

Quantum error correction meets continuous symmetries: fundamental trade-offs and case studies

— Supplementary Information for “Approximate symmetries and quantum error correction” —

Zi-Wen Liu^{1,2,*} and Sisi Zhou^{1,3,4,†}

¹*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*

²*Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, China*

³*Institute for Quantum Information and Matter,*

California Institute of Technology, Pasadena, CA 91125, USA

⁴*Pritzker School of Molecular Engineering, The University of Chicago, Illinois 60637, USA*

(Dated: October 30, 2023)

CONTENTS

SUPPLEMENTARY NOTE 1: OVERVIEW OF SUPPLEMENTARY INFORMATION

SUPPLEMENTARY NOTE 2: APPROXIMATELY COVARIANT APPROXIMATE QEC CODES: QUANTITATIVE CHARACTERIZATIONS OF QEC AND SYMMETRY

SUPPLEMENTARY NOTE 3: TRADE-OFF BETWEEN QEC AND GLOBAL COVARIANCE: CHARGE FLUCTUATION APPROACH

SUPPLEMENTARY NOTE 4: TRADE-OFF BETWEEN QEC AND GLOBAL COVARIANCE: GATE IMPLEMENTATION ERROR APPROACH

SUPPLEMENTARY NOTE 5: LIMITATIONS ON TRANSVERSAL LOGICAL GATES

SUPPLEMENTARY NOTE 6: TRADE-OFF BETWEEN QEC AND LOCAL SYMMETRY MEASURES

SUPPLEMENTARY NOTE 7: CASE STUDIES OF EXPLICIT CODES

SUPPLEMENTARY DISCUSSION

SUPPLEMENTARY METHOD 1: SUMMARY AND COMPARISON OF THE DIFFERENT TRADE-OFF RESULTS

SUPPLEMENTARY METHOD 2: PURIFIED DISTANCE AND DIAMOND DISTANCE BETWEEN ROTATED DEPHASING CHANNELS AND THE IDENTITY

SUPPLEMENTARY METHOD 3: TWO LOWER BOUNDS ON $|\partial_\theta \xi_\theta|^2|_{\theta=0}$.

SUPPLEMENTARY METHOD 4: REFINEMENTS OF PROPOSITION 2, PROPOSITION 3 AND PROPOSITION 5

SUPPLEMENTARY METHOD 5: AN INEQUALITY BETWEEN $\mathfrak{J}(\mathcal{N}_S, H_S)$ AND $\mathfrak{F}(\mathcal{N}_S, H_S)$

SUPPLEMENTARY METHOD 6: UPPER BOUNDS ON $\mathfrak{J}(\mathcal{N}_S, H_S)$ FOR RANDOM LOCAL NOISE AND INDEPENDENT NOISE

SUPPLEMENTARY METHOD 7: COMPUTING THE QEC INACCURACY OF MODIFIED THERMODYNAMIC CODES

SUPPLEMENTARY METHOD 8: REFINEMENT OF THEOREM 23

REFERENCES

* zwliu0@mail.tsinghua.edu.cn

† sisi.zhou26@gmail.com

I. SUPPLEMENTARY NOTE 1: OVERVIEW OF SUPPLEMENTARY INFORMATION

In Supplementary Information, we systematically study the fundamental competition between quantum error correction (QEC) and continuous symmetries, two key notions in quantum information and physics, in a quantitative manner. Three meaningful measures of approximate symmetries in quantum channels and, in particular, QEC codes, based on covariance violation over the entire symmetry group, covariance violation at a local point (closely related to quantum Fisher information), and the violation of charge conservation, respectively, are introduced and studied. Each measure induces a corresponding characterization of approximately covariant codes. We explicate a host of different ideas and techniques that enable us to derive various forms of trade-off relations between the QEC accuracy and all symmetry measures. More specifically, we introduce two frameworks for understanding and analyzing the trade-offs, based on the notions of charge fluctuation and gate implementation error (which may be of interest in their own rights) respectively, and employ methods including the Knill–Laflamme conditions as well as quantum metrology and quantum resource theory for the derivation. From the perspective of fault-tolerant quantum computing, our bounds on symmetry violation indicate limitations on the precision or density of transversally implementable logical gates for general QEC codes, refining the Eastin–Knill theorem. To exemplify nontrivial approximately covariant codes and understand the achievability of the above fundamental limits, we analyze two explicit types of codes: 1) a parametrized extension of the thermodynamic code, which gives a construction of a code family that continuously interpolates between exact QEC and exact symmetry, and 2) the quantum Reed–Muller codes, which represents a prominent example of approximately covariant exact QEC code. We show that both codes can saturate the scaling of the bounds for group-global covariance and charge conservation asymptotically, indicating the near-optimality of both our bounds and codes. **Note that this Supplementary Information is self-contained can be read independently as a companion paper with extended results and technical details, while the main text focuses on the most representative results and the physical motivation behind this study.**

