

Error Thresholds for Arbitrary Pauli Noise

JOHANNES BAUSCH

CQIF, DAMTP
University of Cambridge
Cambridge CB3 0WA, UK
jkrb2@cam.ac.uk

FELIX LEDITZKY

JILA & CTQM
University of Colorado Boulder
Boulder, USA
felix.leditzky@jila.colorado.edu

September 2019

Abstract

The error threshold of a one-parameter family of quantum channels is defined as the largest noise level such that the quantum capacity of the channel remains positive. This in turn guarantees the existence of a quantum error correction code for noise modeled by that channel. Discretizing the single-qubit errors leads to the important family of Pauli quantum channels; curiously, multipartite entangled states can increase the threshold of these channels beyond the so-called hashing bound, an effect termed superadditivity of coherent information. In this work, we divide the simplex of Pauli channels into one-parameter families and compute numerical lower bounds on their error thresholds. We find substantial increases of error thresholds relative to the hashing bound for large regions in the Pauli simplex corresponding to biased noise, which is a realistic noise model in promising quantum computing architectures. The error thresholds are computed on the family of graph states, a special type of stabilizer state. In order to determine the coherent information of a graph state, we devise an algorithm that exploits the symmetries of the underlying graph resulting in a substantial computational speed-up. This algorithm uses tools from computational group theory and allows us to consider symmetric graph states on a large number of vertices. Our algorithm works particularly well for repetition codes and concatenated repetition codes (or cat codes), for which our results provide the first comprehensive study of superadditivity for arbitrary Pauli channels. In addition, we identify a novel family of quantum codes based on tree graphs. The error thresholds of these tree graph states outperform repetition and cat codes in large regions of the Pauli simplex, and hence form a new code family with desirable error correction properties.

Contents

1. Introduction	4
1.1. Main Results	5
1.2. Structure of this Paper	6
2. Preliminaries	6
2.1. Quantum Capacity	6
2.2. Graph States	9
2.3. Coherent Information of Decohered Graph States	14
3. Exploiting Graph Symmetries	17
3.1. Automorphisms of Colored Graphs	17
3.2. Cosets of Graph Colorings	18
3.3. Canonical Colorings	21
3.4. Canonical Images for Colored Graphs	24
3.5. Homomorphic Group Actions	25
3.6. Coherent Information of Decohered Graph States with Symmetries	25
4. Numerical Methods	32
4.1. Visualizing Thresholds	33
5. Results	37
5.1. Repetition Codes	37
5.2. Concatenated Codes	41
5.3. Tree Codes	45
5.4. Exhaustive Search	48
5.5. Rate Comparisons	53
6. Discussion	56
6.1. Noise Models in NISQ Devices	58
6.2. Open Problems	59
A. LU-Equivalence of an Arbitrary Stabilizer State to a Graph State	65
B. Subroutines for Full Coherent Information Algorithm	67
C. Monotonicity of Coherent Information	68

1. Introduction

It is widely believed that quantum error correction is a necessary requirement for quantum computers to maintain coherence in computations. Quantum error-correcting codes protect logical quantum information from environmental noise by encoding it in larger physical systems in a redundant way. Based on the particular architecture of the quantum computing device different mathematical models, or *quantum channels*, are used to describe this environmental noise. The analysis of quantum error correction is greatly simplified by the fact that the full unitary group of possible errors affecting a physical qubit can be discretized to the group of Pauli errors [Sho95; Ste96]: a bit flip X , a phase flip Z , or a combined bit-phase flip Y , where X , Y and Z are the well-known Pauli matrices. The most extensively studied noise model is depolarizing noise, in which the three Pauli errors occur with equal probabilities. However, in many promising quantum computing architectures such as superconducting qubits [Ali+09], quantum dots [Shu+12], or trapped ions [Nig+14], the actual noise model is biased towards dephasing noise consisting of only one type of Pauli error.

A common assumption in the noise model is that each physical qubit is independently affected by the same type of noise modeled by a quantum channel, that is, no memory effects between qubits are considered. In this model, commonly referred to as the independent and identically distributed (i.i.d.) noise model, the ability to protect a quantum system (and the quantum information stored in it) from environmental noise is quantified by the *quantum capacity* $Q(\mathcal{N})$ of a quantum channel \mathcal{N} . A positive quantum capacity of a particular quantum channel implies that error correction against the (i.i.d.) noise modeled by that channel is in principle possible. For a one-parameter family $x \mapsto \mathcal{N}_x$ of quantum channels (such as depolarizing noise above), this leads to the question of determining the *error threshold* (or noise threshold), defined as the supremum over all noise levels x such that the quantum capacity remains positive, i.e. $Q(\mathcal{N}_x) > 0$.

The quantum capacity theorem [Sch96; SN96; Llo97; Sho02; Dev05] gives a regularized entropic formula for the quantum capacity. One consequence of this theorem is that positive quantum capacity of a channel can be certified by finding an entangled state that retains coherence after part of it has been affected by the noise. Here, the coherence is measured by an entropic quantity called the *coherent information*. In contrast to Shannon's classical theory of information [Sha48], it does *not* suffice to compute the coherent information with respect to one copy of the channel, as entangled multipartite quantum states acted upon by i.i.d. copies of the quantum channel may maintain more coherence [SS96; DiV+98; SS07; FW08; Cub+15; Led+18a]. This effect is called *superadditivity* of coherent information, and it is a direct consequence of the multipartite entanglement present in the quantum state.

Finding suitable multipartite entangled states is challenging due to the exponential scaling of the dimension of the underlying Hilbert space.

1.1. Main Results

In this work we study the quantum capacity thresholds of qubit Pauli channels

$$\rho \mapsto p_0\rho + p_1X\rho X + p_2Y\rho Y + p_3Z\rho Z, \quad (1)$$

corresponding to the Pauli error model in terms of the Pauli matrices X, Y, Z introduced above. To this end, we evaluate the coherent information of a Pauli channel acting on graph states [BR01; Hei+04], a special type of stabilizer states [Got97]. We find large threshold increases in the entire Pauli channel simplex, in particular for biased noise models where one type of Pauli error dominates over the other ones. We identify a novel family of quantum codes defined in terms of tree graphs with superadditive coherent information and an increased error threshold. In large regions of the Pauli channel simplex these new codes outperform the known families of (concatenated) repetition codes.

To obtain our results, we make use of the special error-correcting properties of graph states that simplify the computation of the coherent information under Pauli noise [Hei+05]. We devise an algorithm that exploits the symmetries of the underlying graph, providing a substantial speed-up in the computation of the coherent information for suitable graphs. The algorithm relies on tools from computational group theory such as strong generating sets for permutation groups, and makes heavy use of the theory of group actions. The resulting speed-up allows us to consider quantum states on a large number of constituent systems. Moreover, we obtain algebraic expressions for the coherent information in terms of the probability distribution (p_0, p_1, p_2, p_3) characterizing the Pauli channel in (1), which can easily be evaluated numerically for the entire Pauli channel simplex.

Our algorithm is particularly well-suited for the class of repetition codes and concatenated repetition codes (or *cat codes*). These codes are examples of *degenerate* codes where multiple errors are mapped to the same syndrome, and have long been known to increase the thresholds of certain Pauli channels such as the depolarizing channel or the independent bit and phase flip channel [SS96; DiV+98; SS07; FW08]. Our paper comprises the first comprehensive study of the thresholds of these degenerate codes for *arbitrary* Pauli channels. We also identify a novel type of quantum codes given by tree graph states. These codes outperform the best repetition and cat codes in large regions in the Pauli channel simplex, and provide new examples of quantum codes with favorable error-correcting properties.

1.2. Structure of this Paper

This paper is structured as follows. Section 2 contains necessary background information about the quantum capacity of a quantum channel, superadditivity of coherent information, graph states, and their coherence properties under Pauli noise. In Section 3 we derive our main technical tool, an algorithm to compute the coherent information of graph states under Pauli noise that exploits the symmetries of the graph (Algorithms 2 to 4). In Section 4 we describe the numerical methods used throughout our study. Section 5 contains the main results of this paper, a detailed analysis of the error thresholds of Pauli channels using graph states, and an analysis of rates close to the noise threshold. We discuss repetition codes (Section 5.1), concatenated repetition codes (Section 5.2), and a novel code family based on tree graphs (Section 5.3). We also carry out an exhaustive search over connected graphs of up to eight vertices (Section 5.4). We conclude in Section 6 with a discussion of our results and how they are relevant for realistic noise models in near-term quantum devices, and we mention open questions that are subject to further study.

2. Preliminaries

We briefly define the notation used throughout this paper. A quantum system A is associated with a finite-dimensional Hilbert space \mathcal{H}_A , and multipartite systems $ABC \dots$ are described by the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \otimes \dots$ of the respective Hilbert spaces. We denote by $|A| = \dim \mathcal{H}_A$ the dimension of a quantum system A . A quantum state (or density operator) ρ_A on A is a linear positive semidefinite operator with unit trace. A pure quantum state ψ_A is a quantum state of rank 1, which can be associated with a normalized vector (or ket) $|\psi\rangle_A \in \mathcal{H}_A$ such that $\psi_A = |\psi\rangle\langle\psi|_A$. A quantum channel $\mathcal{N}: A \rightarrow B$ is a linear, completely positive and trace-preserving map from the algebra of linear operators on \mathcal{H}_A to the algebra of linear operators on \mathcal{H}_B . We denote by $\mathbb{1}_A$ the identity operator on \mathcal{H}_A , and by $\text{id}_A: A \rightarrow A$ the identity map on the algebra of linear operators on \mathcal{H}_A . All logarithms are taken to base 2. For $n \in \mathbb{N}$, we use the notation $[n] := \{1, \dots, n\}$.

2.1. Quantum Capacity

In information-theoretic terms, the quantum capacity $Q(\mathcal{N})$ of a quantum channel $\mathcal{N}: A \rightarrow B$ characterizes the channel's ability to faithfully transmit quantum information. Here, we define it operationally through the task of entanglement generation. In this task, Alice prepares a bipartite state ρ_{RA^n} in her lab, and sends the A^n -part to Bob through n copies of the quantum channel \mathcal{N} . Bob applies a local decoding operation \mathcal{D} to his share of the state in order to

bring the total state $(\text{id}_R \otimes \mathcal{D} \circ \mathcal{N}^{\otimes n})(\rho_{RA^n})$ close to a maximally entangled state Φ^+ up to some error ε defined in terms of a suitable distance measure. Informally, the quantum capacity is defined as the largest Schmidt Rank of Φ^+ such that ε vanishes asymptotically in the limit $n \rightarrow \infty$.

Formally, let $\varepsilon \in (0, 1)$ and $|\Phi^+\rangle_{RR'} = |R|^{-1/2} \sum_{i=0}^{|R|-1} |i\rangle_R |i\rangle_{R'}$ be a maximally entangled state on a bipartite quantum system. An ε -entanglement generating code for a quantum channel $\mathcal{N}: A \rightarrow B$ is a pair (ρ_{RA}, \mathcal{D}) consisting of an arbitrary quantum state ρ_{RA} and a decoding quantum channel $\mathcal{D}: B \rightarrow R'$ such that

$$F(\Phi_{RR'}^+, (\text{id}_R \otimes \mathcal{D} \circ \mathcal{N})(\rho_{RA})) \geq 1 - \varepsilon,$$

where $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$ denotes the fidelity of two quantum states ρ and σ . Let $r(n, \varepsilon)$ denote the largest dimension $|R|$ of R such that there exists an ε -entanglement generating code for $\mathcal{N}^{\otimes n}$. Then the quantum capacity $Q(\mathcal{N})$ is defined as

$$Q(\mathcal{N}) = \inf_{\varepsilon > 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log r(n, \varepsilon).$$

A positive quantum capacity means that it is theoretically possible to faithfully transmit quantum information through the channel at a positive rate. Alternatively, $Q(\mathcal{N}) > 0$ means that a quantum system can be protected from environmental noise modeled by \mathcal{N} assuming that the noise acts independently and identically distributed (i.i.d.) on each qubit. Both points of view lead to the fundamental questions of determining the value of $Q(\mathcal{N})$ for a given noise model \mathcal{N} , and more generally whether we have $Q(\mathcal{N}) > 0$.

In information-theoretic terms, these questions have found a partial answer in the following *coding theorem* providing a regularized entropic formula for the quantum capacity [Sch96; SN96; Llo97; Sho02; Dev05]:

$$Q(\mathcal{N}) = \sup_{n \in \mathbb{N}} \frac{1}{n} I_c(\mathcal{N}^{\otimes n}), \quad (2)$$

where the *channel coherent information* is defined as

$$I_c(\mathcal{N}) = \max_{|\psi\rangle_{RA}} I_c(\psi, \mathcal{N}) = \max_{|\psi\rangle_{RA}} [S(\mathcal{N}(\psi_A)) - S(\mathcal{N}(\psi_{RA}))], \quad (3)$$

with $S(\rho) = -\text{tr} \rho \log \rho$ denoting the von Neumann entropy of a density operator ρ . In this paper, we refer to the pure input state $|\psi\rangle_{RA}$ in Eq. (3) simply as a *quantum code*.¹ The regularization in Eq. (2) is necessary since the channel coherent information is known to be

¹We refer to [BL18] for a justification of this terminology.

superadditive: there exist channels \mathcal{N} and quantum codes $\psi_{\mathbb{R}A^n}$ such that [SS96; DiV+98; SS07; FW08; Cub+15; Led+18a]

$$\frac{1}{n}I_c(\psi, \mathcal{N}^{\otimes n}) > I_c(\mathcal{N}). \quad (4)$$

A well-known quantum channel exhibiting superadditivity of coherent information as in Eq. (4) is the qubit depolarizing channel [SS96; DiV+98], defined for $p \in [0, 1]$ as

$$\mathcal{D}_p(\rho) = (1-p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z), \quad (5)$$

where X, Y, Z are the Pauli matrices. The depolarizing channel belongs to the class of *Pauli channels* $\mathcal{N}_{\mathbf{p}}$, defined in terms of a probability distribution $\mathbf{p} = (p_0, p_1, p_2, p_3)$ as

$$\mathcal{N}_{\mathbf{p}}(\rho) = p_0\rho + p_1X\rho X + p_2Y\rho Y + p_3Z\rho Z. \quad (6)$$

Superadditivity of coherent information has also been demonstrated for other channels in this class, such as the BB84 channel with $\mathbf{p} = ((1-p)^2, p-p^2, p^2, p-p^2)$ or the Two-Pauli channel with $\mathbf{p} = (1-p, p/2, 0, p/2)$ [SS07; FW08]. One of the main goals of this paper is a comprehensive study of superadditivity for the class of Pauli channels. We note here that for dephasing-type noise of the form $\mathcal{N}_p^D(\rho) = (1-p)\rho + pD\rho D$ with $D \in \{X, Y, Z\}$, the coherent information is additive and hence equal to the quantum capacity [DS05].

Superadditivity of the coherent information renders the quantum capacity of general quantum channels intractable to compute. As a result, the quantum capacity of Pauli channels is unknown except for a few special cases, and one needs to resort to finding lower and upper bounds on $Q(\mathcal{N})$, e.g., in order to certify that $Q(\mathcal{N}) > 0$. A helpful tool in this task is the direct part of the coding theorem Eq. (2), which states that

$$Q(\mathcal{N}) \geq \max_{|\psi\rangle_{\mathbb{R}A^n}} \frac{1}{n}I_c(\psi_{\mathbb{R}A^n}, \mathcal{N}^{\otimes n}) \quad \text{for any } n \in \mathbb{N}. \quad (7)$$

Even for fixed $n \in \mathbb{N}$, solving the optimization problem on the right-hand side of Eq. (7) is generally infeasible due to the exponential growth of the number of parameters in $\psi_{\mathbb{R}A^n}$. Instead, one may try to restrict the optimization to a smaller class of multipartite quantum states $|\psi\rangle_{\mathbb{R}A^n}$ that is numerically tractable. For example, in our previous paper [BL18] we optimized the coherent information in Eq. (7) over neural network states, a class of multipartite quantum states with polynomially many degrees of freedom [CT17]. For sufficiently low $n \in \mathbb{N}$, this ansatz proved successful in efficiently approximating the channel coherent information for noise models such as the dephasing channel [Led+18a].

In this work, we focus on the class of Pauli channels as defined in Eq. (6). The form of these channels naturally suggests a restriction in Eq. (7) to the family of *stabilizer states*,

which are states defined in terms of their invariance properties under (an Abelian group of) many-qubit Pauli operators. In addition, for a special type of stabilizer states called *graph states* the coherent information Eq. (3) turns into a “classical” quantity that can be computed more efficiently. This is the starting point for our analysis of Pauli channels and explained in the next section, in which we first review stabilizer and graph states along with some of their useful properties.

2.2. Graph States

Before we define and discuss graph states, the main object of study in this paper, we briefly introduce stabilizer groups and stabilizer states [Got97].

Let $\mathcal{P} = \{I, X, Y, Z\}$ be the set of Pauli matrices including the identity I on \mathbb{C}^2 , and define the n -qubit Pauli group (or simply Pauli group)

$$\mathcal{P}_n = \{i^k \sigma_1 \otimes \dots \otimes \sigma_n : \sigma_i \in \mathcal{P}, k = 0, 1, 2, 3\}.$$

A *stabilizer group* \mathcal{S} is an Abelian subgroup of \mathcal{P}_n that does not contain $-I^{\otimes n}$. Since operators in \mathcal{S} commute pairwise, they can be simultaneously diagonalized. The *stabilizer subspace* $\mathcal{H}_{\mathcal{S}}$ is defined as the joint eigenspace of the operators in \mathcal{S} corresponding to the eigenvalue 1:

$$\mathcal{H}_{\mathcal{S}} = \{|\psi\rangle \in (\mathbb{C}^2)^{\otimes n} : S|\psi\rangle = |\psi\rangle \text{ for all } S \in \mathcal{S}\}.$$

Let S_1, \dots, S_r with $r \leq n$ be a (minimal) set of *stabilizer generators* for the stabilizer group \mathcal{S} . Then it can be shown that $\dim \mathcal{H}_{\mathcal{S}} = 2^k$ where $k = n - r$. In the language of quantum error-correcting codes, the *stabilizer code* \mathcal{S} (or $\mathcal{H}_{\mathcal{S}}$) encodes k logical qubits into n physical qubits, and is commonly denoted as $[n, k]$.² Let $N_{\mathcal{P}_n}(\mathcal{S}) = \{P \in \mathcal{P}_n : PSP^\dagger \in \mathcal{S}\}$ be the *normalizer* of \mathcal{S} in \mathcal{P}_n . The stabilizer group \mathcal{S} is a normal subgroup of $N_{\mathcal{P}_n}(\mathcal{S})$, and the factor group $N_{\mathcal{P}_n}(\mathcal{S})/\mathcal{S}$ is isomorphic to \mathcal{P}_k , i.e., one can choose $\bar{X}_i, \bar{Z}_i \in N_{\mathcal{P}_n}(\mathcal{S})/\mathcal{S}$ for $i = 1, \dots, k$ satisfying the usual relations for Pauli matrices. The operators \bar{X}_i, \bar{Z}_i are called *logical operators*; they commute with all stabilizers in \mathcal{S} , but act non-trivially on the code space $\mathcal{H}_{\mathcal{S}}$. If $r = n$, then we may choose a pure state $|\psi\rangle$ such that $\mathcal{H}_{\mathcal{S}} = \text{span}(|\psi\rangle)$, and we refer to $|\psi\rangle$ as a *stabilizer state*.

²We will not be concerned with the *distance* of a stabilizer code.

2.2.1. Definition and Properties

We now define graph states. Let $\Gamma = (V, E)$ be a simple undirected graph with vertex set $V = \{1, \dots, n\}$ and edge set E , and identify each vertex with a qubit associated with a Hilbert space \mathbb{C}^2 . The *graph state* $|\Gamma\rangle \in (\mathbb{C}^2)^{\otimes n}$ is defined as

$$|\Gamma\rangle = \prod_{e \in E} CZ_e |+\rangle^{\otimes n},$$

where $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and CZ_e denotes the controlled-Z gate $|0\rangle\langle 0|_i \otimes \mathbb{1}_j + |1\rangle\langle 1|_i \otimes Z_j$ acting on the qubits i and j connected by the edge $e \in E$. Equivalently, the graph state $|\Gamma\rangle$ can be regarded as a stabilizer state defined by the n stabilizer generators

$$S_i = X^i Z^{N_i} \quad \text{for } i = 1, \dots, n, \quad (8)$$

where $N_i := \{k \in V : \{k, i\} \in E\}$ is the neighborhood of the vertex $i \in V$ consisting of all vertices $k \in V$ adjacent to i . In Eq. (8) and throughout the paper, the notation O^U for a subset $U \subset V$ indicates that the single-qubit operator O acts on the vertices contained in U , e.g., for $U = \{2, 3\}$ and $|V| = 5$ we have $O^U = \mathbb{1}_1 \otimes O_2 \otimes O_3 \otimes \mathbb{1}_4 \otimes \mathbb{1}_5$.

By the above observation, every graph state is a stabilizer state. Conversely, any stabilizer state is local-Clifford-equivalent to a graph state, which we show in detail in Appendix A. Here, the set of *local Clifford operations* is defined as those local unitaries $U_1 \otimes \dots \otimes U_n$ mapping the Pauli group \mathcal{P}_n into itself.

The action of any local Clifford operation on a graph state can be decomposed into a series of so-called *local complementations* (LC) [Nes+04]. The local complementation LC_i of a graph $\Gamma = (V, E)$ with respect to a vertex $i \in V$ is obtained by replacing the subgraph induced by the vertex neighborhood N_i with its complement. For example, the graphs \boxtimes and \boxdot can be obtained from each other by a local complementation with respect to the bottom-left vertex. If a quantum channel \mathcal{N} is *covariant* with respect to local Clifford operations C , i.e.,

$$\mathcal{N}^{\otimes n}(C \cdot C^\dagger) = C \mathcal{N}^{\otimes n}(\cdot) C^\dagger \quad \text{for all local Cliffords } C,$$

then two graph states with LC-equivalent graphs Γ and $\Gamma' = LC_i(\Gamma)$ have the same coherent information, $I_c(|\Gamma\rangle, \mathcal{N}^{\otimes n}) = I_c(|\Gamma'\rangle, \mathcal{N}^{\otimes n})$. This holds in particular for the depolarizing channel $\mathcal{N}_{\mathbf{p}}$ with $\mathbf{p} = (1 - p, p/3, p/3, p/3)$, which is in fact covariant with respect to any local unitary operation. Hence, a maximization of the coherent information of a depolarizing channel over all graph states corresponds to a maximization over all stabilizer states.

Graph states form a rich class of multipartite quantum states with many appealing properties; in the following we briefly collect the most important ones needed in this paper. A comprehensive review of graph states and their properties can be found in [Hei+06].

