Numerical and analytical bounds on threshold error rates for hypergraph-product codes

Alexey A. Kovalev, Sanjay Prabhakar, Ilya Dumer, and Leonid P. Pryadko

Numerical and analytical bounds on threshold error rates for hypergraph-product codes

Alexey A. Kovalev and Sanjay Prabhakar
Department of Physics and Astronomy and Nebraska Center for Materials and Nanoscience,
University of Nebraska, Lincoln, Nebraska 68588, USA

Ilya Dumer
Department of Electrical Engineering, University of California, Riverside, California 92521, USA

Leonid P. Pryadko
Department of Physics & Astronomy, University of California, Riverside, California 92521, USA
(Dated: May 21, 2018)

We study analytically and numerically decoding properties of finite rate hypergraph-product quantum LDPC codes obtained from random $(3, 4)$-regular Gallager codes, with a simple model of independent X and Z errors. Several non-trival lower and upper bounds for the decodable region are constructed analytically by analyzing the properties of the homological difference, equal minus the logarithm of the maximum-likelihood decoding probability for a given syndrome. Numerical results include an upper bound for the decodable region from specific heat calculations in associated Ising models, and a minimum weight decoding threshold of approximately 7%.

I. INTRODUCTION

Coherence protection is one of the key technologies required for scalable quantum computation. Quantum error correction is one such technique. It enables scalable quantum computation with a polylogarithmic overhead per logical qubit as long as the accuracy of elementary gates and measurements exceeds certain threshold[1–4]. For a given family of quantum error correcting codes (QECCs), the actual threshold value depends, e.g., on hardware architecture, implementation of the elementary gates, and on the algorithm used for syndrome-based decoding. While these details are ultimately very important, for the purposes of comparing different families of QECCs, one is also interested in the threshold(s) computed in simple "channel" models where errors on different qubits are assumed independent and identically distributed (i.i.d.), with the assumption of perfect syndrome measurement. The resulting threshold is a single number which depends on the chosen algorithm for syndrome-based decoding. Among any decoders, the threshold is maximal for the (exponentially expensive) maximum-likelihood (ML) decoder.

While the original version of the threshold theorem was based on concatenated codes[1, 2, 5, 6], much better thresholds are obtained with topological surface[7–9] and related topological color codes[10, 11]. Even though their thresholds are higher, all surface codes, and, generally, all codes local in D dimensions, are necessarily zero-rate codes[12, 13]. Just like codes obtained by repeated concatenation, such local codes also require an overhead (per logical qubit) that is increasing with the length of the code, so that the overall overhead must increase as the size of the computation grows.

Scalable quantum computation with a *constant* overhead can be potentially achieved[14] using more general quantum LDPC (low density parity-check) codes. These are stabilizer codes, with the property that each stabilizer generator involves a bounded number of qubits. Here, a non-zero fault-tolerant threshold is guaranteed if the distance scales logarithmically or faster with the block length[15, 16]. At the same time, such codes may achieve a finite rate only if their generators remain non-local whenever the qubits are laid out in a Euclidean space of finite dimension[12, 13]. Several infinite code families are known to achieve these requirements[17–21].

The ML decoding probability for a quantum LDPC code can be formally expressed as the average of a ratio of partition functions for two associated random-bond Ising models (RBIM) [7, 22], computed on the Nishimori line[23, 24] in the (p, T) plane, where p is the error probability and T is the temperature. A temperature not on the Nishimori line corresponds to a suboptimal decoder which assumes an incorrect error probability. For topological codes local in D dimensions, the decodable region is a subset of (and possibly coincides with) the thermodynamic phase of RBIM where certain extended topological defects have finite tension[7, 22, 25–30]. For finite-rate codes, decodability requires that the average defect tension be sufficiently large[30].

In this work we analyze error-correcting properties of the finite-rate family of hypergraph-product codes[17] based on random $(3, 4)$-regular Gallager ensemble of classical LDPC codes, in conjunction with the phase diagrams of the two mutually dual associated RBIMs, constructed assuming independent X and Z errors which happen independently with probability p at each qubit (see Fig. 1). More specifically, we use a large-distance subset of the random Gallager codes. Each constructed code is a CSS code[31, 32] with parameters $[[n, k, d]]$.

Here the asymptotic rate $R = k/n = 1/25$, and the distance d scales as a square root of the code length n. The first corresponding Ising model has n interaction terms ("bonds") and $r \leq 12n/25$ spins; each bond is a product of 3 or 4 spin variables, and each spin participates in up to seven bonds. These numbers are higher for the second (dual) model which includes a summation over additional spin variables corresponding to the codewords.
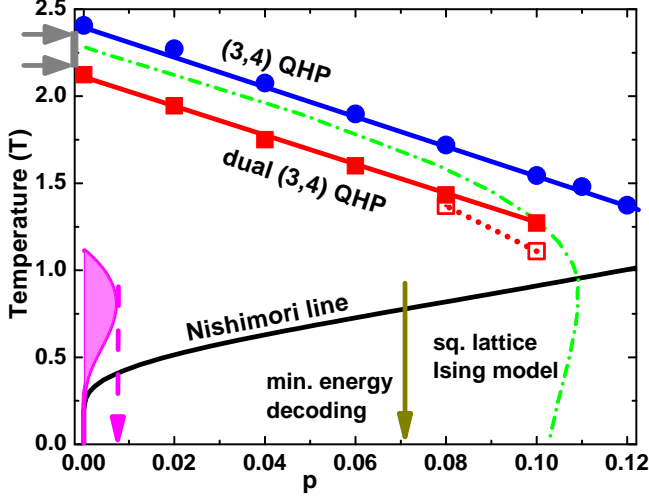

FIG. 1. (Color online) The (p, T) phase diagram of the two random-bond Ising models associated with QHP codes from $(3, 4)$ Gallager ensemble, labeled "$(3, 4)$ QHP" and "dual $(3, 4)$ QHP", where the latter model includes a summation over additional spin variables corresponding to the codewords, see Eq. (5). Here p is the error probability and T the dimensionless temperature. Positions of the specific heat maxima at different p, extrapolated to infinite system size, are shown with solid blue circles and solid red boxes with ad hoc linear fits [red open boxes connected with a dotted line correspond to parabolic extrapolation in Fig. 5(b)]. The two transition lines intersect the $p = 0$ axis in approximately mutually dual temperatures, and are located, respectively, above and below the critical temperature for the square-lattice random-bond Ising model (dashed-dotted green line; data from Ref. 33). The ML decoding problem corresponds to the points on the Nishimori line shown with a solid black line; temperatures away from the Nishimori line correspond to suboptimal decoding. The magenta-shaded region shows the lower bound for the decodable region from Theorem 2; the right-most point of this region coincides with the lower bound obtained by analyzing minimum-energy decoder in Ref. 16 (dashed magenta downward arrow). The forest-green solid vertical arrow shows our numerical estimate for the minimum weight decoding threshold, see Sec. IV B. The temperatures $T_{\max}$ and its dual, $T_{\max}^*$ from Theorem 3 are marked with a pair of horizontal arrows separated by a gray bar on the vertical axis; the lower arrow corresponds to the analytical upper temperature bound of the decodable region. More accurate upper bound from the present data is given by the "dual $(3, 4)$ QHP" line.

Analytically, we construct a lower bound for the decodable region, and study the relation between the decodability and thermodynamical phases of the corresponding Ising model. In particular, this produces a non-trivial upper temperature bound for the decodable region. Numerically, we use Metropolis updates in canonical ensemble simulations and in feedback-optimized parallel tempering Monte Carlo method to compute average specific heat as a function of temperature and the flipped bond probability p; extrapolation to infinite code distance gives the transition temperatures in the two models. We give an argument that it is the transition temperature of the dual model that gives a more accurate estimate of the upper temperature bound of the decodable region. We also use a decoder approximating the minimum-energy decoder to obtain a lower bound for the ML decoding threshold, with the result $p_c \geq p_{\min E} = 7.0\%$.

The rest of the paper is organized as follows. In Sec. II, we give a brief overview of classical and quantum error correcting codes, the Ising models related to ML decoding, and quantum hypergraph-product codes. We give the analytical bounds for the decodable region in Sec. III, with the proofs given in the Appendices. The numerical techniques and the corresponding results are presented in Sec. IV. We summarize our results and give some concluding remarks in Sec. V.

II. BACKGROUND

A. Classical and quantum error correcting codes

A classical binary linear code $\mathcal{C}$ with parameters $[n, k, d]$ is a k-dimensional subspace of the vector space $\mathbb{F}_2^n$ of all binary strings of length n. Code distance d is the minimal weight (number of non-zero elements) of a non-zero string in the code. A code $\mathcal{C} \equiv \mathcal{C}_G$ can be specified in terms of the generator matrix G whose rows are the basis vectors of the code. All vectors orthogonal to the rows of G form the dual code $\mathcal{C}_G^\perp = \{\mathbf{c} \in \mathbb{F}_2^n | G\mathbf{c}^T = \mathbf{0}\}$. The generator matrix P of the dual code, $\mathcal{C}_G^\perp \equiv \mathcal{C}_P$,

$$GP^T = 0, \quad \operatorname{rank} G + \operatorname{rank} P = n, \qquad (1)$$

is also called dual of G, $P = G^*$. It is the parity check matrix of the original code, $\mathcal{C}_G = \mathcal{C}_P^\perp$.

A quantum $[[n, k, d]]$ stabilizer code is a 2^k-dimensional subspace of the n-qubit Hilbert space $\mathbb{H}_2^{\otimes n}$, a common $+1$ eigenspace of all operators in an Abelian *stabilizer group* $\mathscr{S} \subset \mathscr{P}_n$, $-\mathbb{1} \notin \mathscr{S}$, where the n-qubit Pauli group $\mathscr{P}_n$ is generated by tensor products of the X and Z single-qubit Pauli operators. The stabilizer is typically specified in terms of its generators, $\mathscr{S} = \langle S_1, \ldots, S_{n-k} \rangle$. The weight of a Pauli operator is the number of qubits that it affects. The distance d of a quantum code is the minimum weight of an operator U which commutes with all operators from the stabilizer $\mathscr{S}$, but is not a part of the stabilizer, $U \notin \mathscr{S}$. Such operators correspond to the logical qubits and are called logical operators. A Pauli operator $U \equiv i^m X^{\mathbf{v}} Z^{\mathbf{u}}$, where $\mathbf{v}, \mathbf{u} \in \{0, 1\}^{\otimes n}$ and $X^{\mathbf{v}} = X_1^{v_1} X_2^{v_2} \ldots X_n^{v_n}$, $Z^{\mathbf{u}} = Z_1^{u_1} Z_2^{u_2} \ldots Z_n^{u_n}$, can be

mapped, up to a phase, to a binary vector $\mathbf{e} = (\mathbf{u}, \mathbf{v})$. With this map, generators of the stabilizer group are mapped to rows of a generator matrix $\mathcal{G} = (G^x, G^z)$ forming a binary classical linear code[34]. We will also consider the matrix $\mathcal{L}$ obtained in a similar fashion from independent logical operators.