Quantum error correction (QEC) is one of the most important and widely studied ideas in quantum information processing [1–4]. The spirit of QEC is to protect quantum information against noise and errors by suitably encoding logical quantum systems into quantum codes living in a larger physical Hilbert space. Since quantum systems are highly susceptible to noise effects such as decoherence so that errors easily occur, it is clear that QEC is of vital importance to the practical realization of quantum computing and other quantum technologies. Interestingly, besides the enduring efforts on the study of QEC and quantum codes for quantum information processing purposes, in recent years, they are also found to play fundamental roles in many important physical scenarios in e.g., holographic quantum gravity [5, 6] and many-body physics [7–10], and have consequently drawn great interest in physics.

When considering the practical implementation of QEC as well as its connections to physical problems, it is important to take symmetries and conservation laws into account as they are ubiquitous in physical systems. More explicitly, symmetries may constrain the encoders in the way that they must be covariant with respect to the symmetry group, i.e., commute with certain representations of group actions, generating the so-called *covariant codes* [11–14]. A principle of fundamental significance in both quantum information and physics is that (finite-dimensional) covariant codes for continuous symmetries (mathematically modeled by Lie groups)¹ are in a sense fundamentally incompatible with exact QEC [11, 15]. A well known no-go theorem that unfolds this principle from a quantum computation perspective is the Eastin–Knill theorem [15], which indicates that any QEC code covariant with respect to any continuous symmetry group in the sense that the logical group actions are mapped to *transversal* physical actions (that are tensor products on physical subsystems) cannot correct local errors perfectly. An intuitive explanation of this phenomenon is that, due to the conservation laws and transversality, physical subsystems necessarily contain logical charge information that gets leaked into the environment upon errors, so that the perfect recovery of logical information is prohibited. Crucially, transversal actions are highly desirable for the “fault tolerance” [2–4, 16] of practical quantum computation schemes because they do not spread errors across physical subsystems within each code block. It is also worth noting that the transversality property is widely important in physics as a fundamental feature of internal symmetries in many-body scenarios. More specifically, they are normally generated by sums of disjoint local charge observables, or in particular, on-site (transversal with respect to sites). Note that whether the symmetries are on-site is linked to whether they can be gauged or are anomaly-free, which plays important roles in the physics of quantum many-body systems and field theories [17]. In AdS/CFT, transversality also plays fundamental roles [18–20].

Due to the Eastin–Knill theorem, unfortunately, it is impossible to find an exact QEC code that implements a universal set of gates transversally, or namely achieves the full power of quantum computation while maintaining transversality. However, it may still be feasible to perform QEC approximately under these constraints, and a

¹ In what follows, we assume that the associated symmetry group is continuous and that the relevant Hilbert spaces are finite-dimensional when using the term “covariant codes”.

natural task is then to characterize the optimal degree of accuracy. Recently, several such bounds on the QEC accuracy achievable by covariant codes (which give rise to “robust” or “approximate” versions of the Eastin–Knill theorem) as well as explicit constructions of near-optimal covariant codes are found using many different techniques and insights from various areas in quantum information [12–14, 21–27], showcasing the fundamental nature of the problem. Remarkably, covariant codes have also found interesting applications to several important areas in physics already, including quantum many-body physics [9, 10, 14], AdS/CFT correspondence [12, 13, 18, 28, 29], and quantum information [11, 13, 27].

