubits.

The graph $\mathcal{G}$ in our case is the same as in [11]. and has a $O(1)$ degree denoted by $d_{\mathcal{G}}$.

Proposition 1. *Suppose a state is stored using a constant rate expander qLDPC code from [10–12] with parameters (d_A, d_B) and $\delta_A = \delta_B < \frac{1}{32}$, and it is periodically error corrected using the small-set-flip decoder [11, Alg 2]. Suppose that the memory has been subjected to the long-range correlated error discussed above where the average error rate per epoch is $\bar{p}$, $\{g_i : i \in [n]\}$ are c_n-Lipschitz with respect to the Hamming distance with $c_n = O\left(\frac{1}{n^{0.5+\epsilon_g}}\right)$, and the total correlation at i, $V(J_i)$, scales as $o(n^\gamma)$ for $0 \leq \gamma < \epsilon_g \leq 0.5$. Then, for $C_{\mathcal{G}} = \ln\left((d_{\mathcal{G}} - 1)(1 + \frac{1}{d_{\mathcal{G}}-2})^{d_{\mathcal{G}}-2}\right)$, there exists $p_{th} = \exp\left(-\frac{808\cdot(1+\frac{3r}{8})\cdot C_{\mathcal{G}}}{300\cdot r}\right)$ such that for all $\bar{p} < p_{th}$, if the state is retrieved at a time $T \leq \exp(O(n^{\epsilon_g - \gamma}))$, it would have fidelity $1 - o(1)$.*

For the same code parameters used in [11] for numerical evaluations, the threshold value here is similar: 10^{-16}–10^{-15}. Thus, the noise threshold above for long-range correlated errors is comparable to that of local stochastic and i.i.d. errors. Note that γ has to be less than ϵ_g in this proposition and ϵ_g bounded by 0.5 to ensure $\bar{p} = \Omega(1)$. Hence, the slowest correlation decay and the fastest growth of total correlation at a location, for which the proposition ensures a positive threshold for a super-polynomial memory retention time, are $\frac{1}{\sqrt{n}}$ and $\sqrt{n}$ respectively.

The noise threshold in Proposition 1 depends directly on the rate of the chosen code, which is known to have a lower bound of $\frac{(1-r)^2}{1+r^2}$, from [7, 10, 11]. If this constant code rate is small, i.e., r is close to 1,
$p_{th} \gtrapprox \exp\left(-\frac{808 \cdot C_{\mathcal{G}}}{300}\right) \cdot \exp\left(-\frac{1.01 \cdot C_{\mathcal{G}}}{1-\sqrt{2 \cdot \text{code rate}}}\right)$.

Since the code rate decreases with increasing r, the noise threshold p_{th} decreases with higher code rates. This implies the existence of a trade-off between the noise threshold and the constant space overhead. This is not unexpected, since even for classical codes, codes with higher rates are less suited to correct more frequent errors.

Proof of Proposition 1

To prove the main result in Proposition 1, we use the following general theorem, which relates the noise threshold to parameters of the chosen code and the constant-time decoding algorithm.

Theorem 1. *For a 1-dimensional arrangement of qubits, for any time $t \leq \exp(O(n^{\epsilon_g - \gamma}))$, the memory state can be accurately retrieved at t with probability $(1 - o(1))$ if $\bar{p} < \exp\left(-\frac{101 \cdot C_{\mathcal{G}}}{100 \cdot \alpha}\right)$ and the long-range correlated error satisfies the conditions in Proposition 1, where $C_{\mathcal{G}} = \ln\left((d_{\mathcal{G}} - 1)(1 + \frac{1}{d_{\mathcal{G}}-2})^{d_{\mathcal{G}}-2}\right)$. Here $\alpha = \frac{\frac{r}{2}(1-8\delta_A)}{\frac{r}{2}(1-8\delta_A)+1}$.*

□

This theorem states that using the small-set-flip decoder in each error correction phase, the probability of accurate retrieval asymptotically reaches 1 when the error probability is below the given threshold. Next, we use this result to prove the main result in Proposition 1.

Proof of Proposition 1. Correct recovery with probability $1 - o(1)$ ensures that fidelity is no less than $1 - o(1)$. In Theorem 1 when we use $\delta_A < \frac{1}{32}$ the result in Proposition 1 follows. ■

To prove the general noise threshold result in Theorem 1 we generalize the proof technique from [11] to the proposed long-range error model. To do this, we must first establish some additional definitions.