For a more narrow set of CSS codes[31, 32] the stabilizer generators can be chosen as products of either only X or only Z Pauli operators. The corresponding generator matrix is a direct sum, $\mathcal{G} = G_x \oplus G_z$, where rows of the matrices $G \equiv G_x$ and $H \equiv G_z$ are orthogonal, $G_x G_z^T = 0$. For any CSS code, independent logical operators can also be chosen as products of only X or only Z Pauli operators, which gives $\mathcal{L} = L_x \oplus L_z$. Rows of the matrix L_x are orthogonal to rows of G_z, $G_z L_x^T = 0$, and they are linearly independent from rows of G_x. Similarly, rows of the matrix L_z are orthogonal to rows of G_x, and they are linearly independent from rows of G_z. For a CSS code of block length n, these matrices have n columns, and the number of encoded qubits is

$$k = \operatorname{rank} L_x = \operatorname{rank} L_z = n - \operatorname{rank} G_x - \operatorname{rank} G_z. \quad (2)$$

Rows of L_x and L_z, respectively, have weights that are bounded from below in terms of the corresponding CSS distances,

$$d_x \equiv \min_{\mathbf{c} \in \mathcal{C}_{G_z}^\perp \setminus \mathcal{C}_{G_x}} \operatorname{wgt}(\mathbf{c}), \quad d_z \equiv \min_{\mathbf{b} \in \mathcal{C}_{G_x}^\perp \setminus \mathcal{C}_{G_z}} \operatorname{wgt}(\mathbf{b}). \quad (3)$$

The code distance is just $d \equiv \min(d_{G_x}, d_{G_z})$.

In what follows, we concentrate on CSS codes. It will be convenient to assume that matrices L_x and L_z have full row rank (each has exactly k rows), and specifically define the form of the dual matrix G_x^* [see Eq. (1)] as a combination of rows of matrices G_z and L_z, and, similarly, the dual matrix G_z^* as a combination of rows of G_x and L_x. Also, to simplify the notations, it will be convenient to drop the indices x and z and use the matrices $G \equiv G_x$ and $H \equiv G_z$. The corresponding CSS distances (3) will be denoted as $d_G \equiv d_x$ and $d_H \equiv d_z$.

B. Maximum likelihood decoding and random-bond Ising model

Consider a CSS code with generator matrices $G \equiv G_x$ and $H \equiv G_z$, and an error model where bit-flip and phase-flip errors happen independently with the same probability p. In such a case, decoding of X and Z errors can be done separately. In the following, we only consider X errors.

Generally, an X error can be described by a length-n binary vector $\mathbf{e}$; errors obtained by adding linear combinations of rows of G are mutually *degenerate* (equivalent), they act identically on the code. In the absence of measurement errors, one needs to figure out the degeneracy class of the error from the measured *syndrome* vector, $\mathbf{s}^T = H\mathbf{e}^T$. While it is easy to come up with a vector $\mathbf{e}_0$

that satisfies these equations, so do $2^k - 1$ vectors $\mathbf{e}_0 + \mathbf{c}$ obtained by adding inequivalent codewords $\mathbf{c} \in \mathcal{C}_H^\perp \setminus \mathcal{C}_G$. For the maximum-likelihood (ML) decoding, one compares the probabilities of errors in different degeneracy sectors (inequivalent $\mathbf{c}$), and chooses the most likely.

The probability of an error degenerate with $\mathbf{e}$ is obtained as a sum of probabilities of errors $\mathbf{e} + \boldsymbol{\alpha}G$ for different binary $\boldsymbol{\alpha}$. Such a sum can be readily seen[7] to be proportional to the partition function of RBIM in Wegner's form[35],

$$Z_{\mathbf{e}}(G; K_p) = \sum_{S_i = \pm 1} \prod_{b=1}^n e^{K_p(-1)^{e_b} R_b}, \quad (4)$$

where the interaction term for bond b, $R_b = \prod_j S_j^{G_{jb}}$, is defined by the column b of the matrix G, e_b is the corresponding bit in the vector $\mathbf{e}$, and the coupling constant $K_p \equiv 1/T_p$ is the inverse Nishimori temperature[23, 24], $e^{-2K_p} = p/(1-p)$.

Similarly, for a given codeword $\mathbf{c}$, the probability of an error degenerate with $\mathbf{e} + \mathbf{c}$ is proportional to $Z_{\mathbf{e}+\mathbf{c}}(G; K_p)$. Given the syndrome $\mathbf{s} = \mathbf{e}H^T$, the conditional probability that an error degenerate with $\mathbf{e}$ actually happened can be written as the ratio[30]

$$P(\mathbf{e}|\mathbf{s}) = \frac{Z_{\mathbf{e}}(G; K_p)}{\sum_{\mathbf{c}} Z_{\mathbf{e}+\mathbf{c}}(G; K_p)} = \frac{Z_{\mathbf{e}}(G; K_p)}{Z_{\mathbf{e}}(H^*; K_p)}, \quad (5)$$

where the sum in the denominator is proportional to the probability of the syndrome $\mathbf{s}$ to happen. In Eq. (5), H^* is a matrix dual of H, see Eq. (1); for correct normalization, H^* should be constructed from G by adding exactly k rows corresponding to mutually non-degenerate codewords $\mathbf{c} \in \mathcal{C}_H^\perp \setminus \mathcal{C}_G$. The conditional probability (5), with $\mathbf{e} = \mathbf{e}_{\max}(\mathbf{s})$ taken from the most likely degeneracy class for the syndrome $\mathbf{s}$, is the probability of successful ML decoding for the given syndrome. One can then calculate the average probability of successful ML decoding [30],

$$P_{\text{succ}}(G, H; K, p) = [P(\mathbf{e}|\mathbf{e}H^T)]_p, \quad K = K_p, \quad (6)$$

where $[\cdot]_p$ denotes the averaging over error vectors (each set bit $e_b = 1$ occurs independently with probability p).

Notice that if we take a temperature away from the Nishimori line, $T \neq T_p \equiv 1/K_p$, we are using a decoder with an incorrect p, which would result in suboptimal decoding[30]. For an infinite sequence of codes (G_t, H_t), $t \in \mathbb{N}$ with increasing distance, we define the *decodable* region on the p-T plane as such where

$$\lim_{t \to \infty} P_{\text{succ}}(G_t, H_t; K, p) = 1. \quad (7)$$

The overlap of the decodable region with the Nishimori line gives the threshold error rate p_c for ML decoding with the chosen sequence of codes. More generally, the extent of the decodable region away from the Nishimori line can be seen as a measure of the decoding robustness.

Generally, a code with the distance d can correct any $\lfloor (d-1)/2 \rfloor$ errors. If the errors on different (qu)bits happen independently with probability p, a typical error has

4

weight asymptotically close to pn; the existence of a decodable region is guaranteed only if the asymptotic relative distance $\delta = d/n$ is finite. Thus, in general, the decoding threshold satisfies $p_c \geq \delta/2$.

Existence of a finite threshold for (quantum and classical) LDPC codes with sublinear distance scaling where $\delta = 0$ has been established by two of us in Ref. 15. The basic reason for the existence of a threshold is that at small enough $p > 0$, likely error configurations can be decomposed into relatively small connected clusters on the (qu)bit connectivity graph. Specifically, two (qu)bits are considered connected if there is a check (a stabilizer generator) with the support including both positions. For an LDPC code (quantum or classical), the connectivity graph has a bounded degree. Then, formation of the connected error clusters is described by the site percolation process on the connectivity graph; it has a finite threshold[36] $p_{\text{perc}} \geq (\Delta - 1)^{-1}$ for any graph with the maximum degree Δ. Moreover, below this bound, the probability to encounter a large cluster decreases exponentially with the cluster size; this fact may be used to construct a syndrome-based decoder[15].

More accurate lower bounds for decoding thresholds in different error models (including phenomenological error model for syndrome measurement errors) are given in Ref. 16. Consider CSS codes whose generator matrices G_x and G_z have row weights not exceeding some fixed m, and distance scaling logarithmically or faster with n,

$$d \geq D \ln n. \qquad (8)$$

Assuming independent X and Z errors with equal probabilities $p = p_X = P_Z$, the corresponding lower bound reads

$$2[p(1-p)]^{1/2} \geq (m-1)^{-1} e^{-1/D}. \qquad (9)$$

With distance scaling like a power of n, $d \geq An^\alpha$ with $A, \alpha > 0$, one should use $D = \infty$. The bound (9) was obtained by analyzing a minimum energy decoder, which corresponds to $T = 0$.

C. Duality

As demonstrated by Wegner[35], a general Ising model with the partition function (4) has a dual representation, which is a generalization of Kramers-Wannier[37] duality. The same duality has been first established in coding theory by MacWilliams[38] as a relation between weight polynomials of two dual codes. It is convenient to introduce a generalized partition function,

$$Z_{\mathbf{e},\mathbf{m}}(G; K) \equiv \sum_{S_i=\pm 1} \prod_{b=1}^{n} R_b^{m_b} e^{K(-1)^{e_b} R_b}, \qquad (10)$$

that involves binary vectors of "electric" $\mathbf{e}$ and "magnetic" $\mathbf{m}$ charges. Then, the duality reads

$$Z_{\mathbf{e},\mathbf{m}}(G; K) = (-1)^{\mathbf{e}\cdot\mathbf{m}} Z_{\mathbf{m},\mathbf{e}}(G^*; K^*) A(K), \qquad (11)$$

where an $r^* \times n$ matrix G^* is the exact dual of G (dimensions $r \times n$), see Eq. (1), K^* is the Kramers-Wannier dual of K, $\tanh K^* = e^{-2K}$, and the scaling factor depends on the dimensions of the matrices,

$$A(K) = 2^{r-r^*+\text{rank}\,G^*} (\sinh K \cosh K)^{n/2}. \qquad (12)$$

Notice that the electric charges in Eq. (10) define the negative bonds as in Eq. (4), while the magnetic charges select the bonds to be used in an average,

$$\frac{Z_{\mathbf{e},\mathbf{m}}(G; K)}{Z_{\mathbf{e},\mathbf{0}}(G; K)} = \left\langle \prod_{b=1}^{n} R_b^{m_b} \right\rangle = \left\langle \prod_{b=1}^{n} \prod_{j=1}^{r} S_j^{G_{jb} m_b} \right\rangle, \qquad (13)$$

which is the most general form of a spin correlation function that is not identically zero[35].