These existing studies on covariant codes are mostly concerned with the precision of QEC under exact symmetry conditions. Indeed, when symmetry principles arise, they are exactly respected by default. However, especially for continuous symmetries, it is often important or even necessary to consider approximate forms of symmetries or conservation laws in physical and practical scenarios. First of all, realistic quantum many-body systems are often dirty or defective so that the exact symmetry conditions and conservation laws could generally be violated to a certain extent. Furthermore, there are many important situations in physics where non-exact symmetries need to be considered for fundamental reasons. There are various symmetry breaking mechanisms that play key roles in wide-ranging physical scenarios including spontaneous symmetry breaking, anomalies, and non-renormalizable effects [30]. In particle physics, many important symmetries are known to be only approximate [31]. More notably, it has long been believed that global symmetries cannot be exact in a unified theory of quantum mechanics and gravity [31–36] (justified in more concrete terms in AdS/CFT [18, 28]). Considering the need for large quantum systems to boost the advantages of quantum technologies and also the broad connections between QEC and physics, it would be important and fruitful to have a quantitative theory of QEC codes with approximate symmetries, or *approximately covariant* codes. For example, given that the QEC accuracy of exactly covariant codes is limited, one may wonder whether for codes that achieve exact QEC there are “dual” bounds on the degree of symmetry or covariance. It is particularly worth noting that the no-global-symmetry arguments in AdS/CFT indeed have deep connections to covariant codes, and in particular this question [12, 18]. However, our understanding of approximate symmetries, especially characterizations and applications on a quantitative level, is very limited to date.

Our work aims to establish a quantitative theory of the interplay between the degree of continuous symmetries and QEC accuracy, which in particular allows us to understand symmetry violation in exact QEC codes. (Note that our discussion here mainly proceeds in terms of the most fundamental $U(1)$ symmetry which is sufficient to reveal the key phenomena.) To this end, we first formally define three different meaningful measures of symmetry violation, in terms of the violation of covariance conditions globally over the entire symmetry group or locally at a specific point in the group, and the violation of charge conservation respectively, which induce corresponding quantitative notions of approximately covariant codes. Our main results are a series of trade-off bounds between QEC accuracy and the above different symmetry measures under a general condition called *Hamiltonian-in-Kraus-span (HKS) condition* which subsumes transversality in our setup, each of which may suit certain scenarios the best. (For readers’ convenience, we provide in Sec. IX a table that identifies the key theorems and summarizes their respective strengths and weaknesses.) We introduce two concepts—charge fluctuation and gate implementation error—each providing a framework for analyzing the QEC-symmetry trade-off and could be useful in their own rights. Furthermore, our derivations feature ideas and techniques from several different fields. More explicitly, various different forms of the trade-off relations are derived by analyzing the “perturbation” of the Knill–Laflamme conditions [37, 38], as well as by leveraging insights and techniques from the fields of quantum metrology [39, 40] and quantum resource theory [24, 41–43]. Our theory provides a complete understanding of the transition between exact QEC and exact symmetry. On the exact symmetry end, the previous limits on covariant codes (often referred to as “approximate Eastin–Knill theorems” [12, 13, 21, 22]) are recovered, while the exact QEC end provides new lower bounds on various forms of symmetry violation for the commonly studied exact codes. In particular, we use our symmetry bounds to derive fundamental limitations on the set of transversally implementable logical gates for general QEC codes, which represent a new type of improvement of the Eastin–Knill theorem and apply more broadly than previous results along a similar line about stabilizer codes in Refs. [44–48], advancing our understanding of fault tolerance. Then, to solidify our general theory, we present case studies on two explicit code constructions, which can be seen as examples of approximately covariant codes that exhibit certain key features, as well as upper bounds (achievability results) that help understand how strong our fundamental limits are. First, we construct a parametrized code family that interpolates between the two ends of exact QEC and exact symmetry and exhibits a full trade-off between QEC and symmetry, by modifying the so-called thermodynamic code [9, 12]. In the second case study, we analyze the quantum Reed–Muller codes which exhibit nice structures and features and, in particular, have been widely applied for the transversal implementation of certain non-Clifford gates and magic state distillation [49–52]. We find that in both cases the codes can almost saturate the bounds on global covariance and charge conservation (up to constant factors) asymptotically, that is, both the code constructions and bounds are nearly optimal.

Here we present the study in a rigorous and comprehensive manner. In particular, this Supplementary Information contains all technical details of the derivation, thorough discussions of all different approaches, and many additional

results.