Graph State Basis. A helpful tool for working with graph states is the so-called *graph state basis* for $(\mathbb{C}^2)^{\otimes n}$. For a given graph $\Gamma = (V, E)$ with associated graph state $|\Gamma\rangle$, it is defined as $\mathcal{G}(\Gamma) = \{|U\rangle\}_{U \subset V}$, where

$$|U\rangle = Z^U |\Gamma\rangle.$$

To see that $\mathcal{G}(\Gamma)$ is indeed a basis for $(\mathbb{C}^2)^{\otimes n}$, observe that for $U, U' \subset V$ with $U \neq U'$ we have $S = (U \cup U') \setminus (U \cap U') \neq \emptyset$. Hence, the operator $O = Z^U Z^{U'} = Z^S$ has at least one Z operator acting on some qubit, and consequently $\langle \Gamma | O | \Gamma \rangle = 0$ [Hel14]. This can be checked by noting that Z^i commutes with any controlled- Z gate involving the qubit i , and $CZ^2 = \mathbb{1}$. It then follows for $O = Z^S$ with $S \neq \emptyset$ that

$$\langle \Gamma | O | \Gamma \rangle = \langle + |^{\otimes n} \left(\prod_{e \in E} CZ_e \right) O \left(\prod_{e \in E} CZ_e \right) | + \rangle^{\otimes n} = \langle + |^{\otimes n} O | + \rangle^{\otimes n} = 0.$$

Partial Trace Formula. Consider a graph Γ_{AR} on a vertex set $V = V_A \cup V_R$ consisting of vertices (qubits) V_A and V_R belonging to quantum systems A and R , respectively. We denote by $\Gamma_{A \rightarrow A}$ the induced subgraph on V_A of Γ_{AR} . There is a simple formula for the mixed marginal $\rho_A = \text{tr}_R \Gamma_{AR}$, which should not be confused with the pure subgraph state $|\Gamma_{A \rightarrow A}\rangle$. To state the formula, we consider a block form of the adjacency matrix of Γ_{AR} (which we denote by the same symbol in slight abuse of notation):

$$\Gamma_{AR} = \begin{pmatrix} \Gamma_{A \rightarrow A} & \Gamma_{A \rightarrow R} \\ \Gamma_{A \rightarrow R}^T & \Gamma_{R \rightarrow R} \end{pmatrix}, \quad (9)$$

where $\Gamma_{A \rightarrow R}$ is a $|V_A| \times |V_R|$ -matrix encoding the edges between the vertex sets V_A and V_R . Throughout the paper, for an arbitrary set X we identify subsets $Y \subset X$ with binary vectors $\tilde{X} \in \mathbb{F}_2^{|X|}$, where $\tilde{X}_i = 1$ if $i \in X$ and 0 otherwise. In slight abuse of notation, we use the same symbol for subsets and binary vector representations. Abbreviating $\Gamma' = \Gamma_{A \rightarrow R}$, the following formula for the mixed marginal ρ_A on the vertices belonging to A is proved in [Hei+06]:

$$\rho_A = \frac{1}{2^{|V_R|}} \sum_{R' \subset V_R} |\Gamma' R'\rangle \langle \Gamma' R'|_A, \quad (10)$$

where $\Gamma' R' \subset V_A$ is understood as matrix multiplication modulo 2, and the state $|\Gamma' R'\rangle_A = Z^{\Gamma' R'} |\Gamma_{A \rightarrow A}\rangle$ is an element of $\mathcal{G}(\Gamma_{A \rightarrow A})$. Equation (10) shows that the marginal ρ_A of a graph state $|\Gamma_{AR}\rangle$ is diagonal in a preferred basis, the graph state basis $\mathcal{G}(\Gamma_{A \rightarrow A})$ of the subgraph $\Gamma_{A \rightarrow A}$.

2.2.2. Decoherence under Pauli Noise

We now review the decoherence properties of graph states under Pauli noise discussed in [Hei+05]. The crucial observation is that graph states translate any Pauli error into Z-type errors: If qubit i is affected by an X error, then

$$X^i |\Gamma\rangle = X^i S_i |\Gamma\rangle = X^i X^i Z^{N_i} |\Gamma\rangle = Z^{N_i} |\Gamma\rangle = |N_i\rangle,$$

where we used that $S_i = X^i Z^{N_i}$ stabilizes $|\Gamma\rangle$. Furthermore,

$$Y^i |\Gamma\rangle = -i Z^i X^i |\Gamma\rangle = -i Z^i |N_i\rangle = -i |N_i \oplus i\rangle,$$

where \oplus is addition modulo 2 and we identify a subset $U \subset V$ with its binary representation.

The above observation shows that under conjugation with a Pauli error $\sigma \in \mathcal{P}$ the graph state $|\Gamma\rangle\langle\Gamma|$ is mapped onto some $|U\rangle\langle U|$, where $|U\rangle$ is the element of the graph state basis $\mathcal{G}(\Gamma)$ corresponding to the subset $U \subset V$. This holds more generally for any n -qubit Pauli group element $\sigma_n \in \mathcal{P}_n$. Consider now a Pauli channel

$$\mathcal{N}_{\mathbf{p}}(\rho) = p_0 \rho + p_1 X \rho X + p_2 Y \rho Y + p_3 Z \rho Z,$$

defined in terms of a probability distribution $\mathbf{p} = (p_0, p_1, p_2, p_3)$. The above observation shows that under the i.i.d. action of a Pauli channel $\mathcal{N}_{\mathbf{p}}$ the channel output $\sigma = \mathcal{N}_{\mathbf{p}}^{\otimes n}(\Gamma)$ is diagonal with respect to $\mathcal{G}(\Gamma)$, i.e., $\sigma = \sum_{U \subset V} \lambda_U |U\rangle\langle U|$ for some probability distribution $\{\lambda_U\}_{U \subset V}$. A compact formula for the coefficients λ_U is derived in [Hei+05].

In this paper, we use a simple generalization of this formula to the situation where the (i.i.d. action of the) Pauli noise only affects a subset of the qubits. To this end, consider a graph $\Gamma_{\text{AR}} = (V_A \cup V_R, E)$ whose vertices are partitioned into $k \equiv k_A$ system vertices (qubits) V_A and k_R environment vertices (qubits) V_R . By the above discussion, the state $\sigma_{\text{BR}} = (\mathcal{N}_{\mathbf{p}}^{\otimes k} \otimes \text{id}_R)(\Gamma_{\text{AR}})$ is again diagonal in the graph state basis $\mathcal{G}(\Gamma_{\text{AR}})$,

$$\sigma_{\text{RB}} = \sum_{U \subset V_A \cup V_R} \lambda_U |U\rangle\langle U|. \quad (11)$$

Setting $q_i = p_i/p_0$ for $i = 1, 2, 3$, we have the following formula for λ_U with $U \subset V$ as a straight-forward generalization of the one reported in [Hei+05]:

$$\lambda_U = p_0^k \sum_{(U_1, U_2, U_3) \in \mathcal{M}(U)} q_1^{|U_1|} q_2^{|U_2|} q_3^{|U_3|}, \quad (12)$$

where $\mathcal{M}(U)$ is the set of all triples (U_1, U_2, U_3) satisfying

$$\begin{aligned} U_i &\subset V_A && \text{for } i = 1, 2, 3, \\ U_i \cap U_j &= \emptyset && \text{for } 1 \leq i < j \leq 3, \\ U &= \Gamma_{\text{AR}}(U_1 \oplus U_2) \oplus U_2 \oplus U_3. \end{aligned} \quad (13)$$

Here and throughout the paper, \oplus denotes addition mod 2. Note that in Eq. (13) we interpret the U_i as subsets of $V = V_A \cup V_R$, which amounts to padding the binary vectors of length k_A with k_R 0's. The resulting subset U has support in V_R whenever there is an edge in Γ_{AR} from a vertex in $U_1 \oplus U_2$ to a vertex in V_R (that is, the off-diagonal block $\Gamma_{A \rightarrow R}$ in Eq. (9) has an entry 1 in a row corresponding to a vertex in $U_1 \oplus U_2$).

2.2.3. The Partial Trace of a Decohered Graph State.

We now derive a formula for taking the partial trace of a graph state after the i.i.d. action of a Pauli channel. More precisely, let Γ_{AR} be a graph on a vertex set $V = V_A \cup V_R$ with $k \equiv k_A = |V_A|$ and $k_R = |V_R|$, and let $\mathcal{N}_{\mathbf{p}}$ be a Pauli channel. Then by the above discussion, we have

$$\sigma_{BR} = \mathcal{N}_{\mathbf{p}}^{\otimes k}(\Gamma_{AR}) = \sum_{U \subset V} \lambda_U |U\rangle\langle U|, \quad (14)$$

where the coefficients λ_U are given by Eq. (12).

We now apply the partial trace formula Eq. (10) to Eq. (14). Abbreviating $\Gamma' = \Gamma_{A \rightarrow R}$, and for a subset $U \subset V = V_A \cup V_R$ denoting by U_A its restriction to V_A , we have:

$$\begin{aligned} \sigma_B &= \text{tr}_R \sigma_{BR} \\ &= \sum_{U \subset V} \lambda_U \text{tr}_R |U\rangle\langle U| \\ &= \sum_{U \subset V} \lambda_U \text{tr}_R Z^U |\Gamma_{AR}\rangle\langle\Gamma_{AR}| Z^U \\ &= \sum_{U \subset V} \lambda_U Z^{U_A} \text{tr}_R (|\Gamma_{AR}\rangle\langle\Gamma_{AR}|) Z^{U_A} \\ &= \frac{1}{2^{k_R}} \sum_{\substack{U \subset V \\ R' \subset V_R}} \lambda_U Z^{U_A \oplus \Gamma' R'} |\Gamma_{A \rightarrow A}\rangle\langle\Gamma_{A \rightarrow A}| Z^{U_A \oplus \Gamma' R'}, \end{aligned} \quad (15)$$

where we used Eq. (10) in the last step. This can be rewritten as

$$\sigma_B = \sum_{W \subset V_A} \lambda_W |W\rangle\langle W|, \quad (16)$$

where

$$\lambda_W = \frac{1}{2^{k_R}} \sum_{(U, R') \in \mathcal{M}(W)} \lambda_U, \quad (17)$$

$$\mathcal{M}(W) = \{(U, R') : U \subset V, R' \subset V_R, U_A \oplus \Gamma' R' = W\}. \quad (18)$$

Similar to σ_{BR} above, the operator σ_B is again diagonal in a preferred basis, the graph state basis $\mathcal{G}(\Gamma_{A \rightarrow A})$ of the subgraph $\Gamma_{A \rightarrow A}$.

2.3. Coherent Information of Decohered Graph States

Recall from Section 2.1 that the coherent information of a quantum code $|\psi\rangle_{\text{AR}}$ and a channel $\mathcal{N}: A \rightarrow B$ is given by

$$I_c(\psi, \mathcal{N}) = S(\mathcal{N}(\psi_A)) - S(\mathcal{N}(\psi_{\text{AR}})). \quad (19)$$

Consider now a graph state $|\Gamma_{\text{AR}}\rangle$ defined in terms of a graph $\Gamma_{\text{AR}} = (V_A \cup V_R, E)$ with $k \equiv k_A = |V_A|$ and $k_R = |V_R|$, and a Pauli channel $\mathcal{N}_{\mathbf{p}}$. The formulas (11) for $\sigma_{\text{BR}} = \mathcal{N}_{\mathbf{p}}^{\otimes k}(\Gamma_{\text{AR}})$ and (17) for $\sigma_{\text{B}} = \text{tr}_R \sigma_{\text{BR}}$ show that these operators are diagonal with respect to the graph state bases $\mathcal{G}(\Gamma_{\text{AR}})$ and $\mathcal{G}(\Gamma_{\text{A}\rightarrow\text{A}})$, respectively. These observations yield a ‘‘classical’’ algorithm for computing the coherent information Eq. (19), which is described in Algorithm 1 and explained in detail in the following paragraph. A MATLAB implementation of Algorithm 1 can be found at [Led19]. A similar algorithm to efficiently compute the coherent information of Pauli channels acting on quantum codes diagonal in the graph state basis of a given graph was derived in [CJ11].

To start Algorithm 1, we obtain in Line 3 the subsets $U_i \subset V$ satisfying the conditions in Eq. (13) as matrices whose columns are the binary vector representations of U_i . We store their cardinality $|U_i|$ in vectors u_i in Line 4, and determine the matrix $U = \Gamma(U_1 \oplus U_2) \oplus U_2 \oplus U_3$ in Line 5, assigning the first k_A rows of U to a matrix U_a in Line 6. The columns of U and U_a are binary vectors indexing subsets of V , and in Line 7 we convert them to decimal numbers indicating their corresponding position in coefficient vectors λ and μ for the states $\sigma_{\text{BR}} = (\mathcal{N}_{\mathbf{p}}^{\otimes k} \otimes \text{id}_R)(\Gamma_{\text{AR}})$ and $\omega_{\text{B}} = \mathcal{N}_{\mathbf{p}}^{\otimes k}(\Gamma_{\text{A}\rightarrow\text{A}})$, respectively.³ In Line 8 we compute a vector c whose entries are the coefficients $p_0^k q_1^{|U_1|} q_2^{|U_2|} q_3^{|U_3|}$ appearing in Eq. (12). Note that $q_1 \cdot q_2$ denotes entry-wise multiplication. If one of the q_i is zero, we eliminate it from the equation altogether in order to avoid computing 0^0 . For each element in c , we then add up the contributions to the coefficient λ_U indexed by the subset $U \subset V$ in Line 12. In the same loop, we also add up the contributions to the coefficient of U_A in the marginal σ_{B} in Line 13. This procedure computes the eigenvalues λ and μ of the states $\sigma_{\text{BR}} = (\mathcal{N}_{\mathbf{p}}^{\otimes k} \otimes \text{id}_R)(\Gamma_{\text{AR}})$ and $\omega_{\text{B}} = \mathcal{N}_{\mathbf{p}}^{\otimes k}(\Gamma_{\text{A}\rightarrow\text{A}})$, which are diagonal in their respective graph state bases $\mathcal{G}(\Gamma_{\text{AR}})$ and $\mathcal{G}(\Gamma_{\text{A}\rightarrow\text{A}})$.

³Note that the state ω_{B} , obtained from applying the channel to all qubits of the *pure* graph state $|\Gamma_{\text{A}\rightarrow\text{A}}\rangle$ on the vertices V_A corresponding to the subgraph $\Gamma_{\text{A}\rightarrow\text{A}}$, is an auxiliary state that does not feature itself in the formula for the coherent information.

In the second part of Algorithm 1 we compute the marginal σ_B from the intermediate state ω_B based on the following observation: using Eq. (15), we can write σ_B as

$$\begin{aligned}\sigma_B &= \frac{1}{2^{k_R}} \sum_{R' \subset V_R} Z^{\Gamma' R'} \left(\sum_{U \subset V} \lambda_U Z^{U_A} |\Gamma_A \rangle \langle \Gamma_A| Z^{U_A} \right) Z^{\Gamma' R'} \\ &= \frac{1}{2^{k_R}} \sum_{R' \subset V_R} Z^{\Gamma' R'} \left(\sum_{U \subset V} \lambda_U |U_A \rangle \langle U_A| \right) Z^{\Gamma' R'} \\ &= \frac{1}{2^{k_R}} \sum_{R' \subset V_R} Z^{\Gamma' R'} \omega_B Z^{\Gamma' R'}.\end{aligned}$$

We have

$$|U_A \rangle \langle U_A| \mapsto Z^{\Gamma' R'} |U_A \rangle \langle U_A| Z^{\Gamma' R'} = |U_A \oplus \Gamma' R' \rangle \langle U_A \oplus \Gamma' R'|,$$

which amounts to a permutation $A' \mapsto A' \oplus \Gamma' R'$ of the subsets $A' \subset V_A$ and hence a permutation of the coefficients λ_A of σ_B indexed by $A' \subset V_A$. To use this observation, in Lines 15 and 16 we define A' and R' to be matrices whose columns are the subsets of V_A and V_R in binary vector representation, respectively. Furthermore, in Line 17 we define Δ to be the subsets of V_A obtained from moving the V_R -subsets R' into V_A via $\Gamma' = \Gamma_{A \rightarrow R}$, i.e., along the edges connecting vertices in V_A to vertices in V_R .

For a fixed column Δ_j (corresponding to a particular subset of V_R moved into V_A), we then compute the map $A' \mapsto \Delta_j \oplus A'$ for all $A' \subset V_A$ in Line 19, and convert the resulting matrix to a decimal representation of the corresponding subsets in V_A in Line 20. The result of this operation is a vector V_{ind} which can be interpreted as a permutation in S_k with $k = k_A = |V_A|$. The corresponding permutations of the eigenvalue vector μ of ω_B are then added to λ_A (the eigenvalue vector of σ_B) in Line 21. Finally, in Line 23 we compute the Shannon entropies of the resulting eigenvalue vectors λ_A and λ to obtain the coherent information of σ_{RB} .

While Algorithm 1 is a ‘‘classical’’ algorithm computing the probability distributions corresponding to the diagonal states σ_{RB} and σ_B , it nevertheless has exponential scaling due to the loops over subsets of V_A and V_R . In the next section, we deal with this problem by exploiting the symmetries in the underlying graph Γ_{AR} with the goal of speeding up the computation of the coherent information.

Algorithm 1 Algorithm to compute the coherent information $I_c(|\Gamma\rangle, \mathcal{N}_{\mathbf{p}}^{\otimes k})$ of a graph state $|\Gamma\rangle_{\text{AR}}$ defined in terms of a graph $\Gamma = (V_A \cup V_R, E)$ with $k \equiv k_A = |V_A|$, $k_R = |V_R|$, and a Pauli channel $\mathcal{N}_{\mathbf{p}}$ with $\mathbf{p} = (p_0, p_1, p_2, p_3)$. We denote the adjacency matrix of the graph Γ by the same symbol, and set $\Gamma' = \Gamma_{A \rightarrow R}$ (see Eq. (9)). Unless stated otherwise, binary operations on vectors and matrices are carried out entry-wise. The subroutines GETUSUBSETS, GETUSUBSETSCARD, BINARYTODECIMAL, SUBSETS, PERMVECTOR, and SHANNONENTROPY are defined in Appendix B. Due to the concurrent write operations in lines 12 and 13, parallelising the trivial inner loop does not necessarily lead to a runtime speedup.

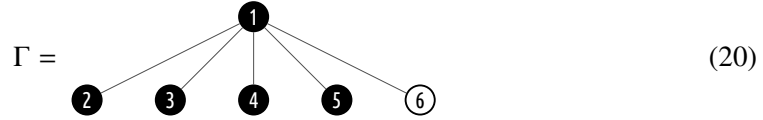
```

1: function  $I_c(|\Gamma\rangle, \mathcal{N}_{\mathbf{p}}^{\otimes k})$ 
2:    $q_i \leftarrow p_i/p_0$  for  $i = 1, 2, 3$ 
3:    $(U_1, U_2, U_3) \leftarrow \text{GETUSUBSETS}(k, k_R)$ 
4:    $(u_1, u_2, u_3) \leftarrow \text{GETUSUBSETSCARD}(k)$ 
5:    $U \leftarrow \Gamma(U_1 \oplus U_2) \oplus U_2 \oplus U_3$ 
6:    $U_a \leftarrow U_{1:k,:}$   $\triangleright U_{1:k,:}$  denotes the submatrix of the first  $k$  rows and all columns.
7:    $v \leftarrow \text{BINARYTODECIMAL}(U)$ ,  $w \leftarrow \text{BINARYTODECIMAL}(U_a)$ 
8:    $c \leftarrow p_0^k \left( q_1^{u_1} \cdot q_2^{u_2} \cdot q_3^{u_3} \right)$ 
9:    $\lambda \leftarrow 0^{2^{k+k_R}}$ ,  $\mu \leftarrow 0^{2^k}$   $\triangleright 0^n$  denotes a zero-vector of length  $n$ .
10:   $\lambda_A \leftarrow 0^{2^k}$ 
11:  for  $j \in \{1, \dots, |c|\}$  do
12:     $\lambda_{v_j} \leftarrow \lambda_{v_j} + c_j$ 
13:     $\mu_{w_j} \leftarrow \mu_{w_j} + c_j$ 
14:  end for
15:   $A' \leftarrow \text{SUBSETS}(k)$ 
16:   $R' \leftarrow \text{SUBSETS}(k_R)$ 
17:   $\Delta \leftarrow \Gamma' R' \pmod{2}$ 
18:  for  $j \in \{1, \dots, 2^{k_R}\}$  do
19:     $V_{\text{str}} \leftarrow \Delta_{:,j} \oplus A'$ 
20:     $V_{\text{ind}} \leftarrow \text{BINARYTODECIMAL}(V_{\text{str}})$ 
21:     $\lambda_A \leftarrow \lambda_A + 2^{-k_R} \text{PERMVECTOR}(\mu, V_{\text{ind}})$ 
22:  end for
23:  return  $\frac{1}{k}(\text{SHANNONENTROPY}(\lambda_A) - \text{SHANNONENTROPY}(\lambda))$ 
24: end function

```

3. Exploiting Graph Symmetries

One crucial observation for quantum codes based on graph states is that we can exploit the symmetries of the underlying graph state in the calculation of the coherent information. Consider e.g. a graph corresponding to a repetition code with $k_A = 5$ system vertices and $k_R = 1$ environment vertex colored in white:⁴



This graph is evidently invariant under any permutation of the vertices $\{2, 3, 4, 5\}$. How is this reflected in the calculation of the coherent information? We will see that we can indeed exploit the graph's symmetries to significantly speed up the performance of calculating the coherent information for a graph state.

In the following sections, we will summarize the graph-theoretic background necessary, and refer the reader to any standard textbook (e.g. [Tru93; Die10]) on graphs wherever we do not explicitly cite any sources.

3.1. Automorphisms of Colored Graphs

In this section we collect some fundamental facts of graph symmetries and introduce the notation used throughout the paper. To start, we analyse the 1-in-5-graph in Eq. (20). As mentioned, any interchange of the vertices 2 to 5 yields precisely the same graph; if we were to ignore the colors this symmetry would also include vertex 6. However, we use the color to signify the difference between a system vertex—which is passed through the channel—and an environment vertex, which is not; we do not want to be able to interchange it with the system.

The *colored graph's automorphism group* is the right concept to capture this type of symmetry. In the example in Eq. (20), it is generated by a list of transpositions,

$$\text{Aut}(\Gamma) = \langle (23), (34), (45) \rangle \subset S_6, \quad (21)$$

where S_6 is the finite symmetric group of 6 elements. More generally, for a graph state we define it to be as follows.

⁴For brevity and consistency we will call this a 1-in-5 code; see Section 5.1.

Definition 1 (Automorphism Group for Graph State). *Let $|\Gamma\rangle$ be a graph state with underlying graph $\Gamma = (V, E)$ and $|V| = k_{AR}$, such that the vertex set $V = (v_1, \dots, v_{k_{AR}})$ has a specific ordering. Let $1 \leq k_A < k_{AR}$. Then*

$$\text{Aut}(|\Gamma\rangle, k_A) := \text{Aut}(\Gamma(k_A)),$$

where $\Gamma(k_A)$ is the vertex-colored graph of Γ such that

$$\text{color}(v_i) = \begin{cases} \text{black} & i \leq k_A \\ \text{white} & \text{otherwise.} \end{cases}$$

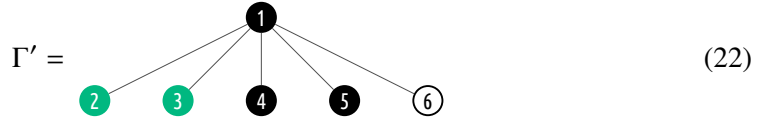
We note that it is generally *not* the case that, just by coloring the environment vertices in distinct colors, the automorphism group factors across this bipartition; a counterexample is a graph state such as



with automorphism group $\langle(45)(23)\rangle$.

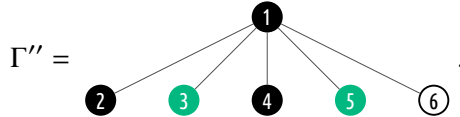
3.2. Cosets of Graph Colorings

Let $\Gamma = (V, E)$ be a graph. To visualize a vertex subset $U \subset V$ we can give them a distinct color, just as we marked the subset of environment vertices in white in the last section. We give an example; if Γ is the six-vertex graph in Eq. (20) with the environment vertex marked in white, we want to select the vertex subset $U = \{2, 3\}$, which yields the following colored graph:



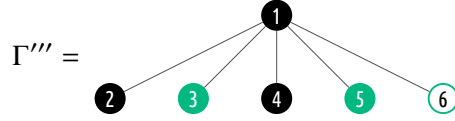
To simplify notation, we write $\Gamma' = \Gamma(k_A, U)$, indicating that all vertices with index beyond k_A are white, and further all vertices in U have an extra set color.

However, under the graph's automorphism group $\text{Aut}(\Gamma(k_A)) = \langle(23), (24), (25)\rangle$ which allows us to treat the originally-black leaves $2, \dots, 5$ interchangeably, this subset coloring is equivalent to marking a different set of vertices such as



Here, we call two colorings equivalent if there exists a permutation $\tau \in \text{Aut}(\Gamma(k_A))$ such that $\tau(\Gamma') = \Gamma''$, where we define $\tau(\Gamma') := (V', E')$ with $V' = \{v_{\tau(i)} : v_i \in V\}$, and E' is defined analogously.

In this case we only wanted to color a subset in one color; in general, we can obviously allow multiple colors, in which case the graph coloring is simply a list of associations such as $\{2 \rightarrow 1, 3 \rightarrow 1, 4 \rightarrow 2\} \subset V \times \mathbb{N}$. The second element is an arbitrary indexing of colors (e.g. 1 is green, 2 is blue etc.). As mentioned, applying a coloring to a graph $\Gamma(k_A, C)$ is then done in such a way that the environment vertices are marked in a different color. Continuing with our example, this means that if $6 \in U$, then



where the environment vertex is filled out in white. As vertex 6 cannot be interchanged with any of the system vertices, Γ''' is not equivalent to Γ' . For consistency with the graph state terminology and clarity of presentation we will generally call a graph $\Gamma(k_A)$ *uncolored*—despite the fact that the environment vertices are singled out; in contrast, $\Gamma(k_A, C)$ is *colored*. Likewise, we refer to the (strictly speaking 2-colored) automorphism group $\text{Aut}(\Gamma(k_A))$ simply as the automorphism group of a graph state.

