Stabilizer codes for Heisenberg-limited many-body Hamiltonian estimation

Santanu Bosu Antu[1,2] and Sisi Zhou[1,3]

[1]Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada
[2]Department of Applied Physics and Yale Quantum Institute, New Haven, Connecticut 06520, USA
[3]Department of Physics and Astronomy, Department of Applied Mathematics, and Institute for Quantum Computing, University of Waterloo, Ontario N2L 3G1, Canada

Estimating many-body Hamiltonians has wide applications in quantum technology. By allowing coherent evolution of quantum systems and entanglement across multiple probes, the precision of estimating a fully connected k-body interaction can scale up to $(n^k t)^{-1}$, where n is the number of probes and t is the probing time. However, the optimal scaling may no longer be achievable under quantum noise, and it is important to apply quantum error correction in order to recover this limit. In this work, we study the performance of stabilizer quantum error correcting codes in estimating many-body Hamiltonians under noise. When estimating a fully connected ZZZ interaction under single-qubit noise, we showcase three families of stabilizer codes—thin surface codes, quantum Reed–Muller codes and Shor codes—that achieve the scalings of $(nt)^{-1}$, $(n^2t)^{-1}$ and $(n^3t)^{-1}$, respectively, all of which are optimal with t. We further discuss the relation between stabilizer structure and the scaling with n, and identify several no-go theorems. For instance, we find codes with constant-weight stabilizer generators can at most achieve the n^{-1} scaling, while the optimal n^{-3} scaling is achievable if and only if the code bears a repetition code substructure, like in Shor code.

1 Introduction

High-precision measurement of parameters, or metrology, is crucial in all branches of science, especially in physics [1–5]. Classically, a common method to estimate an unknown parameter is to repeat the same experiment multiple times. As a consequence of the central limit theorem, the estimation error will scale as $1/\sqrt{t}$, known as the standard quantum limit (SQL), where t is the time of experiment. Quantum mechanics allows us to achieve better precision by adopting the advantage of long-time coherent evolution, which results in an estimation error of order $1/t$, known as the *Heisenberg limit* (HL) [6, 7] which quadratically outperforms the SQL.

The HL, however, can be fragile against quantum noise [8–12]. To address this issue, quantum error correction (QEC) has been proposed as a tool to recover the HL under noise [11–20]. Specifically, to estimate a Hamiltonian parameter under Markovian noise, it was known the HL is achievable if and only if the "Hamiltonian-not-in-Lindblad-span" (HNLS) condition is satisfied [11, 12], in which case a QEC protocol was proposed to recover the HL. Despite the generality of the QEC protocol, its implementation can be challenging in practice, partially due to the lack of efficient encoding and decoding procedure, and the requirement of a noiseless ancillary system as part of the code [12, 19]. Using stabilizer codes [21, 22], a class of QEC codes with natural encoding and decoding methods via stabilizer measurements and no requirement of noiseless ancilla, for quantum metrology can alleviate the above constraints. So far, stabilizer codes have rarely been explored for quantum metrology in previous works [12, 17–20], except for repetition codes for phase estimation against bit-flip noise [13–15].

In multi-probe systems, estimating the strength of Hamiltonians has always been a major topic in quantum metrology [6, 7, 23–30]. The scaling of the estimation error with respect to the number of probes n can achieve up to $1/n$ for 1-local Hamiltonians, using entangled states, e.g., the Greenberger–Horne–Zeilinger (GHZ) state [7] and we will refer to it as the *Heisenberg scaling* (HS)[1].

Interestingly, the estimation with higher-order interactions can go beyond the HS [30–32] (even without entanglement [33]). For example, estimating a k-local Hamiltonian can reach a precision of order $1/n^k$, usually referred to as the *super-Heisenberg scaling* (SHS) when $k > 1$ [30]. Similar to the HL, the HS and SHS are also compromised by the effect of quantum noise. For example, the HS for estimating 1-local Hamiltonians will degrade to $1/\sqrt{n}$ under generic single-qubit noise [9–12]. In general, it is of great interest to ex-

Sisi Zhou: sisi.zhou26@gmail.com

[1]Note that the scaling $1/n$ can also sometimes be referred to as the Heisenberg limit (with respect to the number of probes). However, in this paper we call it the HS to distinguish it from the HL with respect to probing time t.

plore which types of SHS are recoverable for estimating many-body Hamiltonians under different types of quantum noise.

In this work, we consider the achievability of the HL, HS and SHS for estimating a fully connected ZZZ interaction under single-qubit noise using stabilizer codes. First, we will present three classes of $[[n, 1, 3]]$ stabilizer codes—thin surface codes [34–36], quantum Reed-Muller codes [37] and Shor codes [38]—that achieve the HL and HS. This demonstrates the applicability of stabilizer codes in noisy quantum metrology. In particular, the optimal SHS $1/n^3$ can be achieved by Shor codes. Secondly, we will dive into the achievability of the SHS and explore the roles of different code structures, e.g., nonlocality of stabilizer generators, repetition code substructures, etc., that are necessary for achieving different types of SHS. One prominent result is the nonachievability of the SHS when stabilizer generators have constant weights, e.g., for quantum low-density parity-check (LDPC) codes [39]. These no-go results then imply the optimality of the previously mentioned three classes of codes under different constraints on code structures, and may further provide guidance on identifying or constructing stabilizer codes that are suitable for many-body Hamiltonian estimation.

Our work studies Hamiltonian estimation under Markovian noise during time evolution with the novel stabilizer QEC technique. Here we comment on the relationship between our work and previous literature. First, this work focuses on Hamiltonian estimation from time evolution of the Hamiltonian [28–30, 40–47], where input states, quantum controls during the evolution and measurements can be chosen by the observer, a framework different from Hamiltonian estimation from Gibbs states [48–51]. In our case, the input states, quantum controls, and measurements are logical states, quantum error correction, and logical measurements, respectively, and our goal is the achieve the HL with respect to the probing time [28–30]. Second, the noise model we consider is the Markovian noise during the evolution of the Hamiltonian, which is different from other literature that primarily focused on the state preparation, measurement, and gate implementation errors [42–44]. Our results represent the ultimate achievable estimation limits under stochastic noise when perfect quantum controls are available. Third, we treated the Hamiltonian estimation problem as a single-parameter estimation problem [30] where the strength of an all-to-all ZZZ interaction is to be measured and we analyze the scaling of the estimation precision with respect to the number of qubits. This is in contrast to (although related to) the multi-parameter approach [28, 29, 40–47] where the goal is to show identifiability of the Hamiltonian coefficients, and to optimize the sample complexity with respect to the system size and the estimation error. Finally, our work is based on the point estimation approach using the quantum Cramér–Rao bound, where the range of the unknown parameter is assumed to be priorly narrowed down to a small neighbourhood and our goal is to perform high-precision estimation in the neighbourhood. The Cramér–Rao approach characterizes the asymptotic behaviour of parameter estimation when the number of experiments is sufficiently large. This contrasts with the Bayesian approach [52, 53], which selects estimators, measurement schemes, and initial states adaptively based on prior probabilities [40–43] and is more relevant in scenarios where the target estimation precision is a constant and the number of allowed experiments is limited.

2 Formulation of the Problem

Hamiltonian evolution under Markovian noise can be described by the Lindblad master equation [54–56]

$$\frac{d\rho}{dt} = -i[H, \rho] + \sum_a \left(L_a \rho L_a^\dagger - \frac{1}{2}\{L_a^\dagger L_a, \rho\} \right), \quad (1)$$

where ρ is the density matrix of the probe system, H is the Hamiltonian and $\{L_a\}_a$ are the Lindblad operators describing quantum noise. We consider the situation where the probe system consists of n qubits ($n \geq 3$),

$$H = \omega \sum_{1 \leq i < j < k \leq n} Z_i Z_j Z_k + H_0 \coloneqq \omega G + H_0, \quad (2)$$

where H_0 is any (unknown) 2-local Hamiltonian, ω is the strength of the 3-local Hamiltonian G that we want to estimate, and Z_i represents the Pauli-Z operator acting on the i-th qubit (we will use X, Y, Z to denote Pauli-X,Y,Z operators, respectively). The set $\{L_a\}_a$ consists of local operators (i.e., each L_a acts non-trivially on only one qubit). We further assume that the noise is complicated enough that the Lindblad span

$$\mathcal{S} \coloneqq \mathrm{span}_{\mathrm{H}}\{I, L_a, L_{a'}^\dagger, L_a^\dagger L_{a'}\}_{\forall a, a'} \quad (3)$$

$$= \mathrm{span}_{\mathrm{H}}\{P_i P_j\}_{\forall i,j; P \in \{X,Y,Z,I\}} \quad (4)$$

contains all two-local Hermitian operators ($\mathrm{span}_{\mathrm{H}}\{\cdot\}$ denotes the set of all Hermitian operators that can be written as linear combinations of operators in $\{\cdot\}$). It is a natural assumption of quantum noise, and is satisfied by, for example, single-qubit depolarizing noise, single-qubit amplitude damping noise, etc.

In general, the HNLS condition [12] states that it is possible to achieve the HL if and only if $G \notin \mathcal{S}$, which is satisfied in our setting. Due to the HNLS condition [12], our scenario represents one of the *simplest* cases of many-body Hamiltonian estimation where a QEC code can be used to recover the HL under single-qubit noise. This is because, for arbitrary single-qubit errors, the Lindblad span includes all two-body Pauli

terms, and, consequently, the HL cannot be recovered for any two-body Hamiltonian estimation. This motivates our focus on three-body Hamiltonians, which lie just outside of the Lindblad span and therefore allow for Heisenberg-limited estimation using QEC.

Consider the quantum metrological protocol where an initial n-qubit state $\rho(0)$ is prepared, quantum controls are applied sufficiently fast during the evolution, and ω is estimated through quantum measurement on the output state $\rho_\omega(t)$ at time t [11, 12, 16]. The estimation precision $\delta\omega$ can be given by the quantum Cramér–Rao bound [57, 58]

$$\delta\omega \geq 1/\sqrt{\nu F(\rho_\omega(t))}, \tag{5}$$

where ν is the number of repeated experiments and $F(\rho_\omega(t))$ is the quantum Fisher information (QFI) of the state $\rho_\omega(t)$ [59, 60]. The quantum Cramér–Rao bound is saturable as $\nu \to \infty$ [61], and thus the scaling of $F(\rho_\omega(t))$ determines the scaling of $\delta\omega$ with respect to t and n. Specifically, we say that the HL is achieved when $\delta\omega \sim t^{-1}$ or $F(\rho_\omega(t)) = \Theta(t^2)$, the HS is achieved when $\delta\omega \sim n^{-1}$ or $F(\rho_\omega(t)) = \Theta(n^2)$, and the SHS is achieved when $\delta\omega \sim n^{-k}$ or $F(\rho_\omega(t)) = \Theta(n^{2k})$ for any $k > 1$.

In the noiseless case ($L_a = H_0 = 0$), the optimal initial state is the GHZ state

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}(|0^{\otimes n}\rangle + |1^{\otimes n}\rangle), \tag{6}$$

which leads to an output state

$$|\psi_\omega(t)\rangle = \frac{e^{i\omega\frac{n(n-1)(n-2)}{6}}|0^{\otimes n}\rangle + e^{-i\omega\frac{n(n-1)(n-2)}{6}}|1^{\otimes n}\rangle}{\sqrt{2}}, \tag{7}$$

after evolution time t, and the corresponding QFI of the output state $\rho_\omega(t) = |\psi_\omega(t)\rangle\langle\psi_\omega(t)|$ is

$$F(\rho_\omega(t)) = 4\left(\frac{n(n-1)(n-2)}{6}\right)^2 t^2 = \Theta(n^6t^2), \tag{8}$$

which achieves the HL and the SHS.