This Supplementary Information is organized as follows. First, in [Sec. II](#), we review the formalism of QEC and the incompatibility between QEC and continuous symmetries, and also formally define the accuracy of approximate QEC codes as well as the different quantitative characterizations of approximate continuous symmetries associated with QEC codes that will be considered. In [Sec. III](#) and [Sec. IV](#), we introduce the two frameworks based on the notions of charge fluctuation and gate implementation error respectively, under which we discuss a series of different approaches to deriving the trade-off relations between the QEC inaccuracy and the group-global covariance violation. Then in [Sec. V](#), we specifically discuss the application to fault-tolerant quantum computing, deriving general restrictions on the transversally implementable logical gates in QEC codes from the results above. Afterwards, in [Sec. VI](#), we present our results on the trade-off relations between QEC inaccuracy and group-local symmetry measures including the group-local covariance violation and the charge conservation violation. After the above discussion of fundamental limits, in [Sec. VII](#) we study the modified thermodynamic code and quantum Reed–Muller codes, which gives concrete examples of nearly optimal approximately covariant approximate codes in certain cases. Finally, in [Sec. VIII](#) we summarize our study, and discuss important open problems and future directions.

II. SUPPLEMENTARY NOTE 2: APPROXIMATELY COVARIANT APPROXIMATE QEC CODES: QUANTITATIVE CHARACTERIZATIONS OF QEC AND SYMMETRY

Here we formally define the quantitative measures of QEC and symmetry that will be used in our study. Specifically, we first overview the notions of QEC and covariant codes and discuss how to quantify the deviation of general quantum codes from them. In particular, we will define the QEC inaccuracy which quantifies the QEC capability of a quantum code under specific noise. We will also define a measure of group-global symmetry which quantifies the approximate covariance of a code over the entire $U(1)$ group and two measures of group-local symmetry which are linked to the covariance of a code at an exact point in $U(1)$. The trade-off relations between QEC and these symmetry measures will be thoroughly studied in later sections.

Note that in this Supplementary Information, we extensively use distance metrics defined based on the purified distance [53, 54], which are well-behaved and commonly used in the quantum information literature. Notably, the choice of purified distance directly relates the local covariance violation to the well known quantum Fisher information (QFI) [55–58]. In principle, one may also consider other metrics. We shall also discuss the situations where one uses the diamond distance [59], another standard channel distance measure.

A. Approximate quantum error correction

QEC functions by encoding the logical quantum system in some quantum code living in a larger physical system with redundancy, so that a limited number of errors can be corrected to recover the original logical information. A quantum code is defined by an encoding quantum channel $\mathcal{E}_{S \leftarrow L}$ from a logical system L to a physical system S , and it perfectly protects the logical information against a physical noise \mathcal{N}_S if and only if there exists a recovery channel $\mathcal{R}_{S \leftarrow L}$ such that

$$\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L} = \mathbb{1}_L. \quad (1)$$

In particular, when $\mathcal{E}_{S \leftarrow L}$ is isometric, $\mathcal{N}_S(\cdot) = \sum_{i=1}^r K_{S,i}(\cdot) K_{S,i}^\dagger$ and Π is the projection onto the code subspace in the physical system, such a recovery channel exists if and only if the Knill–Laflamme (KL) conditions, $\forall i, j, \Pi K_{S,i}^\dagger K_{S,j} \Pi \propto \Pi$ [37], hold.

In many scenarios, a quantum code is still useful in protecting quantum information when it only achieves approximate QEC, namely, $\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}$ is close to but not exactly equal to $\mathbb{1}_L$. To characterize the inaccuracy of an approximate QEC code, we will use the channel fidelity and the Choi channel fidelity, defined by

$$f(\Phi_1, \Phi_2) = \min_{\rho} f((\Phi_1 \otimes \mathbb{1})(\rho), (\Phi_2 \otimes \mathbb{1})(\rho)), \quad (2)$$

$$\bar{f}(\Phi_1, \Phi_2) = f((\Phi_1 \otimes \mathbb{1})(\Psi), (\Phi_2 \otimes \mathbb{1})(\Psi)), \quad (3)$$

respectively, where $f(\rho, \sigma) = \text{Tr}(\sqrt{\rho^{1/2} \sigma \rho^{1/2}})$ is the fidelity of quantum states, $|\Psi\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i\rangle |i\rangle$ is the maximally entangled state and $\Psi = |\Psi\rangle \langle \Psi|$. Here the inputs ρ and Ψ lie in a bipartite system consisting of the original system $\Phi_{1,2}$ acting on and a reference system as large as the original. Correspondingly, one can define the purified distance of states $P(\rho, \sigma) = \sqrt{1 - f(\rho, \sigma)^2}$, the purified distance of channels $P(\Phi_1, \Phi_2) = \sqrt{1 - f(\Phi_1, \Phi_2)^2}$ and the Choi purified