Definition 2. *α-subset and $MaxConn_\alpha$ [11]*

Consider a graph $\mathcal{G}$ with a bounded degree. With respect to $\mathcal{G}$, we can define $\alpha \in (0; 1]$ and let $X, Y \subseteq \mathcal{V}$. X is an α-subset of Y if

$$|X \cap Y| \geq \alpha |X|.$$

We also define the integer $MaxConn_\alpha(Y)$ as:

$$MaxConn_\alpha(Y) = \max\{|X| : X \text{ is connected in } \mathcal{G} \text{ and is an } \alpha\text{-subset of } Y\}$$

These quantities have direct implications for the correction capabilities of quantum expander codes, as is extensively shown in [11].

Theorem 2. *The error after running the small-set-flip decoder for $(t-1)$ time-steps is represented by E_t^e (effective error at time t). Then, for $t \leq \exp\left(c_4 n^{b_4}\right)$, under the same conditions for α and $\bar{p}$ as in Theorem 1,*

$$\mathbb{P}\Big[MaxConn_\alpha(E_t^e) > c'\sqrt{n}\Big] \leq \exp\left(-c_4 n^{b_4}\right),$$

for $\alpha \ln \frac{1}{\bar{p}} \geq \frac{101}{100} C_{\mathcal{G}}$, where $b_4 = \epsilon_g - \gamma$, $c_4 = \Omega(1)$, and $C_{\mathcal{G}} = \ln\left((d_{\mathcal{G}} - 1)(1 + \frac{1}{d_{\mathcal{G}}-2})^{d_{\mathcal{G}}-2}\right)$.

□

This probability is directly tied to the number of errors that can be successfully corrected by our model, a property that will be elaborated on in the following proof.

Proof of Theorem 1. By Proposition 3.9 in [11], when $MaxConn_\alpha(E_t^e)$ is less than the number of adversarial errors that can be corrected by the small set flip algorithm, the errors in E_t^e are corrected by the original small set flip decoder (Alg. 2 in [11]).

The number of adversarial errors that can be corrected by the original small set flip decoder is at least $c'\sqrt{n}$, where $c' = \frac{r(1-8\delta_A)}{4+2r(1-8\delta_A)} \gamma_A \frac{r^2}{\sqrt{1+r^2}}$, by Proposition 4.16 in [12].

By Theorem 2, $MaxConn_\alpha(E_t^e)$ is less than $c'\sqrt{n}$ with probability $\geq 1 - \exp\left(-c_4 n^{b_4}\right)$ for $t \leq \exp\left(c_4 n^{b_4}\right)$. Hence, for $t < \exp\left(c_4 n^{b_4}\right)$, the state in the memory can be retrieved using the original small set flip decoder with probability $1 - o(1)$.

From [11], for a particular δ_A, the maximum possible value of α is $\frac{\frac{r}{2}(1-8\delta_A)}{(1-8\delta_A)\frac{r}{2}+1}$.

■

For proving Theorem 2 we have to introduce certain quantities similar to those in [11].

Definition 3. *For a graph $\mathcal{G}$, define $\mathcal{C}_s(\mathcal{G})$ as the set of connected sets of size s in $\mathcal{G}$.*

For $v \in V$, $X \subseteq V$, $E \subseteq V$ and $s \in \mathbb{N}$, define:
$\mathcal{C}_s(v) = \{X \in \mathcal{C}_s(\mathcal{G}) : v \in X\}$, *connected sets containing* v,

$\partial X = \{$ *neighbors of* $X\} \setminus X$,
$A(E, \alpha, s, v) = \{X \in \mathcal{C}_s(v) : |X \cap E| \geq \alpha|X|, \partial X \cap E = \varnothing\}$

Next, we state two lemmas and a theorem that we subsequently use to prove Theorem 2.

Lemma 1. *For $s = \omega(\log n)$,*

$$|\mathcal{C}_S(v)| \leq e^{C_{\mathcal{G}} s},$$

where $C_{\mathcal{G}} = \ln\Big((d_{\mathcal{G}} - 1)(1 + \frac{1}{d_{\mathcal{G}}-2})^{d_{\mathcal{G}}-2}\Big)$.

Definition 4.

$$A_J^t := \Big\{ J_{i,u}, i \in [n], u \leq t : \forall i \in [n], \forall u \leq t, \\ g_i(\{J_{i,u}\}) \leq \bar{p} + \frac{1}{\log n}) \Big\}.$$

Lemma 2. $\mathbb{P}(A_J^t) \geq 1 - \exp\big(-c_4 n^{b_4}\big)$ *for* $t \leq \exp\big(c_4 n^{b_4}\big)$, *where* $b_4 = \epsilon_g - \gamma$ *and* $c_4 = \Omega(1)$.