With single-qubit errors, the QEC protocol utilizes a QEC code that perfectly corrects them, i.e., $\Pi S\Pi \propto \Pi$ for all $S \in \mathcal{S}$, where Π is the projection onto the code subspace [62, 63]. In the fast control limit (i.e., fast quantum controls can be applied at an arbitrarily high frequency), the evolution of $\rho_{\text{L}} = \Pi\rho\Pi$ is given by

$$\frac{d\rho_{\text{L}}}{dt} = -i\omega[G_{\text{eff}}, \rho_{\text{L}}], \tag{9}$$

where $G_{\text{eff}} = \Pi G\Pi \not\propto \Pi$. (Note that in the noiseless case, $\Pi = I$ and the discussion below also holds.) The QFI is maximized when the initial state is chosen as $|\psi(0)\rangle = \frac{1}{\sqrt{2}}(|\lambda_{\min}\rangle + |\lambda_{\max}\rangle)$, that leads to

$$F(\rho_\omega(t)) = (\lambda_{\max} - \lambda_{\min})^2 t^2 =: (\Delta G_{\text{eff}})^2 t^2, \tag{10}$$

which achieves the HL, where $\lambda_{\min,\max}$ are the smallest (and largest) eigenvalues that correspond to the eigenstates $|\lambda_{\min,\max}\rangle$ of G_{eff} that are restricted in the code subspace. $\Delta G_{\text{eff}} := \lambda_{\max} - \lambda_{\min}$. Note that the above implies that it is sufficient to consider a two-level logical system, as we can always, without loss of generality, restrict the code subspace to $\text{span}\{|\lambda_{\min}\rangle, |\lambda_{\max}\rangle\}$.

Among all possible ancilla-assisted QEC codes, the optimal QEC protocol [12] can achieve $(\Delta G_{\text{eff}})_{\text{opt}} = 2\|G - \mathcal{S}\|$, where $\|G - \mathcal{S}\|$ is the distance between G and $\mathcal{S}$ in terms of the operator norm. Specifically in our case (see Appx. A),

$$(\Delta G_{\text{eff}})_{\text{opt}} = 2\min_{\text{2-local } S}\|G - S\| = \frac{n^3}{12} + O(n^2). \tag{11}$$

Comparing this to the noiseless case, where

$$\Delta G = \frac{n^3}{3} + O(n^2), \tag{12}$$

we see that the optimal QFI for estimating ZZZ interaction under generic qubit noise is $F(\rho_\theta(t)) = \Theta(n^6t^2)$, which bears the same scaling as the noiseless one.

Although QEC can in principle recover the HL and the SHS for estimating the 3-local Hamiltonian G from the above discussion, it is unclear how the QEC protocol (with encoding, decoding and measurement operations) can be implemented in reality. In this work, we consider stabilizer QEC for the task of estimating ZZZ interactions, which is the most popular family of QEC and was demonstrated recently in experiments across multiple platforms [64–66]. One advantage of it is the standard encoding, decoding and measurement procedures via syndrome measurements [21] that were currently unavailable in the known QEC protocols for metrology. In addition, as stabilizer codes have already been theoretically extensively studied in previous literature for quantum memory, quantum computation, etc., it is natural to explore their applications in other noisy information processing tasks like quantum metrology and identify the code properties that are crucial for many-body Hamiltonian estimation.

3 Achieving the HL, HS and SHS

As discussed in the previous section, the HL is achievable using a QEC code if and only if $G_{\text{eff}} = \Pi G\Pi$ acts non-trivially in the code subspace. In this section, we first establish a relation between ΔG_{eff} and the logical operators of the stabilizer code, defined by an abelian subgroup $\mathscr{S}$ of the n-qubit Pauli group $\{\pm 1, \pm i\} \times \{I, X, Y, Z\}^{\otimes n}$ which stabilizes the code space, and then present three families of stabilizer codes that achieve both the HL and the HS (or SHS).

3.1 Relation between QFI and $Z^{\otimes 3}$-type Logical Operators

Here we use the stabilizer formalism to establish a relation between the QFI and the number of $Z^{\otimes 3}$-*type logical operators*, i.e., operator $Z_iZ_jZ_k$ for some i, j, k that maps logical states to logical states in a non-trivial way, and this relation will be used throughout the remainder of the paper.

Lemma 1. *Consider using an* $[[n, 1, 3]]$[2] *stabilizer code to estimate* ω *in Eq. (1) and Eq. (2). Let* ℓ *be the number of* $Z_iZ_jZ_k$ *operators appearing in* G *that are logical operators of the code. Then the output QFI*

$$F(\rho_\omega(t)) \leq 4\ell^2t^2. \tag{13}$$

The equality holds when the stabilizer code satisfies

Property 1. *Any* Z*-type stabilizer (i.e., stabilizer that consists of only* I *and* Z*) has a positive sign*[3].

Proof. Let $\mathscr{S}$ and $\mathscr{N}(\mathscr{S})$ be the stabilizer and normalizer sets of the stabilizer code, respectively. A term $Z_iZ_jZ_k$ of G belongs to one of the following three classes: (1) $Z_iZ_jZ_k \in \mathscr{S}$, (2) $Z_iZ_jZ_k \notin \mathscr{N}(\mathscr{S})$, and (3) $Z_iZ_jZ_k \in \mathscr{N}(\mathscr{S}) \setminus \mathscr{S}$. The operator is called a *logical operator* if it belongs to class (3). By Eq. (10), the QFI is proportional to ΔG_{eff} where $G_{\text{eff}} = \Pi G \Pi = \sum_{ijk} \Pi Z_iZ_jZ_k \Pi$ with Π being the projection onto the code space. We separately calculate the contributions of the three types of terms to the ΔG_{eff}.

(1) Suppose $Z_iZ_jZ_k \in \mathscr{S}$. It immediately follows that $Z_iZ_jZ_k\Pi = \Pi$ such that $\Pi Z_iZ_jZ_k\Pi = \Pi$. Hence, the terms that are in $\mathscr{S}$ do not contribute to ΔG_{eff}.

(2) Suppose that $Z_iZ_jZ_k \notin \mathscr{N}(\mathscr{S})$. Then, there is an operator $Q \in \mathscr{S}$ that anti-commutes with $Z_iZ_jZ_k$. So, $\Pi Z_iZ_jZ_k\Pi = \Pi QZ_iZ_jZ_k\Pi = -\Pi Z_iZ_jZ_kQ\Pi = -\Pi Z_iZ_jZ_k\Pi$, which leads to $\Pi Z_iZ_jZ_j\Pi = 0$. Hence, the terms that are not in the normalizer group also do not contribute to the QFI.

(3) Suppose that $Z_iZ_jZ_k \in \mathscr{N}(\mathscr{S}) \setminus \mathscr{S}$. As all the ZZZ terms commute with each other, the non-trivial logical operators must be different representatives of the same logical operator up to a ± 1 coefficient. In addition, $\Pi Z_iZ_jZ_k\Pi$ must correspond to a logical Pauli operator in the two-dimensional code subspace, which means $\Delta(\Pi Z_iZ_jZ_k\Pi) = 2$.

Hence, if the number of $Z^{\otimes 3}$-type logical operators in G is ℓ, the QFI is given by $F(\rho_\omega(t)) = (\Delta G_{\text{eff}})^2t^2 \leq 4\ell^2t^2$. When Property 1 holds, all $Z^{\otimes 3}$-type logical operators must correspond to the same logical operator with the same sign, and the equality holds. □

We note that any stabilizer code can be modified to satisfy Property 1 by only changing the signs of its stabilizer generators. Thus, we always assume, without the loss of generality, that Property 1 holds in all discussions below. For example, the surface code, the Reed–Muller code and the Shor code we consider below all naturally satisfy Property 1. Given a stabilizer code, to achieve Property 1, we first find a set of its stabilizer generators which has the maximal number of Z-type stabilizers. That means the rest of the stabilizer generators in the set cannot generate a Z-type stabilizer. This set can be efficiently obtained by Gaussian elimination. Then we modify all signs of these Z-type stabilizer generators to the positive sign. For example, if we have a stabilizer generator $-ZZIII$, we switch the sign to $+1$ and define a new stabilizer code where $-ZZIII$ is replaced by $ZZIII$. This guarantees Property 1 holds because any Z-type stabilizer can only be expressed as products of Z-type stabilizer generators and thus also must has a positive sign.

Lemma 1 implies that the HL is achievable for any stabilizer codes with $Z^{\otimes 3}$-type logical operators. The scaling of the QFI with respect to n is determined by the scaling of ℓ with respect to n. In the following parts of this section, we will apply Lemma 1 to demonstrate the achievability of the HS ($F(\rho_\omega(t)) = \Theta(n^2)$) and SHS ($F(\rho_\omega(t)) = \Theta(n^4), \Theta(n^6)$) using three different families of stabilizer codes, or more specifically, Calderbank–Shor–Steane (CSS) codes [67, 68].

Note that Lemma 1 automatically applies to the estimation of higher-order $Z^{\otimes k}$ terms, i.e. when $G = \sum_{\{i_k\}\subseteq\{1,2,\ldots,n\}} Z_{i_1}Z_{i_2}\cdots Z_{i_k}$ with $k \geq 3$, as the proof of Lemma 1 is still correct when we replace $Z_iZ_jZ_k$ with higher-order Z-type operators $Z_{i_1}Z_{i_2}\cdots Z_{i_k}$.

3.2 Thin Surface Code

First, we consider a special kind of surface code [34–36] that we will refer to as the *thin surface code*, which corrects single-qubit Z errors. Fig. 1 shows the $[[n, 1, 3]]$ thin surface code with $n = 33$. The qubits are placed on the edges of a $(2\times(n-3)/5)$ lattice. On the primary (solid) lattice, the X-type stabilizer generators are $X^{\otimes 4}$ plaquette operators in the bulk and on the smooth edges, and $X^{\otimes 3}$ plaquette operators on the rough edges. Similarly, the Z-type stabilizer generators are $Z^{\otimes 4}$ and $Z^{\otimes 3}$ plaquette operators on the dual (dashed) lattice,

The quantum code encodes one logical qubit into n physical qubits. A logical X operator is a string of X operators connecting the two rough edges on the primary lattice. Similarly, a logical Z operator is a string of Z operators connecting the two smooth edges on the dual lattice.

[2] We use $[[n, k]]$ ($[[n, k, d]]$) to represent a stabilizer code that encodes k logical qubits into n physical qubits (with distance d).

[3] For example, $ZZIII$ has a positive sign; $-ZZIII$ has a negative sign. ($\pm iZZIII$ cannot be in any stabilizer group because their squares are $-I$.)

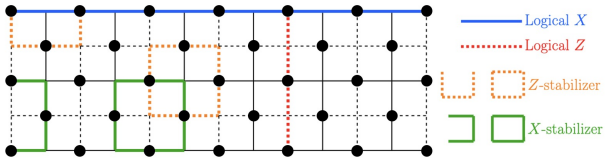

Figure 1: $[[33,1,3]]$ thin surface code. Physical qubits are placed on the edges of the solid lattice (including the start and end points of the horizontal solid lines). X/Z-type stabilizer and logical operators are represented by solid/dashed lines that pass through the involved qubits.

For the $[[n,1,3]]$ thin surface code, the $Z^{\otimes 3}$-type logical operators must correspond to one vertical $Z^{\otimes 3}$ strings on the dual lattice. This is because multiplying any one of them by a stabilizer can either increase its weight, or transform it to another $Z^{\otimes 3}$-type logical operator. By a simple counting, the number of such $Z^{\otimes 3}$ strings for an n-qubit code is $(n+2)/5$. Therefore, by Lemma 1, we have the following result.