D. Quantum hypergraph-product codes

In this work we specifically focus on the quantum hypergraph product (QHP) codes[17, 39], an infinite family of quantum CSS codes which includes finite-rate LDPC codes with distance scaling as a square root of the block length. A general QHP code is defined in terms of a pair of binary matrices $\mathcal{H}_1$ and $\mathcal{H}_2$ with dimensions $r_1 \times n_1$ and $r_2 \times n_2$. The corresponding stabilizer generators are formed by two blocks constructed as Kronecker products[40],

$$\begin{aligned} G_x &= (E_2 \otimes \mathcal{H}_1, \mathcal{H}_2 \otimes E_1), \\ G_z &= (\mathcal{H}_2^T \otimes \widetilde{E}_1, \widetilde{E}_2 \otimes \mathcal{H}_1^T), \end{aligned} \qquad (14)$$

where E_i and $\widetilde{E}_i$, $i = 1, 2$, are unit matrices of dimensions given by r_i and n_i; the matrices G_x and G_z have $r_1 r_2$ and $n_1 n_2$ rows, respectively. Clearly, the ansatz (14) guarantees that the rows of G_x and G_z are orthogonal, $G_x G_z^T = 0$. The block length of such a quantum code is the number of columns, $n \equiv r_2 n_1 + r_1 n_2$.

We are using the construction originally proposed in Ref. 17, namely, $\mathcal{H}_2 = \mathcal{H}_1^T$, where $\mathcal{H}_1$ is assumed to have a full row rank. If the binary code with the check matrix $\mathcal{H}_1$, $\mathcal{C}_{\mathcal{H}_1}^\perp$, has parameters $[n_1, k_1, d_1]$, the corresponding QHP code has the parameters[17, 39] $[[n, k, d]]$, where $n = n_1^2 + r_1^2$, $k = k_1^2$, $d = d_1$, and $r_1 = n_1 - k_1$.

We should mention that the QHP construction with $\mathcal{H}_2 = \mathcal{H}_1^T$ is weakly self-dual, meaning that the matrices G_x and G_z in Eq. (14) can be transformed into each other by row and column permutations. As a result, in particular, for any binary matrix $\mathcal{H}_1$, the decoding probabilities (6) in X and Z sectors must coincide, $P_{\text{succ}}(G_x, G_z; K, p) = P_{\text{succ}}(G_z, G_x; K, p)$.

A family of quantum LDPC codes with distance scaling as a square root of the block size can be obtained, e.g., by taking $\mathcal{H}_1$ from a random ensemble of classical LDPC codes, which are known to have finite rates k_1/n_1 and finite relative distances d_1/n_1, and removing any linearly-dependent rows. We specifically consider the

ensemble $\mathbb{B}(\ell, m)$ of regular (ℓ, m)-LDPC codes with column weight ℓ and row weight m originally introduced by Gallager[41, 42]. For each code in $\mathbb{B}(\ell, m)$, its parity-check matrix $\mathcal{H}$ of size $r \times n$ is divided into ℓ horizontal blocks $\mathcal{H}_1, \ldots, \mathcal{H}_\ell$ of size $\frac{r}{\ell} \times n$. Here the first block $\mathcal{H}_1$ consists of m unit matrices of size $\frac{r}{\ell} \times \frac{r}{\ell}$. Any other block $\mathcal{H}_i$ is obtained by some random permutation $\pi_i(n)$ of n columns of H_1. Thus, all columns in each block H_i have weight 1. This ensemble achieves the best asymptotic distance for a given designed code rate $1 - \ell/m$ among the LDPC ensembles studied to date[43]. In practice, it often happens that one or few rows of thus constructed $\mathcal{H}_1$ are linearly dependent, which gives a code with a larger rate, $R \equiv k/n \geq 1 - \ell/m$. It is easy to check, however, that asymptotic at $n \to \infty$ rate equals the designed rate, $R \to 1 - \ell/m$. For the ensemble $\mathbb{B}(3, 4)$ used in this work, the asymptotic relative distance $\delta = d/n$ is $\delta_{3,4} \approx 0.112$ [43].

For brevity, we will refer to QHP codes constructed using the matrices $\mathcal{H}_2 = \mathcal{H}_1^T$ from Gallager $\mathbb{B}(\ell, m)$ ensemble (with any linearly dependent rows dropped) as (ℓ, m) QHP codes.

III. ANALYTICAL BOUNDS

Partition function (4) scales exponentially with the system size; one more commonly works with the corresponding logarithm, the (dimensionless) free energy

$$F_\mathbf{e}(G; K) = -\ln Z_\mathbf{e}(G; K), \qquad (15)$$

which is an extensive quantity, meaning that it scales linearly with the system size. Alternatively, one can also use the free energy density (per bond), $f_\mathbf{e}(G; K) = F_\mathbf{e}(G; K)/n$, which usually has a well-defined thermodynamical limit. The logarithm of the ML decoding probability (5), up to a sign, equals the *homological difference*,

$$\Delta F_\mathbf{e}(G, H; K) \equiv F_\mathbf{e}(G; K) - F_\mathbf{e}(H^*; K). \qquad (16)$$

At $\mathbf{e} = 0$, this quantity satisfies the inequalities

$$0 \leq \Delta F_\mathbf{0}(G, H; K) \leq k \ln 2, \qquad (17)$$

where the lower and the upper bounds are saturated, respectively, in the limits of zero and infinite temperatures. Combining duality (11) with Griffiths-Kelly-Sherman[44, 45] (GKS) inequalities for spin averages we also obtain

$$\Delta F_\mathbf{e}(G, H; K) - \Delta F_\mathbf{0}(G, H; K) \geq 0. \qquad (18)$$

In addition, also at $\mathbf{e} = 0$, the duality (11) gives

$$\Delta F_\mathbf{0}(G, H; K) = k \ln 2 - \Delta F_\mathbf{0}(H, G; K^*), \qquad (19)$$

where $\tanh K^* = e^{-2K}$, and k is the dimension of the CSS code, see Eq. (2). The proof of these expressions is given in Appendix A.

The relation of the homological difference averaged over the disorder, $[\Delta F_\mathbf{e}]_p$, and the corresponding quantity normalized per unit bond, $[\Delta f_\mathbf{e}]_p \equiv [\Delta F_\mathbf{e}(G, H; K)]_p/n$, to decoding with asymptotic probability one, see Eq. (7), is given by the following Lemma (proved in App. B).

Lemma 1. *For a sequence of quantum CSS codes defined by pairs of matrices (G_t, H_t), $t \in \mathbb{N}$, where $G_t H_t^T = 0$, given a finite $K > 0$ and an error probability $p \geq 0$,*
(a) $\lim_{t \to \infty} [\Delta F_\mathbf{e}(G_t, H_t; K)]_p = 0$ implies the point (p, K) to be in the decodable region;
(b) $\liminf_{t \to \infty} [\Delta f_\mathbf{e}(G_t, H_t; K)]_p > 0$ implies the point (p, K) to be outside of the decodable region.

A. Lower bound for decodable region

Here, we use part (a) of Lemma 1 to establish an existence bound for the decodable region. Specifically, we construct an upper bound for $[\Delta F_\mathbf{e}(G, H; K)]_p \geq 0$ in a finite system, and use it to show the existence of a non-trivial region where $[\Delta F_\mathbf{e}]_p \to 0$, as long as the distance scales logarithmically or faster with the block length n, see Eq. (8). In Appendix C we prove:

Theorem 2. *Consider a sequence of quantum CSS codes $\mathcal{Q}(G_t, H_t)$, $t \in \mathbb{N}$, of increasing lengths n_t, where row weights of each G_t and H_t do not exceed a fixed m, and the code distances $d_t \geq D \ln n_t$, with some $D > 0$. Then the sequence $\Delta F_t \equiv [\Delta F_\mathbf{e}(G_t, H_t; K)]_p$, $t \in \mathbb{N}$, converges to zero in the region*

$$(m-1)[e^{-2K}(1-p) + e^{2K}p] < e^{-1/D}. \qquad (20)$$

The rightmost point of this region, the maximum value $p = p_\mathrm{bnd}$ where Eq. (20) has a solution, satisfies the equation $2(m-1)[p_\mathrm{bnd}(1 - p_\mathrm{bnd})]^{1/2} = e^{-1/D}$. The same bound was obtained previously in Ref. 16 using estimates based on minimum-energy decoding which corresponds to $T = 0$. Thus, present bound does not improve the existing lower bound for ML decoding threshold.

Further, the entire region (20) lies at temperatures $T = 1/K$ above the Nishimori line (see Fig. 1). In particular, at the right-most cusp of this region, the temperature $T_\mathrm{bnd} = 1/K_\mathrm{bnd}$ is exactly twice the Nishimori temperature at p_bnd. The importance of Theorem 2 is that we got a sense of the robustness of suboptimal decoding, where the ML decoder assumes a value of p larger than the actual one.

B. Upper temperature bound for decodable region

Here we combine part (b) of Lemma (1) with duality (11) to establish an upper temperature bound for the decodable region for a sequence of codes with asymptotic rate R. We first argue that existence of a low-temperature *homological* region where $\Delta f_\mathbf{0}(G, H; K) \to 0$, by duality, implies the existence of a high-temperature

dual homological region where $\Delta f_{\mathbf{0}}(H, G; K) \to R \ln 2$, and thus $[\Delta f_{\mathbf{e}}(H, G; K)]_p \geq R \ln 2$ at any $p \geq 0$. Further, the derivative of $f_{\mathbf{e}}$ with respect to K is just the energy per bond, with negative sign; its magnitude does not exceed one. Therefore, there should be some minimal distance between the upper temperature bound $K_0^{(1)}(G, H)$ of the homological region and the lower temperature bound $K_0^{(2)}(H, G)$ at $p = 0$ of the dual homological region. This gives an upper temperature bound for the homological region. By part (b) of Lemma (1), the same bound also works as an upper bound for the decodable region at any $p > 0$. These arguments give (see App. D):

Theorem 3. *Consider a sequence of CSS codes defined by pairs of finite binary matrices with mutually orthogonal rows, $G_t H_t^T = 0$, $t \in \mathbb{N}$, where row weights of G_t and H_t do not exceed a fixed m, the sequence of CSS distances $d_t = \max(d_{H_t}, d_{G_t})$ is strictly increasing with t, $d_{t+1} > d_t$, and the sequence of rates $R_t \equiv k_t/n_t$ converges, $\lim_{t \to \infty} R_t = R$. Then, assuming equal probabilities of X and Z errors, the upper temperature boundary of the decodable region, $T_{\max} = 1/K_{\max}$, satisfies the inequality*

$$K_{\max} - K_{\max}^* \geq R \ln 2. \tag{21}$$

Explicitly, this gives an upper temperature bound for the location of the ML-decodable region for any CSS code family with asymptotic rate R,

$$e^{2K_{\max}} \geq \frac{1 + r + \sqrt{(1+r)^2 + r}}{2}, \quad r \equiv 2^{2R} \geq 1. \tag{22}$$

In the case $R = 0$ this bound corresponds to the self-dual point, which equals to the upper bound of the decodable region of the square-lattice toric code (ferromagnetic phase of the square-lattice Ising model).

IV. NUMERICAL RESULTS

A. Justification

We first note that numerically, it is only possible to analyze systems of finite size. Numerical techniques used for predicting asymptotic large-size properties, such as finite-size scaling, are only good as long as such properties exist and change with the system size in a regular manner. For example, even though we know the existence of a non-zero decoding threshold, it is not a priori clear that the finite-size data would show a well defined crossing point, as seen on Fig. 2.