Theorem 3. *For the effective error at time-step t, given by E_t^e, for any $X \subset [n]$,*

$$\mathbb{P}(\exists X \in \mathcal{C}_s(v) : |X \cap E_t^e| > \alpha|X| \;\big|\; A_J^t) \\ \leq \mathbb{P}(\exists X \in \mathcal{C}_s(v) : |X \cap E_t| > \alpha|X| \;\big|\; A_J^t) + t\exp\big(-C_1\sqrt{n}\big),$$

for $\alpha \ln \frac{1}{\bar{p}} \geq \frac{101}{100} C_{\mathcal{G}}$ *and* $C_1 = \frac{c' C_{\mathcal{G}}}{100}$. □

Proof of Theorem 2. Here we prove Theorem 2 using the quantities defined above and Lemma 1 and Theorem 3.

Starting from the first equation from page 13 of [11], we have

$$\mathbb{P}\Big[MaxConn_\alpha(E_t^e) > \theta\Big] \leq \sum_{s \geq \theta} \sum_{v \in V} \mathbb{P}\Big[A(E_t^e, \alpha, s, v) \neq \varnothing\Big]$$

Considering only the summand:

$$\mathbb{P}\Big[A(E_t^e, \alpha, s, v) \neq \varnothing\Big] \\ = \mathbb{P}\Big[\exists X \in \mathcal{C}_s(v) : |X \cap E_t^e| > \alpha|X|, \partial X \cap E_t^e = \varnothing\Big] \\ \leq \mathbb{P}\Big[\exists X \in \mathcal{C}_s(v) : |X \cap E_t^e| > \alpha|X|\Big] \\ \leq \mathbb{P}\Big[\exists X \in \mathcal{C}_s(v) : |X \cap E_t^e| > \alpha|X| \Big| A_J^t\Big] + (1 - \mathbb{P}(A_J^t))$$

The last step comes from the fact that for any random variables A and B where B is binary, $P(A) = P(B)P(A|B) + (1 - P(B))P(A|\bar{B})$. Omitting the $P(B)$ and $P(A|\bar{B})$ terms gives us the required inequality, since probabilities are upper bounded by 1.

Both the summed quantities in the last expression can now be bounded individually. By Lemma 2, we get: $1 - \mathbb{P}(A_J^t) \leq \exp\big(-c_4 n^{b_4}\big)$.

Now,

$$\mathbb{P}\Big[\exists X \in \mathcal{C}_s(v) : |X \cap E_t^e| > \alpha|X| \;\Big|\; A_J^t\Big] \\ \leq \mathbb{P}\Big[\exists X \in \mathcal{C}_s(v) : |X \cap E_t| > \alpha|X| \;\Big|\; A_J^t\Big] + t\exp\big(-C_1\sqrt{n}\big) \\ \leq \sum_{X \in \mathcal{C}_s(v)} \mathbb{P}\Big[|X \cap E_t| > \alpha|X| \;\Big|\; A_J^t\Big] + t\exp\big(-C_1\sqrt{n}\big),$$

by invoking Theorem 3 and union bound. The first term can further be bounded by

$$|\mathcal{C}_s(v)| \mathbb{P}\Big[|X \cap E_t| > \alpha|X| \;\Big|\; A_J^t\Big]$$

Given any $J_{i,k} \in A_J^t$ for $k \leq t$, $\{L_{i,k}\}$ are independent Bernoulli. Their probabilities being 1 are upper-bounded by $\bar{p} + \frac{1}{\log n}$. By invoking simple Bernoulli coupling, the probability that $(|X \cap E_k| > \alpha|X|)$ conditioned on $J_{i,k} \in A_J^t$ is upper-bounded by the probability of the same event when $\{L_{i,k}\}$ are i.i.d. Bernoulli$(\bar{p} + \frac{1}{\log n})$. Thus, by the Chernoff bound [23], the probability of $|X \cap E_k| > \alpha|X|$ conditioned on $J_{i,k} \in A_J^t$ is upper-bounded by $\exp\Big(-|X| \text{ KL}(\alpha||\bar{p} + \frac{1}{\log n})\Big)$. Here, $\text{KL}(a, b)$ for $a, b \in (0, 1)$ is the Kullback-Leibler divergence between two Bernoulli distributions a and b.