Coming back to our example, how many different equivalent graph colorings for V are there? To answer this question rigorously, we define the following quantity.

Definition 2 (Number of Equivalent Graph Colorings). *For a graph state $|\Gamma\rangle$ with underlying graph $\Gamma = (V, E)$ and $k_A < k_{AR} = |V|$, and a coloring $C \subset V \times \mathbb{N}$, the number of equivalent graph colorings for C is defined as*

$$\#(\Gamma, k_A, C) := |\{\Gamma' : \exists \tau \in \text{Aut}(\Gamma(k_A)) \text{ s.t. } \tau(\Gamma') = \Gamma(k_A, C)\}|.$$

In our example graph Γ' in Eq. (22), it is easy to see that $\#(\Gamma, 5, U) = 6$. But how many is it in general? This question is answered most concisely using the theory of group actions. A (left) action of a group G on a set X is a map $G \times X \ni (g, x) \mapsto g \cdot x$ satisfying $g \cdot (h \cdot x) = (gh) \cdot x$ for all $g, h \in G$, $x \in X$, and $e \cdot x = x$ for all $x \in X$, where $e \in G$ denotes the identity element in G . Whenever there is no source of confusion we simply write $gx \equiv g \cdot x$ for $g \in G$ and $x \in X$. We use the notation ${}_G X$ to indicate an action of G on X . For each $x \in X$ we define the *orbit* $Gx = \{gx : g \in G\}$ and the *stabilizer group*

$G_x = \{g \in G : gx = x\} \leq G$. We will make heavy use of the *orbit-stabilizer theorem* relating the two objects:

$$|Gx| = [G : G_x] = \frac{|G|}{|G_x|}, \quad (23)$$

and the elements of the orbit Gx are in bijective correspondence with a transversal of G_x in G . The set of orbits of points $x \in X$ under the action of G forms a partition of X which we denote by X/G .

Returning to graph colorings, we now consider the action of the group $G = \text{Aut}(\Gamma(k_A))$ on the set of graph colorings, and the stabilizer group $\text{Stab}(G, C) = G_C$ of a coloring C , i.e., the subgroup of G such that for all elements $\tau \in \text{Stab}(G, C)$ we have $\tau(C) = C$. The orbit-stabilizer theorem (23) now gives the following result.

Lemma 3. *Let $G = \text{Aut}(\Gamma(k_A))$ be the graph state's automorphism group. Then $\#(\Gamma, k_A, C) = |G|/|\text{Stab}(G, C)|$.*

Now that we know which graph colorings are equivalent, we are interested in the question whether we can enumerate all non-equivalent graph colorings. To formalize this, let $[c]^k$ be the set of all possible colorings of a graph Γ , where either $k = k_A$ (padded with no colors for the environment vertices), or $k = k_{AR}$. In view of the discussion of group actions above, we observe the following:

Lemma 4. *The set $\Xi(\Gamma, c, k)$ of non-equivalent colorings of c colors of a graph Γ is equal to the set of orbits of all colorings under the action of the graph state's automorphism group. Formally,*

$$\Xi(\Gamma, c, k) = [c]^k / \text{Aut}(\Gamma(k_A)).$$

We will often drop the third argument in the coloring set $\Xi(\Gamma, c) = \Xi(\Gamma, c, k_A)$ for brevity. Can we explicitly list all colorings for a graph state? It turns out that this question is closely linked to the active field of research of finding canonical representatives within each coset $\omega \in \Xi(\Gamma, c, k)$. This in turn is closely related to the field of graph isomorphisms, i.e. deciding whether two graphs Γ_1 and Γ_2 are equivalent under vertex permutations.

Definition 5 (Canonical Map and Coloring). *A canonical map for a graph Γ and c colors is a map*

$$\text{Can}: \Xi(\Gamma, c, k) \longrightarrow [c]^k,$$

such that $\text{Can}(\omega) = C_\omega$ for some coloring $C_\omega \in \omega$; we also call $\text{Can}(\omega)$ a canonical representative of the coloring ω . A canonical coloring is then the set of canonical representatives $\text{Can}(\Xi(\Gamma, c, k))$.

Given the last definition, an immediate question is whether we can efficiently list such a canonical coloring. Assuming a canonical map Can is available, it is obvious that for small graphs a filtering approach—iterating over all colorings $[c]^k$ and collecting the results in a hash set—is a viable option. However, this method scales exponentially in k_A : for e.g. $c = 4$ and $k = 20$, we have to filter 2^{40} different colorings. Another caveat is that this only works as long as the hash set of previously-determined cosets fits into the working memory, as we need to check whether a specific coloring has been found before.

A different problem is that Can might not, in fact, be available, or very expensive to compute (as in our case). It turns out that a series of significantly faster algorithms are known to address this problem: one by McKay and Piperno [MP14] which we call VCoLG ,⁵ and another algorithm by Borie [Bor13] which we denote SGSCoLG . Both are based on checking whether a coloring is already canonical or not, which is much easier from a computational perspective. For the sake of completeness we explain SGSCoLG in detail in the following section.

3.3. Canonical Colorings

The algorithm SGSCoLG described by Borie [Bor13] is based on introducing a total order on the colorings, which provides a way of testing whether a specific choice of order is already maximal. In conjunction with a breadth-first search respecting said order, this allows for early pruning of what is essentially a backtracking algorithm. Because it is significantly more memory-efficient to traverse the colorings depth-first, we alter the algorithm by exchanging the tree traversal method in the following; the rest of the procedure closely follows [Bor13]. As Borie mentions, this algorithm is an extension of McKay’s canonical graph labeling algorithm described in [HR09].

The automorphism group $G = \text{Aut}(\Gamma)$ of a graph $\Gamma = (V, E)$ is a permutation group, i.e. a subgroup of the symmetric group $S_{|V|}$. For a permutation group $G \leq S_n$ acting on $[n]$, a *base* is a set of points $\{\beta_1, \dots, \beta_m\}$ with $\beta_j \in [n]$ such that the identity in G is the only group element stabilizing all base points. Given a base $\{\beta_1, \dots, \beta_m\}$, we can define the *group stabilizer chain*

$$\{\text{id}\} = G_m < G_{m-1} < \dots < G_1 < G_0 = G \quad (24)$$

where

$$G_i = \{\tau \in G : \tau(\beta_j) = \beta_j \ \forall j \leq i\}.$$

⁵This is derived from nauty’s toolkit [MP14], which provides a program for enumerating vertex colorings.

In words, the stabilizer chain is a sequence of nested subgroups of G , such that each subgroup G_i leaves all base points β_1, \dots, β_i invariant. Since the action base $\{\beta_i\}$ is not unique the stabilizer chain is not unique either; unless stated otherwise, in the following we use the *standard base* $\{\beta_1, \dots, \beta_n\}$ with $\beta_i = i$, where i corresponds to the graph's i^{th} vertex, and such that the system vertices are sorted to the front.

As shown in the definition below, a stabilizer chain naturally yields a list of generators for the group which respect a specific ordering of the points called a Strong Generating Set. It has many uses in computational group theory algorithms and can be efficiently calculated using the Schreier-Sims algorithm [Ser03].

Definition 6 (Strong Generating Set). *Let G be a finite permutation group with stabilizer chain corresponding to some action base $\{\beta_i\}$ as in Eq. (24). Then a strong generating set $S \subseteq G$ is a set of generators for G satisfying $\langle S \cap G_i \rangle = G_i$, i.e., the i^{th} stabilizer in the chain is generated by $S \cap G_i$ for all i . The strong generating set inherits its base-dependence from the stabilizer chain.*

Using such a strong generating set, it is straightforward to calculate the orbit of a point β_{i+1} under the action of the stabilizer group G_i . In turn, this allows us to determine a transversal for G_i/G_{i+1} in each link of the stabilizer chain in Eq. (24) relative to a base $\{\beta_1, \dots, \beta_m\}$. This can be obtained via the bijection between the orbit $\{\sigma\beta_{i+1} : \sigma \in G_i\} = \{o_1, \dots, o_l\}$ and the cosets of G_i/G_{i+1} : a transversal of the latter is obtained from cycling through o_1, \dots, o_l and for each o_j selecting an element $\tau_{ij} \in G_i$ such that $\tau_{ij}\beta_{i+1} = o_j$ for all $1 \leq j \leq l = [G_i : G_{i+1}]$.

For our applications, we specialize this to the standard base $\beta_i = i$: For all $j \geq i + 1$ in the orbit $G_i(i + 1)$, we can track which permutation $\tau_{ij} \in G_i$ maps $i + 1 \mapsto j$, i.e. $\tau_{ij}(i + 1) = j$. This yields a transversal for G_i/G_{i+1} with respect to the stabilizer chain of G relative to the standard base, which is an example of a so-called Strong Generating System.

Definition 7 (Strong Generating System). *A strong generating system for a finite permutation group with respect to some base $\{\beta_i\}$ is a list of transversals for a strong generating set of the stabilizer chain $\{G_i\}_i$ for base $\{\beta_i\}$. For the standard base $\beta_i = i$ we denote the strong generating system with $T(G) := \{t_1, \dots, t_k\}$, where each t_i is a list of permutations representing the transversal of G_i/G_{i+1} .*

A strong generating system will be tremendously useful when deciding whether a graph coloring is already canonical. In order to exploit it, we define a specific canonical function Can_{lex} : a coloring is canonical under Can_{lex} if it is maximal within its coset, with respect to the natural lexicographical ordering.

Algorithm 2 Is given graph coloring canonical with respect to Can_{lex} ?

```

1: function IsCANONICAL( $C \in \mathbb{N}^{\times k}$ ,  $T = \{t_1, \dots, t_k\}$  : transversal)
2:    $queue \leftarrow \{c\}$ 
3:   for  $i \in \{1, \dots, k\}$  do
4:      $set \leftarrow \{\}$ 
5:     for  $w \in queue$  do
6:       for  $child \in t_i \cdot w$  do                                 $\triangleright$  elementwise image of  $w$ 
7:         if  $c_{1:i} <_{\text{lex}} child_{1:i}$  then                             $\triangleright$  compare prefix of length  $i$ 
8:           return false
9:         end if
10:        if  $c_{1:i} = child_{1:i}$  then
11:           $set = set \cup \{child\}$ 
12:        end if
13:      end for
14:    end for
15:     $queue \leftarrow set$ 
16:  end for
17:  return true
18: end function

```

More specifically, for two strings of colors $v, w \in [c]^k$ we define $v <_{\text{lex}} w$ to be true if v comes before w with respect to a standard lexicographical ordering; \leq_{lex} is defined analogously. This ordering allows us to define a subroutine IsCANONICAL given in Algorithm 2, and enables us to determine if a given coloring is maximal with respect to Can_{lex} .

The intuition behind the algorithm is the following: Instead of computing the entire orbit of a given coloring C , we explore the orbit of the prefix of C up to some i , denoted $C_{1:i}$. In order to do so, we use a transversal t_i which can map any element at position $1, \dots, i$ to the i^{th} site, and require the indices C_i to be in little-Endian order. The way a stabilizer chain acts thus respects lexicographic number ordering, as the high-Endian bits (i.e. those with higher indices) are acted on and compared first.

For such a child element $child$, if $child_{1:i} >_{\text{lex}} C_{1:i}$, then already $child >_{\text{lex}} C$ and C was not canonical. If $child_{1:i} <_{\text{lex}} C_{1:i}$, we drop said child (because no permutation of the lower-Endian sites can ever make it larger than C) and proceed with the next element. Finally, if the prefixes are equal, we cannot say anything: we have to check further bits and enqueue the child to be processed by t_{i+1} etc. The benefit is clear: instead of calculating the full group G and checking all possible permutations of C under it—where we note that $|G|$ can

be enormous—we at most calculate the orbit of C under G ; the worst-case complexity of the algorithm thus has the promise to be significantly better than a naïve approach. Finally, we only need to keep track of a much smaller list of elements within each loop since we compare the element’s lexicographical order position-by-position; this means we try to only ever explore a small portion of the orbit of C .

The `IsCANONICAL` subroutine allows us to compute one representative of each canonical coloring using any backtracking technique. Borie used a breadth-first approach; we found that the unpredictable memory consumption and the associated memory allocations made this very inefficient in practice, so we replace it with depth-first pruning, given in Algorithm 3.

3.4. Canonical Images for Colored Graphs

A closely-related question to the one answered in `IsCANONICAL` in Algorithm 2—namely whether a given graph’s coloring C is already canonically-ordered—is whether we can map C to its canonical image, i.e. whether we can compute $\text{Can}(C)$ efficiently. At first glance this problem might look like it is independent of the graph Γ that gives rise to the permutation group G based on which we want to decide canonicity. In fact, it turns out that there is a package for `GAP` that does just that, based on a given permutation group [Jef+19]. The fundamental issue with this approach is that it is slower than making use of the extra information available—the graph—and isomorphism-based algorithms developed for the latter.⁶

The way we address this particular question is by drawing on graph isomorphism research, in particular the question of finding canonical images of colored graphs. Indeed, the way many graph isomorphism tests work in practice is by defining a canonical function on graphs, just like for colorings; we then have that two graphs $\Gamma_1 \simeq \Gamma_2$ if and only if their canonical images $\text{Can}(\Gamma_1) = \text{Can}(\Gamma_2)$, where for simplicity of notation we use `Can` as for colorings (which will never be ambiguous). For colored graphs, this of course implies both that the uncolored graph is canonically mapped, and that the coloring itself is in a specific order dictated by `Can`.

There is a series of tools that can be used to calculate such canonical graph images, foremost `nauty` and `Traces` by McKay and Piperno [MP14] and `bliss` by Junttila and Kaski [JK07], to which e.g. `IGraph` is linked.⁷ All of them use rigorously-proven underlying algorithms for finding canonical graph images, and we refer the reader to the respective papers for more details. While they solve a seemingly larger problem—graph isomorphisms instead of

⁶Of course another factor is that `GAP` is not as fast as C. However, even a re-implementation will be unlikely to perform as well as the graph-based approach we are about to introduce.

⁷Other tools are `saucy` and `conauto`; we used neither of them.

canonical colorings—we found this approach to be extremely fast even for moderately-sized graphs of up to 50 vertices. For a discussion on the tested graphs vs. size of their respective automorphism groups we refer the reader to Section 5.

For a given colored graph Γ with c colors, the canonical graph image approach gives us a map

$$\Gamma(k_A, C) \mapsto \text{Can}(\Gamma)(k_A, \text{Can}(C)) \quad \text{where} \quad \text{Can}(C) \in \Xi(\Gamma, c, k). \quad (25)$$

Here we assumed that the graph canonical map $\text{Can}(\Gamma)$ respects the partition between system and environment vertices (which one can usually achieve in practice). Eq. (25) thus induces a function

$$\text{CANONICALIMAGE}_{\Gamma(k_A)}: [c]^k \longrightarrow \Xi(\Gamma, k_A, c).$$

In case the graph in the context is clear we leave out the subscript, and we emphasize that, for our purposes, it will not be important that $\text{CANONICALIMAGE}(C) = \text{Can}_{\text{lex}}(C)$.

3.5. Homomorphic Group Actions

The final ingredient for the main result of this section, an algorithm computing the coherent information of a graph state that exploits the graph's symmetries, is the concept of homomorphic group actions. Following [Ker99], two group actions ${}_G X$ and ${}_H Y$ are called *homomorphic* if there exists an epimorphism $\eta: G \rightarrow H$ (i.e., a surjective group homomorphism) and a surjective map $\theta: X \rightarrow Y$, such that $\theta(\tau x) = \eta(\tau)\theta(x)$ for all $\tau \in G$ and $x \in X$. In this situation we have the following result.

Lemma 8 ([Ker99]). *Let ${}_G X$ and ${}_H Y$ be homomorphic group actions under θ , and let T be a transversal of Y/H . Then for all $\omega \in X/G$ there exists a unique $y \in T$ such that $\omega \cap \theta^{-1}(y) \neq \emptyset$.*

In other words, Lemma 8 says that if a map θ induces a group action homomorphism, it is never the case that elements from an orbit $\omega \in X/G$ are mapped to different orbits in Y/H . Of course it is still possible that two distinct $\omega, \omega' \in X/G$ are mapped to the same orbit $\tau \in Y/H$.

3.6. Coherent Information of Decohered Graph States with Symmetries

We will now use the theory developed over the last few sections to significantly reduce the amount of calculations going into calculating the λ -vectors in Eq. (12), which is our starting point. Take a graph $\Gamma = (V, E)$ where the vertex set is partitioned into system and

environment vertices $V = V_A \cup V_R$ with $|V_A| = k_A$, $|V_R| = k_R$, and for a Pauli channel with noise parameters (p_0, p_1, p_2, p_3) for which $p_0 = 1 - \sum_i p_i$ let $q_i := p_i/p_0$. For some $U \subset V$, we then have

$$\lambda_U = p_0^{k_A} \sum_{(U_1, U_2, U_3) \in \mathcal{M}(U)} q_1^{|U_1|} q_2^{|U_2|} q_3^{|U_3|},$$

where $\mathcal{M}(U)$ is defined in (13). Each triple (U_1, U_2, U_3) of disjoint subsets of V can be identified with a base 4 string via the map

$$(U_1, U_2, U_3) \xrightarrow{\sim} 1 \cdot U_1 + 2 \cdot U_2 + 3 \cdot U_3 \in [4]^{k_A}. \quad (26)$$

Let $G = \text{Aut}(\Gamma(k_A))$, which acts on $[c]^k$ by permuting indices. Using the identification in Eq. (26), we can define a map

$$\theta' : [4]^{k_A} \longrightarrow [2]^{k_{AR}}, \quad \theta'(s) = \Gamma(U_1 \oplus U_2) \oplus U_2 \oplus U_3 \quad (27)$$

where we implicitly assume the U_i to be padded with zeros within the brackets. This leads to the following technical lemma.

Lemma 9. *The map $\theta := \theta'(\cdot)|_A$ with θ' as in Eq. (27) induces a surjective map between $X := [4]^{k_A}/G$ and $Y := [2]^{k_A}/G$. Moreover, letting $n(U) := |\theta^{-1}(U)|$, we have for all $\tau \in Y$ that there exists an $n_0 \in \mathbb{N}$ such that $n(U) = n_0$ for all $U \in \tau$.*

Proof. We first note that θ is trivially surjective. Now let $\sigma \in G$. Then for some $s \simeq (U_1, U_2, U_3) \in [4]^{k_A}$, we have

$$\begin{aligned} \theta'(\sigma s) &= \Gamma(\sigma U_1 \oplus \sigma U_2) \oplus \sigma U_2 \oplus \sigma U_3 \\ &= \sigma \Gamma \sigma^{-1}(\sigma U_1 \oplus \sigma U_2) \oplus \sigma U_2 \oplus \sigma U_3 \\ &= \sigma [\Gamma(U_1 \oplus U_2) \oplus U_2 \oplus U_3] \\ &= \eta(\sigma) \theta'(s) \end{aligned}$$

for $\eta \equiv \text{id}$, which is a trivial epimorphism $\eta : G \mapsto G$. Hence, ${}_G[4]^{k_A}$ and ${}_G[2]^{k_A}$ are homomorphic group actions via θ , and the first claim follows from Lemma 8 and the fact that restriction $\cdot|_A$ and G commute.

We now prove the second claim of the lemma. By Lemma 8, for a fixed $\omega \in X$ we have $\theta(\omega) \subset \tau$ for a particular $\tau \in Y$. In fact, ω covers τ uniformly via θ . To see this, let x be a representative of ω , i.e., $\omega = \{\sigma x : \sigma \in G\}$, and let $\sigma \in G_x$, the stabilizer group of x . Then, $\theta(x) = \theta(\sigma x) = \sigma \theta(x)$, and hence also $\sigma \in G_{\theta(x)}$. This means that we have a chain of subgroups $G_x \leq G_{\theta(x)} \leq G$. Define now $p = [G_{\theta(x)} : G_x]$ and $q = [G : G_{\theta(x)}]$, and choose

left transversals $\{a_1 G_x, \dots, a_p G_x\}$ of G_x in $G_{\theta(x)}$ and $\{b_1 G_{\theta(x)}, \dots, b_q G_{\theta(x)}\}$ of $G_{\theta(x)}$ in G . Then $\{b_i a_j G_x : i \in [q], j \in [p]\}$ is a left transversal of G_x in G , and we have

$$\theta(b_i a_j x) = b_i a_j \theta(x) = b_i \theta(x) \quad \text{for all } i \in [q], j \in [p], \quad (28)$$

since $a_j \in G_{\theta(x)}$. The left transversal $\{b_i a_j G_x : i \in [q], j \in [p]\}$ is in bijective correspondence with the elements of ω by the orbit-stabilizer theorem, and so is $\{b_1 G_{\theta(x)}, \dots, b_q G_{\theta(x)}\}$ with the elements of τ . Eq. (28) shows that each element $b_i \theta(x)$ in τ is hit by exactly p elements $b_i a_j x$ in ω , and by we have

$$p = [G_{\theta(x)} : G_x] = \frac{|G_{\theta(x)}|}{|G_x|} = \frac{|G|/|G_x|}{|G|/|G_{\theta(x)}|} = \frac{|\omega|}{|\tau|}.$$

The proof is concluded by noting that for any $\tau \in Y$ and $U \in \tau$ we have

$$n(U) = \frac{1}{|\tau|} \sum_{\substack{\omega \in X \\ \omega \cap \theta^{-1}(U) \neq \emptyset}} |\omega|,$$

which again follows from Lemma 8. \square

This finally leads us to the following main result, which allows us to calculate λ_U in a symmetry-aware fashion.

Theorem 10. *Let $X := [4]^{k_A}/G$, and $Y' := [2]^{k_{AR}}/G$, where $G = \text{Aut}(\Gamma(k_A))$. Then Eq. (11) becomes*

$$\sigma_{\text{RB}} = \sum_{\tau \in Y'} \lambda_{\tau} \Pi_{\tau}$$

where

$$\lambda_{\tau} := p_0^{k_A} \sum_{\substack{\omega \in X \\ \theta'(\omega) \subset \tau}} \frac{|\omega|}{|\tau|} q_1^{|\omega|} q_2^{|\omega|} q_3^{|\omega|},$$

$$\Pi_{\tau} := \sum_{U \in \tau} |U \chi U|.$$

Proof. Since $|\sigma U_i| = |U_i|$, we have that $\lambda_{\sigma U} = \lambda_U$ for any $\sigma \in G$. We can therefore break up Y' into its covering by the cosets induced from the graph's automorphism group, i.e.

$$\sum_U \lambda_U |U \chi U| = \sum_{\tau \in Y'} \sum_{U \in \tau} \lambda_U |U \chi U| = \sum_{\tau \in Y'} \lambda_{\tau} \Pi_{\tau}.$$

By Lemma 9 we further know that, for each $\omega \in X$ mapped into $\tau \in Y'$ via θ' , there exist precisely $|\omega|/|\tau|$ elements in ω that map to the same $U \in \tau$. The claim follows. \square

We emphasize that the sum in the expression for λ_τ in Theorem 10 can be calculated over a traversal of X , and it is irrelevant which one we choose. Indeed, as mentioned at the end of Section 3.4, this precisely corresponds to the choice of coloring iterator which can follow a different ordering than the canonical map used to collect the nonzero terms within λ_τ .

Each λ_U in Eq. (11) thus occurs with a certain multiplicity $|\tau|$ and weight $|\omega|/|\tau|$; calculating the polynomial only once and keeping track of either quantity thus has the promise of saving a tremendous amount of computational time and memory. Effectively, we have constructed a sparse representation of σ_{RB} , which compresses identical expressions within the sum that defines λ_U into one term, *and* compresses repeated λ_U .

Yet in order to make full use of the symmetries, we also need to be able to efficiently calculate the partial trace of σ_{RB} . Showing how this can be achieved given the sparse representation of σ_{RB} is our next main result.