distance of channels $\bar{P}(\Phi_1, \Phi_2) = \sqrt{1 - \bar{f}(\Phi_1, \Phi_2)^2}$ [53, 54, 60]. The (*worst-case*) *QEC inaccuracy* and the *Choi QEC inaccuracy* for approximate QEC codes are then defined as

$$\varepsilon(\mathcal{N}_S, \mathcal{E}_{S \leftarrow L}) := \min_{\mathcal{R}_{L \leftarrow S}} P(\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L), \quad (4)$$

$$\bar{\varepsilon}(\mathcal{N}_S, \mathcal{E}_{S \leftarrow L}) := \min_{\mathcal{R}_{L \leftarrow S}} \bar{P}(\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L), \quad (5)$$

respectively. The Choi inaccuracy reflects the average-case behavior in the sense that $\bar{\varepsilon} = \sqrt{\frac{d_L+1}{d_L}} \varepsilon_{\text{avg}}$ [61, 62], where $\varepsilon_{\text{avg}} := \sqrt{1 - \max_{\mathcal{R}_{L \leftarrow S}} \int d\psi f^2(\psi, \mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}(\psi))}$ in which the integral is over the Haar random pure logical states. For simplicity, we will not explicitly write down the arguments of ε or $\bar{\varepsilon}$ (and of many other measures defined later) when they are unambiguous.

In the above, we used the channel purified distances as channel distance measures. As mentioned, we may also consider the the diamond distance $D_\diamond(\Phi_1, \Phi_2)$ induced by the diamond norm of channels [59, 63]:

$$D_\diamond(\Phi_1, \Phi_2) := \max_{\rho} \frac{1}{2} \|(\Phi_1 \otimes \mathbb{1})(\rho) - (\Phi_2 \otimes \mathbb{1})(\rho)\|_1 \quad (6)$$

$$= \max_{|\psi\rangle} \frac{1}{2} \|(\Phi_1 \otimes \mathbb{1})(|\psi\rangle\langle\psi|) - (\Phi_2 \otimes \mathbb{1})(|\psi\rangle\langle\psi|)\|_1. \quad (7)$$

where $\|\cdot\|_1$ is the nuclear (trace) norm. Naturally, the diamond distance version of QEC inaccuracy is defined as

$$\varepsilon_\diamond := \min_{\mathcal{R}_{L \leftarrow S}} D_\diamond(\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L). \quad (8)$$

It is easy to see that lower bounds on ε (that we derive below) directly indicate lower bounds on ε_\diamond . According to the Fuchs–van de Graaf inequality $1 - f(\rho, \sigma) \leq \frac{1}{2} \|\rho - \sigma\|_1$ [64], we have

$$D_\diamond(\Phi_1, \Phi_2) \geq 1 - f(\Phi_1, \Phi_2) \geq \frac{1}{2} P(\Phi_1, \Phi_2)^2. \quad (9)$$

In the case of our interest where the second channel is the identity, the above inequality can be further improved using $1 - f(\rho, |\psi\rangle\langle\psi|)^2 \leq \frac{1}{2} \|\rho - |\psi\rangle\langle\psi|\|_1$:

$$D_\diamond(\Phi_1, \mathbb{1}) \geq 1 - f(\Phi_1, \mathbb{1})^2 = P(\Phi_1, \mathbb{1})^2. \quad (10)$$

Therefore, $\varepsilon_\diamond \geq \varepsilon^2$.