Similarly, a well-defined thermodynamical limit is known to exist for bulk quantities like magnetization or specific heat for Ising models on lattices that are local in D dimensions, simply because the corrections due to the boundary scale as the surface area, which scales as a sublinear power of the volume[46]. Well-defined infinite-size limit (although not necessarily universal) also exists

if one considers a sequence of models on increasing subgraphs of an infinite graph, where the boundary spins are either "free" (the couplings connecting them to outside are set to zero), or "wired" (the outside couplings are set to infinity). The existence of a thermodynamical limit in each of these cases follows from the GKS inequalities[44, 45], which require that spin correlations change monotonously with the system size (increase for wired and decrease for free boundary conditions).

The problem we are considering is different from either case, as a sequence of matrices G_t (or the corresponding bipartite graphs) defines a sequence of finite few-body Ising models without boundaries. Further, the finite asymptotic rate of the considered code family guarantees the absence[12, 13] of a D-dimensional layout of the qubits with local stabilizer generators, at any finite D. The only rigorous result, proved in a companion paper[47], is that a well defined limit for average free energy density for (ℓ, m) QHP codes exists for any p in a finite region around the infinite temperature and, by duality, for $p = 0$, in a finite region around the zero temperature. This follows from the absolute convergence of the corresponding high-temperature series (HTS) established using the bound on high-order cumulants[48], and from the fact that a large random bipartite graph with vertex degrees ℓ and m has few short cycles. The local Benjamini-Schramm limit[49] of such a graph is a bipartite tree, meaning that the asymptotic coefficients of the high-temperature series expansion to any finite order can be computed by analyzing only the clusters present when $\mathcal{H}_1$ in Eq. (14) corresponds to such a tree. The corresponding argument is a direct generalization of that in Refs. 50 and 51, where the existence of a well defined limit for free energy density was analyzed for general models with up to two-body interactions.

One consequence of this argument is that in the asymptotic limit, we do not expect much difference between the use of matrices $\mathcal{H}_1$ from the full Gallager $\mathbb{B}(\ell, m)$ ensemble, and the corresponding subset where for each size we pick only the matrices which result in the largest distance d_1 of the classical code $\mathcal{C}_{\mathcal{H}_1}^\perp$. On the other hand, we expect that the use of such matrices should significantly improve the convergence in the high- and low-temperature regions where the corresponding series converge: with larger distance, a larger number of coefficients of the series would match those for the infinite-size system.

Unfortunately, even though the corresponding series can be analytically continued beyond the convergence radius, this does not guarantee the existence of a well-defined limit for thermodynamical quantities at all temperatures, as would be required to formally justify the use of finite size scaling. Therefore, numerical results presented in the following sections represent numerical trends in systems of relatively small size; they do not necessarily guarantee the existence of well defined transition(s).
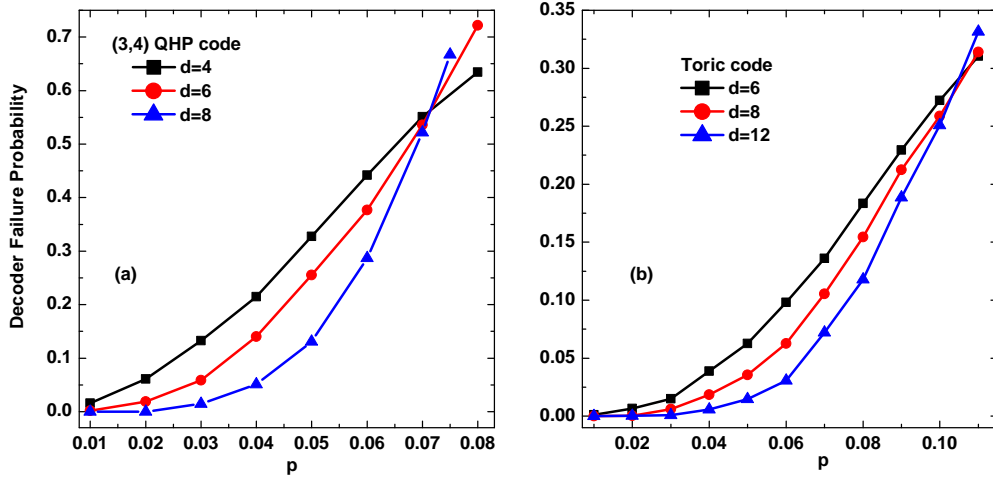
FIG. 2. (Color online) Decoder failure probability as a function of bit-flip error probability for three $(3,4)$ QHP codes (left) and the rotated toric codes (right). The decoding was performed over 1024 error realizations for $(3,4)$ QHP codes and over 4096 error realizations for toric codes. The corresponding decoding pseudothreshold is close to $p = 7.0\%$ for $(3,4)$ QHP codes and to 10.4% for toric codes.

B. Approximate minimum-weight decoding

To obtain an empirical lower bound for the ML decoding threshold, we constructed a cluster-based decoder using the approach suggested in Refs. [15, 16] (see Sec. II B). Specifically, given the syndrome vector $\mathbf{s}$, we construct a list of *irreducible clusters* up to the chosen cut-off weight w_1. Each irreducible cluster should correct some syndrome bits without introducing new ones, and it should not contain a subcluster with the same property. As explained in Sec. II B, with an LDPC code where the stabilizer weight is bounded, and for large enough w_1, we expect this list with high probability to include all clusters present in the connected-cluster decomposition of the actual error. The actual decoding is done by solving a minimum-weight set cover problem: among the subsets of the cluster list with the property that every non-zero syndrome bit be covered exactly once, we want to find such that the sum of the cluster weights be minimal. This latter problem is solved in two steps: first, by running the `LinearProgramming` over integers in Mathematica[52] to arrive at a valid solution with a reasonably small weight, and then by trying to minimize the weight further with the help of a precomputed list of non-trivial irreducible codewords[16]. In our calculations, for each disorder realization we generated irreducible clusters of weight up to $w_1 = 10$, and, for each code, the list of irreducible codewords of weight up to $w_2 = 19$.

Without the limits on the clusters' and codewords' weights, this procedure would be equivalent to minimum-weight decoding. Unfortunately, the corresponding complexity grows prohibitively (exponentially with the size of the code). Nevertheless, for smaller codes we were able to choose large enough w_1 and w_2 to estimate the minimum-weight decoding threshold, as seen from the

convergence of the corresponding decoding probabilities.

The decoding complexity is determined by the sum of those for the construction of the cluster list and for solving the weighted set cover problem. The construction of the cluster list was analyzed in detail in Refs. 53–55. In particular, if the maximum weight of a stabilizer generator is m, the corresponding complexity is $N_1 \sim n(m-1)^{w_1-1}$. At small enough p, the probability of a large cluster decays exponentially with its weight. Thus, in most cases, maximum cluster size scales logarithmically with the code length n, and a sufficient cluster list can be prepared with the cost polynomial in n.

On the other hand, the weighted set cover problem is NP-complete[56]; the corresponding cost is exponential in the length L of the cluster list. Generally, this problem is equivalent to an integer linear programming (LP) problem. To find a valid (but not necessarily the minimal) solution, we use a call to the built-in Mathematica function `LinearProgramming`. While the details of its implementation are proprietary, it is our understanding that an integer solution is found by first solving the corresponding problem over reals using an algorithm with polynomial complexity, and then finding the nearest integer point in the LP polytope. With rare exception (few instances over the entire set of our simulations where we had to record decoder failure due to calculation time-out), `LinearProgramming` returns a valid solution $\mathbf{e}$ which satisfies the constraints but does not necessarily have the smallest weight.

To reduce the weight further, we used a version of the approach used previously[16] to construct an analytical bound for minimum-energy decoding threshold. Notice that the minimum-weight (same as minimum-energy) solution $\mathbf{e_{min}}$ produces the same syndrome as $\mathbf{e}$, thus $\Delta \mathbf{e} = \mathbf{e} - \mathbf{e_{min}}$ produces the zero syndrome, and in gen-

eral can be decomposed into a sum of *irreducible code-words*[16], $\Delta \mathbf{e} = \mathbf{e}_1 + \ldots + \mathbf{e}_s$, such (a) that the supports of different $\mathbf{e}_j$ do not overlap, (b) each of $\mathbf{e}_j$ is a valid codeword, in the sense that it produces a zero syndrome, and (c) any $\mathbf{e}_j$ cannot be decomposed further into a sum of non-overlapping codewords. Such a decomposition is not necessarily unique. It is easy to see[16] that weight of $\mathbf{e} + \mathbf{e}_j$ for any $j \in \{1, 2, \ldots, s\}$ must not exceed that of $\mathbf{e}$. Thus, if we have a list of all non-trivial, $\mathbf{c}_j \not\simeq \mathbf{0}$, irreducible codewords, the equivalence class of the minimum-weight solution can be found by adding those $\mathbf{c}_j$ that reduce the weight of $\mathbf{e}$, until the weight can no longer be reduced.

Notice that with a complete list of non-trivial irreducible codewords, the degeneracy class of the minimum weight solution can be correctly identified from any vector $\mathbf{e}$ which produces the correct syndrome. In practice, since the weights of the irreducible codewords in our list are limited, the decoding success probability increases with the reduced weight of the initial vector $\mathbf{e}$.

Overall, for each code in our simulations, the majority of the computational time was spent on preparing the list of non-trivial irreducible codewords with weights $w \leq w_2 = 19$.

The results of the described threshold simulations are presented in Fig. 2 (a) and (b), respectively, for $(3, 4)$ QHP codes and for the toric codes. More precisely, in Fig. 2(a), we show the fraction of decoder failures for QHP codes constructed from three large-distance classical codes from Gallager $\mathbb{B}(3, 4)$ ensemble, for different values of bit flip probability p. The codes used have parameters $[[80, 16, 4]]$, $[[356, 36, 6]]$, and $[[832, 64, 8]]$; they were constructed from binary codes with parameters $[8, 4, 4]$, $[16, 6, 6]$, and $[24, 8, 8]$. Code $[[1921, 121, 10]]$ obtained from the binary code $[36, 11, 10]$ turned out too large for the present decoding technique; the corresponding data is not included in Fig. 2(a).

In our calculations, we used $w_1 = 10$ and $w_2 = 19$, which was sufficient for convergence of the average decoding probability for $p \leq 0.08$ used in the simulations. The well defined crossing point in Fig. 1(a) indicates a (pseudo)threshold for decoding of $(3, 4)$ QHP codes in the vicinity of 7.0%. Convergence of the average decoding probability with increasing w_1 and w_2 is an indication that this value is a good estimate of the minimum-weight decoding threshold.

For comparison, in Fig. 2(b), we show the corresponding results for the rotated toric codes[57] with the parameters $[[d^2, 2, d]]$, with $d = 6$, 8 and 12, where the crossing point is close to 10.4%, the minimum-weight decoder threshold obtained using the minimum-weight matching algorithm[7].

Notice that both for $(3, 4)$ QHPs and for the toric code, the obtained threshold estimates are much larger than the corresponding analytical lower bounds from Ref. 16, 0.70% and 2.8%, respectively.