Theorem 11. *We use the same notation for X and Y' as in Theorem 10, and we let $Y = [2]^{k_A}/G$. Then Eq. (16) becomes*

$$\sigma_{\text{B}} = \sum_{\tau \in Y} \sum_{b \in [2]^{k_{\text{R}}}} \lambda_{\tau,b}^{\text{pre}} \Pi_\tau,$$

where

$$\lambda_{\tau,b}^{\text{pre}} := \sum_{\substack{\tau' \in Y' \\ \tau'|_{\text{A}} = h(\tau,b)}} \frac{|\tau'|}{|\tau'|_{\text{A}}} \lambda_{\tau'},$$

$$h(a,b) := \Gamma_{\text{A} \rightarrow \text{R}} \cdot b \oplus a \quad \text{for} \quad \Gamma = \begin{pmatrix} \Gamma_{\text{A} \rightarrow \text{A}} & \Gamma_{\text{A} \rightarrow \text{R}} \\ \Gamma_{\text{R} \rightarrow \text{A}} & \Gamma_{\text{R} \rightarrow \text{R}} \end{pmatrix}.$$

Proof. As a first step, we show that $h(a,b)$ transforms covariantly under the group's action on the system vertices. For some permutation $\sigma \in G$ and denoting with $[a, 0]$ and $[0, b]$ the padded vectors of a and b in $[2]^{k_{\text{AR}}}$, restriction to the system vertices $\cdot|_{\text{A}}$ and the group action of $G = \text{Aut}(\Gamma(k_{\text{A}}))$ commute, since system and environment vertices are never interchanged by construction of G . We thus have

$$\begin{aligned} h(\sigma(a), b) &= \Gamma_{\text{A} \rightarrow \text{R}} \cdot b \oplus \sigma(a) \\ &= (\Gamma \cdot [0, b] \oplus \sigma[a, 0])|_{\text{A}} \\ &= \sigma(\sigma^{-1}\Gamma \cdot [0, b] \oplus [a, 0])|_{\text{A}} \\ &= \sigma(h(a, \sigma^{-1}(b))), \end{aligned}$$

where we used $\sigma^{-1}\Gamma\sigma = \Gamma$ in the last step. This shows that demanding $\tau'|_A = h(\tau, b)$ is a well-defined expression, since

$$\sum_{b \in [2]^{k_R}} \lambda_{\tau, \sigma^{-1}(b)}^{\text{pre}} \Pi_\tau = \sum_{b \in [2]^{k_R}} \lambda_{\tau, b}^{\text{pre}} \Pi_\tau.$$

Our starting point is the partial trace algorithm for Eqs. (16) and (17), described in Section 2.3:

$$\sigma_B = \sum_{a \in [2]^{k_A}} \sum_{b \in [2]^{k_R}} \mu_{a,b}^{\text{pre}} |a\rangle\langle a| \quad (29)$$

where

$$\begin{aligned} \mu_{a,b}^{\text{pre}} &= \sum_{\substack{U \in [2]^{k_{AR}} : \\ U|_A = h(a,b)}} \lambda_U \\ &= \sum_{\tau' \in Y'} \sum_{\substack{U \in \tau' \\ U|_A = h(a,b)}} \lambda_U \\ &= \sum_{\substack{\tau' \in Y' : \\ \tau'|_A \ni h(a,b)}} \frac{|\tau'|_A|}{|\tau'|} \sum_{\substack{U \in \tau' : \\ U|_A = h(a,b)}} \lambda_U \\ &= \sum_{\substack{\tau' \in Y' : \\ \tau'|_A \ni h(a,b)}} \frac{|\tau'|}{|\tau'|_A} \underbrace{\frac{1}{|\tau'|} \sum_{U \in \tau'} \lambda_U}_{\equiv \lambda_{\tau'}}. \end{aligned}$$

Eq. (29) therefore becomes

$$\sigma_B = \sum_{\tau \in Y} \sum_{a \in \tau} \sum_{b \in [2]^{k_R}} \sum_{\substack{\tau' \in Y' \\ \tau'|_A \ni h(a,b)}} \frac{|\tau'|}{|\tau'|_A} \lambda_{\tau'}$$

and the claim then follows from covariance of h under G . \square

Finally, we translate Theorems 10 and 11 into an algorithm, using the coloring and canonical map sub-procedures developed in the last few sections. The pseudocode listing can be found in Algorithm 4.

Algorithm 3 All canonical colorings of c colors for graph Γ .

```

1: function COLORINGCHILDREN( $C \in \mathbb{N}^k, c \in \mathbb{N}$ )
2:   for  $i \in \{k, \dots, 1\}$  do
3:     if  $c_i \neq 0$  then
4:       if  $C_i < c$  then
5:         yield  $C + |i\rangle$  ▷ unit vector  $|i\rangle$  of length  $k_A$ 
6:       end if
7:     break
8:   end if
9:   yield  $C + |i\rangle$ 
10: end for
11: end function
12: function EXPLORECOLORINGSDF( $root \in \mathbb{N}^k, \text{PRUNE: function}$ )
13:   for  $child \in \text{COLORINGCHILDREN}(root)$  do
14:     if  $\neg \text{PRUNE}(child)$  then
15:       yield  $child$ 
16:       EXPLORECOLORINGSDF( $child, \text{PRUNE}$ )
17:     end if
18:   end for
19: end function
20: function CANONICALCOLORINGS( $\Gamma: \text{graph}, k_A \in \mathbb{N}, c \in \mathbb{N}$ )
21:    $G_A \leftarrow \text{Aut}(\Gamma(k_A))|_A$  ▷ action restricted to system vertices
22:    $T \leftarrow \text{STRONGGENERATINGSYSTEM}(G_A)$ 
23:    $root \leftarrow (0, \dots, 0)$  ▷ vector length  $k_A$ 
24:    $\text{PRUNE} \leftarrow \text{ISCANONICAL}(\cdot, T)$ 
25:   EXPLORECOLORINGSDF( $root, \text{PRUNE}$ )
26: end function

```

Algorithm 4 Symmetry-aware σ_{RB} and σ_B from Theorems 10 and 11. The λ hashtables have a tuple value type $P \times \mathbb{N}$, where the first component stores a polynomial in the q_i , and the second entry is its multiplicity. Parallelising the inner loops leads to a speedup since the concurrent operations are much less costly than the rest of the loop body.

```

1: function SYMMETRICLAMBDA( $\Gamma = (V, E) : \text{graph}, k_A \in \mathbb{N}$ )
2:    $\text{Can} \leftarrow (\cdot) \mapsto \text{CANONICALIMAGE}_{\Gamma(k_A)}(\cdot)$ 
3:    $\text{MULTIPLICITY} \leftarrow (\cdot) \mapsto \#(\Gamma, k_A, \cdot)$ 
4:    $\lambda, \lambda_a, \lambda^{\text{pre}} \leftarrow \{\}$  ▷ empty hashtables of tuples
5:   for  $(U_1, U_2, U_3) \in \text{CANONICALCOLORINGS}(\Gamma, k_A, 4)$  do
6:      $C = U_1 + 2U_2 + 3U_3$  ▷ coloring  $\in [4]^{k_A}$ 
7:      $m_4 \leftarrow \text{MULTIPLICITY}(C)$ 
8:      $U \leftarrow \Gamma(U_1 \oplus U_2) \oplus U_2 \oplus U_3$ 
9:      $m_2 \leftarrow \text{MULTIPLICITY}(U)$ 
10:     $m'_2 \leftarrow \text{MULTIPLICITY}(U|_A)$ 
11:     $idx \leftarrow \text{Can}(U)$ 
12:     $idx' \leftarrow \text{Can}(U|_A)$ 
13:     $p \leftarrow q_1^{|U_1|} q_2^{|U_2|} q_3^{|U_3|}$ 
14:     $(poly, m) \leftarrow \&\lambda_{idx}$  ▷ reference to hashtable entry
15:     $poly \leftarrow poly + \frac{m_4}{m_2} p$ 
16:     $m \leftarrow m_2$  ▷  $m = 0$ , or  $m = m_2$  already
17:     $(poly, m) \leftarrow \&\lambda_{idx'}^{\text{pre}}$ 
18:     $poly \leftarrow poly + \frac{m_4}{m_2} p$ 
19:     $m \leftarrow m'_2$  ▷  $m = 0$ , or  $m = m'_2$  already
20:  end for
21:  for  $a \in \text{CANONICALCOLORINGS}(\Gamma, k_A, 2)$  do
22:    for  $b \in [2]^{k_R}$  do
23:       $b2a \leftarrow \Gamma_{R \rightarrow A} \cdot b \oplus a$ 
24:       $idx \leftarrow \text{Can}(a)$ 
25:       $idx' \leftarrow \text{Can}(b2a)$ 
26:       $(poly, m) \leftarrow \&\lambda_{a,idx}$ 
27:       $(poly', m') \leftarrow \&\lambda_{idx'}^{\text{pre}}$ 
28:       $poly \leftarrow poly + poly'$ 
29:       $m \leftarrow m'$  ▷  $m = 0$ , or  $m = \text{MULTIPLICITY}(b2a) = m'$  already
30:    end for
31:  end for
32:  return  $(\lambda, \lambda_a)$ 
33: end function

```

4. Numerical Methods

As part of this project we developed a suite called `CoffeeCode` ([Bau19]) that implements Algorithms 2 to 4 described in Section 3. For special cases we replace the coloring iterator `SGSCoLg` with `VCOLg`; we found that depending on the automorphism group of the graph one or the other performs faster. A MATLAB implementation of Algorithm 1 in Section 2.2 (which is not exploiting graph symmetries) is available at [Led19].

To obtain our main results in Section 5 we used the suite `CoffeeCode` in the following way. Starting from a sparse expression for the λ_U as explained in Section 3.6, we can substitute the parameters q_i with the channel’s noise tuple \mathbf{p} to obtain a numerical expression for the coherent information. Of particular interest to us in this context is the question for which noise parameters of the associated Pauli channels the coherent information for a given graph state code is positive—i.e., determining the noise threshold above which the coherent information is zero.

As explained in Section 2.1, it is known that for specific Pauli channels various codes exhibit superadditivity. For 1-parameter families $x \mapsto \mathcal{N}_x$ of channels such as the depolarizing, BB84 or 2-Pauli channel, a special case of superadditivity of coherent information occurs when the *noise threshold*, the largest x such that $Q(\mathcal{N}_x) > 0$, is higher than the *single-letter threshold*, the largest x such that $I_c(\mathcal{N}_x) > 0$. Naturally, what we mean by “higher” in this context depends on the ordering of a family of channels. For instance, unbiased noise, modeled by the qubit depolarizing channel \mathcal{D}_p with parameters as given in Eq. (5), $\mathbf{p} = (1 - p, p/3, p/3, p/3)$ for some $p \in [0, 1]$, exhibits a natural order: a larger p implies a more likely X, Y or Z flip. Similarly the 2-Pauli channel $\mathbf{p} = (1 - p, p/2, 0, p/2)$ or BB84 channel $\mathbf{p} = ((1 - p)^2, p - p^2, p^2, p - p^2)$ are 1-parameter families with a natural ordering for which the definition of threshold makes intuitive sense.

The notion of threshold becomes more involved when looking at the general case of Pauli channels, where we have three independent noise parameters p_1, p_2 and p_3 . Can we define a meaningful ordering $\mathbf{p} > \mathbf{p}'$ for arbitrary Pauli channels? The most natural choice appears to be an unbiased one which allows a comparison only if the *ratio* of noise types is equivalent: if

$$\mathbf{p}_x := (1 - x, xp_1, xp_2, xp_3) \quad \text{for} \quad \sum_i p_i = 1 \text{ and } x \in [0, 1], \quad (30)$$

then we define $\mathbf{p}_x \geq \mathbf{p}_y$ if and only if $x \geq y$, and similarly for strict inequality; if the p_i comprising \mathbf{p}_x and \mathbf{p}_y are not identical, the channels are incomparable. This introduces a partial order on the set of all Pauli channels, with the intuitive understanding that a larger x in \mathbf{p}_x simply means “more of the same noise”. Increased thresholds for this parametrization thus only occur within families of Pauli channels $\mathcal{N}_{\mathbf{p}_x}$ parametrized by x ,

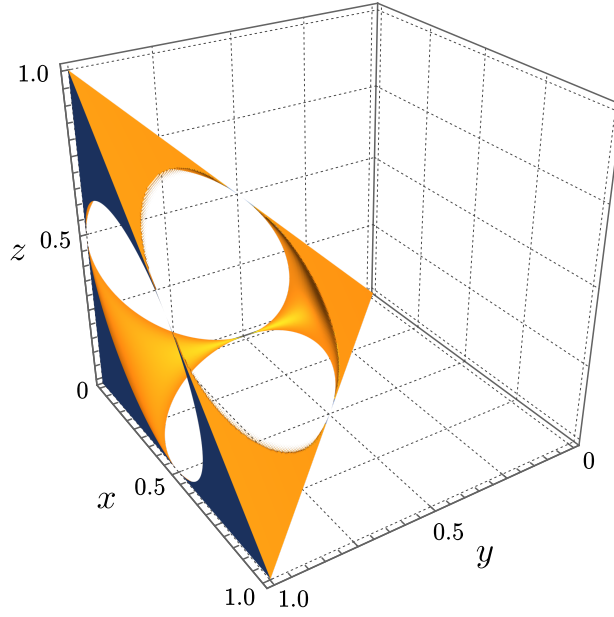


Figure 1: Set of all non-antidegradable Pauli channels, as derived from Eq. (31). The sector close to $p_i = 0$ in the lower left corner is equivalent to the three others: If e.g. z -type noise has $p_3 > 0.5$, i.e. z flips occur with more than 50% probability, one can first apply a z flip, mapping that sector back to $p_3 < 0.5$. For this reason we restrict the analysis of this paper to the sector around $p_i = 0$.

and there is one family for each probability distribution (p_1, p_2, p_3) . Note that the above partial ordering includes the depolarizing channel $x \mapsto (1 - x, x/3, x/3, x/3)$ and the 2-Pauli channel $x \mapsto (1 - x, x/2, 0, x/2)$, but *not* the BB84-channel $x \mapsto ((1 - x)^2, x - x^2, x^2, x - x^2)$.

4.1. Visualizing Thresholds

In order to visualize the coherent information threshold for a graph state code $|\Gamma\rangle$ for all possible Pauli channels, we chose the parametrization given in Eq. (30). The advantage of this parametrization is that for fixed p_1, p_2, p_3 the family $x \mapsto \mathcal{N}_{\mathbf{p}_x}$ has a unique noise threshold in the interval $x \in [0, 1/2]$, which we prove in Lemma 12 below.

We first recall that every quantum channel $\mathcal{N}: A \rightarrow B$ can be written as $\mathcal{N}(\rho_A) = \text{tr}_E V \rho_A V^\dagger$ in terms of an isometry $V: \mathcal{H}_A \rightarrow \mathcal{H}_B \otimes \mathcal{H}_E$ and an auxiliary Hilbert space \mathcal{H}_E usually referred to as the environment. A complementary channel $\mathcal{N}^c: A \rightarrow E$ modeling the leakage of information to the environment is defined as $\mathcal{N}^c(\rho_A) = \text{tr}_B V \rho_A V^\dagger$. A channel $\mathcal{N}: A \rightarrow B$ is called *antidegradable*, if there exists another quantum channel $\mathcal{A}: E \rightarrow B$

such that $\mathcal{N} = \mathcal{A} \circ \mathcal{N}^c$. For antidegradable channels \mathcal{N} , we have $I_c(\sigma, \mathcal{N}) \leq 0$ for any quantum state σ , and also $Q(\mathcal{N}) = 0$, which we show in Appendix C.

For qubit-qubit quantum channels \mathcal{N} there is a necessary and sufficient condition for antidegradability [ML09; Che+14]. Specialized to Pauli channels $\mathcal{N}_{\mathbf{p}}$ with $\mathbf{p} = (p_0, p_1, p_2, p_3)$, this condition reads as follows (see e.g. [PP16]):

$$1 \geq 2 \left(p_0^2 + p_1^2 + p_2^2 + p_3^2 \right) - 8\sqrt{p_0 p_1 p_2 p_3} \quad (31)$$

Checking the inequality (31) in the whole probability simplex $\{(p_1, p_2, p_3): p_i \geq 0, \sum_i p_i = 1\}$ yields the region of antidegradable Pauli channels. Since the quantum capacity is zero for those channels, we focus our analysis on the region of non-antidegradable Pauli channels, depicted in Fig. 1. Evidently, this region consists of four symmetric ‘sectors’, and we focus in the following on the sector closest to the origin. This sector includes the set of channels $\mathcal{N}_{\mathbf{p}_x}$ for $x \in [0, 1/2]$, where $\mathbf{p}_x = (1 - x, xp_1, xp_2, xp_3)$ is the parametrization of Pauli channels introduced above. This follows since $\mathcal{N}_{\mathbf{p}_{1/2}}$ is antidegradable for any (p_1, p_2, p_3) (see Appendix C). Moreover, as x is increased from 0 to 1/2 the coherent information of a fixed input state has a unique root. This is shown in the following lemma, which we prove in Appendix C.

Lemma 12. *Let ψ_{RA} be a state with $I_c(\psi, \mathcal{N}_{\mathbf{p}_0}^{\otimes k}) = S(\psi_A) > 0$ and $I_c(\psi, \mathcal{N}_{\mathbf{p}_{1/2}}^{\otimes k}) < 0$. Then the function $x \mapsto I_c(\psi, \mathcal{N}_{\mathbf{p}_x}^{\otimes k})$ has at most one root in $[0, 1/2]$.*

This result serves the following purpose in our numerical analysis. Instead of evaluating $f(x) := I_c(|\Gamma_{AR}\rangle, \mathcal{N}_{\mathbf{p}_x}^{\otimes k_A})$ for an evenly-spaced set of points within the interval, the unique root promised by Lemma 12 allows us to use a technique borrowed from raymarching game engines, which essentially performs binary search to zone in on the single root of a function within the interval parametrizing the ray. Instead of a cost $\propto 1/\epsilon$ in the precision ϵ of the threshold this approach reduces the number of required steps to $\propto \log(1/\epsilon)$. For our purposes we chose $\epsilon = 2^{-20}$, which yields approximately 7.5 decimal digits of precision; higher accuracy is readily reached.

To cover all Pauli channels, we choose a high resolution covering of their parameter space derived from spherical coordinates,

$$(p_1, p_2, p_3) \in \{(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) : \theta, \phi \in 0, \delta, 2\delta, \dots, \pi/2\}.$$

For $\delta = 2^{-10}\pi$ this yields a net of 512×512 rays of length 1/2 starting at the origin, each for a separate channel family.⁸ The zero threshold of the coherent information thus yields a surface

⁸We remark that this parametrization distributes overproportionally many points towards the north pole of the parameter space where $\cos \theta$ is small, i.e. small Z error.

in the Pauli channel simplex, as e.g. depicted in Fig. 2. In order to visualize superadditivity, we additionally color each position on the surface depending on how far away the point lies from the single-letter threshold along the same ray. For a detailed explanation of the surface plot types see Figs. 2 and 5.

The single-letter threshold for a general Pauli channel $\mathcal{N}_{\mathbf{p}}$ with $\mathbf{p} = (p_0, p_1, p_2, p_3)$ is determined by the so-called *hashing bound* [Ben+96], which gives the optimal single-letter coherent information $I_c(\mathcal{N}_{\mathbf{p}})$ of a Pauli channel:

$$I_c(\mathcal{N}_{\mathbf{p}}) = 1 - H(\mathbf{p}), \quad (32)$$

where $H(\mathbf{p}) = -\sum_i p_i \log p_i$ is the Shannon entropy of \mathbf{p} . The coherent information in (32) is achieved by the graph state $|\Gamma\rangle$ where $\Gamma = \circ\text{---}\bullet$ and $\mathcal{N}_{\mathbf{p}}$ acts on the solid vertex. For the x -parametrization $x \mapsto \mathbf{p}_x = (1 - x, xp_1, xp_2, xp_3)$ with a fixed probability distribution (p_1, p_2, p_3) introduced in Eq. (30), the single-letter threshold of $\mathcal{N}_{\mathbf{p}_x}$ is determined by the (unique) root of

$$I_c(\mathcal{N}_{\mathbf{p}_x}) = 1 - H(\mathbf{p}_x) = 1 - h(x) - xH((p_1, p_2, p_3)),$$

where $h(x) = -x \log x - (1 - x) \log(1 - x)$ denotes the binary entropy.

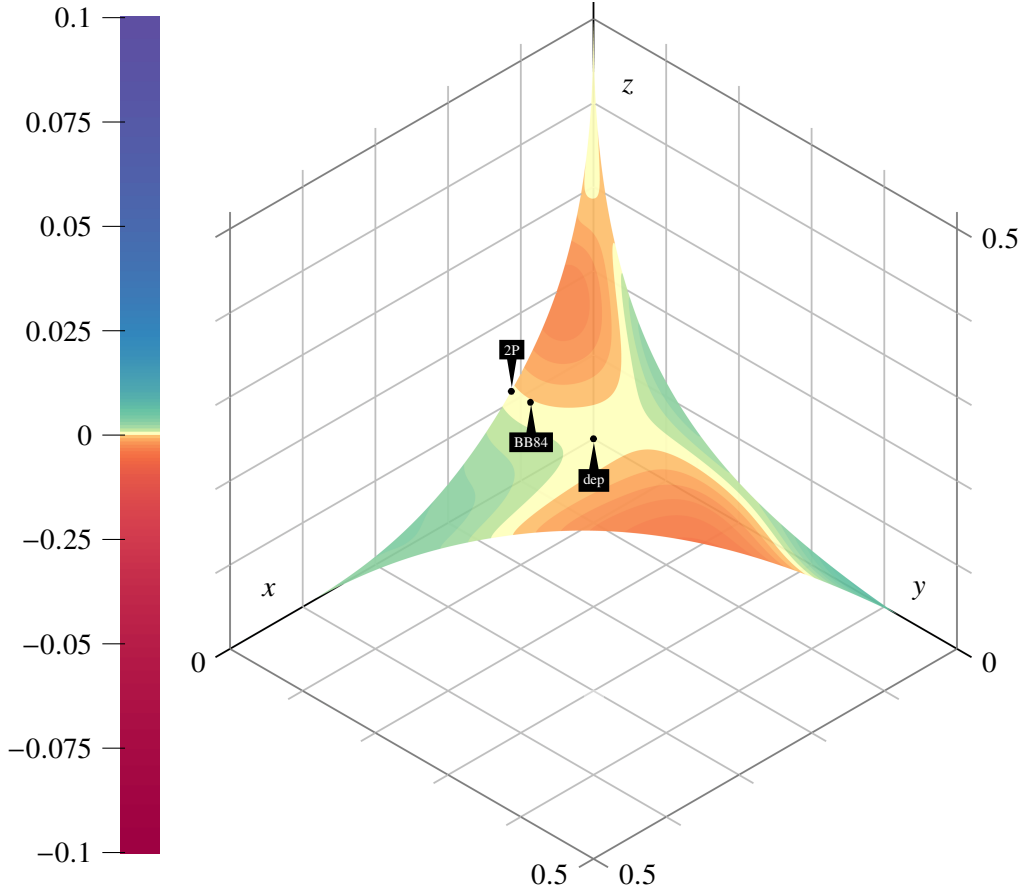


Figure 2: Surface plot of the tree graph code T_{15} vs. the 1-in-5 repetition code; the T_{15} -code is shown in Fig. 8, and the 1-in-5-code is shown in Eq. (20). Labeled are the points in the space of all Pauli channels where the families of depolarizing (dep), BB84 and 2-Pauli (2P) channels cut the threshold surface of the T_{15} code. The color indicates the superadditivity of the code family along the x -parametrized ray starting at the origin according to Eq. (30). The color scale for all plots of this type is identical throughout the paper, as well as the axes' extent; as such, to simplify the plots, we will generally leave out all but the axes labels.

5. Results

In this section we present our main results of the paper. Using the parametrization $x \mapsto (1 - x, xp_1, xp_2, xp_3)$ for some probability distribution (p_1, p_2, p_3) detailed in Section 4, we investigate error thresholds for the quantum capacity of Pauli channels in the entire Pauli channel simplex. Following the discussion in Section 2.1, this is achieved by choosing (families of) test input states ψ_{RA} on k_A input qubits and determining the supremum over all x for which $I_c(\psi, \mathcal{N}_{\mathbf{p}_x}^{\otimes k_A}) > 0$.