Thus, $\mathbb{P}\Big[|X \cap E_k| > \alpha|X| \;\Big|\; A_J^t\Big]$ is upper-bounded by $\exp\Big(-|X| \text{ KL}(\alpha||\bar{p} + \frac{1}{\log n})\Big)$. Note that for a fixed a and b close to 0, the dominant term in $\text{KL}(a, b)$ is $a \ln \frac{1}{b}$.

By Lemma 1, we have $|\mathcal{C}_s(v)| \leq \exp(s\, C_{\mathcal{G}})$ for $s = \omega(\log n)$, where $C_{\mathcal{G}}$ is a constant that depends on the degree of $\mathcal{G}$ only. Hence, for $\bar{p}$ small enough such that $\alpha \ln \frac{1}{\bar{p}} \geq \frac{101}{100} C_{\mathcal{G}}$, $\mathbb{P}\Big[A(E_t^e, \alpha, s, v) \neq \varnothing \;\big|\; A_J^t\Big]$ vanishes exponentially as $\exp\Big(-\frac{C_{\mathcal{G}}}{100} s\Big)$.Hence,

$$\mathbb{P}(MaxConn_\alpha(E_t^e) > c'\sqrt{n})) \\ \leq \sum_{s \geq \theta} \sum_{v \in V} \mathbb{P}\Big[A(E_t^e, \alpha, s, v) \neq \varnothing\Big] \\ \leq n \cdot n \cdot \Big(t \exp\Big(-\frac{c' C_{\mathcal{G}}}{100}\sqrt{n}\Big) + \exp\big(-c_4 n^{b_4}\big)\Big)$$

■

Proof of Theorem 3. Given A_J^t, let us repeat some steps of the proof in Theorem 2 for $t = 1$ and get the bound on $\mathbb{P}\Big[\exists X \in \mathcal{C}_s(v) : |X \cap E_1^e| > \alpha|X| \;\Big|\; A_J^t\Big]$ for $s = c'\sqrt{n}$ as $\exp(-C_1\sqrt{n})$ for $C_1 = \frac{c' C_{\mathcal{G}}}{100}$. Thus, with probability $1 - \exp(-C_1\sqrt{n})$, $E_2^e = E_2$ since all errors are corrected at $t = 1$ if $MaxConn_\alpha(E_1) \leq c'\sqrt{n}$.

Repeat the same procedure at $t = 2$ and so on. This, by union bound, implies that till t, with probability $1 - t \cdot \exp(-C_1\sqrt{n})$, we have $E_t^e = E_t$ given A_J^t. ■

Proof of Lemma 2. We first prove that $\mathbb{P}(g_i(\{J_{i,u}\}) \leq (1+\frac{1}{\log n})) \geq 1-\exp\left(-c_4 n^{b_4}\right)$. Then, the lemma follows using the union bound and the fact that $n\exp\left(-c_4 n^{b_4}\right) = \exp\left(-(c_4 - 1/\log n)\ n^{b_4}\right) \approx \exp\left(-c_4 n^{b_4}\right)$ for all sufficiently large n.

We use the result from Theorem 1 in [27] to bound $\mathbb{P}\left(g_i(\{J_{i,u}\}) \leq \bar{p} + \frac{1}{\log n}\right)$, where the definition of elements of the coupling matrix D is:

$$D_{i,j} = \max_{a,b}\ \mathbb{P}^J_{i,a,b}\left(J_j^{(1)} \neq J_j^{(2)}\right)$$

where $\mathbb{P}^J_{i,a,b}$ is the maximal coupling of the conditional laws $(J_{>i}|J_{<i}, J_i = a)$ and $(J_{>i}|J_{<i}, J_i = b)$. As $\{J_i\}$ is 1D Markov, the conditional laws become $(J_{>i}|J_i = a)$ and $(J_{>i}|J_i = b)$. Using the coupling definition of total variation distance (denoted by TV below) and the fact that $J_{i,t}$ are spatially Markov at any t (not temporally Markov) we get:

$$D_{i,j} = \sup_{a,b}\ \|\mathbb{P}(J_j|J_i = a) - \mathbb{P}(J_j|J_i = b)\|_{TV}.$$

This quantity has been defined before as correlation in the model and its equivalence with $\mathrm{cov}(L_i, L_j)$ was discussed. The rest follows from the following lemma and by taking a union bound over time. ■

Lemma 3. *Assuming $D_{i,j}$ decays no slower than $\frac{1}{dist(i,j)}$ and g_i is c_n-Lipschitz, where c_n is $O\left(\frac{c}{n^{0.5+\epsilon_g}}\right)$, we have*

$$\mathbb{P}\Big(g_i(\{J_{i,u}\}) - \mathbf{E}[g_i(\{J_i\})] \geq \frac{1}{\log n}\Big) \leq \exp\left(-\frac{1}{c} n^{(2-\delta_g)\epsilon_g}\right)$$

for any $\delta_g > 0$.