C. Monte Carlo simulations and the phase diagram

In this section we analyze numerically the low-disorder portion of the phase diagram of the two random-bond Ising models corresponding to the ML decoding of $(3, 4)$ QHP codes with i.i.d. bit-flip errors. For a CSS code with generators $G = G_x$ and $H = G_z$, the corresponding Ising models have the free energies $F_\mathbf{e}(G; K)$ and $F_\mathbf{e}(H^*; K)$, see Eqs. (4) and (15), where $\mathbf{e}$ is the binary error vector whose non-zero bits indicate the flipped bonds, and $K = 1/T$ is the inverse temperature. These models, respectively, correspond to the numerator and the denominator of the conditional ML decoding probability (5).

The parameters of the four $(3, 4)$ QHP codes used in the simulations are described in the previous section. For simulation efficiency, we attempted to minimize the weights of the rows of the matrices H^*. To this end, starting with the matrix $G' = G$, we added one row at a time, corresponding to one of the minimum-weight vectors in $\mathcal{C}_H^\perp \setminus \mathcal{C}_{G'}$, where G' is the previously constructed matrix. As a result, the row weights of each matrix H^* did not exceed $\max(7, d_G)$.

To calculate the averages, we performed feedback optimized parallel tempering Monte Carlo simulations[58, 59], as well as the usual simulated annealing. In both cases we used standard Metropolis updates.

For both models, the observed scaling of the height of the specific heat maxima with n, and the hysteresis which we could not eliminate for larger codes, are consistent with the discontinuous transitions. We also observe that the use of the parallel tempering method does not improve the convergence significantly; we attribute this to the discontinuity of the phase transition.

Samples of the computed specific heat (per bond) for the $(3, 4)$ QHP models, $C(T) = (\langle E^2 \rangle - \langle E \rangle^2)/(nT^2)$, where E is the energy, n is the number of bonds, and T is the temperature, are shown in Fig. 3 (comparing different values of p separately for distances $d = 4$, 6, and 8). The specific heat values shown in Fig. 4 have been additionally divided by the number of bonds n; the corresponding maximum values are weakly increasing with the code distance for $p = 0$, see Fig. 4(a), and weakly decreasing for $p = 2\%$ and 10%, see Figs. 4(b) and 4(c). Such a slow dependence on the system size is consistent with a 1st order transition, where one expects $C(T_c) \propto n$. In Fig. 4(c) we also compare the data obtained using parallel tempering and the usual annealing. These data were obtained after 1×10^7 Monte Carlo sweeps for QHP codes of distance $d = 4$ and $d = 6$, and 5×10^7 Monte Carlo sweeps for codes of distances $d = 8$ and $d = 10$, for each of 128 (in some cases 256) realizations of disorder at every p.

When the positions of the specific heat maxima are plotted as a function of $1/d^2$ (asymptotically, $d^2 \propto n$, although such a relation does not hold for the small codes used in the simulations), the corresponding points are seated close to a straight line, see Fig. 5(a). Respectively, we used the linear fit $T_{\max}(d, p) = T_c(p) + A/d^2$ to
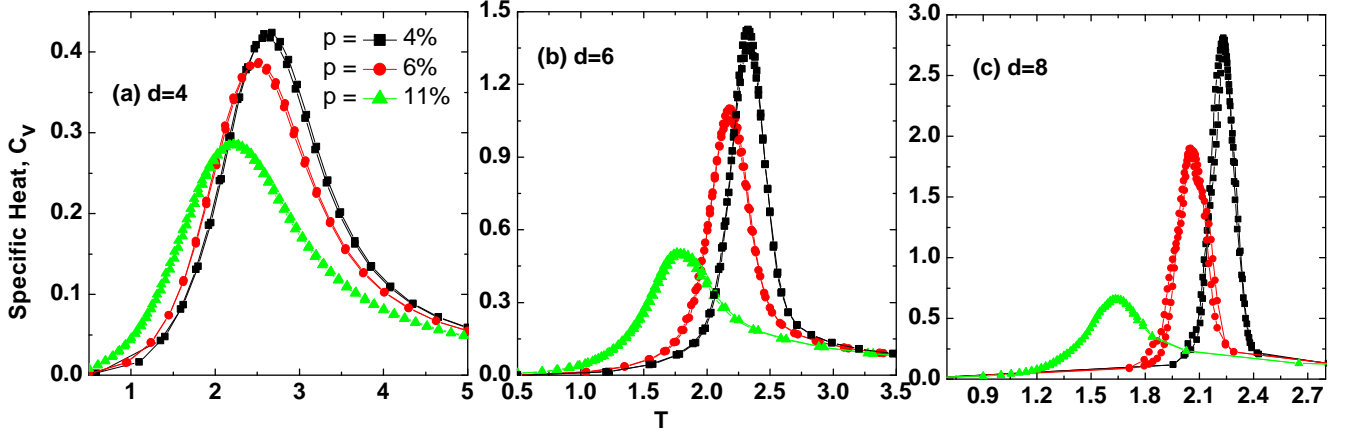
FIG. 3. (Color online) Specific heat C vs. dimensionless temperature T for $(3,4)$ QHP models with distances (a) $d = 4$, (b) $d = 6$, and (c) $d = 8$, at p values as indicated on panel (a). Each curve contains data points from the feedback optimized parallel tempering simulation where ordered and disordered configurations are used as initial states. The peak positions are extrapolated to infinite distance to obtain the transition temperatures, see Fig. 5.
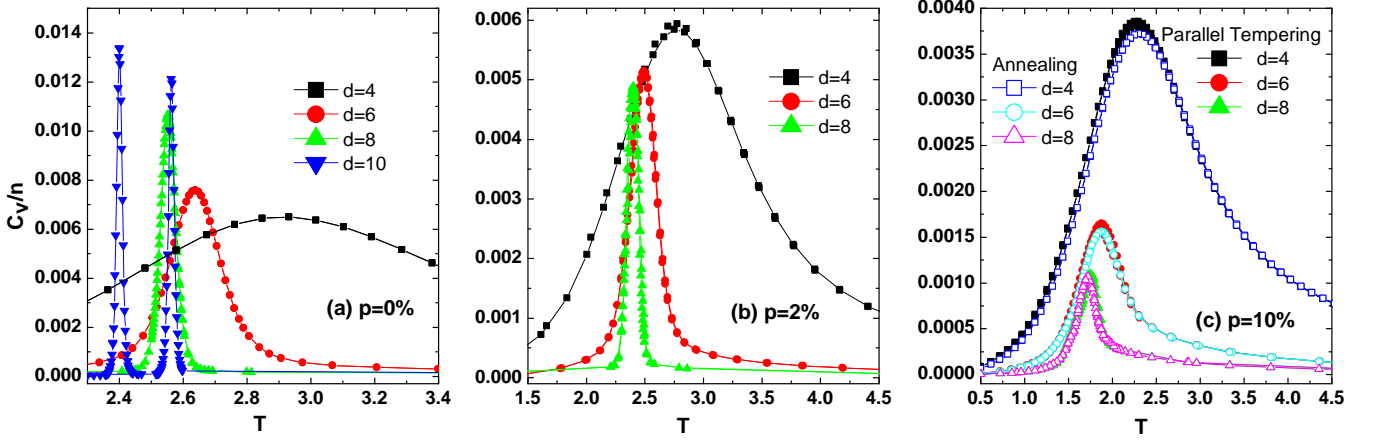

FIG. 4. (Color online) Specific heat C divided by the number of bonds n vs. dimensionless temperature T for $(3,4)$ QHP models at (a) $p = 0$, (b) $p = 2\%$, and (c) $p = 10\%$, with code distances as indicated in the captions. Each curve contains data points from the feedback optimized parallel tempering simulation. The annealing plot contains data points from upward and downward temperature sweeps. Relatively weak variation of the peak height with system size is indicative of a discontinuous transition. Open symbols in plot (c) show the data obtained with annealing, which agrees with the parallel tempering data (filled symbols).

extrapolate our finite-size data and extract more accurate critical point of the transition, $T_c(p)$, as a function of the flipped bond probability p. The resulting phase boundary is shown in Fig. 1 with solid blue circles, along with the solid blue line which is the linear fit to the data. The fit indicates that for $p \leq 0.12$, the phase transition temperature $T_c(p)$ is approximately linear in p. It is also clear from Fig. 1 that at every p, the phase boundary for this model is higher than the corresponding line for the square-lattice Ising model, plotted with dot-dashed line using the data from Ref. 33.

The analysis for the dual $(3,4)$ QHP models was performed similarly (specific heat data not shown). The positions of the specific heat maxima as a function of $1/d^2$

for different values of p are shown in Fig. 5(b), along with the corresponding linear fits. Notice that the points at $p = 10\%$ show significant curvature which cannot be attributed to the statistical errors alone. By this reason we also tried a parabolic fit, which resulted in a substantially lower extrapolated $T_c = 1.11$ compared with 1.27 ± 0.05 from the linear fit. In comparison, at $p = 8\%$, parabolic fit gives $T_c = 1.37$, which is not as significantly reduced compared to the linear fit result of 1.43 ± 0.02.

The extrapolated positions of the specific heat maxima are plotted in Fig. 1 with solid red boxes, along with a solid red line which is the ad hoc linear fit to the data. (The two extrapolated values obtained from parabolic fits in Fig. 5(b) are shown in Fig. 1 with open red boxes.)
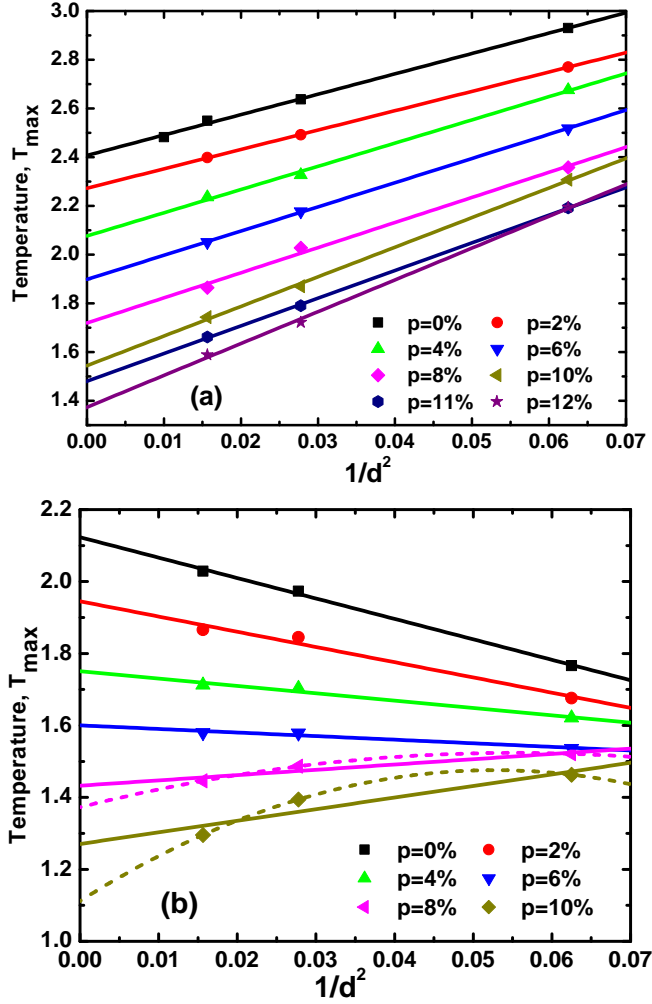
FIG. 5. (Color online) (a) Finite size scaling of the dimensionless temperatures $T_{\max}$ where specific heat reaches the maximum for $(3,4)$ QHP models vs. the inverse square of the code distance, with the fraction of flipped bonds p as indicated. For the point $p=0$ and $d=10$ where we could not eliminate the hysteresis, the average temperature was used. (b) Same for the dual $(3,4)$ QHP models. To accommodate the increased curvature, for $p=0.08$ and 0.10 we also used parabolic fits, $T_{\max} = T_c + A/d^2 + B/d^4$, which results in a significantly reduced extrapolated value of T_c for $p=0.10$. The T_c values from parabolic fits are shown in Fig. 1 with open red symbols.