In the following subsections we investigate three families of code states defined in terms of graph states. The first one is the family of repetition codes (or GHZ states) $|0\rangle^{\otimes n} + |1\rangle^{\otimes n}$. Despite being rather simplistic error-correcting codes, repetition codes have long been known to exhibit superadditivity of coherent information; in particular, they increase the error threshold for Pauli channels such as the depolarizing channel [SS96; DiV+98; SS07; FW08; BL18] or the BB84 (or independent bit and phase flip) channel [SS07; FW08], and non-Pauli channels such as the dephasing channel [Led+18a]. The second family of codes we investigate comprises concatenated repetition codes, or *cat codes*, which are obtained from concatenating a Z -type repetition code with an X -type repetition code. With the right choice of repetition code lengths, these codes have been shown to increase the error thresholds of depolarizing and BB84 channel even further [DiV+98; SS07; FW08]. Both repetition codes and cat codes have large symmetry groups, enabling us to obtain substantial speed-ups in computing their coherent information using Algorithms 2 to 4 described in Section 3. Moreover, our analysis of these code families covers the *entire* Pauli channel simplex, in contrast to previous works that only dealt with particular channels. Finally, we identify a new family of codes based on tree graphs. Their large symmetry groups again provide a speed-up via our algorithm, and the resulting error thresholds are better than the ones obtained from repetition and cat codes in large regions of the Pauli channel simplex. In addition to determining the error thresholds of the above code families, we analyze their rate superadditivity below the threshold as well.

We also carry out an exhaustive search on all graphs with up to 6 system vertices and 4 environment vertices. However, due to the number of possible graph configurations this exhaustive search quickly becomes infeasible both in terms of runtime and memory.

5.1. Repetition Codes

Repetition codes are the most simple type of error correcting code. Represented as a graph state, they correspond to star graphs with k_{AR} vertices; one ray of the star is the environment vertex. The 1-in-5 code in Eq. (20) is an example of a repetition code.

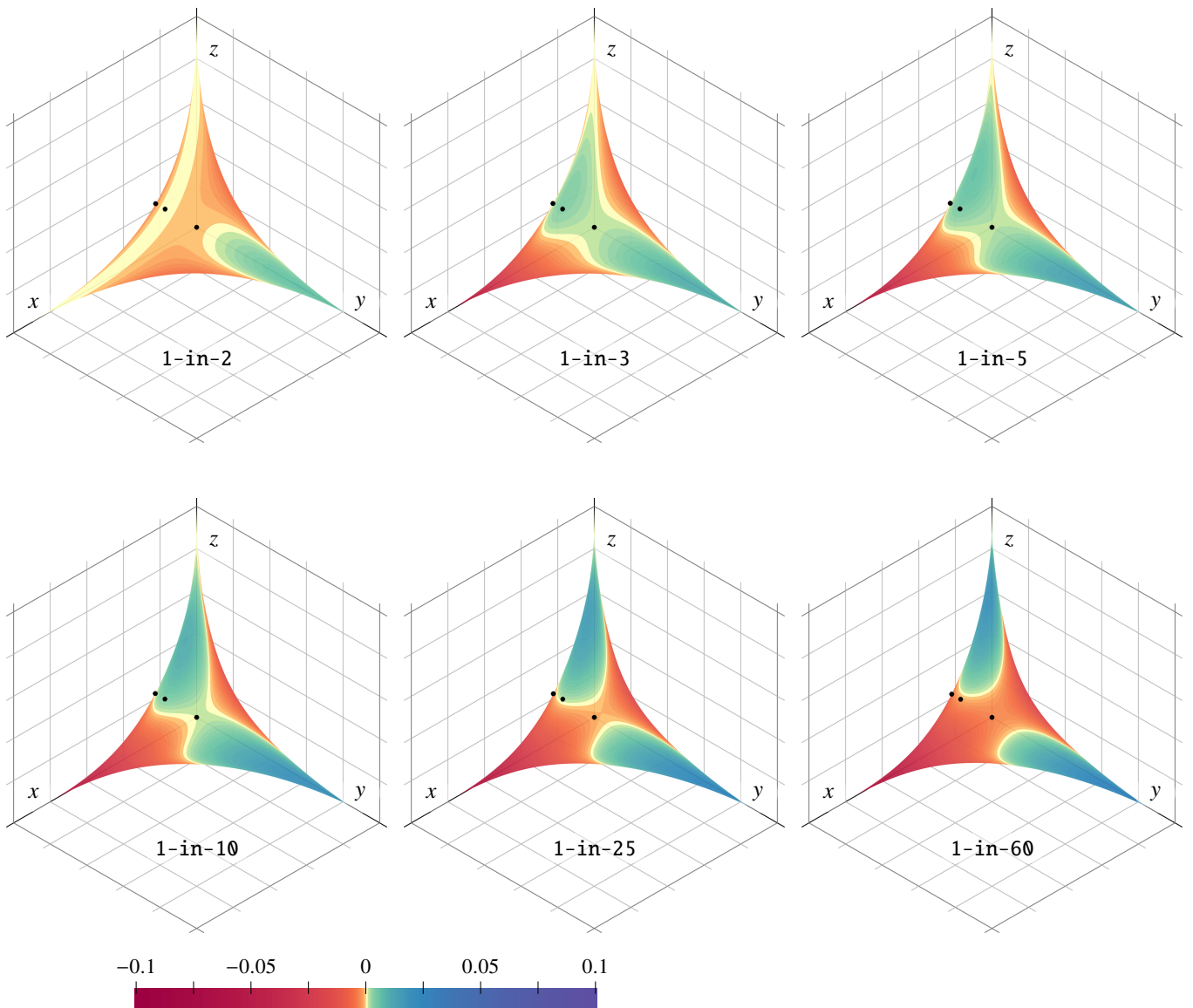


Figure 3: Thresholds for repetition codes; plotted are the codes 1-in- k_A for $k_A \in \{2, 3, 5, 10, 25, 60\}$ from top left to bottom right. The axes' extent is the interval $[0, 1/2]$, and the three black dots (from center to left) are the loci of depolarizing, BB84, and 2-Pauli channel families, respectively (for details see Fig. 2). The color scale indicates the distance between noise threshold and single-letter threshold.

It is immediately obvious that star graphs have a large symmetry group. If Γ_n denotes an 1-in- n repetition code, and using the same vertex enumeration as in Eq. (20)—i.e. such that the root has index 1, and the environment vertex has index n —its automorphism group is

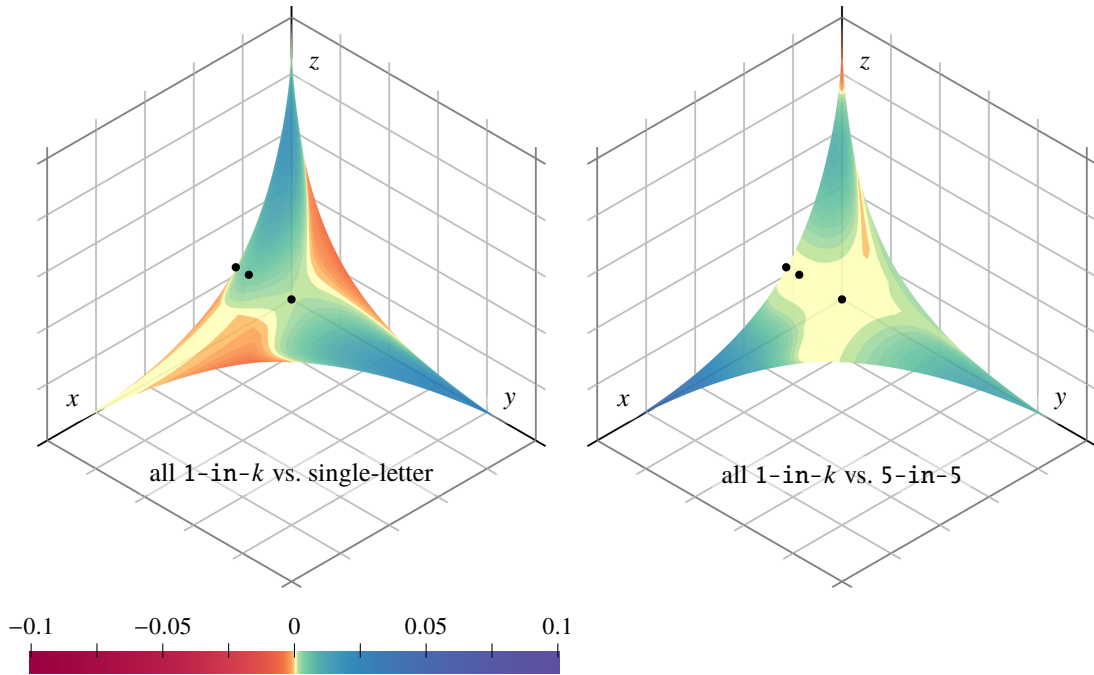


Figure 4: All repetition codes up to $k_A \leq 60$. The axes' extent is the interval $[0, 1/2]$, and the three black dots (from center to left) are the loci of depolarizing, BB84, and 2-Pauli channel families, respectively—for details see Fig. 2. Shown is the relative threshold of the best code in the code family vs. the single-letter threshold on the left, and vs. the 5-in-5 cat code (defined in Eq. (36)) on the right. In both cases, the color scale indicates the distance between the respective thresholds.

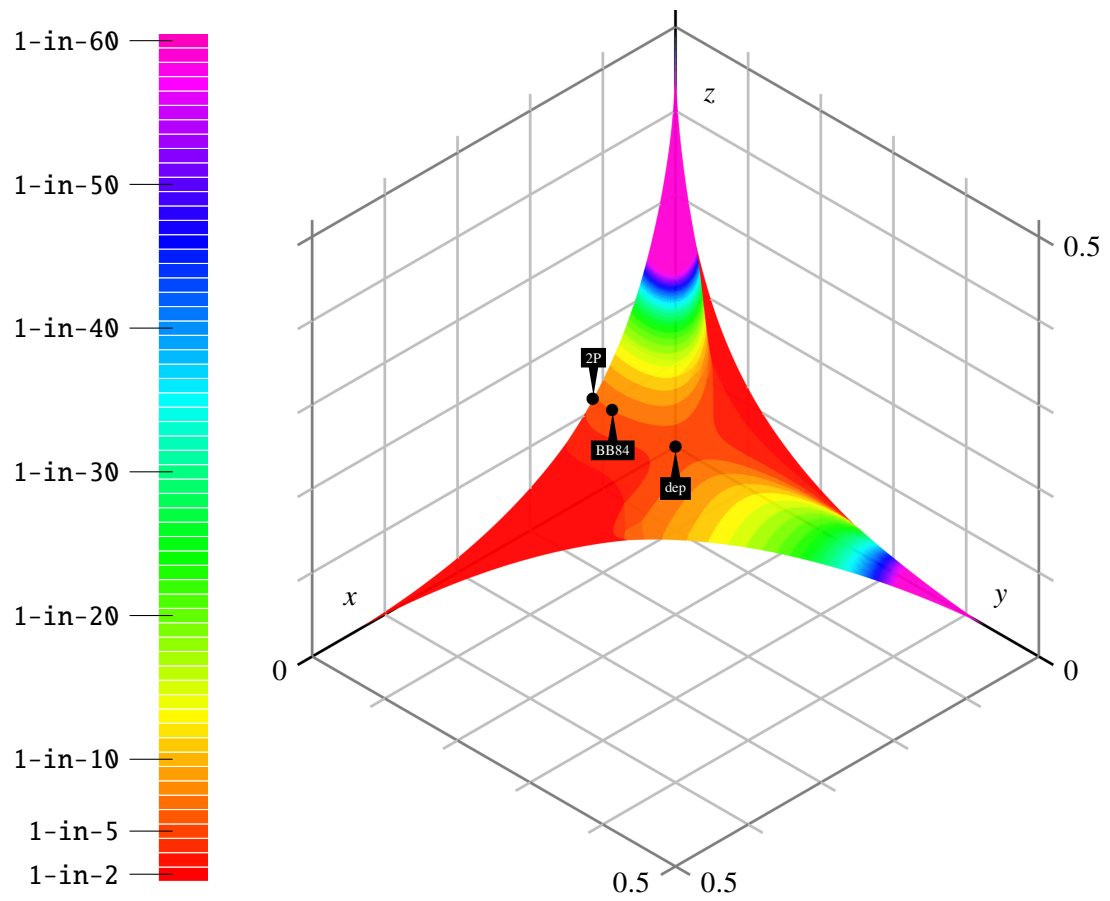


Figure 5: Continuation of Fig. 4. The plot shows the same surface as before, only this time indexed by colors referring to the highest CI threshold repetition code, starting at red for 1-in-2 up to magenta for 1-in-60.

given by $\text{Aut}(\Gamma_n) = S_{\{2, \dots, n-1\}}$, and hence $|\text{Aut}(\Gamma_n)| = (n-2)!$. It is this fact together with the very simple orderings of the canonical image and coloring algorithms in Section 3.6 that make repetition codes particularly amenable to our numerical methods, allowing us to evaluate codes on ≈ 100 system qubits in total without much difficulty.

In Fig. 3, we plot the coherent information threshold for various repetition codes. Visible is how for growing k_A the code shifts from superadditivity in the Y and Z -directions to superadditivity in the X and Z -directions. However, it is known that for unbiased noise—i.e. a depolarizing channel—the 1-in-5 code has the optimal threshold under all repetition codes. We also point out that while the smaller codes (apart from the 1-in-2 code) are not symmetric, for growing k_A the code family has an emerging approximate mirror symmetry across the plane dividing the X and Z axes.

Treating all repetition codes together as one “family” of codes, we compare the largest thresholds of this code family with the single-letter thresholds and a 5-in-5 code (see Section 5.2) in Fig. 4. The left-hand plot of Fig. 4 shows that repetition codes have better thresholds than the single-letter threshold for Z - and Y -biased noise and unbiased noise, but not for X -biased noise. Moreover, the right-hand plot shows that repetition codes provide advantage over the 5-in-5 code for all biased noise types, but not for the region around the depolarizing channel (unbiased noise), as observed in [DiV+98]. It is also interesting to see which repetition code is optimal for which type of noise shifts when the noise becomes more and more biased to one type, see Fig. 5.

We note here that the apparent difference between the three types of biased noise in Figs. 3 to 5 (as well as all plots in the subsequent sections) is merely due to the particular basis choice in the definition of graph states (see Section 2.2.1). To see this, note that for a given graph $\Gamma = (V, E)$ the stabilizer generators S_i of the associated graph state can be chosen as

$$S_i = R^i T^{N_i}$$

for any choice of $R, T \in \{X, Y, Z\}$ with $R \neq T$; different choices for R and T yield different relabelings of the axes in e.g. Fig. 3. In this paper, we use the canonical choices $R = X$ and $T = Z$.

5.2. Concatenated Codes

Concatenated codes, or *cat codes*, arise from concatenating a Z -type repetition code $|0\rangle^{\otimes n_1} + |1\rangle^{\otimes n_1}$ with an X -type repetition code $|+\rangle^{\otimes n_2} + |-\rangle^{\otimes n_2}$, where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. We call such a code an n_1 -in- n_2 code. The well-known 9-qubit Shor code [Sho95] is a 3-in-3 code. Cat codes are stabilizer codes and hence local-Clifford-equivalent to graph states; in our paper, we simply refer to any state of this equivalence class as cat code.

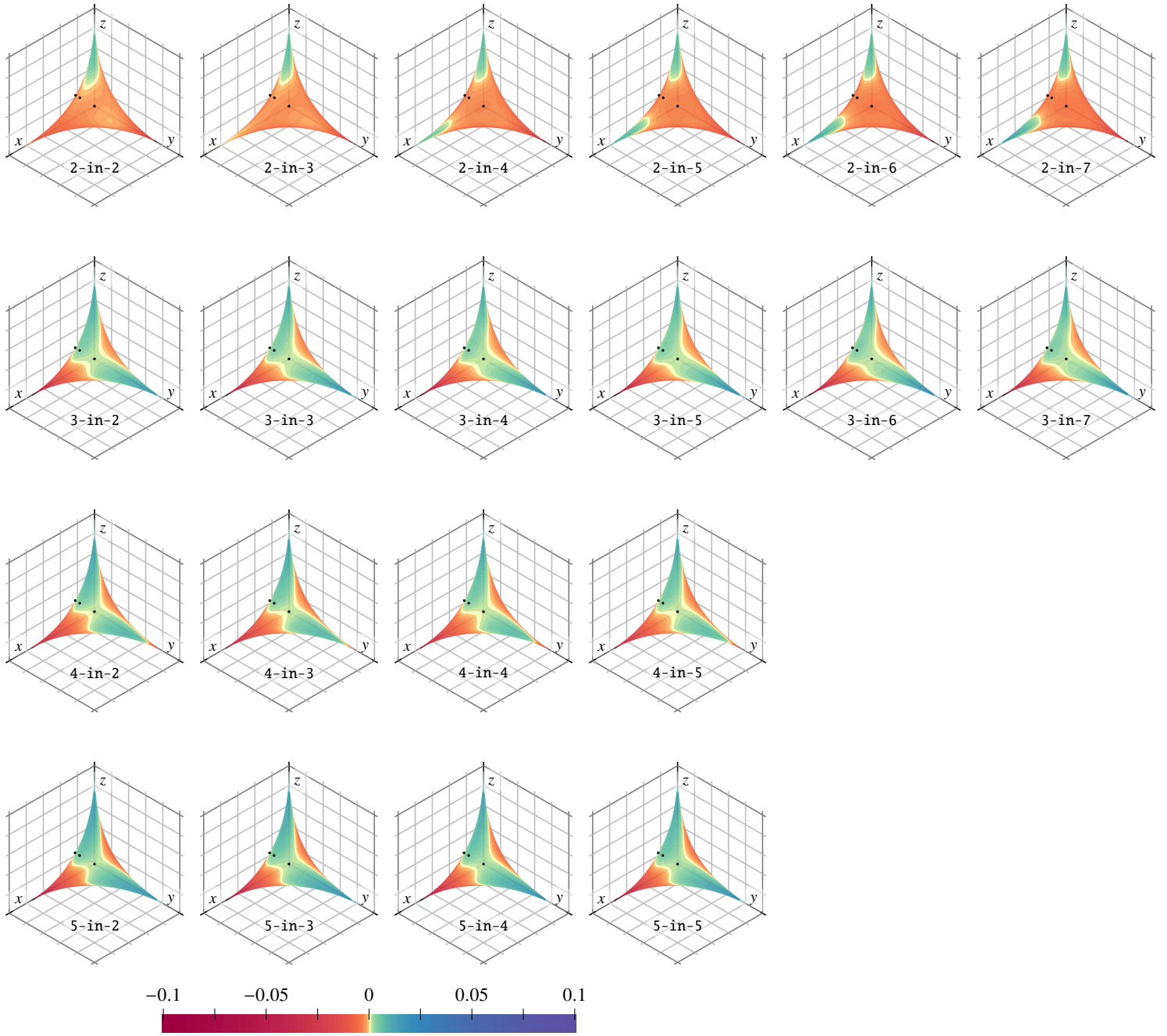


Figure 6: Cat codes 4×7 table. The axes' extent is the interval $[0, 1/2]$, and the three black dots (from center to left) are the loci of depolarizing, BB84, and 2-Pauli channel families, respectively—for more details see Fig. 2. The color scale indicates the distance between noise threshold and single-letter threshold.

Before giving the construction for the underlying graph of a cat code, we briefly review the procedure of concatenation of stabilizer codes. Let $[n_1, k]$ and $[n_2, 1]$ be qubit stabilizer codes with generators $S = \{s_1, \dots, s_{n_1-k}\}$ and $T = \{t_1, \dots, t_{n_2-1}\}$, respectively. We call S and T the *outer* and *inner* code, respectively. Furthermore, denote by \bar{X} and \bar{Z} the logical operators of the inner code T , and denote by $\{|\bar{0}\rangle, |\bar{1}\rangle\}$ the computational basis of the encoded qubit, i.e., the eigenstates of \bar{Z} corresponding to $+1$ and -1 , respectively. Then the concatenated stabilizer code $[n_1 n_2, k]$ is a stabilizer code whose generators are n_1 copies of the $n_2 - 1$ generators of the inner code acting on the physical qubits within blocks of size n_2 , plus the $n_1 - k$ generators of the outer code acting on the *logical* qubit of the inner code, for a total of $n_1(n_2 - 1) + n_1 - k = n_1 n_2 - k$ generators. The latter $n_1 - k$ stabilizers are obtained by the replacements $X \leftarrow \bar{X}$ and $Z \leftarrow \bar{Z}$ in every $s \in S$.

As an example, consider Shor's 9-qubit code, which is a $[3, 1]$ Z -repetition code with stabilizers $Z^1 Z^2, Z^1 Z^3$ and logical operators $\bar{X} = X^1 X^2 X^3, \bar{Z} = Z^1 Z^2 Z^3$, concatenated with a $[3, 1]$ X -repetition code with stabilizers $X^1 X^2, X^1 X^3$ and logical operators $\bar{X} = X^1 X^2 X^3, \bar{Z} = Z^1 Z^2 Z^3$. This results in a $[9, 1]$ code by the above discussion, with inner and outer code stabilizers

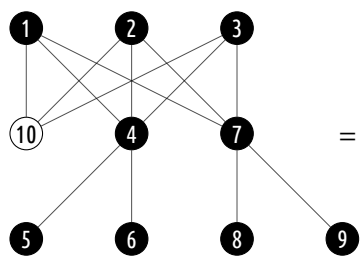
$$\begin{aligned} Z^1 Z^2, \quad Z^1 Z^3, \quad Z^4 Z^5, \quad Z^4 Z^6, \quad Z^7 Z^8, \quad Z^7 Z^9 \\ X^1 X^2 X^3 X^4 X^5 X^6, \quad X^1 X^2 X^3 X^7 X^8 X^9 \end{aligned} \quad (33)$$

and logical operators

$$\bar{X} = X^1 X^2 X^3 X^4 X^5 X^6 X^7 X^8 X^9 \quad \bar{Z} = Z^1 Z^2 Z^3 Z^4 Z^5 Z^6 Z^7 Z^8 Z^9. \quad (34)$$

To use the Shor code as an input state to $k_A = 9$ copies of a Pauli channel, we encode it in half of a maximally entangled state $|\phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_R |\bar{0}\rangle_A + |1\rangle_R |\bar{1}\rangle_A)$, where $|\bar{0}\rangle_A$ and $|\bar{1}\rangle_A$ are the ± 1 eigenstates of the logical \bar{Z} in (34), respectively. The state $|\phi\rangle$ is again a stabilizer state with stabilizer generators $X^1 \bar{X}^l$ and $Z^1 \bar{Z}^l$ (where we denote the logical qubit encoded in the 9-qubit code by l). Hence, we add the all X - and all Z -stabilizers (defined on 10 qubits) to the list in (33) and turn this stabilizer state into an LU-equivalent graph state using the procedure outlined in Appendix A. This yields the graph Γ_{Shor} together with its adjacency

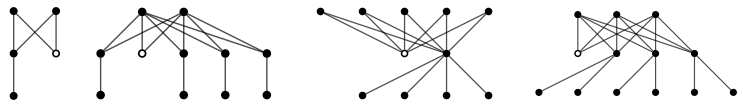
matrix (where we replaced 0 with \cdot for readability) for the graph state LU-equivalent to the Shor code:



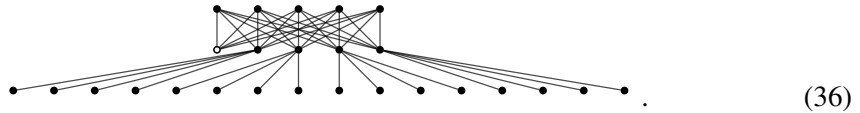
$$= \Gamma_{\text{Shor}} = \begin{pmatrix} \cdot & \cdot & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & 1 \\ 1 & 1 & 1 & \cdot & 1 & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & 1 & \cdot & \cdot & \cdot & \cdot & 1 & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ 1 & 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \quad (35)$$

The procedure above can be generalized to arbitrary n_1 -in- n_2 concatenated (“cat”) codes as follows. We take a single environment vertex v_R and disjoint sets of system vertices A with $|A| = n_1$, B with $|B| = n_2 - 1$, and C^i for $i = 1, \dots, n_2 - 1$ with $|C^i| = n_1 - 1$. We fully connect $\{v_R\}$ with A , and A with B , such that the vertices $\{v_R\} \cup A$ form a complete bipartite graph with B . Finally, we fully connect the i^{th} vertex b^i in B with C^i . Overall the graph has $n_1 + (n_2 - 1) + (n_2 - 1)(n_1 - 1) = n_1 n_2$ system vertices, and a single environment vertex, v_R . For $n_1 = 1$ this construction yields a star graph on $n_2 + 1$ vertices, and hence a 1-in- n_2 -cat code is a repetition code as defined in Section 5.1.