The Proof for the above lemma can be found in the appendix.

Conclusion

In this paper, we proved a positive noise threshold for square-root distance expander qLDPC codes against long-range correlated errors. This is the first step towards answering the question: can constant overhead sublinear distance codes offer fault-tolerance against highly correlated errors? The analytical noise threshold has a relatively simple expression in terms of the code rate when the code rate is small. This offers a simple tradeoff between the space overhead for these codes and the achievable noise threshold. The important practical question of local syndrome extraction for qLDPC codes was recently answered affirmatively [14, 15]. A future quest of practical significance would thus be to understand the performance of qLDPC codes with local syndrome extraction against long-range correlated errors.

Acknowledgments

AC thanks ANRF India (formerly, SERB India) for support through grant CRG/2023/005345 and the Ministry of Education, India.

[1] B. M. Terhal and G. Burkard, Fault-tolerant quantum computation for local non-markovian noise, Physical Review A—Atomic, Molecular, and Optical Physics **71**, 012336 (2005).

[2] P. Aliferis, D. Gottesman, and J. Preskill, Quantum accuracy threshold for concatenated distance-3 codes, arXiv preprint quant-ph/0504218 (2005).

[3] D. Aharonov, A. Kitaev, and J. Preskill, Fault-tolerant quantum computation with long-range correlated noise, Physical review letters **96**, 050504 (2006).

[4] S. Bagewadi and A. Chatterjee, Effect of correlated errors on quantum memory, Phys. Rev. A **112**, 022422 (2025).

[5] H. Kunsch, S. Geman, and A. Kehagias, Hidden markov random fields, The annals of applied probability **5**, 577 (1995).

[6] A. Leverrier and G. Zémor, Quantum tanner codes, in *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, 2022) pp. 872–883.

[7] J.-P. Tillich and G. Zémor, Quantum ldpc codes with positive rate and minimum distance proportional to the square root of the blocklength, IEEE Transactions on Information Theory **60**, 1193 (2013).

[8] A. A. Kovalev and L. P. Pryadko, Fault tolerance of quantum low-density parity check codes with sublinear distance scaling, Physical Review A—Atomic, Molecular, and Optical Physics **87**, 020304 (2013).

[9] D. Gottesman, Fault-tolerant quantum computation with constant overhead, Quantum Information & Computation **14**, 1338 (2014).

[10] A. Leverrier, J.-P. Tillich, and G. Zémor, Quantum expander codes, in *2015 IEEE 56th Annual Symposium on Foundations of Computer Science* (IEEE, 2015) pp. 810–824.

[11] O. Fawzi, A. Grospellier, and A. Leverrier, Efficient decoding of random errors for quantum expander codes, in *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing* (ACM, 2018) p. 521–534.

[12] A. Grospellier, *Constant time decoding of quantum expander codes and application to fault-tolerant quantum computation*, Theses, Sorbonne Université (2019).

[13] S. Bravyi, A. W. Cross, J. M. Gambetta, D. Maslov, P. Rall, and T. J. Yoder, High-threshold and low-overhead fault-tolerant quantum memory, Nature **627**, 778 (2024).