Theorem 1. *The $[[n,1,3]]$ thin surface code achieves a QFI of $4(n+2)^2t^2/25 = \Theta(n^2t^2)$.*

We also note that the result above can readily generalize to estimating k-body terms in Hamiltonian (when k is a constant independent of n and n is sufficiently large). We observe that the $[[n,1,3]]$ thin surface code has $\Theta(n^{\lfloor k/3 \rfloor})$ $Z^{\otimes k}$-type logical operators, implying a QFI of $\Theta(n^{2\lfloor k/3 \rfloor}t^2)$. To prove this, we can divide the entire n qubits into $\Theta(n)$ constant-size groups, e.g. groups of qubits in a 5×5 grid, with each group sufficiently large to accommodate at least one $Z^{\otimes 3}$-type logical operator, $Z^{\otimes 3}$-type, $Z^{\otimes 4}$-type and $Z^{\otimes 5}$-type stabilizers. On one hand, we note that $\Omega(n^{\lfloor k/3 \rfloor})$ different $Z^{\otimes k}$-type logical operators can be constructed by picking $\lfloor k/3 \rfloor$ groups out of the entire $\Theta(n)$ groups and choose one $Z^{\otimes 3}$-type logical operator, $\lfloor k/3 \rfloor - 2$ $Z^{\otimes 3}$-type stabilizers and one $Z^{\otimes k-3\lfloor k/3 \rfloor+3}$-type stabilizer from each group. On the other hand, the number of $Z^{\otimes k}$-type logical operators is at most $O(n^{\lfloor k/3 \rfloor})$ because any $Z^{\otimes k}$-type logical operator can be supported on at most $\lfloor k/3 \rfloor$ groups, and the number of subsets containing at most $\lfloor k/3 \rfloor$ groups is $\Theta(n^{\lfloor k/3 \rfloor})$.

3.3 Quantum Reed-Muller Code

The stabilizer group of a CSS code is generated by operators from two subsets S_x and S_z, where the subset S_x (S_z) consists of stabilizer generators that are tensor products of X (Z) and I. Here, we consider *quantum Reed–Muller (QRM) codes* that are CSS codes whose S_x and S_z are obtained from the parity-check matrices of two *classical Reed–Muller (RM) codes* [37, 69].

We first review the definition of classical RM codes [70]. Define $\boldsymbol{v}_0^T = (1,1,\ldots,1) =: \mathbf{1}_{2^m}^T$ as the 2^m-tuple row vector consisting entirely of 1s. Throughout this section, we will use the notation $\mathbf{0}_p$ to denote a vector of p zeros, and $\mathbf{1}_p$ to denote a vector of p ones. For any integer i with $1 \le i \le m$, define the 2^m-tuple row vector $\boldsymbol{v}_i^T = (\mathbf{0}_{2^{i-1}}^T | \mathbf{1}_{2^{i-1}}^T | \mathbf{0}_{2^{i-1}}^T | \ldots | \mathbf{1}_{2^{i-1}}^T)$, where "$|$" denotes the concatenation of row vectors. The generator matrix of $RM(1,m)$, i.e., the 1st order RM code of length 2^m, is given by

$$\mathsf{G}(1,m) = \begin{pmatrix} —\boldsymbol{v}_0^T— \\ —\boldsymbol{v}_1^T— \\ \vdots \\ —\boldsymbol{v}_m^T— \end{pmatrix}. \tag{14}$$

The r-th order RM code, denoted by $RM(r,m)$, is constructed by taking the Boolean products (denoted by $*$) of the codewords of a 1st-order RM code. In particular, the generator matrix $\mathsf{G}(r,m)$ of $RM(r,m)$ is obtained by taking the rows of $RM(1,m)$ along with all the distinct $2,3,\ldots,r$-fold Boolean products of the rows of $\mathsf{G}(1,m)$. For example, the set of rows of the generator matrix $\mathsf{G}(2,3)$ is $\{\boldsymbol{v}_0^T, \boldsymbol{v}_1^T, \boldsymbol{v}_2^T, \boldsymbol{v}_3^T, \boldsymbol{v}_1^T * \boldsymbol{v}_2^T, \boldsymbol{v}_1^T * \boldsymbol{v}_3^T, \boldsymbol{v}_2^T * \boldsymbol{v}_3^T\}$, where $\mathsf{G}(1,3)$ corresponds to the first four rows. It can be shown that $RM(r,m)$ is a $[2^m, N(m,r), 2^{m-r}]$ code, where $N(m,r) := 1 + \binom{m}{1} + \binom{m}{2} + \cdots + \binom{m}{r}$ and $RM(r,m)^\perp = RM(m-r-1,m)$ [70].

The QRM codes are derived from shortened RM codes $\overline{RM}(r,m)$ whose generating matrix $\overline{\mathsf{G}}(r,m)$ is obtained by deleting the first row and the first column of $\mathsf{G}(r,m)$. The following three properties of a shortened RM code $\overline{RM}(r,m)$ will be useful to us. Their proofs are given in Appx. B.

(1) $\overline{RM}(1,m)$ is strictly contained in $\overline{RM}(m-2,m)$.

(2) The vector $\mathbf{1}_{2^m-1}$ and those corresponding to the rows of $\overline{\mathsf{G}}(m-2,m)$ form a basis for $\overline{RM}(1,m)^\perp$.

(3) The number of codewords in $\overline{RM}(1,m)^\perp$ with weight (2^m-4) is $\frac{1}{6}(4^m - 3\cdot 2^m + 2)$.

Here, we focus on $QRM(1,m)$ which is an $[[n=2^m-1,1,3]]$ code (see a definition of general $QRM(r,m)$ in e.g., [71]). The set of X-type stabilizer generator S_x is obtained from the rows of $\overline{\mathsf{G}}(1,m)$ by substituting 1s by Xs and 0s by Is in the tensor product forms of the operators. The elements of S_z are similarly obtained from $\overline{\mathsf{G}}(m-2,m)$ by replacing 1s by Zs and 0s by Is. Due to Property (1), the basis codewords of the CSS code $QRM(1,m)$ are given by $|z + \overline{RM}(1,m)\rangle = \frac{1}{2^m}\sum_{y\in\overline{RM}(1,m)} |z+y\rangle$ for $z \in \overline{RM}(m-2,m)$ [22]. Note that $Z^{\otimes n}$ commutes with all the X-type stabilizer generators since each generator is of even weight; however, $Z^{\otimes n}$ is not in the stabilizer group by Property (2). Thus, $Z^{\otimes n}$ is a logical operator and one can choose a specific basis in the logical subspace such that it is a logical Z operator, i.e.,

$$\overline{Z} := Z^{\otimes n} = Z^{\otimes(2^m-1)}. \tag{15}$$

Following Lemma 1, we obtain the scaling of the QFI with respect to n by calculating the number of weight-3 representatives of the logical Pauli-Z operator.

Theorem 2. *The $[[n=2^m-1,1,3]]$ $QRM(1,m)$ code achieves a QFI of $n^2(n-1)^2t^2/9=\Theta(n^4t^2)$.*

Proof. Operators equivalent to the logical Z operator in Eq. (15) are obtained by multiplying $\overline{Z}$ with different stabilizers. Since the weight of $\overline{Z}$ is 2^m-1, it follows that the number of $Z^{\otimes 3}$-type logical operators ℓ is the same as the number of $Z^{\otimes(n-3)}=Z^{\otimes(2^m-4)}$ stabilizers. This amounts to calculating the number of weight-(2^m-4) codewords in the shortened code $\overline{RM}(m-2,m)$.

Below we prove the number of weight-(2^m-4) codewords in $\overline{RM}(m-2,m)$ is equal to that in $\overline{RM}(1,m)^{\perp}$. Using Property (3), this implies that $\ell=\frac{1}{6}(4^m-3\cdot 2^m+2)=\frac{1}{6}n(n-1)$. Combining with Lemma 1, the theorem is then proven.

To prove the statement above, let $V=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\ldots,\boldsymbol{v}_q\}$ be the set of weight-(2^m-4) vectors in $\overline{RM}(m-2,m)$. Using Property (2), it follows that $V\cup\{\mathbf{1}_n+\boldsymbol{u}_i:1\le i\le p\}$ is the set of weight-(2^m-4) vectors in $\overline{RM}(1,m)^{\perp}$, where $U=\{\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_p\}$ is the set of weight-3 vectors in $\overline{RM}(m-2,m)$. We know that the weights of the codewords of the unshortened $RM(m-2,m)$ are even[4]. Therefore, the weight-3 codewords of $\overline{RM}(m-2,m)$ are exactly the shortened versions of the weight-4 codewords of $RM(m-2,m)$ that have 1 as the first entry. So, for each $\boldsymbol{u}_i\in U$, we have a distinct weight-4 vector $(1|\boldsymbol{u}_i^T)^T\in RM(m-2,m)$. This vector is of the form $(1|\boldsymbol{u}_i^T)^T=\mathbf{1}_{2^m}+\boldsymbol{w}$, where $\boldsymbol{w}\in RM(m-2,m)$ is a weight-(2^m-4) vector whose first entry is 0. Hence, $\boldsymbol{w}^T=(0|\boldsymbol{v}_j^T)$ for some $\boldsymbol{v}_j\in V$. This implies that $\boldsymbol{v}_j=\mathbf{1}_n+\boldsymbol{u}_i$. Therefore, $V\cup\{\mathbf{1}_n+\boldsymbol{u}_i:1\le i\le p\}=V$, thus completing the proof. □

The QRM code achieves the SHS, which outperforms the thin surface code. As we will see, we can achieve a better scaling using Shor codes.

3.4 Shor Code and Generalization

We now consider the n-qubit Shor code, which is a simple generalization of the 9-qubit Shor code [38], where $n=3n_r$ is divisible by 3. The code is constructed by concatenating a phase flip code $\left(\frac{|0\rangle\pm|1\rangle}{\sqrt{2}}\right)^{\otimes 3}$ of distance three with an n_r-qubit repetition code. (Note that here a concatenation of code 1 and code 2 means encoding every physical qubit of code 1 using code 2.) The logical basis states are given by

$$|0_{\mathrm{L}}\rangle=\frac{1}{\sqrt{8}}\left(|0\rangle^{\otimes n_r}+|1\rangle^{\otimes n_r}\right)^{\otimes 3}, \tag{16a}$$

[4]Can be shown using the Macwilliams identity [70] (see Appx. B).

Name	Operator
g_1	$Z\ Z$
g_2	$\ \ Z\ Z$
$\vdots$	$\ \ \ \ \ddots$
g_{n_r-1}	$\ \ \ \ \ \ \ Z\ Z$
g_{n_r}	$\ \ \ \ \ \ \ \ \ Z\ Z$
$\vdots$	$\ \ \ \ \ \ \ \ \ \ \ \ddots$
g_{2n_r-2}	$\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ Z\ Z$
g_{2n_r-1}	$\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ Z\ Z$
$\vdots$	$\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ddots$
g_{3n_r-3}	$\ Z\ Z$
g_{3n_r-2}	$X\ X\ X\ \cdots\ X\ X\ \ X\ X\ \cdots\ X\ X$
g_{3n_r-1}	$\ X\ X\ \cdots\ X\ X$

Table 1: A set of stabilizer generators of the Shor code. The code is divided into three blocks of length n_r. Each block supports (n_r-1) Z-type stabilizers, and two consecutive Z-type stabilizers overlap on one qubit within each block. Two Z-type stabilizers from different blocks don't overlap. There are two X-type stabilizer generators g_{3n_r-2} and g_{3n_r-1}, each with support on a string of $2n_r$ qubits. The supports of the two X-type generators overlap on the middle block of qubits.

$$|1_{\mathrm{L}}\rangle=\frac{1}{\sqrt{8}}\left(|0\rangle^{\otimes n_r}-|1\rangle^{\otimes n_r}\right)^{\otimes 3}. \tag{16b}$$

A particularly nice set of stabilizer generators for the code is given in Table 1. It is clear that one logical qubit is encoded into $n=3n_r$ physical qubits.