For the Shor code in Eq. (35) above, a 3-in-3-cat code, we have $A = \{1, 2, 3\}$, $B = \{4, 7\}$, and $C^1 = \{5, 6\}$, $C^2 = \{8, 9\}$. We give a series of other examples: the 2-in-2, 2-in-5, 5-in-2 and 3-in-4-cat codes have the graphs



Larger codes follow the same pattern; the 5-in-5 code, for instance, has 25 system vertices, and its graph is



It is evident that the symmetry group of cat codes factors in a clean fashion with respect to these vertex subsets: if Γ_{n_1, n_2} denotes the graph, we have

$$\text{Aut}(\Gamma_{n_1, n_2}) = S_A \times \left(S_{M^i} \cup \prod_{i=1}^{n_2-1} S_{C^i} \right),$$

where S_{M^i} denotes the symmetric group on the vertex set of the branches between B and the C^i , i.e. $M^i := \{\{b^j\} \cup C^i : j = 1, \dots, n_2 - 1\}$. A straightforward calculation thus gives the size of the automorphism group as $|\text{Aut}(\Gamma_{n_1, n_2})| = n_1!(n_2 - 1)!(n_1 - 1)!(n_2 - 1)^{n_1}$.

Fig. 6 shows a 4×7 -matrix plotting the threshold surfaces of n_1 -in- n_2 cat codes for $2 \leq n_1 \leq 5$ (along the rows) and $2 \leq n_2 \leq 7$ (along the columns). Note that 1-in- n_2 cat codes are in fact repetition codes, which we discussed in Section 5.1. It is evident from Fig. 6 that 2-in- n_2 cat codes provide no advantage over the single-letter optimal code except for heavily Z -biased noise and X -biased noise (for $n_2 \geq 4$). Interestingly, the advantage for X -biased noise disappears for $n_1 \geq 3$, and the increased threshold region extends from Z -biased noise over the unbiased center (corresponding to the depolarizing channel) to the Y -biased noise corner. In particular, the cat codes with $n_1 = 3, 5$ provide substantial threshold increases. For unbiased noise (and the region around it), the increase peaks (in our analysis) at the 5-in-5 cat code, which provides the optimal threshold for depolarizing noise up to $k_A = 25$ system vertices ([DiV+98]; see also Table 1).

5.3. Tree Codes

Another family of codes for which we have a sufficiently large symmetry group are graph states associated with tree graphs. In particular, we consider shallow trees which due to their restricted variety of local neighborhoods exhibit large numbers of equivalent vertices. While analyzing thresholds for the depolarizing channel alone, we found a particularly promising class of trees to be those of depth 2, i.e. a star graph, where each of the spikes is branching at most once. For all those tested 2-level graphs, we found that a single environment vertex attached at the root of the tree gives the highest thresholds in the case of depolarizing noise: for $5 < k_A < 25$, those gave the highest thresholds of any codes we tested of up to the given system vertex count (see Table 1).

The size of the automorphism group for a tree can be computed recursively, starting at the root node; instead of giving a general formula we point out that the best-performing 2-level codes in Fig. 8 appear to have a particularly large symmetry group, as all the branches are equivalent.

The most promising 2-level codes we found are given in Fig. 8 with a threshold comparison against the single-letter case and a 5-in-5 code in Fig. 7. Evidently, the 2-level codes perform

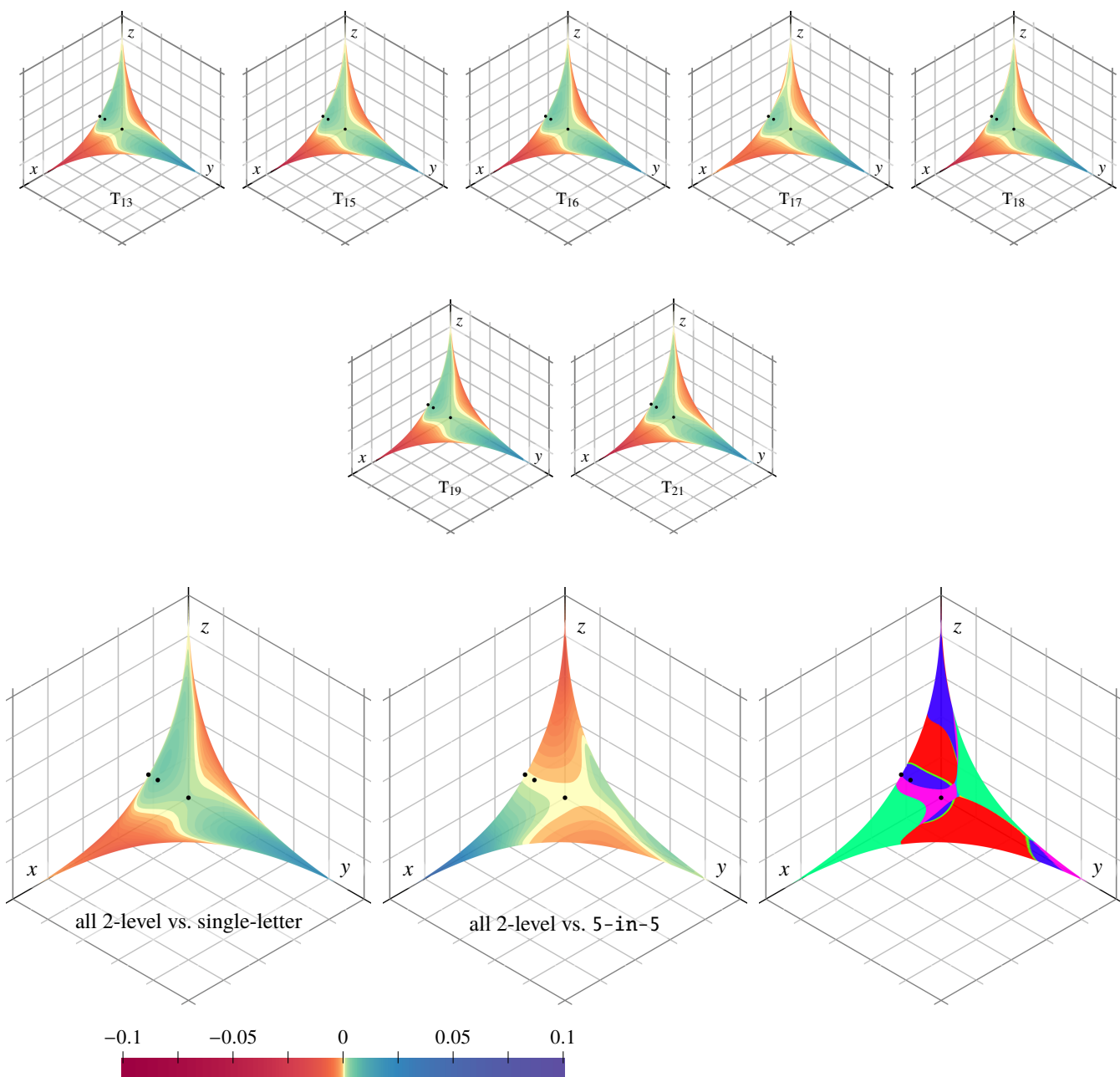


Figure 7: 2-level codes. The axes' extent is the interval $[0, 1/2]$, and the three black dots (from center to left) are the loci of depolarizing, BB84, and 2-Pauli channel families, respectively—for more details see Fig. 2. The first two rows show the noise thresholds vs. the single-letter threshold for the codes defined in Fig. 8. In the third row, the first two plots are colored according to the relative threshold of the best code in the entire tree graph code family vs. single-letter (left) and vs. the 5-in-5 cat code (middle). In all these plots, the color scale indicates the distance between the respective thresholds. The rightmost plot in the third row shows the same surface indexed by colors referring to the highest CI threshold graph state codes; for the corresponding color legend see Fig. 8.

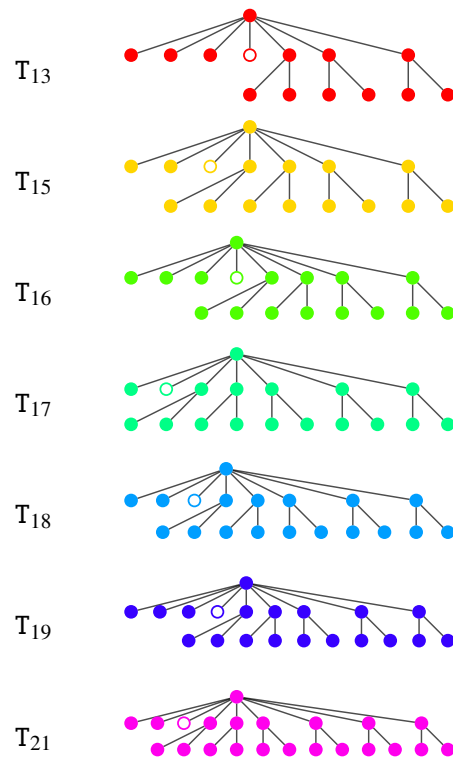


Figure 8: Color legend for 2-level codes in Fig. 7. The system vertices are filled with solid color, the (single) environment vertex is filled white.

best k_A	code	threshold p at depolarizing noise
1	single-letter	0.1898
5	1-in-5	0.190356
13	T ₁₃	0.19040260
15	T ₁₅	0.19043358
16	T ₁₆	0.19045595
	T ₁₇	0.19042111
18	T ₁₈	0.19049974
	T ₁₉	0.19049451
21	T ₂₁	0.19054941
25	5-in-5	0.190561

Table 1: A selection of cat and 2-level codes and their depolarizing noise thresholds. The naming for the 2-level codes is given in Fig. 8. In bold are the best known thresholds for any code with $\leq k_A$ system vertices. The 5-in-5 threshold is confirmed in [DiV+98; FW08].

k_A	2	3	4	5	6	7
# $_{\Gamma}$ of graphs [A00]	1	6	112	11117	11716571	$> 164 \times 10^9$
# $_{ \Gamma\rangle}$ of graph states (Eq. (37))	42	2474	809830	$> 298 \times 10^6$	$> 152 \times 10^{12}$	100×10^{18}
filtered graph states	32	256	3072	43008	917504	27951104

Table 2: Number of graph states to analyse in the exhaustive search before and after filtering out redundancies.

particularly well for X -biased noise, outperforming both the optimal single-letter code (for smaller bias towards X -noise) and the 5-in-5 cat code (for larger bias). A comparison of all code families in Fig. 19 shows that tree graphs outperform both repetition and cat codes for a large region of X -biased noise in the Pauli channel simplex. We also list the thresholds of the 2-level codes for depolarizing noise in Table 1, comparing it to the 1-in-5 and 5-in-5 codes.

5.4. Exhaustive Search

As a final analysis, we performed a brute-force search over all small graph states and collected their respective CI thresholds. This is a challenging problem: the number of connected graphs on k vertices $\#_{\Gamma}(k)$ grows exponentially in k_A , see Table 2. On top of that, any subset

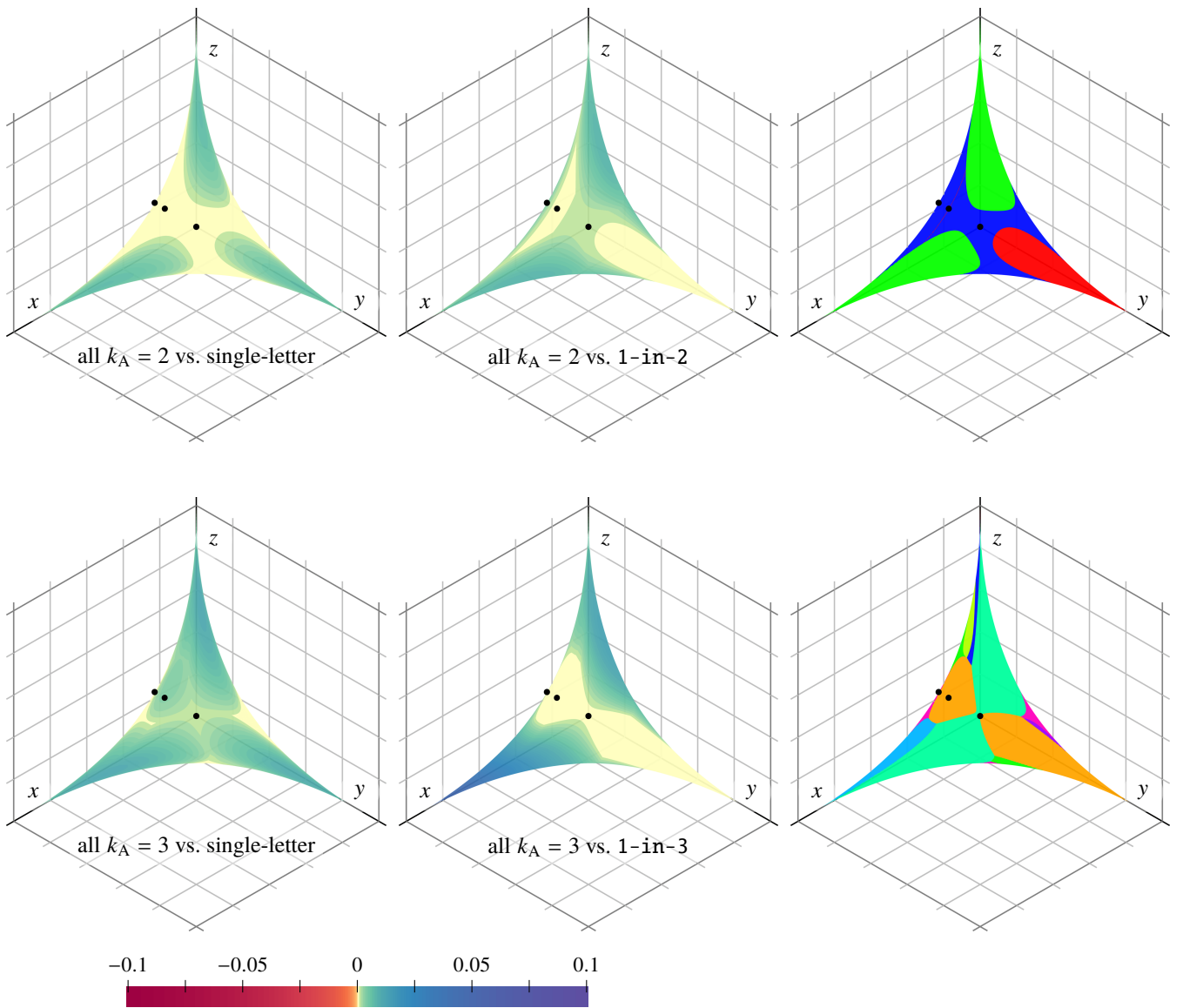


Figure 9: All possible graph codes with connected graphs for $k_A \in \{2, 3\}$, continued in Fig. 10 for $k_A \in \{4, 5\}$ —in particular, the plots for $k_A = 3$ do *not* contain the smaller graph codes. The axes' extent is the interval $[0, 1/2]$, and the three black dots (from center to left) are the loci of depolarizing, BB84, and 2-Pauli channel families, respectively—for more details see Fig. 2. The first two columns are colored according to the relative threshold of the best code in the code family: the first column shows the threshold vs. the single-letter threshold, the second column vs. the threshold of a 1-in- k repetition code. The color scale indicates the distance between noise threshold and single-letter threshold. The rightmost column shows the same surfaces indexed by colors referring to the highest CI threshold graph state codes; for the corresponding color legend see Figs. 11 and 12, respectively.

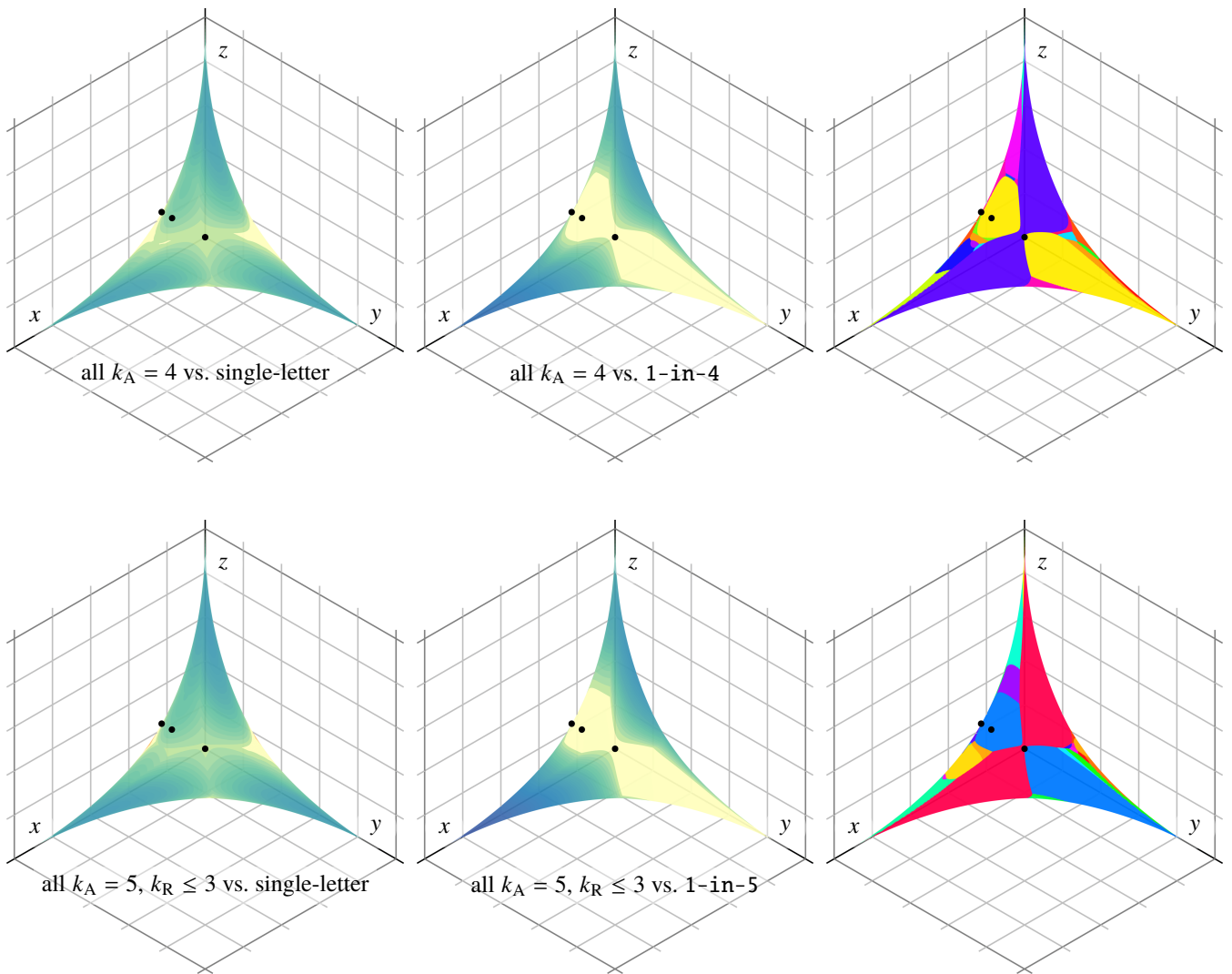


Figure 10: Continuation of Fig. 9, for $k_A \in \{4, 5\}$ and $k_{AR} \leq 8$ (in both cases). The color legends for the last column are in Figs. 13 and 14.

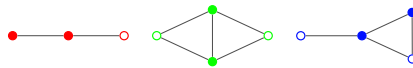


Figure 11: Color legend for all $k_A = 2$ codes in Fig. 9. The system vertices are filled with solid color, the environment vertices in white.

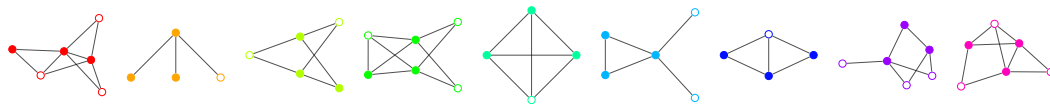


Figure 12: Color legend for all $k_A = 3$ codes in Fig. 9. The system vertices are filled with solid color, the environment vertices in white.

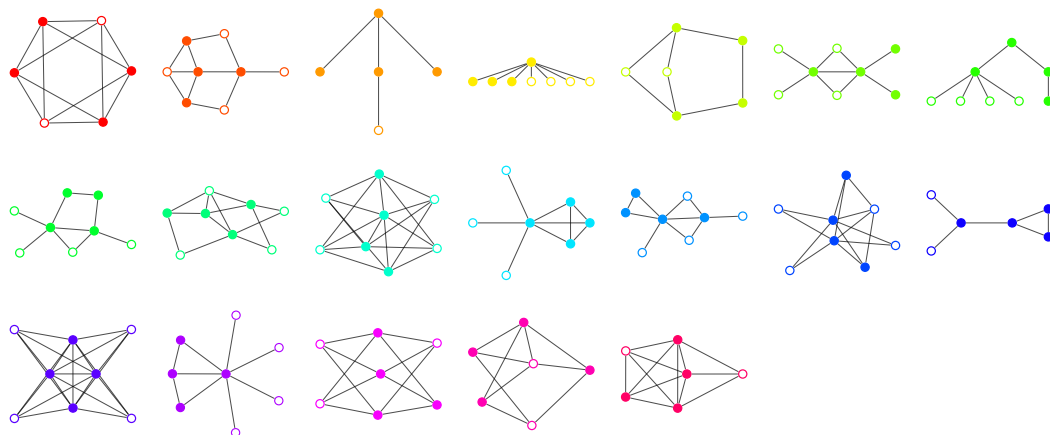


Figure 13: Color legend for all $k_A = 4$ codes in Fig. 10. The system vertices are filled with solid color, the environment vertices in white.

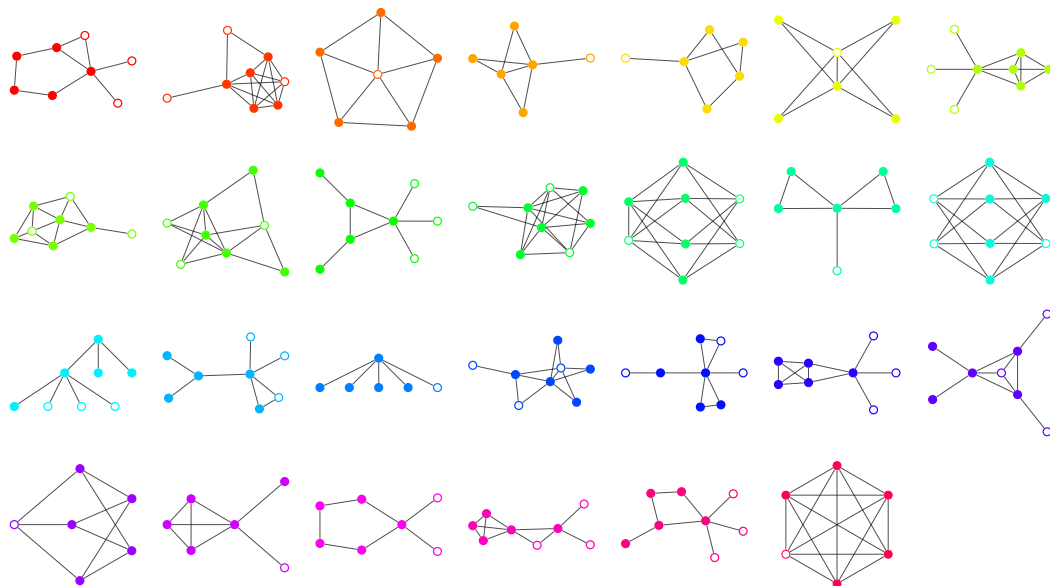


Figure 14: Color legend for all $k_A = 5$ codes with $k_R \leq 3$ in Fig. 10. The system vertices are filled with solid color, the environment vertices in white.

of the total vertices can act as system vertices. This means that with $k_A < k_{AR} \leq 2k_A$ vertices overall, there are

$$\#_{|\Gamma\rangle} = \sum_{k=k_A+1}^{2k_A} \binom{k}{k_A} \#_{\Gamma}(k) \quad (37)$$

different graph states to consider; as listed in Table 2, already for $k_A = 5$ this is more than 10^8 states.

However, since $\Gamma_{R \rightarrow R}$ does not factor into the expressions for λ and λ_A in Eqs. (11) and (16), we are overcounting: edges within the environment play no role. Furthermore, permuting system and environment vertices in a synchronous fashion also leaves the CI invariant. Finally, permutations *only* within the environment—keeping the edges to the system vertices attached—are redundant as well.