The corresponding line is approximately parallel to that for the $(3,4)$ QHP model. As expected (see Sec. II C), the points at $p=0$ are located close to mutually dual positions. For the dual model, the extrapolation gives $T_c(H^*, p=0) \approx 2.12$, which is close to $T_c^*(G, p=0) = 2.14$ obtained from $T_c(G, p=0) \approx 2.41$.

Empirically, the transition temperatures in the two dual models are different. Under this condition[47] $\big($more precisely, assuming that large-system free energy density $[f_{\mathbf{e}}(H^*, K)]_p$ be non-singular at and below the lowest-temperature singular point of $[f_{\mathbf{e}}(G, K)]_p\big)$ the transition

temperature of the dual $(3,4)$ QHP model should coincide with the homological transition, where $[\Delta f_{\mathbf{e}}(G, H; K)]_p$ reaches the lower bound of 0 [cf. Eqs. (17) and (18)]. Above this temperature, $[\Delta f_{\mathbf{e}}]_p > 0$. Thus, according to part (b) of Lemma 1, the critical point $T_c(H^*, p)$ of the dual model (red squares in Fig. 1) gives an upper bound for the decodable phase of $(3,4)$ QHP codes.

Fig. 1 also shows several analytical bounds for the decodable region. Magenta-shaded region corresponds to the lower bound for the decodable region given by Eq. (20) with $m=7$. Its rightmost point is at the same p as the energy-based analytical bound from Ref. [16], the corresponding point is indicated on the horizontal axis by the magenta arrow. The lower bound for the decodable region (pseudothreshold for energy-based decoding) is shown with the blue vertical arrow. Finally, a pair of gray arrows separated by the gray bar on the vertical axis show the bound $T_{\max}$ from Theorem 3 and the corresponding dual temperature, $T_{\max}^* > T_{\max}$. As expected[47], the transitions temperatures for $(3,4)$ QHP and the dual $(3,4)$ QHP models at $p=0$ are outside of this interval.

V. CONCLUSION

In conclusion, we have studied error correction properties of the finite-rate family of quantum hypergraph product codes obtained from $\mathbb{B}(3,4)$ Gallager ensemble of classical binary codes, by combining the threshold calculation using a cluster-based decoder approximating minimum-energy decoding with the analysis of the phase diagram of the associated spin models (Fig. 1). Rigorous analytical bounds for the decodable region are constructed by analyzing the properties of the homological difference (16), equal to the logarithm of the conditional decoding probability with the negative sign.

The estimated minimum-weight decoding threshold error rate for this code family is in the vicinity of 7.0%. This estimate is not so far from the perfect matching algorithm threshold of 10.4% for the toric codes [7], and is much higher compared to the analytic lower bound of 0.7% obtained in Ref. [16].

The most striking feature of the phase diagram of the associated spin models originating from the finite asymptotic rate $[R = 1/25$ for $(3,4)$ QHP codes] is the deviation of the transition lines from the self-dual temperature at $p=0$. In fact, the transitions temperatures of the two dual models deviate from each other throughout the small-p region we studied. We expect the multicritical points, where the corresponding transition lines intersect the Nishimori line, also to be different, contrary to the implicit assumption in Ref. 30.

Notice that the horizontal position p_{bnd} of the rightmost point of the region where Theorem 2 guarantees decodability with asymptotic probability one (magenta-shaded region in Fig. 1) coincides with the analytic lower bound for the energy-based decoding from Ref. 16. While

the former region is entirely located above the Nishimori line, the minimum-energy decoding threshold corresponds to $T = 0$. A point on the Nishimori line correspond to maximum-likelihood decoding at the corresponding p. This guarantees that the portion of the Nishimori line for $p \leq p_{\text{bnd}}$ is also inside the decodable region. It is reasonable to expect that for $p \leq p_{\text{bnd}}$, the entire interval of temperatures below the bound of Theorem 2 would be in the decodable region. However, construction of the corresponding analytical bound is still an open problem.

Appendix A: Proof of Eqs. (17) to (19).

(**i**) The lower bound in Eq. (17),

$$0 \leq \Delta F_{\mathbf{0}}(G, H; K) \leq k \ln 2, \tag{17}$$

is trivial to prove, since $Z_{\mathbf{0}}(H^*; K)$ is a sum of positive terms which include every term present in $Z_{\mathbf{0}}(G; K)$. To prove the upper bound, notice that for any $\mathbf{e} \in \mathbb{F}_2^n$, $Z_{\mathbf{e}}(G; K) \leq Z_{\mathbf{0}}(G; K)$; this can be proved by comparing the corresponding expansions in powers of $\tanh K$. The expression for $Z_{\mathbf{0}}(H^*; K) = \sum_{\mathbf{c}} Z_{\mathbf{c}}(G; K)$ includes the summation over 2^k distinct defect vectors $\mathbf{c}$, thus $Z_{\mathbf{0}}(H^*; K) \leq 2^k Z_{\mathbf{0}}(G; K)$, which gives the upper bound in Eq. (17).

(**ii**) The inequality

$$\Delta F_{\mathbf{e}}(G, H; K) - \Delta F_{\mathbf{0}}(G, H; K) \geq 0 \tag{18}$$

is derived with the help of the duality (11) which maps the l.h.s. into the difference of the logarithms of the averages,

$$\Delta F_{\mathbf{e}} - \Delta F_{\mathbf{0}} = \ln \frac{Z_{\mathbf{e}}(H^*; K)}{Z_{\mathbf{0}}(H^*; K)} - \ln \frac{Z_{\mathbf{e}}(G; K)}{Z_{\mathbf{0}}(G; K)}$$
$$= \ln \langle R^{\mathbf{e}} \rangle_{H; K^*} - \ln \langle R^{\mathbf{e}} \rangle_{G^*; K^*};$$

the difference is non-negative by the GKS second inequality[44, 45] (average in the first term can be obtained from that on the right by applying an infinite field at the k additional spins).

(**iii**) The duality relation

$$\Delta F_{\mathbf{0}}(G, H; K) = k \ln 2 - \Delta F_{\mathbf{0}}(H, G; K^*). \tag{19}$$

is a simple consequence of Eq. (11) with $\mathbf{e} = \mathbf{m} = \mathbf{0}$ and the definition of the dual matrices G^*, H^*. Let r_G and r_H denote the numbers of rows in G and H, respectively. By construction, the dual matrices G^* and H^* have $r_{G^*} = r_H + k$ and $r_{H^*} = r_G + k$ rows, and their ranks are $\operatorname{rank} G^* = n - \operatorname{rank} G$, $\operatorname{rank} H^* = n - \operatorname{rank} H$. We have,

$$\frac{Z_{\mathbf{0}}(H^*, K)}{Z_{\mathbf{0}}(G, K)} = \frac{2^{r_H^* - r_H + \operatorname{rank} H}}{2^{r_G - r_G^* + \operatorname{rank} G^*}} \frac{Z_{\mathbf{0}}(H, K^*)}{Z_{\mathbf{0}}(G^*, K^*)}$$
$$= 2^k \frac{Z_{\mathbf{0}}(H, K^*)}{Z_{\mathbf{0}}(G^*, K^*)}.$$

Eq. (19) is obtained by taking the logarithm.

Appendix B: Proof of Lemma 1

Lemma 1. *For a sequence of quantum CSS codes defined by pairs of matrices (G_t, H_t), $t \in \mathbb{N}$, where $G_t H_t^T = 0$, given a finite $K > 0$ and an error probability $p \geq 0$,*
(a) $\lim_{t \to \infty} [\Delta F_{\mathbf{e}}(G_t, H_t; K)]_p = 0$ implies the point (p, K) to be in the decodable region;
(b) $\liminf_{t \to \infty} [\Delta f_{\mathbf{e}}(G_t, H_t; K)]_p > 0$ implies the point (p, K) to be outside of the decodable region.

Proof. Part (a) immediately follows from the convexity of the exponential function,

$$\left[P(\mathbf{e}|\mathbf{e}H^T) \right]_p \geq \exp \left[\ln \frac{Z_{\mathbf{e}}(G; K)}{Z_{\mathbf{e}}(H^*; K)} \right]_p = e^{-\Delta F_p(G, H; K)}.$$

Part (b) follows from the trivial bounds on the partition function, $2^r e^{-Kn} \leq Z_{\mathbf{e}}(G; K) \leq 2^r e^{Kn}$, where G is an $r \times n$ matrix. This gives a lower bound for the conditional probability (5),

$$\ln P(\mathbf{e}|\mathbf{s}) \geq \ln \left(\frac{2^r e^{-Kn}}{2^{r+k} e^{Kn}} \right) = -n(2K + R) \ln 2. \tag{B1}$$

Now, for some $\delta > 0$, let us say that a "good" disorder configuration $\mathbf{e}$ corresponds to $P(\mathbf{e}|\mathbf{e}H^T) \geq 1 - \delta$, to obtain

$$[\Delta F_{\mathbf{e}}]_p = -[\ln P(\mathbf{e}|\mathbf{e}H^T)]_p$$
$$\leq nM P_{\text{bad}} + (1 - P_{\text{bad}}) \ln \frac{1}{1 - \delta}$$
$$\leq nM P_{\text{bad}} + \ln \frac{1}{1 - \delta}, \tag{B2}$$

where $M = (2K + R) \ln 2$ is the constant in the r.h.s. of Eq. (B1), and $P_{\text{bad}} = 1 - P_{\text{good}}$ is the net probability to encounter a bad configuration. A similar chain of inequalities gives an upper bound for P_{bad}:

$$P_{\text{succ}} = [P(\mathbf{e}|H^T\mathbf{e})]_p$$
$$\leq P_{\text{good}} + (1 - P_{\text{good}})(1 - \delta)$$
$$= 1 - (1 - P_{\text{good}})\delta; \text{ thus}$$
$$1 - P_{\text{succ}} \geq (1 - P_{\text{good}})\delta = P_{\text{bad}}\,\delta,$$