One set of representatives of the logical X and logical Z operators is

$$\overline{X}:=\left(Z\otimes I^{\otimes(n_r-1)}\right)^{\otimes 3}, \tag{17a}$$

$$\overline{Z}:=X^{\otimes n}. \tag{17b}$$

Note that the logical operator $\overline{X}$ in Eq. (17a) is a term in $G=\sum_{i<j<k}Z_iZ_jZ_k$. Thus, all $Z^{\otimes 3}$-type logical operators must be representatives of the logical X operator. From the definition of the logical basis states in Eq. (16), it is clear that $Z_iZ_jZ_k$ is a logical operator if and only if the three Z operators are supported on different blocks. There are $n_r^3=(n/3)^3$ such logical operators. Using Lemma 1, we have the following result.

Theorem 3. *The $[[n,1,3]]$ Shor code achieves a QFI of $4n^6t^2/27=\Theta(n^6t^2)$.*

We also note that the above discussion on the Shor code can be generalized to estimating $Z^{\otimes k}$ terms in Hamiltonian for any constant $k\ge 3$. For example, when we choose $n_r:=n/k$ (assuming n is divisible by k) and assume n is sufficiently large, then the $[[n,1,k]]$ code

$$|0_{\mathrm{L}}\rangle=\frac{1}{\sqrt{2^k}}\left(|0\rangle^{\otimes n_r}+|1\rangle^{\otimes n_r}\right)^{\otimes k}, \tag{18a}$$

$$|1_{\mathrm{L}}\rangle=\frac{1}{\sqrt{2^k}}\left(|0\rangle^{\otimes n_r}-|1\rangle^{\otimes n_r}\right)^{\otimes k}, \tag{18b}$$

has $\Theta(n^k)$ $Z^{\otimes k}$-type logical operators, implying a QFI of $\Theta(n^{2k}t^2)$, which achieves the optimal SHS scaling, the same as the noiseless case.

Inspired by the repetition code substructure in the Shor code, we have a more general statement:

Theorem 4. *Let an $[[n,1]]$ code C be the concatenation of a $[[q,1]]$ stabilizer code C' that has at least one $Z^{\otimes 3}$-type logical operator and can correct single-qubit Pauli-Z errors and the $[[n/q,1]]$ repetition code, where q is a constant integer and n is divisible by q. Then C achieves a QFI of $\Theta(n^6t^2)$.*

Proof. We need to prove (1) C has $\Theta(n^3)$ $Z^{\otimes 3}$-type logical operators and (2) C is an $[[n,1,3]]$ stabilizer code.

Suppose that $Z_1Z_2Z_3$ is a logical operator for C', it then implies $Z_iZ_jZ_k$ are logical operators for C whenever $1 \le i \le n_r$, $n_r+1 \le j \le 2n_r$, $2n_r+1 \le k \le 3n_r$ where $n_r := n/q$. Note that in C, qubits with labels from q_1n_r+1 to $(q_1+1)n_r$ correspond to the q_1-th qubit in C' expanded from the repetition code concatenation for any q_1. The number of different $Z_iZ_jZ_k$ above are $n_r^3 = \Theta(n^3)$, proving statement (1).

To prove statement (2), we need to show C can correct any single-qubit error, i.e., any operator S_iS_j for $S = I, X, Z$ must either belong to the stabilizer group of C or anti-commute with one stabilizer of C. First, note that for any i, X_i must anti-commute with Z_iZ_k for any $k \neq i$ and $(\lceil i/n_r \rceil - 1)n_r + 1 \le k \le \lceil i/n_r \rceil n_r$ which is a stabilizer of C. It implies any S_iS_j anti-commutes with some stabilizer of C whenever one of them is a Pauli-X operator. On the other hand, any n-qubit operator S_iS_j with $S = I, Z$ can be mapped to a q-qubit operator $S_{\lceil i/n_r \rceil}S_{\lceil j/n_r \rceil}$ acting on C' which either belongs to the stabilizer group of C' or anti-commutes with one stabilizer of C' because C' corrects single-qubit Pauli-Z errors. It implies the required condition is also satisfied for S_iS_j with $S = I, Z$. Statement (2) is then proven. □

Theorem 3 is a special case of Theorem 4 where C' is the three-qubit phase-flip code. As discussed in Sec. 2, the above codes provide the optimal SHS scaling we can achieve for estimating the ZZZ interaction under single-qubit noise (in the setting of Eq. (1) and Eq. (2)). Note that Theorem 4 can also be straightforwardly generalized to estimating $Z^{\otimes k}$ terms for $k > 3$ if we replace $Z^{\otimes 3}$ in the statement of Theorem 4 by $Z^{\otimes k}$ and $\Theta(n^6t^2)$ by $\Theta(n^{2k}t^2)$.

4 Necessary Conditions for Achieving the SHS

In the previous section, we discussed three examples of quantum codes that asymptotically achieve $(nt)^{-1}$, $(n^2t)^{-1}$, and $(n^3t)^{-1}$ precision scalings for the ZZZ interaction estimation problem. The three families of codes have special stabilizer structures: the thin surface code has a set of low-weight stabilizer generators; the QRM code lacks a complete set of low-weight stabilizer generators; and the Shor code has local low-weight Z-type generators even though the X-type generators scaled linearly with the length of the code. In this section, we address the question of what properties a stabilizer code must have to achieve different precision scalings with respect to n, or equivalently, different scalings of the number of $Z^{\otimes 3}$-type logical operators with respect to n.

4.1 Unachievability of the SHS

We proved in Theorem 1 that the thin surface code can be used to achieve $1/n$ precision scaling in ω. In fact, this scaling can more generally apply to stabilizer codes with constant weight generators that can correct arbitrary single qubit errors. We want to find an upper bound on the number of $Z^{\otimes 3}$-type logical operators for such codes, which will consequently give us the best possible precision scaling.

Theorem 5. *Let C be an $[[n,1,3]]$ stabilizer code such that there exists a set of stabilizer generators with weights at most w. Then, ℓ, the number of $Z^{\otimes 3}$-type logical operators of C is at most $2w(w+1)n/3$. In particular, if w is a constant independent of n, then $\ell = O(n)$.*

Proof. Let S be a stabilizer generator set of C such that the weight of every $g \in \mathsf{S}$ satisfies $\mathrm{wt}(g) \le w$. Without loss of generality, we assume there are no Z_i-type stabilizers. Otherwise, we can always remove from the entire set of n qubits the set of qubits on which a single Pauli-Z are stabilizers, i.e., qubits that are always in states $|0\rangle$. Then the remaining $\tilde{n}$ qubits give an $[[\tilde{n},1,3]]$ stabilizer code $\tilde{C}$ satisfying the same weight constraint without Z_i-type stabilizers. Any $Z^{\otimes 3}$-type logical operator in C must be fully supported on the remaining $\tilde{n}$ qubits and is thus also a $Z^{\otimes 3}$-type logical operator in $\tilde{C}$. Then by proving the theorem for $\tilde{C}$, the theorem for C also holds.

We define the *degree* of a qubit as the number of $Z^{\otimes 3}$-type logical operators that are supported on it. Without loss of generality, assume that the first qubit (i.e., the qubit labeled by 1) has the highest degree. Define a graph $\mathfrak{G}$ with $(n-1)$ vertices corresponding to the qubits labeled by $2, 3, \ldots, n$. There is an edge between the i-th and j-th vertices if $Z_1Z_iZ_j$ is a logical operator of the code. We define the *order* of a vertex as the number of edges connected to it.

Consider the two separate cases: (1) There is some vertex with order $> w$, and (2) The order of every vertex is $\le w$.

Case 1: Suppose that the order of vertex i_1 is greater than w. Then, there are vertices $i_2, i_3, \ldots, i_{w+2}$ that are connected to vertex i_1, as shown in Fig. 2. The product of any two $Z^{\otimes 3}$-type

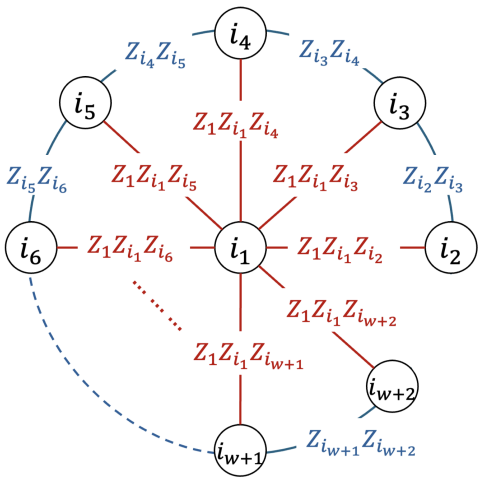

Figure 2: A subgraph of $\mathfrak{G}$ where vertex i_1 has an order greater than w. Here, the red edges correspond to the $Z^{\otimes 3}$-type logical operators. Since the product of two such logical operators is a stabilizer, it follows that $Z_{i_j}Z_{i_{j+1}}$ is a stabilizer for all $2 \leq j \leq w+1$. We graphically denote these stabilizers by blue lines (note that these are not edges in $\mathfrak{G}$).

logical operators is a stabilizer because they must be representatives of a same logical operator. It then follows that $Z_{i_j}Z_{i_{j+1}}$ is a stabilizer for all $2 \leq j \leq w+1$. Now, since Z_{i_2} is not a stabilizer and must be an error operator that anti-commutes with some stabilizer generator, there is some stabilizer generator g with X or Y acting on the qubit labeled by i_2. Because g commutes with the stabilizer $Z_{i_2}Z_{i_3}$, it follows that g has X or Y support on the qubit labeled by i_3 as well. Similarly, we argue that g must have X or Y support on the qubits labeled by $i_4, i_5, \ldots, i_{w+2}$. Then, $\mathrm{wt}(g) > w$, which is a contradiction. Hence, this case is not possible for C.

Case 2: Suppose that the order of each vertex is $\leq w$. Let d_1 be the number of logical $Z^{\otimes 3}$ operators with support on the first qubit, i.e., the number of edges in $\mathfrak{G}$. First, we will prove that for any subset of vertices V that is a minimum *vertex cover*[5] of $\mathfrak{G}$, the size of V satisfies $|V| \geq \lfloor \frac{d_1}{2w} \rfloor$.

The *matching* of a graph is defined as a subset of its edges such that no two edges share a vertex. Let M be a maximum matching of $\mathfrak{G}$. Due to the disjointness of the edges in M, each vertex in V covers at most one edge of the maximum matching. Therefore, $|V| \geq |M|$. So, it is sufficient to prove that $|M| \geq \lfloor \frac{d_1}{2w} \rfloor$. Note that for $w = 1$, the edges are disjoint; hence, $|M| = d_1 \geq \lfloor \frac{d_1}{2w} \rfloor$. Now, suppose $w > 1$. To generate an arbitrary matching, we consider the set of all edges. First, we pick a random edge from it and remove all the edges that share endpoints with it from the set. Since the highest order of the vertices is $\leq w$, the number of edges removed is at most $2(w-1)$. We choose another random edge from the set and continue this process until the set is a matching. In the end, we will obtain a matching of size $\geq \lfloor \frac{d_1}{2(w-1)+1} \rfloor \geq \lfloor \frac{d_1}{2w} \rfloor$. Hence, $|M| \geq \lfloor \frac{d_1}{2w} \rfloor$, which, in turn, implies that $|V| \geq \lfloor \frac{d_1}{2w} \rfloor$ for all w.