B. Measuring approximate symmetries of QEC codes

Symmetries of quantum codes manifest themselves in the covariance of the encoder with respect to symmetry transformations. For the case of current interest, the symmetry transformations on the logical and physical systems are, respectively, $U_{L,\theta} = e^{-iH_L\theta}$ generated by a logical Hamiltonian (charge observable) H_L , and $U_{S,\theta} = e^{-iH_S\theta}$ generated by a physical Hamiltonian (charge observable) H_S ², both representations of the $U(1)$ Lie group periodic with a common period τ . The transversality property of symmetry transformations (gate actions) corresponds to the 1-local form of H_S , namely, $H_S = \sum_{l=1}^n H_{S_l}$ where each term H_{S_l} acts locally on physical subsystem S_l . We say a quantum code is covariant (with respect to such $U(1)$ representations given by H_L and H_S), if

$$\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L,\theta} = \mathcal{E}_{S \leftarrow L,\theta} \circ \mathcal{U}_{L,\theta}, \quad \forall \theta \in \mathbb{R}. \quad (11)$$

The definitions of covariant codes can be easily extended to general compact Lie groups [12, 13]. We also assume H_L and H_S to be both non-trivial, i.e., not a constant operator. Note that applying constant shifts on H_L and H_S do not change the definition of Eq. (11) and we will often use this property below.

² Here H_L and H_S are generators of $U(1)$ representations, or “charge observables”, and should not be confused with the intrinsic system Hamiltonians governing the system dynamics.

As mentioned, the covariance of quantum codes is often incompatible with their error-correcting properties and approximate notions of covariance may play important roles in wide-ranging scenarios. For example, here the Eastin–Knill theorem indicates that codes that can perfectly correct local noise cannot simultaneously be covariant with respect to non-trivial 1-local H_S [15]. More generally, exact QEC is known to be incompatible with exact covariance as long as

$$H_S \in \text{span}\{K_{S,i}^\dagger K_{S,j}, \forall i, j\}, \quad (12)$$

which we refer to as the *Hamiltonian-in-Kraus-span (HKS) condition*, holds [22, 40]. The HKS condition holds for many typical scenarios, including the one mentioned above where \mathcal{N}_S represents single-erasure noise (where one subsystem chosen uniformly at random is erased) and H_S is 1-local. When the HKS condition does not hold, examples of exactly covariant QEC codes exist, e.g., when $\mathcal{N}_S = \mathbb{1}$ (noiseless dynamics), when H_S is a Pauli-X operator and \mathcal{N}_S is dephasing noise [65, 66], and when \mathcal{N}_S is single-erasure noise but H_S is 2-local [67]. We shall assume that the HKS condition holds for the quantum codes considered in our work. We also emphasize that there exist examples of exact QEC codes covariant with respect to discrete symmetry groups [11], so the assumption of continuous groups is important.

Besides quantum computation, approximately symmetries and covariant codes are potentially useful in quantum gravity and condensed matter physics, as discussed in the main text. To formally characterize and study approximate covariance, an important first step is to find reasonable ways to quantify it. We now do so.

1. Group-global covariance violation

The first, most important type of measure is based on the global covariance violation over the entire symmetry group. Codes that are approximately covariant with respect to H_L and H_S in such a global sense should have small covariance violation for all θ . We define the *group-global³ covariance violation* and the *Choi group-global covariance violation* by

$$\delta_{\text{group}} := \max_{\theta} P(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}), \quad (13)$$

$$\bar{\delta}_{\text{group}} := \max_{\theta} \bar{P}(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}), \quad (14)$$

respectively. Intuitively, they measure the maximum deviation of the encoding channel $\mathcal{E}_{S \leftarrow L}$ from the exact covariance condition Eq. (11) in the entire symmetry group. It is known that δ_{group} and ε cannot be simultaneously zero in non-trivial situations, and previous works [12–14, 21–23, 25] mostly focus on deriving lower bounds on ε for exactly covariant codes ($\delta_{\text{group}} = 0$). We will present bounds that involve δ_{group} which reveal the trade-off between QEC and global covariance, derived via two notions we introduce called the charge fluctuation and gate implementation error. This extends the scope of previous consideration to general codes including exact QEC codes.

Similar to the case of QEC inaccuracy, we can also consider the diamond distance and define

$$\delta_{\text{group},\diamond} := \max_{\theta} D_{\diamond}(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}). \quad (15)$$

Again, lower bounds on $\delta_{\text{group},\diamond}$ that we derive below directly indicate lower bounds on δ_{group} . Using Eq. (9), we directly see that $\delta_{\text{group},\diamond} \geq \delta_{\text{group}}^2/2$. In particular, when $\mathcal{E}_{S \leftarrow L}$ is isometric, we have $\delta_{\text{group},\diamond} = \delta_{\text{group}}$, using the fact that $\frac{1}{2} \|\lvert\psi_1\rangle\langle\psi_1\rvert - \lvert\psi_2\rangle\langle\psi_2\rvert\|_1 = P(\lvert\psi_1\rangle\langle\psi_1\rvert, \lvert\psi_2\rangle\langle\psi_2\rvert)$.