Combining with Eq. (B2), this gives for the success probability (6), at a fixed $0 < \delta < 1$:

$$
\begin{aligned}
1 - P_{\text{succ}} &\geq P_{\text{bad}}\,\delta \\
&\geq \delta\,\frac{[\Delta F_{\mathbf{e}}]_p + \ln(1-\delta)}{nM} \\
&\stackrel{n\to\infty}{=} \delta\,\frac{[\Delta f_{\mathbf{e}}]_p}{(2K+R)\ln 2} > 0,
\end{aligned}
\qquad (\text{B3})
$$

which limits P_{succ} from above, away from one. $\square$

Appendix C: Proof of Theorem 2

Theorem 2. *Consider a sequence of quantum CSS codes $\mathcal{Q}(G_t, H_t)$, $t \in \mathbb{N}$, of increasing lengths n_t, where row weights of each G_t and H_t do not exceed a fixed m, and the code distances $d_t \geq D\ln n_t$, with some $D > 0$. Then the sequence $\Delta F_t \equiv [\Delta F_{\mathbf{e}}(G_t, H_t; K)]_p$, $t \in \mathbb{N}$, converges to zero in the region*

$$
(m-1)[e^{-2K}(1-p) + e^{2K}p] < e^{-1/D}. \qquad (20)
$$

The statement of the theorem immediately follows from the positivity of $\Delta F_{\mathbf{e}}(G, H; K)$, see Eq. (17), and the following Lemma:

Lemma 4. *Consider a pair of Ising models defined in terms of matrices G and H with orthogonal rows, such that the matrix H has a maximum row weight m. Let d_G denote the CSS distance (3), the minimum weight of a defect $\mathbf{c} \in \mathcal{C}_H^\perp \setminus \mathcal{C}_G$. Denote $C \equiv e^{-2K}(1-p) + e^{2K}p$, and assume that $(m-1)C < 1$. Then, the disorder-averaged homological difference (16) satisfies*

$$
[\Delta F(G, H; K)]_p \leq n\,\frac{(m-1)^{d_G}C^{d_G+1}}{1 - (m-1)C}. \qquad (\text{C1})
$$

Proof. It is convenient to represent the partition function (4) in the form

$$
Z_{\mathbf{e}}(G; K) = e^{Kn}\sum_{\boldsymbol{\varepsilon}\simeq\mathbf{0}} e^{-2K\,\mathrm{wgt}(\mathbf{e}+\boldsymbol{\varepsilon})},
$$

where the notation $\boldsymbol{\varepsilon} \simeq \mathbf{0}$ indicates that $\boldsymbol{\varepsilon}$ is in the trivial degeneracy class, that is, it can be represented as a linear combination of rows of G, $\boldsymbol{\varepsilon} = \boldsymbol{\alpha}G$, and $\mathrm{wgt}(\mathbf{e}+\boldsymbol{\varepsilon})$ is the total number of flipped bonds with the spins $S_i = (-1)^{\alpha_i}$. In comparison,

$$
Z_{\mathbf{e}}(H^*; K) = e^{Kn}\sum_{\boldsymbol{\varepsilon}:H\boldsymbol{\varepsilon}^T=0} e^{-2K\,\mathrm{wgt}(\mathbf{e}+\boldsymbol{\varepsilon})};
$$

here the summation is over all vectors $\boldsymbol{\varepsilon} \in \mathbb{F}_2^n$ which are orthogonal to the rows of H. Let us consider a decomposition of any such binary vector $\boldsymbol{\varepsilon}$ into *irreducible* components[16], $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2 + \ldots$, where supports of different vectors in the decomposition do not overlap, $\boldsymbol{\varepsilon}_i \cap \boldsymbol{\varepsilon}_j = \emptyset$ if $i \neq j$. The requirement is that each component $\boldsymbol{\varepsilon}_i$ be orthogonal to the rows of H, and cannot be

further decomposed into a sum of non-overlapping zero-syndrome vectors (such a decomposition is not necessarily unique). Now, group all of the components which are trivial, $\boldsymbol{\varepsilon}_i \simeq \mathbf{0}$, into the vector $\boldsymbol{\varepsilon}''$, and the non-trivial components into the vector $\boldsymbol{\varepsilon}'$, so that $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}' + \boldsymbol{\varepsilon}''$, where $\boldsymbol{\varepsilon}' \cap \boldsymbol{\varepsilon}'' = \emptyset$, vector $\boldsymbol{\varepsilon}'$ is a sum of non-trivial non-overlapping codewords $\mathbf{c}_j \in \mathcal{C}_H^\perp \setminus \mathcal{C}_G$, and the remainder is trivial, $\boldsymbol{\varepsilon}'' \simeq \mathbf{0}$.

Given such a decomposition for each vector $\boldsymbol{\varepsilon} \in \mathcal{C}_H^\perp$, we can construct an upper bound for the ratio,

$$
\begin{aligned}
\frac{Z_{\mathbf{e}}(H^*; K)}{Z_{\mathbf{e}}(G; K)} &= \frac{\displaystyle\sum_{\boldsymbol{\varepsilon}:H\boldsymbol{\varepsilon}^T=0} e^{-2K\,\mathrm{wgt}(\mathbf{e}+\boldsymbol{\varepsilon})}}{\displaystyle\sum_{\boldsymbol{\varepsilon}\simeq\mathbf{0}} e^{-2K\,\mathrm{wgt}(\mathbf{e}+\boldsymbol{\varepsilon})}} \\
&\leq \sum_{\boldsymbol{\varepsilon}'}\frac{\displaystyle\sum_{\boldsymbol{\varepsilon}''\simeq\mathbf{0}:\boldsymbol{\varepsilon}''\cap\boldsymbol{\varepsilon}'=\emptyset} e^{-2K\,\mathrm{wgt}(\mathbf{e}+\boldsymbol{\varepsilon}'+\boldsymbol{\varepsilon}'')}}{\displaystyle\sum_{\boldsymbol{\varepsilon}''\simeq\mathbf{0}:\boldsymbol{\varepsilon}''\cap\boldsymbol{\varepsilon}'=\emptyset} e^{-2K\,\mathrm{wgt}(\mathbf{e}+\boldsymbol{\varepsilon}'')}},
\end{aligned}
$$

where the outside summation is over $\boldsymbol{\varepsilon}'$, a sum of non-overlapping irreducible codewords, and (for a given $\boldsymbol{\varepsilon}'$) we reduced the denominator by dropping the terms which overlap with $\boldsymbol{\varepsilon}'$, to match the corresponding sum in the numerator. The ratios for each $\boldsymbol{\varepsilon}'$ can now be trivially calculated in terms of the weight of $\mathbf{e}$ in the support of $\boldsymbol{\varepsilon}'$, which we denote as $\mathrm{wgt}(\mathbf{e}')$. We have

$$
\frac{Z_{\mathbf{e}}(H^*; K)}{Z_{\mathbf{e}}(G; K)} \leq \sum_{\boldsymbol{\varepsilon}'} e^{-2K[\mathrm{wgt}(\boldsymbol{\varepsilon}') - 2\,\mathrm{wgt}(\mathbf{e}')]},
$$

and the corresponding average

$$
\left[\frac{Z_{\mathbf{e}}(H^*; K)}{Z_{\mathbf{e}}(G; K)}\right]_p \leq \sum_{\boldsymbol{\varepsilon}'} C^{\mathrm{wgt}(\boldsymbol{\varepsilon}')},
$$

where the constant $C \equiv (1-p)e^{-2K} + pe^{2K}$. The summation is over sums of irreducible non-overlapping codewords, $\boldsymbol{\varepsilon}' = \mathbf{c}_1 + \mathbf{c}_2 + \ldots + \mathbf{c}_m$; we can further increase the r.h.s. if we allow the overlaps between the codewords, to obtain

$$
\left[\frac{Z_{\mathbf{e}}(H^*; K)}{Z_{\mathbf{e}}(G; K)}\right]_p \leq \exp\left(\sum_{\mathbf{c}} C^{\mathrm{wgt}(\mathbf{c})}\right),
$$

where the summation is now done over irreducible codewords. The bound for $[\Delta F_{\mathbf{e}}(G, H; K)]_p$ is obtained using the concavity of the logarithm,

$$
\left[\ln\frac{Z_{\mathbf{e}}(H^*; K)}{Z_{\mathbf{e}}(G; K)}\right]_p \leq \ln\left[\frac{Z_{\mathbf{e}}(H^*; K)}{Z_{\mathbf{e}}(G; K)}\right]_p \leq \sum_{\mathbf{c}} C^{\mathrm{wgt}(\mathbf{c})}.
$$

The final step is to bound the number of irreducible codewords by the number of the vectors orthogonal to the rows of H of weight d_G or larger. For the number N_w of vectors in $\mathcal{C}_H^\perp$ of weight w one has[16, 53] $N_w \leq n(m-1)^w$; summation over $w \geq d_G$ gives Eq. (C1). $\square$

Appendix D: Proof of Theorem 3

Theorem 3. *Consider a sequence of CSS codes defined by pairs of finite binary matrices with mutually orthogonal rows, $G_t H_t^T = 0$, $t \in \mathbb{N}$, where row weights of G_t and H_t do not exceed a fixed m, the sequence of CSS distances $d_t = \max(d_{H_t}, d_{G_t})$ is strictly increasing with t, $d_{t+1} > d_t$, and the sequence of rates $R_t \equiv k_t/n_t$ converges, $\lim_{t \to \infty} R_t = R$. Then, assuming equal probabilities of X and Z errors, the upper temperature boundary of the decodable region, $T_{\max} = 1/K_{\max}$, satisfies the inequality*

$$K_{\max} - K_{\max}^* \ge R \ln 2. \qquad (21)$$

Proof. By Eq. (18), to establish the upper bound, we can work at $p = 0$. Let $T_1 = 1/K_1$ and $T_2 = 1/K_2$ respectively be the upper boundaries of the homological regions

such that for $\Delta f_{\mathbf{0}}(G, H; K) = 0$ and $\Delta f_{\mathbf{0}}(H, G; K) = 0$. By duality (19), $\Delta f_{\mathbf{0}}(G, H; K_2^*) = R \ln 2$. On the other hand, the derivative of $f_{\mathbf{0}}(G; K)$ with respect to K is the average energy per bond,

$$\partial_K f_{\mathbf{0}}(G; K) = -n^{-1} \sum_b \langle R_b \rangle_{G;K};$$

using the GKS inequalities we obtain

$$0 \le \langle R_b \rangle_{H^*;K} \le \langle R_b \rangle_{G;K} \le 1.$$

This implies the derivative of $-\Delta f_{\mathbf{0}}(G, H; K)$ with respect to K must be in the interval $(0, 1)$. Consequently, $K_1 - K_2^* \ge R \ln 2$. Similar arguments with G and H interchanged gives $K_2 - K_1^* \ge R \ln 2$. If we define $K_{\max} = \max(K_1, K_2)$, then it satisfies Eq. (21). $\qquad \square$

[1] P. W. Shor, "Fault-tolerant quantum computation," in *Proc. 37th Ann. Symp. on Fundamentals of Comp. Sci.*, IEEE (IEEE Comp. Soc. Press, Los Alamitos, 1996) pp. 56–65, quant-ph/9605011.