Now, the code must have a stabilizer generator g that is X or Y on the first qubit because each Z_i must anti-commute with some stabilizer generator. Suppose there is an edge between vertices labeled by i and j, and so $Z_1Z_iZ_j$ is a logical operator. Since g must commute with $Z_1Z_iZ_j$, it follows that g has X or Y support on either the i-th or j-th qubit. Therefore, the vertices corresponding to the qubits that support X or Y of g form a vertex cover of $\mathfrak{G}$. Hence, $\mathrm{wt}(g) \geq \lfloor \frac{d_1}{2w} \rfloor$. However, by the constant weight assumption, $w \geq \mathrm{wt}(g)$, which implies that $\lfloor \frac{d_1}{2w} \rfloor \leq w$. So, $d_1 \leq 2w(w+1)$. Recall that d_1 is the number of logical $Z^{\otimes 3}$ operators with support on the first qubit (qubit with the highest degree). Hence, the total number of $Z^{\otimes 3}$-type logical operators satisfies $\ell \leq d_1 n/3 \leq 2w(w+1)n/3$. The first inequality here is from Lemma 2 that we will prove below separately. □

Here we prove a lemma where we establish a simple relation between the number of $Z^{\otimes 3}$-type logical operators and the degrees of qubits, which was used in the proof above and will also be used in later sections.

Lemma 2. *Suppose the number of $Z^{\otimes 3}$-type logical operators in an n-qubit stabilizer code is ℓ. If d_i is the degree of the i-th qubit of the code, then*

$$\max\{d_i\}_{i=1}^n \geq \frac{3\ell}{n}. \tag{19}$$

Proof. Recall that the degree of a qubit is defined to be the number of $Z^{\otimes 3}$-type logical operators that are supported on it. Without loss of generality, assume that the first qubit has the highest degree, i.e., $d_1 = \max\{d_i\}_{i=1}^n$. The sum of the weights of all $Z^{\otimes 3}$-type logical operators is 3ℓ. This must be equal to the sum of the degrees of the qubits. So, $\sum_{i=1}^n d_i = 3\ell$. But $\sum_{i=1}^n d_i \leq nd_1$. Hence, it follows that $nd_1 \geq 3\ell$, and so $\max\{d_i\}_{i=1}^n \geq \frac{3\ell}{n}$. □

From Theorem 5 and Lemma 1, the best possible scaling of the estimation precision for the ZZZ interaction using a stabilizer code with constant-weight stabilizer generators is the HS of n^{-1}. This family of stabilizer codes include quantum LDPC codes which are an important class of stabilizer codes as a candidate to achieve fault tolerance [39]. The thin surface code is one class of quantum LDPC codes, and by Theorem 1, it obtains the optimal scaling.

Theorem 5 also applies to cases where w increases with respect to n. For example, one corollary of Theorem 5 is when $w = o(n)$, the optimal SHS n^{-3} cannot

[5] A vertex cover is a subset of the vertices such that at least one endpoint of each edge is in the subset. A minimum vertex cover is such a subset with the minimum number of vertices.

be achieved. Note that it is also a corollary of Theorem 7 that we will present later in Sec. 4.3.

4.2 Unachievability of the Optimal SHS

We now relax the constant-weight assumption on the stabilizer generators, in which case we want to establish a different no-go result that precludes a stabilizer code from achieving a precision scaling better than $1/n^2$.

A notable feature of the Shor code is its $Z^{\otimes 2}$-type stabilizer generators, as shown in Table 1. The following theorem demonstrates that the existence of such stabilizers is necessary for achieving the optimal SHS of n^{-3}.

Theorem 6. *Let C be an $[[n,1]]$ stabilizer code. If C does not have a $Z^{\otimes 2}$-type stabilizer, then the number of $Z^{\otimes 3}$-type logical operators is $\ell \leq n(n-1)/6 = O(n^2)$.*

Proof. Let $\mathcal{A} = \{A_1, A_2, \dots A_\ell\}$ be the set of $Z^{\otimes 3}$-type logical operators. Let d_i be the degree of the i-th qubit. Without loss of generality, we make the following three assumptions: (1) $d_1 = \max\{d_i\}_{i=1}^n$, (2) $\mathcal{A}' = \{A_i\}_{i=1}^{d_1}$ is the set of operators with support on the first qubit, and (3) The first n' qubits is the minimum set of qubits that contains the supports of operators in $\mathcal{A}'$. All three assumptions can be satisfied by relabeling the qubits. The operators in $\mathcal{A}$ are different representatives of the same logical operator, and so the product of any two $A_i, A_j \in \mathcal{A}'$ is a stabilizer. If there is no $Z^{\otimes 2}$-type stabilizer, then each pair of operators in $\mathcal{A}'$ only overlap on the first qubit. Hence, each operator in $\mathcal{A}'$ is supported on two distinct qubits other than the first qubit. Therefore, $n' = 2d_1 + 1$. However, since $n' \leq n$, we have $2d_1 + 1 \leq n$. Combining this inequality with Eq. (19) in Lemma 2, we get

$$\frac{3\ell}{n} \leq d_1 \leq \frac{n-1}{2}. \tag{20}$$

Therefore, $\ell \leq \frac{n(n-1)}{6}$, and so if there are no $Z^{\otimes 2}$-type stabilizers, then $\ell = O(n^2)$. □

The scaling of QRM codes derived in Theorem 2 agrees with this result. One can show using the enumerator polynomial in Eq. (31) of Appx. B that QRM codes do not have any $Z^{\otimes 2}$-type stabilizer. Therefore, by Theorem 6 and Theorem 2, the QRM codes achieve the optimal precision scaling $1/n^2$ among all stabilizer codes without $Z^{\otimes 2}$-type stabilizers.

Finally, we discuss generalization to the estimation of higher-order interactions $Z^{\otimes k}$ for $k \geq 3$. We claim that for any $1 \leq k_0 < k-1$, for stabilizer codes without $Z^{\otimes i}$-type stabilizers for all $i \leq 2k_0$, the number of $Z^{\otimes k}$-logical operators is at most $O(n^{k-k_0})$ which is a factor of n^{k_0} away from the optimal scaling. It implies ZZ-type stabilizers are necessary to achieve the optimal scaling $O(n^k)$. Theorem 6 represents the special case where $k = 3$ and $k_0 = 1$.

To see why the generalized statement is true, we use the generalized Ray-Chaudhuri–Wilson theorem [72] which was a celebrated intersection theorem in combinatorics. Consider $\mathscr{F}$, a family of subsets of $\{1, 2, \dots, n\}$. If $|F| = k$ for any $F \in \mathscr{F}$ and $|E \cap F| < k - k_0$ for any distinct subsets $E, F \in \mathscr{F}$. Then the generalized Ray-Chaudhuri–Wilson theorem states that $|\mathscr{F}| \leq \binom{n}{k-k_0} = \Theta(n^{k-k_0})$. In our setting, we pick $\mathscr{F}$ to be the family of subsets of qubits $\{1, 2, \dots, n\}$ such that $F \in \mathscr{F}$ means F is the support of a $Z^{\otimes k}$-type logical operator. When $|E \cap F| = i$, it implies the existence of a $Z^{\otimes 2(k-i)}$-type stabilizer which is equal to two $Z^{\otimes k}$-type logical operators supported on E and F. The non-existence of $Z^{\otimes j}$-type stabilizers with $j \leq 2k_0$ implies that $2(k - |E \cap F|) > 2k_0$ and that $|E \cap F| < k - k_0$. Then we have an upper bound of $\Theta(n^{k-k_0})$ on $|\mathscr{F}|$, i.e., the number of $Z^{\otimes k}$-type logical operators.

4.3 Necessary Condition for Achieving the Optimal SHS

As we discussed, the existence of a $Z^{\otimes 2}$-type stabilizer is necessary for achieving the optimal SHS $\sim n^{-3}$. Note that the stabilizer group of Shor codes that achieve the optimal SHS contains a chain of $Z^{\otimes 2}$-type stabilizers. Motivated by that example, we provide a necessary condition for a stabilizer code that achieves this optimal SHS. Here, we prove that if the optimal SHS is achievable by a stabilizer code, then the stabilizer group contains such a chain of linear length.

Theorem 7. *Let C_n be a family of $[[n,1,3]]$ stabilizer codes. Suppose that the number of $Z^{\otimes 3}$-type logical operators is $\ell = \Theta(n^3)$. Then, for all n, there is a subset of qubits indexed by $\{i_1, \dots, i_{n_r}\}$ such that*

(1) $\{Z_{i_1}Z_{i_2}, Z_{i_2}Z_{i_3}, \dots, Z_{i_{n_r-1}}Z_{i_{n_r}}\}$ is a subset of the stabilizers of C_n;

(2) For any $1 \leq k \leq r$, Z_{i_k} is not a stabilizer of C_n;

(3) $n_r = \Theta(n)$.

Proof. Consider C_n. Let $\mathcal{A} = \{A_1, A_2, \dots, A_\ell\}$ be the set of $Z^{\otimes 3}$-type logical operators. Let d_i denote the degree of the i-th qubit. Without loss of generality, we make the same three assumptions we made in Theorem 6: (1) $d_1 = \max\{d_i\}_{i=1}^n$, (2) The set of operators with support on the first qubit is $\mathcal{A}' = \{A_i\}_{i=1}^{d_1}$, and (3) The first n' qubits is the minimum set of qubits that contains the supports of operators in $\mathcal{A}'$.

Define a graph $\mathfrak{G}$ with $(n'-1)$ vertices $\{v_2, v_3, \dots, v_{n'}\}$ corresponding to the qubits indexed by $2, 3, \dots, n'$, where an edge connects v_i and v_j if $Z_1Z_iZ_j \in \mathcal{A}'$. In total, there are $d_1 = |\mathcal{A}'|$ edges in the graph. Since $\ell = \Theta(n^3)$, Lemma 2 implies that $d_1 = \Omega(n^2)$. However, $\mathfrak{G}$ has $(n'-1)$ vertices,

and so it has at most $d_1 = O(n'^2) = O(n^2)$ edges. Thus, $d_1 = \Theta(n^2)$. Therefore, we can assume $\mathfrak{G}$ has $(n'-1) \geq bn$ vertices with $d_1 \geq an^2$ edges for some $a, b > 0$ if we choose n to be sufficiently large.

Note that the product of an even number of commuting logical operators ($Z^{\otimes 3}$-type operators here to be specific) in $\mathcal{A}$ is a stabilizer. Therefore, if there is an even path (a path consisting of an even number of edges) between v_i and v_j in $\mathfrak{G}$, then Z_iZ_j is a stabilizer. We will show that we can always identify a subset of vertices of size $\Theta(n)$ such that there exists an even path between any pair of adjacent vertices.

By Lemma S1, $\mathfrak{G}$ has a connected component $\mathfrak{G}'$ with $n_1' = \Theta(n)$ vertices and $d_1' = \Theta(n^2)$ edges. Let the number of edges in $\mathfrak{G}'$ be $d_1' \geq (c/2)n_1'^2$ for some constant $c \in (0,1)$. Suppose $\mathfrak{G}'$ has $\geq c_1 n_1'$ vertices whose orders are $\leq c_2 n_1'$, where $c_1, c_2 \in (0,1)$. Then, the sum of the orders of the vertices is at most $c_1 n_1' \cdot c_2 n_1' + (1-c_1)n_1' \cdot n_1'$. The sum of the orders of the vertices of a graph is twice the total number of edges. Therefore, $2 \cdot (c/2)n_1'^2 \leq c_1 c_2 n_1'^2 + (1-c_1)n_1'^2$, and so

$$c \leq 1 - c_1 + c_1 c_2. \tag{21}$$

Note that if we choose $c_1 = 1 - \frac{c}{2}$ and $c_2 = \frac{c}{3}$, we have $c \leq 1 - c_1 + c_1 c_2 = 1 - \left(1 - \frac{c}{2}\right) + \left(1 - \frac{c}{2}\right)\frac{c}{3} < \frac{c}{2} + \frac{c}{3} < c$, which is a contradiction. Therefore, $\mathfrak{G}'$ has more than $(1-c_1)n_1' = \frac{c}{2}n_1'$ vertices with orders greater than $\frac{c}{3}n_1'$.