2. Group-local (point) covariance violation

One may wonder if the incompatibility between QEC and continuous symmetries can be relieved when we relax the requirement from exact global covariance to exact local covariance, i.e., when we require only the code covariance for θ inside a small neighborhood of a point θ_0 , satisfying $\mathcal{U}_{S,\theta_0} \circ \mathcal{E}_{S \leftarrow L} = \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta_0}$. Unfortunately, the no-go results also extend to the local case, meaning that a non-trivial QEC code cannot be exactly covariant even in an arbitrarily small neighborhood of θ_0 . Without loss of generality, we assume $\theta_0 = 0$ because we can always redefine $\mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta_0}$

³ We shall refer to “group-global” and “group-local” as “global” and “local”, respectively, for simplicity, as is common in e.g. estimation theory after their definitions. They should not be confused with the geometric notions commonly used in physical contexts.

to be the new encoding channel such that the code is covariant at $\theta = 0$. To characterize the local covariance, we introduce the *group-local (point) covariance violation* defined by

$$\delta_{\text{local}} := \sqrt{2\partial_{\theta}^2 P(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^{\dagger}, \mathcal{E}_{S \leftarrow L})^2|_{\theta=0}} = \sqrt{F(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^{\dagger})|_{\theta=0}}. \quad (16)$$

Here $F(\Phi_{\theta})$ is the quantum Fisher information (QFI) defined using the second order derivative of the purified distance $F(\Phi_{\theta}) = 2\partial_{\theta'}^2 P(\Phi_{\theta}, \Phi_{\theta'})^2|_{\theta'=\theta}$ which characterizes the amount of information of θ one can extract from Φ_{θ} around point θ [68]. Correspondingly, the QFI of quantum states is defined by $F(\rho_{\theta}) = 2\partial_{\theta'}^2 P(\rho_{\theta}, \rho_{\theta'})^2|_{\theta'=\theta}$ [57, 69] which characterizes the amount of information of θ one can extract from ρ_{θ} around point θ and we have $F(\Phi_{\theta}) = \max_{\rho} F((\Phi_{\theta} \otimes \mathbb{1})(\rho))$. Note that the QFI defined here using the purified distance is usually called the SLD QFI and there are other types of QFIs, e.g., the RLD QFI [70–72] which we will encounter later in Sec. IV B 2. When $\delta_{\text{local}} = 0$, the code is locally covariant up to the lowest order of $d\theta$. We shall see later that for any $\delta_{\text{local}} < \Delta H_L$ (we will use $\Delta(\cdot)$ to denote the difference between the maximum and minimum eigenvalues of (\cdot)), there is a non-trivial lower bound on ε , leading to a trade-off relation between QEC and local covariance.

3. Charge conservation violation

The correspondence between symmetries and conservation laws is a landmark result of modern physics. Inspired by this correspondence, we can define another intuitive measure of the symmetry violation by the degree of charge deviation. It can be shown that for an isometric encoding channel $\mathcal{E}_{S \leftarrow L}$ the covariance condition Eq. (11) is equivalent to

$$(\mathcal{E}_{S \leftarrow L})^{\dagger}(H_S) = H_L - \nu \mathbb{1}, \quad (17)$$

for some $\nu \in \mathbb{R}$, where \mathcal{E}^{\dagger} is the dual channel of \mathcal{E} satisfying $\text{Tr}(H\mathcal{E}(\rho)) = \text{Tr}(\mathcal{E}^{\dagger}(H)\rho)$ for any H and ρ . Since H_S and H_L represent the charge observables in the physical and logical systems, Eq. (17) implies that the eigenstates of H_L are mapped to the corresponding eigenstates of H_S after the encoding operation [12], indicating the charge conservation nature of the encoding map. The charge conservation law can also be understood through the relation $\text{Tr}(H_S \mathcal{E}_{S \leftarrow L}(\rho)) = \text{Tr}(H_L \rho) - \nu$ for any ρ , where ν represents a universal constant offset in the charge. To measure the degree to which the charge conservation law is violated, we consider the following quantity which we call the *charge conservation violation* (also defined in Ref. [12]):

$$\delta_{\text{charge}} := \Delta(H_L - (\mathcal{E}_{S \leftarrow L})^{\dagger}(H_S)). \quad (18)$$