[2] E. Knill, R. Laflamme, and W. H. Zurek, "Resilient quantum computation," Science **279**, 342 (1998).

[3] A. Yu. Kitaev, "Quantum computations: algorithms and error correction," Uspekhi Mat. Nauk **52**, 53112 (1997), [Russian Math. Surveys, 52:6 (1997), 11911249].

[4] D. Aharonov and M. Ben-Or, "Fault-tolerant quantum computation with constant error," in *STOC '97 Proceedings of the twenty-ninth annual ACM symposium on Theory of computing, El Paso, Texas, USA – May 04-06, 1997* (ACM New York, NY USA, 1997) pp. 176–188.

[5] Austin G. Fowler, Charles D. Hill, and Lloyd C. L. Hollenberg, "Quantum-error correction on linear-nearest-neighbor qubit arrays," Phys. Rev. A **69**, 042314 (2004).

[6] P. Aliferis, D. Gottesman, and J. Preskill, "Quantum accuracy threshold for concatenated distance-3 codes," Quantum Inf. Comput. **6**, 97–165 (2006), quant-ph/0504218.

[7] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, "Topological quantum memory," J. Math. Phys. **43**, 4452 (2002).

[8] David S. Wang, Austin G. Fowler, Ashley M. Stephens, and Lloyd Christopher L. Hollenberg, "Threshold error rates for the toric and planar codes," Quantum Inf. Comput. **10**, 456–469 (2010), arXiv:0905.0531.

[9] David S. Wang, Austin G. Fowler, and Lloyd C. L. Hollenberg, "Surface code quantum computing with error rates over 1%," Phys. Rev. A **83**, 020302 (2011).

[10] David S. Wang, Austin G. Fowler, Charles D. Hill, and Lloyd Christopher L. Hollenberg, "Graphical algorithms and threshold error rates for the 2d color code," Quantum Inf. Comput. **10**, 780–802 (2010), arXiv:0907.1708.

[11] A. J. Landahl, J. T. Anderson, and P. R. Rice, "Fault-tolerant quantum computing with color codes," (2011), presented at QIP 2012, December 12 to December 16, arXiv:1108.5738.

[12] Sergey Bravyi and Barbara Terhal, "A no-go theorem for a two-dimensional self-correcting quantum memory based on stabilizer codes," New Journal of Physics **11**, 043029 (2009).

[13] S. Bravyi, D. Poulin, and B. Terhal, "Tradeoffs for reliable quantum information storage in 2D systems," Phys. Rev. Lett. **104**, 050503 (2010), 0909.5200.

[14] D. Gottesman, "Fault-tolerant quantum computation with constant overhead," Quant. Information and Computation **14**, 1338–1371 (2014), 1310.2984.

[15] A. A. Kovalev and L. P. Pryadko, "Fault tolerance of quantum low-density parity check codes with sublinear distance scaling," Phys. Rev. A **87**, 020304(R) (2013).

[16] I. Dumer, A. A. Kovalev, and L. P. Pryadko, "Thresholds for correcting errors, erasures, and faulty syndrome measurements in degenerate quantum codes," Phys. Rev. Lett. **115**, 050502 (2015), 1412.6172.

[17] J.-P. Tillich and G. Zemor, "Quantum LDPC codes with positive rate and minimum distance proportional to $\sqrt{n}$," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)* (2009) pp. 799–803.

[18] N. Delfosse and G. Zémor, "Quantum erasure-correcting codes and percolation on regular tilings of the hyperbolic plane," in *Information Theory Workshop (ITW), 2010 IEEE* (2010) pp. 1–5.

[19] Nicolas Delfosse, "Tradeoffs for reliable quantum information storage in surface codes and color codes," in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on* (IEEE, 2013) pp. 917–921.

[20] A. A. Kovalev and L. P. Pryadko, "Quantum Kronecker sum-product low-density parity-check codes with finite rate," Phys. Rev. A **88**, 012311 (2013).

[21] Larry Guth and Alexander Lubotzky, "Quantum error correcting codes and 4-dimensional arithmetic hyperbolic manifolds," Journal of Mathematical Physics **55**, 082202 (2014), arXiv:1310.5555.

[22] H. Bombin, "Topological subsystem codes," Phys. Rev. A **81**, 032301 (2010).

[23] H. Nishimori, "Exact results and critical properties of the ising model with competing interactions," Journal of Physics C: Solid State Physics **13**, 4071 (1980).

[24] Hidetoshi Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction* (Clarendon Press, Oxford, 2001).

[25] Chenyang Wang, Jim Harrington, and John Preskill, "Confinement-higgs transition in a disordered gauge theory and the accuracy threshold for quantum memory," Annals of Physics **303**, 31 – 58 (2003).

[26] H. G. Katzgraber, H. Bombin, and M. A. Martin-Delgado, "Error threshold for color codes and random three-body Ising models," Phys. Rev. Lett. **103**, 090501 (2009).

[27] H. Bombin, Ruben S. Andrist, Masayuki Ohzeki, Helmut G. Katzgraber, and M. A. Martin-Delgado, "Strong resilience of topological codes to depolarization," Phys. Rev. X **2**, 021004 (2012).

[28] H. G. Katzgraber and R. S. Andrist, "Stability of topologically-protected quantum computing proposals as seen through spin glasses," Journal of Physics: Conference Series **473**, 012019 (2013).

[29] Pejman Jouzdani, E. Novais, I. S. Tupitsyn, and Eduardo R. Mucciolo, "Fidelity threshold of the surface code beyond single-qubit error models," Phys. Rev. A **90**, 042315 (2014).

[30] A. A. Kovalev and L. P. Pryadko, "Spin glass reflection of the decoding transition for quantum error-correcting codes," Quantum Inf. & Comp. **15**, 0825 (2015), arXiv:1311.7688.

[31] A. R. Calderbank and P. W. Shor, "Good quantum error-correcting codes exist," Phys. Rev. A **54**, 1098–1105 (1996).

[32] A. M. Steane, "Simple quantum error-correcting codes," Phys. Rev. A **54**, 4741–4751 (1996).

[33] Creighton K. Thomas and Helmut G. Katzgraber, "Simplest model to study reentrance in physical systems," Phys. Rev. E **84**, 040101 (2011).

[34] A. R. Calderbank, E. M. Rains, P. M. Shor, and N. J. A. Sloane, "Quantum error correction via codes over GF(4)," IEEE Trans. Info. Theory **44**, 1369–1387 (1998).

[35] F. Wegner, "Duality in generalized Ising models and phase transitions without local order parameters," J. Math. Phys. **2259**, 12 (1971).

[36] J. M. Hammersley, "Comparison of atom and bond percolation processes," J. Math. Phys. **2**, 728 (1961).

[37] H. A. Kramers and G. H. Wannier, "Statistics of two-dimensional ferromagnet. Part I," Phys. Rev. **60**, 252–262 (1941).

[38] J. MacWilliams, "A theorem on the distribution of weights in a systematic code," Bell System Tech. J **42**, 79–84 (1963).

[39] Jean-Pierre Tillich and Gilles Zémor, "Quantum LDPC codes with positive rate and minimum distance proportional to the square root of the blocklength," IEEE Transactions on Information Theory **60**, 1193–1202 (2014).

[40] A. A. Kovalev and L. P. Pryadko, "Improved quantum hypergraph-product LDPC codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)* (2012) pp. 348–352, arXiv:1202.0928.

[41] R. Gallager, "Low-density parity-check codes," IRE Trans. Inf. Theory **8**, 21–28 (1962).

[42] Robert G. Gallager, *Low-Density Parity-Check Codes* (M.I.T. Press, Cambridge, Mass., 1963).

[43] S. Litsyn and V. Shevelev, "On ensembles of low-density parity-check codes: asymptotic distance distributions," IEEE Trans. Inf. Theory **48**, 887–908 (2002).

[44] Robert B. Griffiths, "Correlations in Ising ferromagnets. I," J. Math. Phys. **8**, 478 (1967).

[45] D. G. Kelly and S. Sherman, "General Griffiths' inequalities on correlations in Ising ferromagnets," Journal of Mathematical Physics **9**, 466–484 (1968).

[46] R. Griffiths, "Rigorous results and theorems," in *Phase Transitions and Critical Phenomena*, Vol. 1. Exact results, edited by C. Domb and M. S. Green (Acad. Press, London, 1972) Chap. 2, pp. 7–109.

[47] Yi Jiang, A. A. Kovalev, I. Dumer, and L. P. Pryadko, "Free energy analyticity bounds for few-body random-bond ising models with extensive homology," (2018), unpublished.

[48] Valentin Féray, Pierre-Loïc Méliot, and Ashkan Nikeghbali, *Mod-ϕ convergence I: Normality zones and precise deviations* (Springer International Publishing, Cham, 2016) arXiv:1304.2934v4.

[49] I. Benjamini and O. Schramm, "Recurrence of distributional limits of finite planar graphs," Electron. J. Probab. **6**, 23/1–13 (2001), math/0011019.

[50] Christian Borgs, Jennifer Chayes, Jeff Kahn, and László Lovász, "Left and right convergence of graphs with bounded degree," Random Structures & Algorithms **42**, 1–28 (2013), 1002.0115.

[51] L. M. Lovász, "A short proof of the equivalence of left and right convergence for sparse graphs," Eur. J. Comb. **53**, 1–7 (2016).

[52] Wolfram Research, Inc., *Mathematica*, Champaign, Illinois, version 11.2 ed. (2017).

[53] I. Dumer, A. A. Kovalev, and L. P. Pryadko, "Numerical techniques for finding the distances of quantum codes," in *Information Theory Proceedings (ISIT), 2014 IEEE International Symposium on* (IEEE, Honolulu, HI, 2014) pp. 1086–1090.

[54] I. Dumer, A. A. Kovalev, and L. P. Pryadko, "Distance verification for LDPC codes," in *Proceedings of ISIT 2016 - IEEE International Symposium on Information Theory* (IEEE, 2016) to appear, 1605.02410.

[55] I. Dumer, A. A. Kovalev, and L. P. Pryadko, "Distance verification for classical and quantum LDPC codes," IEEE Trans. Inf. Th. **63**, 4675–4686 (2017).

[56] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, 1st ed., Series of Books in the Mathematical Sciences (W. H. Freeman, 1979).

[57] H. Bombin and M. A. Martin-Delgado, "Optimal resources for topological two-dimensional stabilizer codes: Comparative study," Phys. Rev. A **76**, 012305 (2007).

[58] Helmut G. Katzgraber, Simon Trebst, David A. Huse, and Matthias Troyer, "Feedback-optimized parallel tempering monte carlo," Journal of Statistical Mechanics: Theory and Experiment **2006**, P03018 (2006).

[59] Aleksander Kubica, Michael E. Beverland, Fernando Brandao, John Preskill, and Krysta M. Svore, "Three-dimensional color code thresholds via statistical-mechanical mapping," (2017), unpublished, arXiv:1708.07131.