[14] N. Berthusen, D. Devulapalli, E. Schoute, A. M. Childs, M. J. Gullans, A. V. Gorshkov, and D. Gottesman, Toward a 2d local implementation of quantum low-density parity-check codes, PRX Quantum **6**, 010306 (2025).
[15] C. A. Pattison, A. Krishna, and J. Preskill, Hierarchical memories: Simulating quantum ldpc codes with local gates, Quantum **9**, 1728 (2025).
[16] S. Gu, C. A. Pattison, and E. Tang, An efficient decoder for a linear distance quantum ldpc code, in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023 (Association for Computing Machinery, New York, NY, USA, 2023) p. 919–932.
[17] D. Aharonov and M. Ben-Or, Fault-tolerant quantum computation with constant error, in *Proceedings of the 29th Annual ACM Symposium on Theory of Computing (STOC)* (1997) pp. 176–188.
[18] D. AHARONOV and M. BEN-OR, Fault-tolerant quantum computation with constant error rate, SIAM journal on computing **38**, 1207 (2008).
[19] O. Fawzi, A. Grospellier, and A. Leverrier, Constant overhead quantum fault tolerance with quantum expander codes, Communications of the ACM **64**, 106 (2020).
[20] H. Yamasaki and M. Koashi, Time-efficient constant-space-overhead fault-tolerant quantum computation, Nature Physics **20**, 247 (2024).
[21] M. McEwen, L. Faoro, K. Arya, A. Dunsworth, T. Huang, S. Kim, B. Burkett, A. Fowler, F. Arute, J. C. Bardin, *et al.*, Resolving catastrophic error bursts from cosmic rays in large arrays of superconducting qubits, Nature Physics **18**, 107 (2022).
[22] G. Inc, Suppressing quantum errors by scaling a surface code logical qubit, Nature **614**, 676 (2023).
[23] S. Boucheron, G. Lugosi, and P. Massart, *Concentration inequalities: A nonasymptotic theory of independence* (Oxford university press, 2013).
[24] B. M. Terhal, Quantum error correction for quantum memories, Reviews of Modern Physics **87**, 307 (2015).
[25] M. A. Nielsen and I. Chuang, Quantum computation and quantum information (2002).
[26] S. Tamiya, M. Koashi, and H. Yamasaki, Polylog-time-and constant-space-overhead fault-tolerant quantum computation with quantum low-density parity-check codes, arXiv preprint arXiv:2411.03683 (2024).
[27] J.-R. Chazottes, P. Collet, C. Külske, and F. Redig, Concentration inequalities for random fields via coupling, Probability Theory and Related Fields **137**, 201 (2007).

Appendix A: Proof of Lemma 1

Note that for each v, $|\mathcal{C}_s(v)| \leq |\mathcal{C}_s(\mathcal{G})|$, and from [11] we have $|\mathcal{C}_s(\mathcal{G})| \leq n\left((d_{\mathcal{G}}-1)(1+\frac{1}{d_{\mathcal{G}}-2})^{d_{\mathcal{G}}-2}\right)^s$. Hence, $|\mathcal{C}_s(v)| \leq e^{C_{\mathcal{G}}\, s}$ for $s = \omega(\log n)$,
where $C_{\mathcal{G}} = (d_{\mathcal{G}}-1)(1+\frac{1}{d_{\mathcal{G}}-2})^{d_{\mathcal{G}}-2}$.

Appendix B: Proof of Lemma 3

From the work of Chazottes et al. [27], we get

$$\mathbb{P}\left\{g - \mathbb{E}g \geq t\right\} \leq \exp\left(-\frac{2t^2}{\|D\|^2_{\ell^2(\mathbb{N})}\,\|\delta g\|^2_{\ell^2(\mathbb{N})}}\right),$$

where D is the matrix with elements $D_{i,j}$ and δg is an n-dimensional vector as defined in [27]. It follows from [27] that if g is c_n-Lipschitz, then each component of δg is at most c_n.

Consider $V(J_i) = \Theta(n^\gamma)$ for $\gamma > 0$. Note that the row-sum of D is upper-bounded by $V(J_i) = \frac{n^\gamma}{\gamma}$. Hence, upon scaling by that factor, the matrix would be a sub-stochastic matrix whose eigenvalues are no greater than 1. The following steps then hold:

$$\begin{aligned}
&\mathbb{P}\left(g_i(\{J_{i,u}\}) \leq \bar{p} + \frac{1}{\log n}\right)\\
=&\mathbb{P}\left(g_i(\{J_{i,u}\}) - \mathbb{E}[g_i(\{J_{i,u}\}] \leq \frac{1}{\log n}\right)\\
\leq&\exp\left(-\frac{2\left(\frac{1}{\log n}\right)^2}{n^{2\gamma}c_n^2 n}\right) = \exp\left(-\frac{2\left(\frac{1}{\log n}\right)^2}{n^{2\gamma}\frac{c^2}{n^{1+2\epsilon_g}}n}\right)\\
=&\exp\left(-c\frac{n^{2\epsilon_g}}{(n^\gamma)^2(\log n)^2}\right)\\
\leq&\exp\Big(-cn^{(2-\delta_g)(\epsilon_g-\gamma)}\Big) \approx \exp\Big(-cn^{2(\epsilon_g-\gamma)}\Big)
\end{aligned}$$

The rest follows from the fact that polylogarithmic terms are smaller than any polynomial of arbitrarily small positive degree.
In the other case, the row-sum of D is $O(1)$, and hence the bound also follows.