Reindex the vertices in descending order of their vertex orders, that is v_2 has the highest order, v_3 has the second-highest order and so on. Let T be the set of the first $\lfloor \frac{c}{6}n_1' \rfloor$ vertices. Each vertex in T has an order $> \frac{c}{3}n_1'$, which, in turn, is greater than $2|T|$. Therefore, for each vertex $v_i \in T$, we can assign a unique $\tilde{v}_i \notin T$ that is connected to v_i by a single edge such that $\tilde{v}_i \neq \tilde{v}_j$ for $i \neq j$. Now, starting from $i = 2$, suppose that there is no even path between v_i and v_{i+1}. Hence, all the paths from v_i to v_{i+1} consist of an odd number of edges. Then, by construction, there is an even path between v_i and $\tilde{v}_{i+1}$, i.e., the path from v_i to v_{i+1} concatenated with the single edge from v_{i+1} to $\tilde{v}_{i+1}$. We replace v_{i+1} by $\tilde{v}_{i+1}$. Otherwise when there is already an even path between v_i and v_{i+1}, we do no replacement. Similarly, we apply the same replacement rules for $i = 3, 4, \cdots, \lfloor \frac{c}{6}n_1' \rfloor + 1$. After all the appropriate replacements, all the adjacent vertices in T are connected by even paths. Therefore, each pair of adjacent qubits support a $Z^{\otimes 2}$-type stabilizer. Furthermore, each such stabilizer is independent of the others. Hence, we conclude that, by reversing the reindexings, there exists a set of qubits labeled by $\{i_1, i_2, \ldots, i_{n_r}\}$ of size $n_r = \lfloor \frac{c}{6}n_1' \rfloor = \Theta(n)$ such that $\{Z_{i_1}Z_{i_2}, Z_{i_2}Z_{i_3}, \ldots, Z_{i_{n_r-1}}Z_{i_{n_r}}\}$ is a subset of stabilizers.

From the discussion above, for any $1 \leq k \leq r$, there exists some qubit j_k such that $Z_1 Z_{i_k} Z_{j_k}$ is a logical operator in C_n because Z_{i_k} must belong to some connected component of $\mathfrak{G}$. It implies Z_{i_k} is not a stabilizer of C_n because otherwise $Z_1 Z_{j_k} = Z_1 Z_{i_k} Z_{j_k} Z_{i_k}$ is a logical operator of C_n, contradicting with the fact that the code distance is 3. □

Note that having $\{Z_{i_1}Z_{i_2}, \ldots, Z_{i_{n_r-1}}Z_{i_{n_r}}\}$ as a part of the stabilizers of C guarantees the codewords on qubits $\{i_1, \ldots, i_{n_r}\}$ must be in the span of $\{|0^{\otimes n_r}\rangle, |1^{\otimes n_r}\rangle\}$. Furthermore, as $\{Z_{i_k}\}_{1 \leq k \leq r}$ are not stabilizers, for any choice of stabilizer generators there must be one of them that acts non-trivially as X or Y on all qubits in $\{i_1, \ldots, i_{n_r}\}$. It implies the necessary existence of a weight-$\Theta(n)$ stabilizer generator when the optimal SHS is achieved. Theorem 7 shows the repetition code substructure described above is necessary in achieving the optimal SHS.

5 Conclusion and Outlook

In conclusion, we studied how different stabilizer codes perform when estimating the parameter ω encoded in the 3-local Hamiltonian in Eq. (2) under single-qubit noise. We also derived what properties stabilizer codes must satisfy in order to asymptotically achieve the SHS or the optimal SHS. Our work suggests more investigation towards the direction of exploiting the stabilizer codes for many-body quantum metrology.

The crux of the solution of our problem was the relation between the QFI and the number of logical $Z^{\otimes 3}$-type operators of the stabilizer code, which connects quantum metrology to the properties of stabilizer codes. Based on it, some of our results can be straightforwardly generalized to probing higher-order interactions under stronger quantum noise. For example, the $[[n, 1, 3]]$ thin surface code, QRM code and Shor code discussed in Sec. 3 all have natural generalizations to $[[n, 1, d > 3]]$ codes, that may be analogously applied in estimating higher-order interactions.

Many interesting questions remain open. For example, our results assumed the achievement of the HL, based on which we analyzed the achievability of the SHS. It would be interesting to consider whether better scalings with respect to the number of qubits will be available if we abandon the requirement of the HL. Another question is to ask how QEC protocols behave in many-body Hamiltonian estimation when taking fault-tolerance into consideration (see e.g., [73]). Our current analysis was based on fast and precise QEC, and including measurement noise or circuit-level noise is left for future work. Our approach only considers the estimation problem for a single parameter; however we anticipate the study of QEC-assisted multiparameter estimation as a future direction.

Acknowledgements

S.Z. thanks Victor V. Albert and Michael Vasmer for helpful discussions. S.A. and S.Z. acknowledge fund-

ing provided by Perimeter Institute for Theoretical Physics, a research institute supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Colleges and Universities.

References

[1] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Advances in quantum metrology. *Nat. Photonics.*, 5(4):222, 2011.

[2] C. L. Degen, F. Reinhard, and P. Cappellaro. Quantum sensing. *Rev. Mod. Phys.*, 89(3):035002, Jul 2017.

[3] Luca Pezzè, Augusto Smerzi, Markus K. Oberthaler, Roman Schmied, and Philipp Treutlein. Quantum metrology with nonclassical states of atomic ensembles. *Rev. Mod. Phys.*, 90(3):035005, Sep 2018.

[4] Stefano Pirandola, Bhaskar Roy Bardhan, Tobias Gehring, Christian Weedbrook, and Seth Lloyd. Advances in photonic quantum sensing. *Nat. Photonics.*, 12(12):724, 2018.

[5] Lin Jiao, Wei Wu, Si-Yuan Bai, and Jun-Hong An. Quantum Metrology in the Noisy Intermediate-Scale Quantum Era. *Advanced Quantum Technologies*, page 2300218, 2023.

[6] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum-Enhanced Measurements: Beating the Standard Quantum Limit. *Science*, 306(5700):1330–1336, 2004.

[7] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum Metrology. *Phys. Rev. Lett.*, 96(1):010401, Jan 2006.

[8] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac. Improvement of Frequency Standards with Quantum Entanglement. *Phys. Rev. Lett.*, 79(20):3865–3868, Nov 1997.

[9] B. M. Escher, R. L. de Matos Filho, and L Davidovich. General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology. *Nat. Phys.*, 7(5):406, 2011.

[10] Rafał Demkowicz-Dobrzański, Jan Kołodyński, and Mădălin Guţă. The elusive Heisenberg limit in quantum-enhanced metrology. *Nat. Commun.*, 3:1063, 2012.

[11] Rafał Demkowicz-Dobrzański, Jan Czajkowski, and Pavel Sekatski. Adaptive Quantum Metrology under General Markovian Noise. *Phys. Rev. X*, 7(4):041009, Oct 2017.

[12] Sisi Zhou, Mengzhen Zhang, John Preskill, and Liang Jiang. Achieving the Heisenberg limit in quantum metrology using quantum error correction. *Nat. Commun.*, 9(1):78, 2018.

[13] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin. Quantum Error Correction for Metrology. *Phys. Rev. Lett.*, 112(15):150802, Apr 2014.

[14] G. Arrad, Y. Vinkler, D. Aharonov, and A. Retzker. Increasing Sensing Resolution with Error Correction. *Phys. Rev. Lett.*, 112(15):150801, Apr 2014.

[15] W. Dür, M. Skotiniotis, F. Fröwis, and B. Kraus. Improved Quantum Metrology Using Quantum Error Correction. *Phys. Rev. Lett.*, 112(8):080801, Feb 2014.

[16] Pavel Sekatski, Michalis Skotiniotis, Janek Kołodyński, and Wolfgang Dür. Quantum metrology with full and fast quantum control. *Quantum*, 1:27, September 2017.

[17] David Layden, Sisi Zhou, Paola Cappellaro, and Liang Jiang. Ancilla-Free Quantum Error Correction Codes for Quantum Metrology. *Phys. Rev. Lett.*, 122(4):040502, Jan 2019.

[18] Sisi Zhou and Liang Jiang. Optimal approximate quantum error correction for quantum metrology. *Physical Review Research*, 2(1):013235, 2020.

[19] Sisi Zhou and Liang Jiang. Asymptotic theory of quantum channel estimation. *PRX Quantum*, 2(1):010343, 2021.

[20] Sisi Zhou, Argyris Giannisis Manes, and Liang Jiang. Achieving metrological limits using ancilla-free quantum error-correcting codes. *Physical Review A*, 109(4):042406, 2024.

[21] Daniel Gottesman. *Stabilizer codes and quantum error correction*. California Institute of Technology, 1997.

[22] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*, volume 2. Cambridge University Press Cambridge, 2001.

[23] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen. Spin squeezing and reduced quantum noise in spectroscopy. *Phys. Rev. A*, 46(11):R6797–R6800, Dec 1992.

[24] J. J. Bollinger, Wayne M. Itano, D. J. Wineland, and D. J. Heinzen. Optimal frequency measurements with maximally correlated states. *Phys. Rev. A*, 54:R4649–R4652, Dec 1996.

[25] D. Leibfried, M. D. Barrett, T. Schaetz, J. Britton, J. Chiaverini, W. M. Itano, J. D. Jost, C. Langer, and D. J. Wineland. Toward Heisenberg-Limited Spectroscopy with Multiparticle Entangled States. *Science*, 304(5676):1476–1478, 2004.

[26] Brendon L Higgins, Dominic W Berry, Stephen D Bartlett, Howard M Wiseman, and Geoff J Pryde. Entanglement-free Heisenberg-limited phase estimation. *Nature*, 450(7168):393, 2007.

[27] Raphael Kaubruegger, Denis V Vasilyev, Marius Schulte, Klemens Hammerer, and Peter Zoller. Quantum variational optimization of Ramsey in-

terferometry and atomic clocks. *Physical Review X*, 11(4):041045, 2021.

[28] Hsin-Yuan Huang, Yu Tong, Di Fang, and Yuan Su. Learning many-body Hamiltonians with Heisenberg-limited scaling. *Physical Review Letters*, 130(20):200403, 2023.

[29] Alicja Dutkiewicz, Thomas E O'Brien, and Thomas Schuster. The advantage of quantum control in many-body Hamiltonian learning. *Quantum*, 8:1537, 2024.

[30] Sergio Boixo, Steven T Flammia, Carlton M Caves, and John M Geremia. Generalized limits for single-parameter quantum estimation. *Physical Review Letters*, 98(9):090401, 2007.

[31] Mathieu Beau and Adolfo del Campo. Nonlinear quantum metrology of many-body open systems. *Physical Review Letters*, 119(1):010403, 2017.

[32] Jan Czajkowski, Krzysztof Pawłowski, and Rafał Demkowicz-Dobrzański. Many-body effects in quantum metrology. *New Journal of Physics*, 21(5):053031, 2019.

[33] Sergio Boixo, Animesh Datta, Steven T Flammia, Anil Shaji, Emilio Bagan, and Carlton M Caves. Quantum-limited metrology with product states. *Physical Review A—Atomic, Molecular, and Optical Physics*, 77(1):012317, 2008.

[34] A Yu Kitaev. Fault-tolerant quantum computation by anyons. *Annals of physics*, 303(1):2–30, 2003.

[35] Eric Dennis, Alexei Kitaev, Andrew Landahl, and John Preskill. Topological quantum memory. *Journal of Mathematical Physics*, 43(9):4452–4505, 2002.

[36] Austin G Fowler, Matteo Mariantoni, John M Martinis, and Andrew N Cleland. Surface codes: Towards practical large-scale quantum computation. *Physical Review A—Atomic, Molecular, and Optical Physics*, 86(3):032324, 2012.