These observations suggest the following counting:

1. List all non-isomorphic system graphs on k_A vertices; call this set S .
2. For each $\Gamma_A \in S$, we iterate over all binary matrices of size $k_A \times k_A + 1$; if the i^{th} row has an entry in column j , we draw an edge from the i^{th} system vertex to the j^{th} environment vertex.
3. There are no edges within the environment only.

This procedure can be optimized further by deleting isomorphic duplicates from the final list of graphs (note that even though we started with all non-isomorphic system graphs, we will generate redundant links between system and environment). Yet even without this final optimization, the number of graphs to search over shrinks dramatically (cf. Table 2).

With this method, we managed to exhaustively search all graphs up to $k_A = 5$ system vertices, while restricting the overall vertex count $k_{AR} \leq 8$ —i.e. we performed a complete exhaustive search for $k_A \leq 4$, and an exhaustive search for $k_A = 5$ and $k_R \leq 3$. The constraining factor at this point is definitely memory, as e.g. the λ and λ_A expressions take ≈ 3.5 TB of memory in the latter case.

We have plotted the CI thresholds for our results for $k_A \in \{2, 3, 4, 5\}$ in Figs. 9 and 10. Noteworthy is that around the depolarizing channel the repetition codes perform best (given by both the complete and star graphs in Figs. 12 and 14). Deviating from unbiased noise, and depending on whether ZY or ZX noise dominates, the local Clifford invariance (which renders star and complete graphs equivalent) has to be broken; one of the choices performs better.

We note that since we restrict our search to connected graphs of one specific instance of k_A , product codes—such as e.g. products of the optimal single-letter code—are not included in our search. This explains why the code families as displayed can be worse than the single-letter threshold.

5.5. Rate Comparisons

So far we have focused on a special form of superadditivity of coherent information where the error threshold is increased relative to the single-letter threshold. This is a fundamental question in quantum information theory since positive quantum capacity of a channel guarantees the theoretical ability to correct errors (modeled by that channel), albeit at a potentially miniscule rate. From a practical point of view, one is also interested in the *magnitude* of the rate itself, which governs the overhead of an error correction code, i.e., how many physical qubits are needed to protect logical qubits from noise. In this section we study how the rates of repetition codes, cat codes, and tree graph codes compare to the optimal single-letter rates given in (32) *below* the threshold.

In order to visualize rates within the Pauli simplex displayed in Fig. 1, we evaluate the CI rate on layers aligned with the y axis, but with a specific geared ratio of $x = fz$ for various chosen z as shown in Fig. 18.

Figs. 15 to 17 show rate comparisons for the three code families—repetition, concatenated, and tree codes—vs. the single-letter CI. We have seen in the sections above that repetition codes exhibit the largest threshold increases overall, especially for noise biased towards a high $X : Z$ ratio. It is therefore not surprising that the rates of said codes—as compared to the single-letter rate—reach an advantages of ≈ 0.01 in a significant volume below the noise threshold. The volume of superadditivity decreases for the family of cat codes, and is smallest for tree graph codes.

The relatively small region of superadditivity in Figs. 15 to 17 is not surprising, as all code families studied in this paper (repetition, cat, and tree graph codes) are rank-2 codes: they have a single purifying qubit, and hence the marginal input state on the k_A channel qubits has rank at most 2 by Schmidt decomposition, and a maximal entropy of 1. These low-rank codes are well-suited for increasing noise thresholds, albeit at the expense of a poor rate of at most $1/k_A$ for noiseless channels and decreasing further for noisy channels. An interesting question is whether the region of superadditivity relative to the single-letter rate in Figs. 15 to 17 can be increased using higher-rank graph states, i.e., graphs with more than one purifying qubit. This analysis would likely extend superadditivity effects to the mid-noise regime away from the threshold. In the low-noise regime, where the channel is close to a

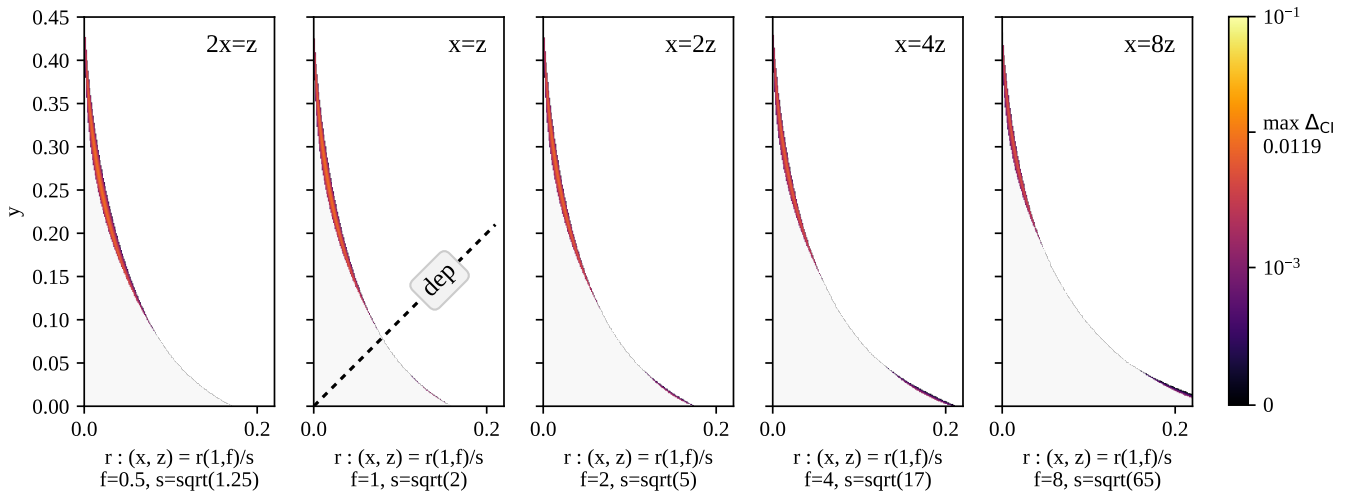


Figure 15: Rate comparison of all repetition codes 1-in- k for $k = 2, \dots, 60$ vs. the optimal single-letter rate. Shown is the CI difference of the two codes within planes that cut through the Pauli simplex (shown in Fig. 18), where the planes are parametrized by $(p_1, p_2, p_3) = (r/s, y, rf/s)$ —i.e. they align with the y axis, and are slanted with a specific ratio in the other two directions via $x = fz$ for specific choices of $f > 0$ shown in the top right corner of each plot. We only draw the area where the rate difference is positive; the color scale and maximum rate difference are given on the right. Shaded in light grey is the area where the repetition code CI is non-negative but below the single-letter rate. For $x = z$, the dashed line indicates the family of depolarizing channels.

noiseless channel, the results of [Led+18b] show that superadditivity can only be limited, and the quantum capacity is essentially given by the single-letter coherent information (up to small corrections in the noise parameter). We leave an in-depth investigation of the mid-noise regime for future work.

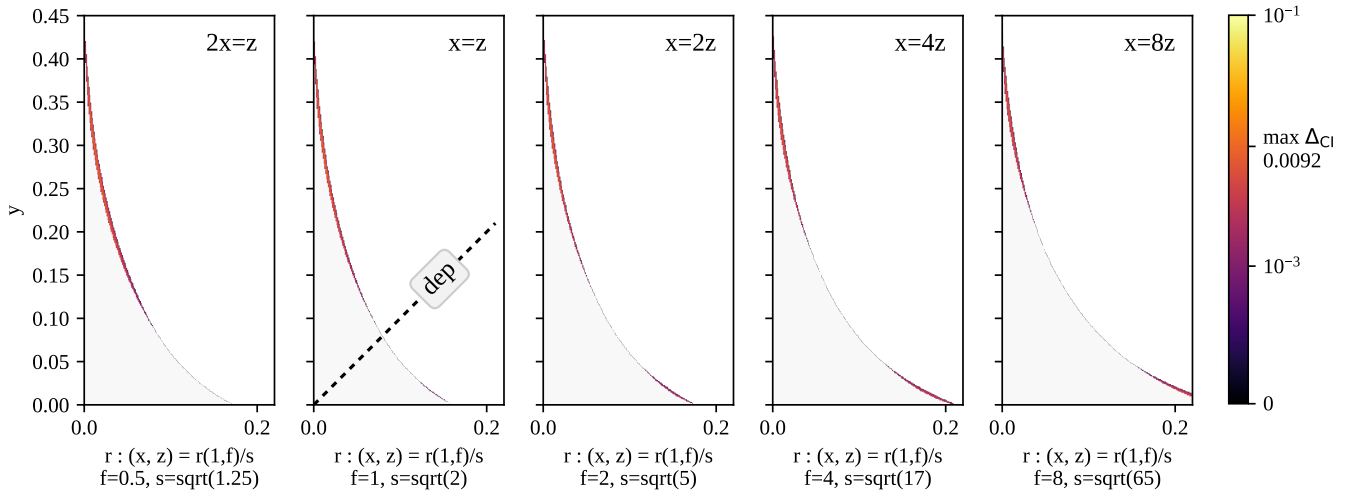


Figure 16: All concatenated codes from Fig. 6 vs. single-letter rate. For a plot description see Fig. 15.

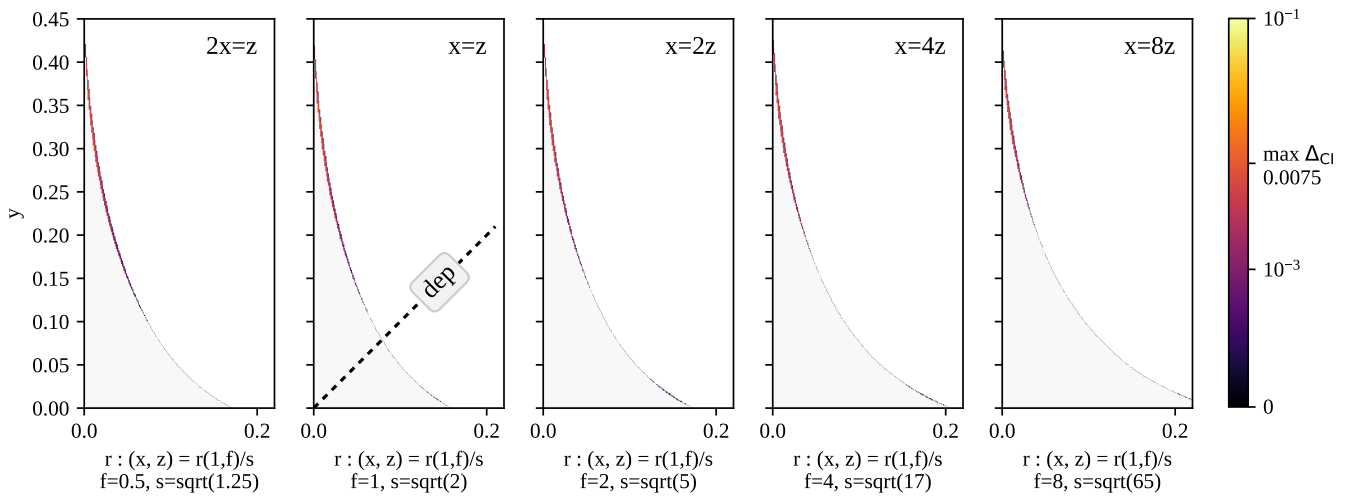


Figure 17: All tree codes from Fig. 7 vs. single-letter CI. For a plot description see Fig. 15.

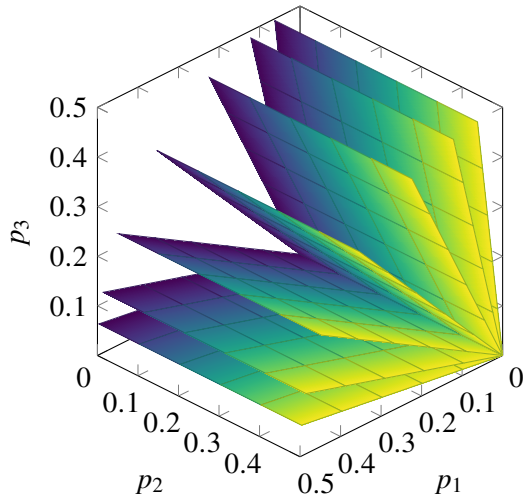


Figure 18: Layer cuts through the simplex of Pauli channels; on the surfaces shown, we evaluate the CI for various graph state codes. The results are shown in Figs. 15 to 17.

6. Discussion

In this work, we studied various classes of graph state codes in the context of Pauli noise channels, and how their coherent information (CI) thresholds relate to the noise parameters. For graphs with symmetries, we derived an algorithm to obtain an analytical expression for the CI, exploiting the graph state’s automorphism group to significantly speed up the computation (Theorems 10 and 11); the speedup is most pronounced for graphs with large symmetry groups.

To quantitatively assess the codes’ error correction potential in various noise regimes, we numerically evaluated the CI threshold over the entire simplex of Pauli channels for individual noise probabilities $\leq 1/2$. To this end, we developed a numerical bisection method that allows us to compute these thresholds in a quick fashion, based on the analytic CI expression obtained before.

The most well-studied Pauli channels are unbiased depolarizing noise, the BB84 channel (corresponding to independent bit and phase flips), and 2-Pauli noise (where only two out of the three possible Pauli errors occur). For these channel models significant superadditivity of coherent information is achieved e.g. by nested concatenation of simpler codes [DiV+98; SS07; FW08]. For more general biased noise, we find much more significant superadditivity; for example, repetition codes achieve a maximum possible superadditivity of $\delta p \approx 0.027$ relative to the single-letter threshold for heavily biased noise. This is almost two orders of magnitude beyond the superadditivity of $\delta p \approx 0.0007$ achieved by the 5-in-5 code. We find

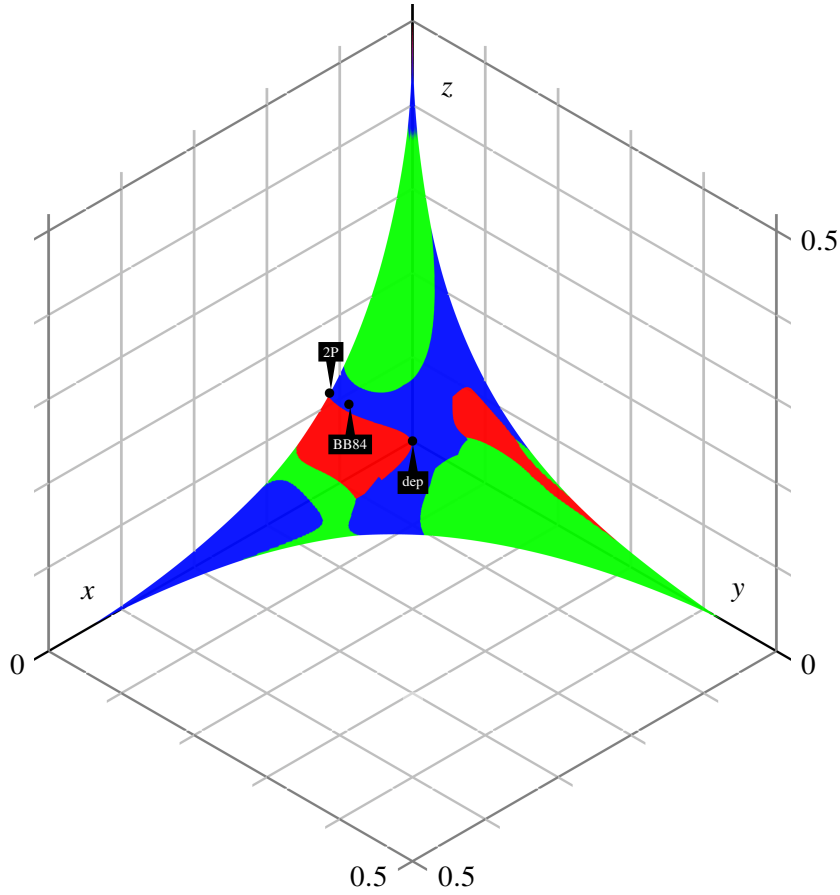


Figure 19: Repetition (green), concatenated (blue) and 2-level (red) code family thresholds compared. Especially around the 2-Pauli, BB84 and depolarizing channels biased towards x noise our new code family outperforms the other two.

a maximum superadditivity of ≈ 0.018 for concatenated codes, and ≈ 0.019 for the codes based on 2-level tree graphs introduced in this paper. All maximum threshold increases were reached for y -biased noise, i.e., the region where Y errors dominate. Note however that a different assignment of the graph state stabilizers achieves arbitrary rotations and reflections in the xyz -hyperplane (as explained in Section 5.1), and thus our results apply to any type of biased noise.

For a given stabilizer state affected by i.i.d. Pauli noise, error syndrome measurements give way to a decomposition of the coherent information in terms of *induced* single-qubit Pauli channels whose bias is determined by the structure of the stabilizer group [DiV+98; SS07]. On the other hand, it is less obvious to determine which code performs best for a given Pauli

channel. Our results presented in Section 5 allow us to answer this question for the code families (repetition, cat, tree graph codes) discussed in this paper, which we summarize in Fig. 19. This plot shows that around depolarizing, BB84 and 2-Pauli noise the concatenated code family has the highest threshold; towards the Y and Z legs repetition codes are best. Yet for large intermediate regions in the ZX and ZY plane, the 2-level codes perform best.

6.1. Noise Models in NISQ Devices

The context of codes we discussed in this paper is that of communication, i.e. where entanglement is generated by sending part of a graph state through a noisy channel. States that exhibit positive coherent information after having been affected by noise can e.g. be used to extract a secret key in quantum key distribution protocols [BB84; Bru98; SP00; Lo01; Smi+08; KR08]. More generally, a central goal in designing quantum storage and quantum computation devices is to maintain coherence of entangled states affected by environmental noise. Usually, the noise is assumed to be uniform over the Pauli errors X , Y , Z , i.e., unbiased (depolarizing) noise.

A fundamental question is how reasonable this assumption of unbiased noise is. For all but the most trivial tasks, data will have to be passed through a quantum device which, if not fully error-corrected, will be noisy. This is not an issue if the device itself features full error correction: data is then guaranteed to stay in a coherent state throughout. Yet the near-term prospects of quantum computing—dubbed the NISQ (noisy intermediate-scale quantum) era [Pre18]—comprise devices of 10-100 qubits of various interaction topologies, and with circuit depths limited to ≈ 100 gates. On such devices, achieving full-fledged error correction is unfeasible, and will remain a challenge for some time to come.

The way these limited near-term capabilities factor into the noise model is twofold. First, because there is no error correction, individual qubits will experience different levels and different types of noise originating from cross-talk across the chip, e.g. depending on whether a qubit is located towards the center of the chip, or at an edge. On average, it is thus conceivable that the noise channel modeling the device is not unbiased depolarizing noise, but dominated to some extent by a different Pauli error ratio—or even non-Pauli noise obtained from slightly perturbing Pauli channels; an analysis of the error thresholds of such close-to-Pauli channels was partly carried out in [Jac17].

Secondly, the way physical qubits are implemented factors into the induced gate noise. Any gate-level unitary on qubits is truly a hardware-level quantum channel that implements the coupling by some control pulse; depending on how the coupling is implemented, it is not obvious why the qubit-level noise should be unbiased. In fact, many proposed architectures

for quantum computation—such as superconducting qubits [Ali+09], quantum dots [Shu+12], or trapped ions [Nig+14; Osp+08]—have noise biased towards one of the Pauli errors.

For such biased noise models, it is known that one can obtain significantly higher error correction thresholds than for unbiased noise, e.g. with a tensor network decoder for the surface code [Tuc+18]. Yet even without error correction there exist circuit level procedures that render biased noise into weak but unbiased noise [AP08; WE16], so it is arguable whether other non-depolarizing noise models are worth the effort to construct tailored codes for them. However, the incurred overhead of such a transpiling procedure for error mitigation is prohibitive for any near-term quantum application with its limited circuit depth: any prolongation of the computation would render the circuit much more noisy than initially, undermining any potential benefit from unbiasing the noise in first place.

So what type of block code is optimal for a specific device and communication task setup? In particular if the overall noise level is close to the zero-CI threshold, this choice will begin to matter. For instance, if the communication task is that of QKD and we choose to implement the BB84 protocol, the quantum devices used for encoding and decoding might shift the overall channel away from an exact BB84 channel, and bias it e.g. towards a dominating Z error. In that case, concatenated codes would be a good choice (see Fig. 19). On the other hand, if X noise gets slightly more pronounced—e.g. a relative increase of $\approx 5\%$ —then picking the T_{21} code is a better option.

6.2. Open Problems

In light of the discussion above about more realistic noise models, it would be interesting to devise an algorithm computing the coherent information of graph states under non-i.i.d. channels, which potentially include memory effects between qubits such as nearest neighbor interactions. Depending on the interaction topology, this type of noise is likely to lack the full permutation symmetry of i.i.d. channels, and hence poses a challenge to exploiting symmetries to speed up the computation of the coherent information as in Algorithm 4 in Section 3.

We conclude with a few open questions that will be the subject of further study. As mentioned in Section 5.5, we focused our analysis on low-rank graph state codes that increase error thresholds relative to the single-letter threshold, noting that they also boost coherent information rates for Pauli channels in a small region below the error threshold. In order to study more general rate increases for Pauli channels in the mid-noise regime, it would be interesting to consider other families of graph states with a larger rank, i.e., for which $k_R \approx k_A$. Our analysis of error thresholds for Pauli channels can also readily be extended to other graph

state families beyond cat codes and tree graph codes. In both cases above, the challenge is to find graphs that have both large symmetry groups and interesting error-correcting properties. While the former is necessary for our algorithm to yield a computational speed-up, the latter guarantees that these graph states achieve superadditivity of coherent information.

Acknowledgements

We would like to thank Graeme Smith for helpful feedback, and Brendan McKay for many enlightening conversations about computational graph theory and his excellent nauty toolkit. J. B. would like to thank the Draper's Research Fellowship at Pembroke College for their support. F. L. acknowledges support from NSF Grant No. PHY 1734006 and the hospitality of DAMTP at the University of Cambridge. We are grateful for an *AI Grant* that enabled us to perform the numerical studies within this paper.

References

- [A00] “A001349”. In: *OEIS Foundation Inc., The On-Line Encyclopedia of Integer Sequences* (2019). url: <https://oeis.org/A001349>.
- [Ali+09] P. Aliferis, F. Brito, D. P. DiVincenzo, J. Preskill, M. Steffen, and B. M. Terhal. “Fault-tolerant computing with biased-noise superconducting qubits: a case study”. In: *New Journal of Physics* 11.1 (2009), p. 013061. arXiv: 0806.0383 [quant-ph].
- [AP08] Panos Aliferis and John Preskill. “Fault-tolerant quantum computation against biased noise”. In: *Physical Review A* 78.5 (2008), p. 052331.
- [Bau19] Johannes Bausch. *CoffeeCode*. 2019. URL: github.com/rumschuettel/CoffeeCode.
- [BL18] Johannes Bausch and Felix Leditzky. “Quantum Codes from Neural Networks”. In: *arXiv preprint* (2018). arXiv: 1806.08781 [quant-ph].
- [BB84] Charles H. Bennett and Gilles Brassard. “Quantum cryptography: Public key distribution and coin tossing”. In: *1984 IEEE International Conference on Computers, Systems and Signal Processing*. Vol. 175. 0. New York. 1984.
- [Ben+96] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. “Mixed-state entanglement and quantum error correction”. In: *Physical Review A* 54.5 (1996), pp. 3824–3851. arXiv: quant-ph/9604024.
- [Bor13] Nicolas Borie. “Generating tuples of integers modulo the action of a permutation group and applications”. In: *Discrete Mathematics and Theoretical Computer Science* (2013), pp. 767–778. arXiv: 1211.6261.
- [BR01] Hans J. Briegel and Robert Raussendorf. “Persistent Entanglement in Arrays of Interacting Particles”. In: *Physical Review Letters* 86.5 (2001), pp. 910–913. arXiv: quant-ph/0004051.
- [Bru98] Dagmar Bruß. “Optimal Eavesdropping in Quantum Cryptography with Six States”. In: *Physical Review Letters* 81.14 (1998), pp. 3018–3021. arXiv: quant-ph/9805019.
- [CT17] Giuseppe Carleo and Matthias Troyer. “Solving the quantum many-body problem with artificial neural networks”. In: *Science* 355.6325 (2017), pp. 602–606. arXiv: 1606.02318 [cond-mat.dis-nn].
- [Che+14] Jianxin Chen, Zhengfeng Ji, David Kribs, Norbert Lütkenhaus, and Bei Zeng. “Symmetric extension of two-qubit states”. In: *Physical Review A* 90.3 (2014), p. 032318. arXiv: 1310.3530 [quant-ph].
- [CJ11] Xiao-yu Chen and Li-zhen Jiang. “Graph-state basis for Pauli channels”. In: *Physical Review A* 83.5 (2011), p. 052316. arXiv: 1007.4611 [quant-ph].