[37] Andrew M Steane. Quantum reed-muller codes. *IEEE Transactions on Information Theory*, 45(5):1701–1703, 1999.

[38] Peter W Shor. Scheme for reducing decoherence in quantum computer memory. *Physical Review A*, 52(4):R2493, 1995.

[39] Nikolas P Breuckmann and Jens Niklas Eberhardt. Quantum low-density parity-check codes. *PRX Quantum*, 2(4):040101, 2021.

[40] Christopher E Granade, Christopher Ferrie, Nathan Wiebe, and David G Cory. Robust online Hamiltonian learning. *New Journal of Physics*, 14(10):103013, 2012.

[41] Ian Hincks, Thomas Alexander, Michal Kononenko, Benjamin Soloway, and David G Cory. Hamiltonian learning with online Bayesian experiment design in practice. *arXiv:1806.02427*, 2018.

[42] Nathan Wiebe, Christopher Granade, Christopher Ferrie, and David G Cory. Hamiltonian learning and certification using quantum resources. *Physical Review Letters*, 112(19):190501, 2014.

[43] Tim J Evans, Robin Harper, and Steven T Flammia. Scalable bayesian hamiltonian learning. *arXiv:1912.07636*, 2019.

[44] Wenjun Yu, Jinzhao Sun, Zeyao Han, and Xiao Yuan. Robust and efficient Hamiltonian learning. *Quantum*, 7:1045, 2023.

[45] Zhi Li, Liujun Zou, and Timothy H Hsieh. Hamiltonian tomography via quantum quench. *Physical Review Letters*, 124(16):160502, 2020.

[46] Dominik Hangleiter, Ingo Roth, Jonás Fuksa, Jens Eisert, and Pedram Roushan. Robustly learning the Hamiltonian dynamics of a superconducting quantum processor. *Nature Communications*, 15(1):9595, 2024.

[47] Daniel Stilck França, Liubov A Markovich, V. V. Dobrovitski, Albert H. Werner, and Johannes Borregaard. Efficient and robust estimation of many-qubit Hamiltonians. *Nature Communications*, 15(1):311, 2024.

[48] Anurag Anshu, Srinivasan Arunachalam, Tomotaka Kuwahara, and Mehdi Soleimanifar. Sample-efficient learning of quantum many-body systems. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 685–691. IEEE, 2020.

[49] Jeongwan Haah, Robin Kothari, and Ewin Tang. Optimal learning of quantum Hamiltonians from high-temperature Gibbs states. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 135–146. IEEE, 2022.

[50] Eyal Bairey, Itai Arad, and Netanel H Lindner. Learning a local Hamiltonian from local measurements. *Physical Review Letters*, 122(2):020504, 2019.

[51] Xiao-Liang Qi and Daniel Ranard. Determining a local Hamiltonian from a single eigenstate. *Quantum*, 3:159, 2019.

[52] Marcin Jarzyna and Rafał Demkowicz-Dobrzański. True precision limits in quantum metrology. *New Journal of Physics*, 17(1):013010, 2015.

[53] Jesús Rubio, Paul Knott, and Jacob Dunningham. Non-asymptotic analysis of quantum metrology protocols beyond the Cramér–Rao bound. *Journal of Physics Communications*, 2(1):015027, 2018.

[54] Vittorio Gorini, Andrzej Kossakowski, and Ennackal Chandy George Sudarshan. Completely positive dynamical semigroups of N-level systems. *Journal of Mathematical Physics*, 17(5):821–825, 1976.

[55] Goran Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48:119–130, 1976.

[56] Heinz-Peter Breuer and Francesco Petruccione. *The theory of open quantum systems*. OUP Oxford, 2002.

[57] Alexander S Holevo. *Probabilistic and statistical aspects of quantum theory*, volume 1. Springer Science & Business Media, 2011.

[58] Carl W Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1:231–252, 1969.

[59] Samuel L Braunstein and Carlton M Caves. Statistical distance and the geometry of quantum states. *Physical Review Letters*, 72(22):3439, 1994.

[60] Matteo GA Paris. Quantum estimation for quantum technology. *Int. J. Quantum Inf.*, 7(supp01):125–137, 2009.

[61] Matteo Paris and Jaroslav Rehacek. *Quantum state estimation*, volume 649. Springer Science & Business Media, 2004.

[62] Charles H Bennett, David P DiVincenzo, John A Smolin, and William K Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A*, 54(5):3824, 1996.

[63] Emanuel Knill and Raymond Laflamme. Theory of quantum error-correcting codes. *Physical Review A*, 55(2):900, 1997.

[64] Google Quantum AI. Suppressing quantum errors by scaling a surface code logical qubit. *Nature*, 614(7949):676–681, 2023.

[65] Ciaran Ryan-Anderson, Justin G Bohnet, Kenny Lee, Daniel Gresh, Aaron Hankin, J. P. Gaebler, David Francois, Alexander Chernoguzov, Dominic Lucchetti, Natalie C. Brown, et al. Realization of real-time fault-tolerant quantum error correction. *Physical Review X*, 11(4):041058, 2021.

[66] Dolev Bluvstein, Simon J Evered, Alexandra A Geim, Sophie H Li, Hengyun Zhou, Tom Manovitz, Sepehr Ebadi, Madelyn Cain, Marcin Kalinowski, Dominik Hangleiter, et al. Logical quantum processor based on reconfigurable atom arrays. *Nature*, 626(7997):58–65, 2024.

[67] A Robert Calderbank and Peter W Shor. Good quantum error-correcting codes exist. *Physical Review A*, 54(2):1098, 1996.

[68] Andrew M Steane. Error correcting codes in quantum theory. *Physical Review Letters*, 77(5):793, 1996.

[69] Jonas T Anderson, Guillaume Duclos-Cianci, and David Poulin. Fault-tolerant conversion between the steane and reed-muller quantum codes. *Physical Review Letters*, 113(8):080501, 2014.

[70] Florence Jessie MacWilliams and Neil James Alexander Sloane. *The theory of error-correcting codes*, volume 16. Elsevier, 1977.

[71] Bei Zeng, Andrew Cross, and Isaac L Chuang. Transversality versus universality for additive quantum codes. *IEEE Transactions on Information Theory*, 57(9):6272–6284, 2011.

[72] Hunter S Snevily. A generalization of the ray-chaudhuri-wilson theorem. *Journal of Combinatorial Designs*, 3(5):349–352, 1995.

[73] Theodoros Kapourniotis and Animesh Datta. Fault-tolerant quantum metrology. *Physical Review A*, 100(2):022335, 2019.

A Optimal QFI Coefficient

Here we compute the optimal QFI coefficient from Eq. (11) that provides the optimal coefficient of the HL among all possible QEC protocols. First, note that

$$\frac{(\Delta G_{\text{eff}})_{\text{opt}}}{2} = \min_{\text{2-local } S} \left\| \sum_{i<j<k} Z_i Z_j Z_k - S \right\| \tag{22}$$

$$= \min_{\alpha_{ij},\beta_i,\gamma\in\mathbb{R}} \left\| \sum_{i<j<k} Z_i Z_j Z_k - \sum_{i<j} \alpha_{ij} Z_i Z_j - \sum_i \beta_i Z_i - \gamma I \right\|, \tag{23}$$

because any off-diagonal term (in the computational basis) will only increase the operator norm. Next, let $A(\alpha,\beta,\gamma) := \sum_{i<j<k} Z_iZ_jZ_k - \sum_{i<j}\alpha_{ij}Z_iZ_j - \sum_i \beta_i Z_i - \gamma I$, where we use α (or β) to describe the column vector whose elements are α_{ij} (or β_i). We have $X^{\otimes n}A(\alpha,\beta,\gamma)X^{\otimes n} = -A(-\alpha,\beta,-\gamma)$. Thus, $\|2A(0,\beta,0)\| = \|A(\alpha,\beta,\gamma) + A(-\alpha,\beta,-\gamma)\| \le \|A(\alpha,\beta,\gamma)\| + \|A(-\alpha,\beta,-\gamma)\| = 2\|A(\alpha,\beta,\gamma)\|$, where we use the triangle inequality and the unitary invariance of operator norm. It implies that

$$\frac{(\Delta G_{\text{eff}})_{\text{opt}}}{2} = \min_{\beta\in\mathbb{R}^n} \|A(0,\beta,0)\|. \tag{24}$$

Furthermore, let $\mathcal{P}$ represents the set of all permutations (of qubits) and P its element, we have $\|A(0,\text{avg}(\beta),0)\| \le \frac{1}{|\mathcal{P}|}\sum_{P\in\mathcal{P}}\|A(0,P\beta,0)\| = \|A(0,\beta,0)\|$, where $\text{avg}(\beta) = (\bar{\beta},\ldots,\bar{\beta})$, $\bar{\beta} = \frac{1}{n}\sum_i \beta_i$, $P\beta =$

$(\beta_{P(1)}, \beta_{P(2)}, \ldots, \beta_{P(n)})$ and we use $\|A(0, P\beta, 0)\| = \|A(0, \beta, 0)\|$ for any permutation P. It implies

$$\frac{(\Delta G_{\text{eff}})_{\text{opt}}}{2} = \min_{\bar{\beta}\in\mathbb{R}} \|A(0, \text{avg}(\beta), 0)\| = \min_{\bar{\beta}\in\mathbb{R}} \left\| \sum_{i<j<k} Z_i Z_j Z_k - \bar{\beta} \sum_i Z_i \right\|. \tag{25}$$

For any computational state $|\phi(k)\rangle$ that contains k $|1\rangle$s and $(n-k)$ $|0\rangle$s, we have $A(0, \text{avg}(\beta), 0)\,|\phi(k)\rangle = \mu_k\,|\phi(k)\rangle$, where $\mu_k = \big(-\binom{k}{3} + \binom{k}{2}\binom{n-k}{1} - \binom{k}{1}\binom{n-k}{2} + \binom{n-k}{3} - \bar{\beta}(n-2k)\big)$. So,

$$\mu_{k+1} - \mu_k = -4k^2 + 4(n-1)k + (2\bar{\beta} - (n-2)(n-1)). \tag{26}$$

When $2\bar{\beta} \leq -(n-1)$, the value of μ_k increases as k increases, and the maximum and minimum eigenvalues are taken at $k = 0$ and $k = n$. When $2\bar{\beta} > -(n-1)$, there are four extreme points of μ_k taken at $k = 0, n$ and locations near $\frac{(n-1)\pm\sqrt{n-1+2\bar{\beta}}}{2}$. Taking both situations and all extreme points into consideration, it is clear that the optimal $\bar{\beta}$ satisfies $2\bar{\beta} > -(n-1)$ and is taken at a point where the absolute values of the four extreme points are the same. As it is difficult to calculate the exact value of the optimal coefficient, we first conveniently assume $n-1$ is divisible by 4, and take $\bar{\beta} = (n-1)^2/8 - (n-1)/2$. In this case, the extreme points are taken at $k = 0, \frac{n-1}{4}, \frac{3(n-1)}{4}, n$, and the corresponding eigenvalues of $A(0, \text{avg}(\beta), 0)$ are $\mu_k = \frac{(n+7)n(n-1)}{24}, -\frac{(n+1)(n-1)(n-3)}{24}, \frac{(n+1)(n-1)(n-3)}{24}, -\frac{(n+7)n(n-1)}{24}$. It implies

$$\frac{(n+1)(n-1)(n-3)}{24} \leq \frac{(\Delta G_{\text{eff}})_{\text{opt}}}{2} = \min_{\bar{\beta}\in\mathbb{R}} \left\| \sum_{i<j<k} Z_i Z_j Z_k - \bar{\beta} \sum_i Z_i \right\| \leq \frac{(n+7)n(n-1)}{24}. \tag{27}$$

For general n, we have

$$(\Delta G_{\text{eff}})_{\text{opt}} = \frac{n^3}{12} + O(n^2). \tag{28}$$

B Proofs of Properties of Shortened RM Codes

The classical RM codes have numerous interesting properties. In Sec. 3.3, we outlined three key properties of the shortened RM code that are relevant to our discussion of QRM codes. Property (1), i.e., $\overline{RM}(1,m) \subsetneq \overline{RM}(m-2,m)$, is a simple observation since the rows of $\overline{\mathsf{G}}(1,m)$ are the same as the first m rows of $\overline{\mathsf{G}}(m-2,m)$. Here, we prove Property (2) and (3).