- [Cub+15] Toby Cubitt, David Elkouss, William Matthews, Maris Ozols, David Pérez-García, and Sergii Strelchuk. “Unbounded number of channel uses may be required to detect quantum capacity”. In: *Nature Communications* 6 (2015), p. 6739. arXiv: 1408.5115 [quant-ph].
- [Dev05] Igor Devetak. “The private classical capacity and quantum capacity of a quantum channel”. In: *IEEE Transactions on Information Theory* 51.1 (2005), pp. 44–55. arXiv: quant-ph/0304127.
- [DS05] Igor Devetak and Peter W. Shor. “The capacity of a quantum channel for simultaneous transmission of classical and quantum information”. In: *Communications in Mathematical Physics* 256.2 (2005), pp. 287–303. arXiv: quant-ph/0311131 [quant-ph].
- [DiV+98] David P. DiVincenzo, Peter W. Shor, and John A. Smolin. “Quantum-channel capacity of very noisy channels”. In: *Physical Review A* 57.2 (1998), p. 830. arXiv: quant-ph/9706061.
- [Die10] Reinhard Diestel. *Graph Theory*. 4th ed. Springer Berlin Heidelberg, 2010, pp. XVIII, 410.
- [Ell+08] Matthew B. Elliott, Bryan Eastin, and Carlton M. Caves. “Graphical description of the action of Clifford operators on stabilizer states”. In: *Physical Review A* 77.4 (2008), p. 042307. arXiv: quant-ph/0703278.
- [FW08] Jesse Fern and K. Birgitta Whaley. “Lower bounds on the nonzero capacity of Pauli channels”. In: *Physical Review A* 78.6 (2008), p. 062335. arXiv: 0708.1597 [quant-ph].
- [Got97] Daniel Gottesman. “Stabilizer Codes and Quantum Error Correction”. PhD thesis. California Institute of Technology, 1997. arXiv: quant-ph/9705052.
- [HR09] Stephen G. Hartke and A. J. Radcliffe. “McKay’s canonical graph labeling algorithm”. In: *Contemporary Mathematics*. 2009, pp. 99–111.
- [Hei+04] Marc Hein, Jens Eisert, and Hans J. Briegel. “Multiparty entanglement in graph states”. In: *Physical Review A* 69.6 (2004), p. 062311. arXiv: quant-ph/0307130.
- [Hei+05] Marc Hein, Wolfgang Dür, and Hans J. Briegel. “Entanglement properties of multipartite entangled states under the influence of decoherence”. In: *Physical Review A* 71.3 (2005), p. 032350. arXiv: quant-ph/0408165.
- [Hei+06] Marc Hein, Wolfgang Dür, Jens Eisert, Robert Raussendorf, Maarten Van den Nest, and Hans J. Briegel. “Entanglement in Graph States and its Applications”. In: *arXiv preprint* (2006). arXiv: quant-ph/0602096.
- [Hel14] Wolfram Helwig. “Multipartite Entanglement: Transformations, Quantum Secret Sharing, Quantum Error Correction”. PhD Thesis. University of Toronto, 2014.

- [Jac17] Tyler B. Jackson. “Degenerate Codes and Capacities of Quantum Channels”. PhD thesis. Department of Mathematics and Statistics, University of Guelph, 2017. Available at: <https://atrium.lib.uoguelph.ca/xmlui/handle/10214/10362>.
- [Jef+19] Christopher Jefferson, Markus Pfeiffer, Rebecca Waldecker, and Eliza Jonauskyste. *Images*. 2019. url: <http://gap-packages.github.io/images>.
- [JK07] Tommi Junttila and Petteri Kaski. “Engineering an efficient canonical labeling tool for large and sparse graphs”. In: *Proceedings of the Ninth Workshop on Algorithm Engineering and Experiments and the Fourth Workshop on Analytic Algorithms and Combinatorics*. Ed. by David Applegate, Gerth Stølting Brodal, Daniel Panario, and Robert Sedgewick. SIAM, 2007, pp. 135–149.
- [Ker99] Adalbert Kerber. *Applied Finite Group Actions*. Vol. 19. Algorithms and Combinatorics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999.
- [KR08] Oliver Kern and Joseph M. Renes. “Improved One-way Rates for BB84 and 6-state Protocols”. In: *Quantum Information and Computation* 8.8 (Sept. 2008), pp. 756–772. arXiv: [0712.1494](https://arxiv.org/abs/0712.1494) [quant-ph].
- [Led19] Felix Leditzky. MATLAB code for Algorithm 1. 2019. Available at: <https://github.com/felixled/graph-states-coherent-info>.
- [Led+18a] Felix Leditzky, Debbie Leung, and Graeme Smith. “Dephasure Channel and Superadditivity of Coherent Information”. In: *Physical Review Letters* 121.16 (2018), p. 160501. arXiv: [1806.08327](https://arxiv.org/abs/1806.08327) [quant-ph].
- [Led+18b] Felix Leditzky, Debbie Leung, and Graeme Smith. “Quantum and Private Capacities of Low-Noise Channels”. In: *Physical Review Letters* 120.16 (2018), p. 160503. arXiv: [1705.04335](https://arxiv.org/abs/1705.04335) [quant-ph].
- [Lin74] Göran Lindblad. “Expectations and entropy inequalities for finite quantum systems”. In: *Communications in Mathematical Physics* 39.2 (1974), pp. 111–119.
- [Llo97] Seth Lloyd. “Capacity of the noisy quantum channel”. In: *Physical Review A* 55.3 (1997), p. 1613. arXiv: [quant-ph/9604015](https://arxiv.org/abs/quant-ph/9604015).
- [Lo01] Hoi-Kwong Lo. “Proof of Unconditional Security of Six-state Quantum Key Distribution Scheme”. In: *Quantum Information and Computation* 1.2 (Aug. 2001), pp. 81–94. arXiv: [quant-ph/0102138](https://arxiv.org/abs/quant-ph/0102138).
- [MP14] Brendan D. McKay and Adolfo Piperno. “Practical graph isomorphism, II”. In: *Journal of Symbolic Computation* 60 (2014), pp. 94–112.
- [ML09] Geir Ove Myhr and Norbert Lütkenhaus. “Spectrum conditions for symmetric extendible states”. In: *Physical Review A* 79.6 (2009), p. 062307. arXiv: [0812.3667](https://arxiv.org/abs/0812.3667) [quant-ph].

- [Nes+04] Maarten Van den Nest, Jeroen Dehaene, and Bart De Moor. “Graphical description of the action of local Clifford transformations on graph states”. In: *Physical Review A* 69.2 (2004), p. 022316. arXiv: [quant-ph/0308151](https://arxiv.org/abs/quant-ph/0308151).
- [NC00] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge: Cambridge University Press, 2000.
- [Nig+14] D. Nigg, M. Müller, E. A. Martinez, P. Schindler, M. Hennrich, T. Monz, M. A. Martin-Delgado, and R. Blatt. “Quantum computations on a topologically encoded qubit”. In: *Science* 345.6194 (2014), pp. 302–305. arXiv: [1403.5426](https://arxiv.org/abs/1403.5426) [quant-ph].
- [Osp+08] C. Ospelkaus, C. Langer, J. Amini, K. Brown, D. Leibfried, and D. Wineland. “Trapped-Ion Quantum Logic Gates Based on Oscillating Magnetic Fields”. In: *Physical Review Letters* 101.9 (2008), p. 090502.
- [PP16] Connor Paul-Paddock. “Degradable and Antidegradable Qubit Channels”. MA thesis. Department of Mathematics and Statistics, University of Guelph, 2016. Available at: <https://atrium.lib.uoguelph.ca/xmlui/handle/10214/9970>.
- [Pre18] John Preskill. “Quantum Computing in the NISQ era and beyond”. In: *Quantum* 2 (Aug. 2018), p. 79. arXiv: [1801.00862](https://arxiv.org/abs/1801.00862) [quant-ph].
- [Sch96] Benjamin Schumacher. “Sending entanglement through noisy quantum channels”. In: *Physical Review A* 54.4 (1996), p. 2614. arXiv: [quant-ph/9604023](https://arxiv.org/abs/quant-ph/9604023).
- [SN96] Benjamin Schumacher and Michael A. Nielsen. “Quantum data processing and error correction”. In: *Physical Review A* 54.4 (1996), p. 2629. arXiv: [quant-ph/9604022](https://arxiv.org/abs/quant-ph/9604022).
- [Ser03] Ákos Seress. *Permutation Group Algorithms*. Cambridge: Cambridge University Press, 2003.
- [Sha48] Claude E. Shannon. “A mathematical theory of communication”. In: *Bell System Technical Journal* 27 (1948), pp. 379–423.
- [Sho95] Peter W. Shor. “Scheme for reducing decoherence in quantum computer memory”. In: *Physical Review A* 52.4 (1995), R2493–R2496.
- [Sho02] Peter W. Shor. “The quantum channel capacity and coherent information”. In: *MSRI Workshop on Quantum Computation*. Berkeley, CA, USA, 2002. Available at: <http://www.msri.org/realvideo/ln/msri/2002/quantumcrypto/shor/1/>.
- [SP00] Peter W. Shor and John Preskill. “Simple Proof of Security of the BB84 Quantum Key Distribution Protocol”. In: *Physical Review Letter* 85.2 (2000), pp. 441–444. arXiv: [quant-ph/0003004](https://arxiv.org/abs/quant-ph/0003004).

- [SS96] Peter Shor and John A. Smolin. “Quantum Error-Correcting Codes Need Not Completely Reveal the Error Syndrome”. In: *arXiv preprint* (1996). arXiv: [quant-ph/9604006](#).
- [Shu+12] M. D. Shulman, O. E. Dial, S. P. Harvey, H. Bluhm, V. Umansky, and A. Yacoby. “Demonstration of Entanglement of Electrostatically Coupled Singlet-Triplet Qubits”. In: *Science* 336.6078 (2012), pp. 202–205. arXiv: [1202.1828 \[cond-mat.mes-hall\]](#).
- [SS07] Graeme Smith and John A Smolin. “Degenerate quantum codes for Pauli channels”. In: *Physical Review Letters* 98.3 (2007), p. 030501. arXiv: [quant-ph/0604107](#).
- [Smi+08] Graeme Smith, Joseph M. Renes, and John A. Smolin. “Structured Codes Improve the Bennett-Brassard-84 Quantum Key Rate”. In: *Physical Review Letters* 100.17 (2008), p. 170502. arXiv: [quant-ph/0607018](#).
- [Ste96] Andrew M. Steane. “Error Correcting Codes in Quantum Theory”. In: *Physical Review Letters* 77 (5 1996), pp. 793–797.
- [Tru93] Richard J. Trudeau. *Introduction to graph theory*. Dover Publications, 1993.
- [Tuc+18] David K. Tuckett, Stephen D. Bartlett, and Steven T. Flammia. “Ultrahigh Error Threshold for Surface Codes with Biased Noise”. In: *Physical Review Letters* 120.5 (2018), p. 050505. arXiv: [1708.08474](#).
- [WE16] Joel J. Wallman and Joseph Emerson. “Noise tailoring for scalable quantum computation via randomized compiling”. In: *Physical Review A* 94.5 (2016), p. 052325.

A. LU-Equivalence of an Arbitrary Stabilizer State to a Graph State

In this appendix we explain in detail how to obtain the graph state that is LU-equivalent to a given stabilizer state. In the following, we adopt the notation of [Eil+08].

To show the equivalence, it is convenient to use the binary representation of the Pauli group, in which the Pauli matrices are assigned vectors in \mathbb{Z}_2^2 as follows:

$$r(I) := (0, 0) \quad r(X) := (1, 0) \quad r(Y) := (1, 1) \quad r(Z) := (0, 1). \quad (38)$$

Ignoring phases, multiplication of Pauli matrices corresponds to addition in \mathbb{Z}_2 in the binary representation. Using this correspondence we assign to an element $P \in \mathcal{P}_n$ in the n -qubit Pauli group a vector in $v \in \mathbb{Z}_2^{2n}$ by setting $v_i = r(P_i)_1$ and $v_{i+n} = r(P_i)_2$. As an example, the Pauli operator $X^1 Z^2 Y^4$ corresponds to the vector $v = (1001|0101)$.

For a stabilizer state on n qubits with n stabilizer generators, consider now the $n \times 2n$ *generator matrix* whose rows are the binary representations of the stabilizer generators. Multiplying a generator by another one does not change the stabilizer group; by the discussion above, this corresponds to adding rows modulo 2 in the generator matrix. Moreover, exchanging the labels of the qubits i and j corresponds to simultaneously exchanging the columns i and j and $n + i$ and $n + j$. Using these two operations, every generator matrix corresponding to a set of n stabilizer generators describing a given stabilizer state on n qubits can be brought into the *canonical form* (or *standard form*) [NC00]

$$\left(\begin{array}{cc|cc} I_k & A & B & 0 \\ 0 & 0 & A^T & I_{n-k} \end{array} \right), \quad (39)$$

where B is a binary symmetric $k \times k$ -matrix, A is an arbitrary binary $k \times (n - k)$ -matrix, and I_j denotes the $j \times j$ -identity matrix. The canonical form of a graph state $|\Gamma\rangle$ with stabilizers Eq. (8) is given by

$$\left(I_n \mid \Gamma \right), \quad (40)$$

where by Γ we also denote the *adjacency matrix* of the graph Γ , i.e., $\Gamma_{ij} = 1$ if $\{i, j\} \in E$ and $\Gamma_{ij} = 0$ otherwise.

To convert a given generator matrix in canonical form Eq. (39) to a graph state with generator matrix Eq. (40), we recall the Clifford unitaries

$$H := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad S := \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix},$$

which satisfy the relations $HXH = Z$ (and hence $X = HZH$) and $SYS^\dagger = -X$. Applying the Hadamard matrix H to qubit j corresponds to exchanging the columns j and $n + j$ in the generator matrix. Hence, applying H to the last $n - k$ qubits achieves the transformation

$$\left(\begin{array}{cc|cc} I_k & A & B & 0 \\ 0 & 0 & A^T & I_{n-k} \end{array} \right) \longrightarrow \left(\begin{array}{cc|cc} I_k & 0 & B & A \\ 0 & I_{n-k} & A^T & 0 \end{array} \right).$$

A valid adjacency matrix of a simple graph has 0's on the diagonal, and hence we are left with removing any 1's on the diagonal of B . This can be achieved by applying S to all qubits $j \leq k$ for which $B_{jj} = 1$, and we are done. Since the procedure outlined above only involves single-qubit Clifford unitaries, we can in fact assert that any stabilizer state is local-Clifford (LC) equivalent to a graph state.

B. Subroutines for Full Coherent Information Algorithm

Algorithm 1 makes use of the following subroutines:

- **GETUSUBSETS(k, r)**
Input: $k, r \in \mathbb{N}$
Output: Three $(k+r) \times 4^k$ -matrices, whose columns are the binary vector representations of all possible tuples (U_1, U_2, U_3) satisfying $U_i \subset \{1, \dots, k\}$ and $U_i \cap U_j = \emptyset$ for $i \neq j$. The U_i are regarded as subsets in $\{1, \dots, k+r\}$, and hence the k -binary vectors are padded with r zeroes.
- **GETUSUBSETSCARD(k)**
Input: $k \in \mathbb{N}$
Output: Three row vectors of length 4^k storing the cardinalities of all possible tuples (U_1, U_2, U_3) satisfying $U_i \subset \{1, \dots, k\}$ and $U_i \cap U_j = \emptyset$ for $i \neq j$.
- **BINARYTODECIMAL(M)**
Input: $M \in \mathcal{M}(\mathbb{F}_2)$
Output: Row vector whose i -th entry is the decimal number corresponding to the binary string in the i -th column of M .
- **SUBSETS(k)**
Input: $k \in \mathbb{N}$
Output: $k \times 2^k$ -matrix whose columns are the binary vectors representing the subsets of $\{1, \dots, k\}$.
- **PERMVECTOR(v, π)**
Input: $v \in \mathbb{R}^n, \pi \in S_n$
Output: $v' = (v_{\pi(1)}, \dots, v_{\pi(n)})$
- **SHANNONENTROPY(\mathbf{q})**
Input: Probability distribution $\mathbf{q} = (q_1, \dots, q_n)$
Output: $-\sum_{i=1}^n q_i \log_2 q_i$.

C. Monotonicity of Coherent Information

In order to prove Lemma 12, we introduce the *quantum relative entropy* $D(\rho\|\sigma)$, defined for a quantum state ρ and a positive semidefinite operator σ as

$$D(\rho\|\sigma) = \begin{cases} \text{tr}(\rho \log \rho - \rho \log \sigma) & \text{if } \text{supp } \rho \subset \text{supp } \sigma \\ \infty & \text{otherwise,} \end{cases}$$

where $\text{supp } \rho := (\ker \rho)^\perp$. The relative entropy is jointly convex [Lin74]: For a probability distribution $\{\lambda_i\}$, quantum states $\{\rho_i\}$ and positive semidefinite operators $\{\sigma_i\}$,

$$D\left(\sum_i \lambda_i \rho_i \middle\| \sum_i \lambda_i \sigma_i\right) \leq \sum_i \lambda_i D(\rho_i\|\sigma_i). \quad (41)$$

The coherent information $I_c(\sigma, \mathcal{N})$ of an arbitrary (possibly mixed) state σ_{RA} and a quantum channel $\mathcal{N}: A \rightarrow B$ can be expressed in terms of the quantum relative entropy as

$$I_c(\sigma, \mathcal{N}) = D(\mathcal{N}(\sigma_{RA})\|\mathbb{1}_R \otimes \mathcal{N}(\sigma_A)).$$

Recall that a channel $\mathcal{N}: A \rightarrow B$ with complementary channel $\mathcal{N}^c: A \rightarrow E$ is antidegradable if there exists another quantum channel $\mathcal{A}: E \rightarrow B$ such that $\mathcal{N} = \mathcal{A} \circ \mathcal{N}^c$. By the data processing inequality and duality of the coherent information, antidegradable channels always have non-positive coherent information on pure states ψ :

$$I_c(\psi, \mathcal{N}) = I_c(\psi, \mathcal{A} \circ \mathcal{N}^c) \leq I_c(\psi, \mathcal{N}^c) = -I_c(\psi, \mathcal{N}),$$

and hence $I_c(\psi, \mathcal{N}) \leq 0$ for all ψ . Since $\mathcal{N}^{\otimes n}$ is antidegradable for all $n \in \mathbb{N}$ if \mathcal{N} is, we also have $I_c(\psi_n, \mathcal{N}^{\otimes n}) \leq 0$ for all $n \in \mathbb{N}$ and pure states ψ_n , and hence $Q(\mathcal{N}) = 0$ for antidegradable \mathcal{N} by the quantum capacity theorem in Eq. (2).

Finally, we make the observation that antidegradable channels have non-positive coherent information also with respect to mixed input states. To see this, let $\mathcal{N}: A \rightarrow B$ be antidegradable, let σ_{RA} be a mixed quantum state, and consider a spectral decomposition $\sigma_{RA} = \sum_i p_i |\psi_i\rangle\langle\psi_i|_{RA}$. Then,

$$\begin{aligned} I_c(\sigma, \mathcal{N}) &= D(\mathcal{N}(\sigma_{RA})\|\mathbb{1}_R \otimes \mathcal{N}(\sigma_A)) \\ &= D\left(\sum_i p_i \mathcal{N}(\psi_i) \middle\| \mathbb{1}_R \otimes \sum_i p_i \mathcal{N}(\text{tr}_R \psi_i)\right) \\ &\leq \sum_i p_i D(\mathcal{N}(\psi_i)\|\mathbb{1}_R \otimes \mathcal{N}(\text{tr}_R \psi_i)) \\ &= \sum_i p_i I_c(\psi_i, \mathcal{N}) \\ &\leq 0, \end{aligned}$$

where the first inequality follows from joint convexity of $D(\cdot\|\cdot)$, and the second inequality follows since each $I_c(\psi_i, \mathcal{N}) \leq 0$ due to antidegradability of \mathcal{N} .

We are now ready to give the proof of Lemma 12:

Proof of Lemma 12. Throughout the proof, we abbreviate $\mathcal{N}_x = \mathcal{N}_{\mathbf{p}_x}$ for some fixed probability distribution (p_1, p_2, p_3) . In terms of $\mathbf{p}_{1/2} = (\frac{1}{2}, \frac{1}{2}p_1, \frac{1}{2}p_2, \frac{1}{2}p_3)$, the antidegradability condition (31) becomes

$$1 \geq \frac{1}{2}(1 + p_1^2 + p_2^2 + p_3^2) - 2\sqrt{p_1 p_2 p_3}.$$

Since $p_1^2 + p_2^2 + p_3^2 \leq (p_1 + p_2 + p_3)^2 = 1$ for any $p_1, p_2, p_3 \geq 0$ with $\sum_i p_i = 1$, it follows that every channel \mathcal{N}_x is antidegradable for $x = 1/2$, and hence also $I_c(\psi, \mathcal{N}_{1/2}^{\otimes k}) \leq 0$ for all ψ and $k \in \mathbb{N}$. By assumption, we even have $I_c(\psi, \mathcal{N}_{1/2}^{\otimes k}) < 0$.

The function $f(x) = I_c(\psi, \mathcal{N}_x^{\otimes k})$ is differentiable, and thus in particular continuous. Moreover, we have $I_c(\psi, \mathcal{N}_0^{\otimes k}) = S(\text{tr}_R \psi) > 0$ since ψ is entangled across the bipartition $R|A$ by assumption. Hence, since $I_c(\psi, \mathcal{N}_{1/2}^{\otimes k}) < 0$, the function f has at least one root on the interval $[0, 1/2]$. Let x_0 be the smallest of these roots, i.e., $f(x_0) = 0$ and $f(x) > 0$ for all $x < x_0$.

To prove the claim of the lemma, assume that there is another root x_1 of f in $(x_0, 1/2]$. For $i = 1, \dots, k$ consider the continuous functions

$$g_i : (\sigma, x) \mapsto I_c(\sigma, \mathcal{N}_x^i), \quad (42)$$

where $\mathcal{N}_x^i = \text{id}_1 \otimes \dots \otimes \text{id}_{i-1} \otimes \mathcal{N}_x \otimes \text{id}_{i+1} \otimes \dots \otimes \text{id}_k$. By the joint convexity property (41) of the relative entropy, each function $x \mapsto g_i(\sigma, x)$ is convex for fixed σ . Moreover, for each i and any σ we have $g_i(\sigma, 1/2) \leq 0$, and hence $g_i(\sigma, \cdot)$ either has exactly one root in $x \in [0, 1/2)$ if $g_i(\sigma, 0) \geq 0$, or $g_i(\sigma, x) < 0$ for all $x \in [0, 1/2)$ if $g_i(\sigma, 0) < 0$. In either case, it follows that $g_i(\sigma, x) < 0$ for all $x \in (x_0, 1/2)$. Thus, for $\sigma^{x_0} := (\mathcal{N}_{x_0}^{\otimes k-1} \otimes \text{id})(\psi)$ satisfying $g_k(\sigma^{x_0}, x_0) = f(x_0) = 0$, we have $g_k(\sigma^{x_0}, x_1) < 0$, since by assumption the second root satisfies $x_1 > x_0$.

Consider now the state $\sigma^{x_0, x_1} = [\mathcal{N}_{x_0}^{\otimes k-2} \otimes \text{id} \otimes \mathcal{N}_{x_1}](\psi)$. By the above argument, we have that $0 > g_k(\sigma^{x_0}, x_1) = g_{k-1}(\sigma^{x_0, x_1}, x_0)$. Then by the same argument as before, convexity of $x \mapsto g_{k-1}(\sigma^{x_0, x_1}, x)$ implies that also $g_{k-1}(\sigma^{x_0, x_1}, x_1) < 0$. We can continue this argument inductively to obtain a sequence of states

$$\sigma^{x_0}, \sigma^{x_0, x_1}, \dots, \sigma^{x_0, x_1, \dots, x_1} \equiv [\text{id} \otimes \mathcal{N}_{x_1}^{\otimes k-1}](\psi)$$

to conclude that also $f(x_1) = g_1(\sigma^{x_0, x_1, \dots, x_1}, x_1) = I_c(\psi, \mathcal{N}_{x_1}^{\otimes k}) < 0$, which contradicts the assumption that x_1 is a root of f . \square