Property 2. *The vector $\mathbf{1}_{2^m-1}$ and those corresponding to the rows of $\overline{\mathsf{G}}(m-2,m)$ form a basis for $\overline{RM}(1,m)^\perp$.*

Proof. We know that $RM(1,m)^\perp = RM(m-2,m)$ [70]. For the shortened code, we first delete the all-1 row and then delete the all-0 column. Hence, it immediately follows that $\overline{RM}(m-2,m) \subseteq \overline{RM}(1,m)^\perp$. Recall that $\dim \overline{RM}(1,m) = m = \binom{m}{m-1}$ and $\dim \overline{RM}(m-2,m) = \binom{m}{1} + \binom{m}{2} + \cdots + \binom{m}{m-2}$. Then, $\dim \overline{RM}(1,m) + \dim \overline{RM}(m-2,m) = 2^m - 2$, which is 1 less than the dimension of the column space. Hence, $\overline{RM}(1,m)^\perp$ has one extra basis vector that is not in the row space of $\overline{RM}(m-2,m)$. We claim that $\mathbf{1}_{2^m-1}$ is the additional basis vector for $\overline{RM}(1,m)^\perp$. Since the rows of $\overline{\mathsf{G}}(1,m)$ are of even weight, it is clear that $\mathbf{1}_{2^m-1}$ is orthogonal to each row vector, and so $\mathbf{1}_{2^m-1} \in \overline{RM}(1,m)^\perp$. It remains to show that $\mathbf{1}_{2^m-1} \notin \overline{RM}(m-2,m)$.

For a contradiction, assume that $\mathbf{1}_{2^m-1} \in \overline{RM}(m-2,m)$. Then there are rows $\boldsymbol{v}_{i_1}^T, \boldsymbol{v}_{i_2}^T, \ldots \boldsymbol{v}_{i_p}^T$ of $\overline{\mathsf{G}}(m-2,m)$ such that $\boldsymbol{v}_{i_1}^T + \boldsymbol{v}_{i_2}^T + \cdots + \boldsymbol{v}_{i_p}^T = \mathbf{1}_{2^m-1}^T$. It follows that $(0|\boldsymbol{v}_{i_1}^T)^T + (0|\boldsymbol{v}_{i_2}^T)^T + \cdots + (0|\boldsymbol{v}_{i_p}^T)^T = (0|\mathbf{1}_{\mathbf{2^m-1}}^T)^T \in RM(m-2,m)$. Adding $\mathbf{1}_{2^m} \in RM(m-2,m)$ to this equation, we obtain $\mathbf{1}_{2^m} + (0|\boldsymbol{v}_{i_1}^T)^T + (0|\boldsymbol{v}_{i_2}^T)^T + \cdots + (0|\boldsymbol{v}_{i_p}^T)^T = (1|\mathbf{0}_{2^m-1}^T)^T$, which is a weight-1 vector. However, $RM(m-2,m)$ has a minimum weight of 4 (The minimum weight of $RM(r,m)$ is 2^{m-r} [70]). So, we have a contradiction; hence, property (2) is proven. □

Property 3. *The number of codewords in $\overline{RM}(1,m)^\perp$ with weight (2^m-4) is $\frac{1}{6}(4^m - 3\cdot 2^m + 2)$.*

Proof. First, we consider the weight distribution, i.e., the number of codewords of each weight, of $RM(1,m)$. From that, we will obtain the weight distribution of $\overline{RM}(1,m)$, which, in turn, will give us the weight distribution of $\overline{RM}(1,m)^\perp$.

Let A_k denote the number of codewords of weight k in $RM(1,m)$. It can be shown that $A_0 = 1$, $A_{2^m} = 1$, $A_{2^{m-1}} = 2^{m+1} - 2$, and $A_k = 0$ for all $k \neq 0, 1, 2^{m-1}$ [70]. Note that adding the vector $\mathbf{1}_{2^m}^T$ to any weight-(2^{m-1}) vector of the form $(0|\boldsymbol{v})^T \in RM(1,m)$ results in another weight-(2^{m-1}) vector of the form $(1|\boldsymbol{v}')^T$, where $\boldsymbol{v}$ and $\boldsymbol{v}'$ are (2^m-1)-tuple row vectors. So, there is a one-to-one correspondence (related by the addition of $\mathbf{1}_{2^m}$) between the weight-2^{m-1} vectors with the first entry 0 and the weight-2^{m-1} vectors with the first entry

1. The shortening process gets rid of the all-1 row vector from the generator matrix. Therefore, the number of weight-2^{m-1} vectors in $\overline{RM}(1,m)$ is half of that in $RM(1,m)$. Denoting $\overline{A}_k$ as the number of weight-k codewords in $\overline{RM}(1,m)$, we get $\overline{A}_0 = 1$, $\overline{A}_{2^{m-1}} = 2^m - 1$ and $\overline{A}_k = 0$ for all $k \neq 0, 2^{m-1}$.

The *enumerator polynomial* $W(C; x, y)$ succinctly represents the weight distribution of a classical code C, with the coefficient of $x^l y^{n-l}$ representing the number of codewords of weight l. Based on the values of $\overline{A}_k$, the enumerator polynomial of $\overline{RM}(1,m)$ is

$$W(\overline{RM}(1,m); x, y) = y^{2^m-1} + (2^m - 1)x^{2^{m-1}} y^{2^{m-1}-1}. \tag{29}$$

One can directly obtain the weight distribution of the dual code $C^\perp$ from the enumerator polynomial of C using the *Macwilliams identity*

$$W(C^\perp; x, y) = \frac{1}{|C|} W(C; y - x, x + y), \tag{30}$$

where $|C|$ is the number of codewords in C [70]. Since $|\overline{RM}(1,m)| = 2^m$, applying the Macwilliams identity to $W(\overline{RM}(1,m); x, y)$ yields

$$W(\overline{RM}(1,m)^\perp; x, y) = \frac{1}{2^m}(x+y)^{2^m-1} + \frac{(2^m-1)}{2^m}(y-x)^{2^{m-1}}(x+y)^{2^{m-1}-1}. \tag{31}$$

We want to find the coefficient of $x^{2^m-4}y^3$ on the right side of Eq. (31). Through a simple combinatorial calculation, the coefficient is found to be $\frac{1}{2^m}\binom{2^m-1}{3} + \frac{2^m-1}{2^m}\big[-\binom{2^{m-1}}{3}\binom{2^{m-1}-1}{0} + \binom{2^{m-1}}{2}\binom{2^{m-1}-1}{1} - \binom{2^{m-1}}{1}\binom{2^{m-1}-1}{2} + \binom{2^{m-1}}{0}\binom{2^{m-1}-1}{3}\big]$, which simplifies to $\frac{1}{6}(4^m - 3\cdot 2^m + 2)$. So, the number of weight-(2^m-4) codewords in $\overline{RM}(1,m)^\perp$ is $\frac{1}{6}(4^m - 3\cdot 2^m + 2)$. □

C Graph Theory Lemma

In Sec. 4, we used several graph-theoretic tools to establish the no-go results. Specifically, in Sec. 4.3, we claimed that any graph with a quadratic number of edges has a connected component with a linear number of vertices and a quadratic number of edges. We now prove this claim in the following lemma.

Lemma S1. *Consider a family of graphs $\{\mathfrak{G}^{(n)}\}_n$ where n belongs to an infinitely large set of positive integers. Assume $\mathfrak{G}^{(n)}$ has $\Theta(n)$ vertices and $\Theta(n^2)$ edges. Then there is a family of connected components of $\mathfrak{G}^{(n)}$ with $\Theta(n)$ vertices and $\Theta(n^2)$ edges.*

Proof. For a sufficiently large n, let $\mathfrak{G} = \mathfrak{G}^{(n)}$ be a graph with $\geq an^2$ edges and $\geq bn$ vertices for some $a, b > 0$. Let the connected components of $\mathfrak{G}$ be $\mathfrak{G}_1$, $\mathfrak{G}_2$, $\ldots$, $\mathfrak{G}_k$, and for all i, the number of vertices and edges in $\mathfrak{G}_i$ be $\bar{v}_i$ and $\bar{e}_i$, respectively. We index the connected components in descending order of the number of vertices. Let $\bar{e} = \max\{\bar{e}_i\}_{i=1}^k$. Suppose that the first l connected components have at least $\bar{v} = \lceil\sqrt{2\bar{e}} + 1\rceil$ vertices, i.e., enough vertices to make a complete graph with $\bar{e}$ edges.

Consider another graph $\mathfrak{F}$ with the same number of vertices as $\mathfrak{G}$ constructed in the following steps. Note that we will first partition vertices into different sets in steps (1) and (2) and then add edges within different sets in step (3).

(1) Let $V_1, V_2, \ldots, V_l$ be subsets of vertices from $\mathfrak{G}_1, \ldots, \mathfrak{G}_l$, where for each $\mathfrak{G}_i$ $(1 \leq i \leq l)$, we only select $\bar{v}$ vertices to place in V_i and place the rest of all vertices of $\mathfrak{G}$ in the set V^R.

(2) Partition the set of vertices V^R into subsets V_{l+1}, $V_{l+2}, \ldots V_{k'}, V^{R'}$ such that each V_i $(1 \leq i \leq k')$ has $\bar{v}$ vertices and the remainder set $V^{R'}$ has $\bar{w} < \bar{v}$ vertices.

(3) Add edges between all vertices within each subsets of vertices $V_1, V_2, \ldots, V_{k'}, V^{R'}$ to make them complete subgraphs $\mathfrak{F}_1, \mathfrak{F}_2, \ldots, \mathfrak{F}_{k'}, \mathfrak{R}$, which compose graph $\mathfrak{F}$.

Connected components $\mathfrak{F}_1, \mathfrak{F}_2, \ldots, \mathfrak{F}_{k'}$ of $\mathfrak{F}$ have at least $\bar{e}$ edges each, and $\mathfrak{R}$ has $\bar{w}(\bar{w}-1)/2$ edges. So, the number of edges in $\mathfrak{F}$ is at least $\left(\frac{bn-\bar{w}}{\bar{v}}\right)\bar{e} + \frac{\bar{w}(\bar{w}-1)}{2}$. From the construction, $\mathfrak{F}$ has at least as many edges as $\mathfrak{G}$ does. Therefore,

$$\begin{aligned} &\left(\frac{bn-\bar{w}}{\sqrt{2\bar{e}}}\right)\bar{e} + \frac{\bar{w}(\bar{w}-1)}{2} \geq an^2, \\ \Rightarrow\quad &bn\sqrt{\bar{e}/2} + \sqrt{\bar{e}/2} + 1 \geq an^2, \end{aligned} \tag{32}$$

where we set $\bar{w} = \sqrt{2\bar{e}}+2$ in the second step which increases the left-hand side so that the inequality still holds. Solving for $\bar{e}$, we get $\bar{e} = \Omega(n^2)$. A graph with $\Omega(n^2)$ edges must have $\Omega(n)$ vertices. However, the number of vertices in any connected component of $\mathfrak{G}$ is $O(n)$, which implies that $\bar{e} = O(n^2)$. Therefore, $\bar{e} = \Theta(n^2)$ and the number of vertices in a component with $\bar{e}$ edges is $\Theta(n)$. Hence, $\mathfrak{G}$ has a connected component with $\Theta(n)$ vertices and $\Theta(n^2)$ edges. □