Note again that $\Delta(\cdot)$ denotes the difference between the maximum and minimum eigenvalues of (\cdot) . It can be easily verified that $\delta_{\text{charge}}/2$ is equal to the difference between physical and logical charges, formally given by $\min_{\nu \in \mathbb{R}} \max_{\rho} |\text{Tr}(H_S \mathcal{E}_{S \leftarrow L}(\rho)) - \text{Tr}((H_L - \nu \mathbb{1})\rho)|$ (a constant offset on the definitions of charges is allowed). For general CPTP encoding maps, δ_{charge} is not always zero for exactly covariant codes [73], unlike δ_{group} and δ_{local} . However, for isometric encoding we always have the following relation between δ_{local} and δ_{charge} :

Proposition 1. *When $\mathcal{E}_{S \leftarrow L}$ is isometric, $\delta_{\text{local}} \geq \delta_{\text{charge}}$.*

Proof. Suppose $\mathcal{E}_{S \leftarrow L}(\cdot) = W(\cdot)W^{\dagger}$ where W is isometric. Then

$$\delta_{\text{charge}} = \Delta(H_L - W^{\dagger}H_S W), \quad (19)$$

and

$$(\delta_{\text{local}})^2 = F(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^{\dagger})|_{\theta=0} = \max_{|\psi\rangle} F(U_{S,\theta} W U_{L,\theta}^{\dagger} |\psi\rangle)|_{\theta=0}. \quad (20)$$

Let

$$|\psi_{\theta}\rangle = U_{S,\theta} W U_{L,\theta}^{\dagger} |\psi\rangle = e^{-iH_S\theta} W e^{iH_L\theta} |\psi\rangle. \quad (21)$$

Then

$$|\partial_{\theta}\psi_{\theta}\rangle = e^{-iH_S\theta} W e^{iH_L\theta} iH_L |\psi\rangle - iH_S e^{-iH_S\theta} W e^{iH_L\theta} |\psi\rangle. \quad (22)$$

The QFI for pure states is given by $F(|\psi_\theta\rangle) = 4(\langle\partial_\theta\psi_\theta|\partial_\theta\psi_\theta\rangle - |\langle\partial_\theta\psi_\theta|\psi_\theta\rangle|^2)$ [69]. Since

$$\begin{aligned}\langle\partial_\theta\psi_\theta|\partial_\theta\psi_\theta\rangle|_{\theta=0} &= \langle\psi|(WH_L - H_S W)^\dagger(WH_L - H_S W)|\psi\rangle \\ &\geq \langle\psi|(WH_L - H_S W)^\dagger WW^\dagger(WH_L - H_S W)|\psi\rangle,\end{aligned}\tag{23}$$

$$\langle\psi_\theta|\partial_\theta\psi_\theta\rangle|_{\theta=0} = \langle\psi|(H_L - W^\dagger H_S W)|\psi\rangle,\tag{24}$$

we have

$$(\delta_{\text{local}})^2 = \max_{\psi} F(|\psi_\theta\rangle)|_{\theta=0}\tag{25}$$

$$\geq \max_{\psi} 4 \langle\psi|(H_L - W^\dagger H_S W)^\dagger(H_L - W^\dagger H_S W)|\psi\rangle - 4 |\langle\psi|(H_L - W^\dagger H_S W)|\psi\rangle|^2 = (\delta_{\text{charge}})^2,\tag{26}$$

proving the result. \square

We shall see later that, similar to the situation of local covariance violation, for any $\delta_{\text{charge}} < \Delta H_L$, there is a non-trivial lower bound on ε . We refer to both δ_{local} and δ_{charge} as local symmetry measures because their values only depend on the approximate covariance of a code in the neighborhood of the point $\theta = 0$. Note that both δ_{local} and δ_{charge} have the same unit as the charges while δ_{group} is dimensionless, i.e., after replacing H_S and H_L with cH_S and cH_L for some constant c , both δ_{local} and δ_{charge} are changed to $c\delta_{\text{local}}$ and $c\delta_{\text{charge}}$ while δ_{group} is unchanged.

4. Remarks on non-compact groups and infinite-dimensional codes

In the above discussion, we assumed compact Lie groups and finite-dimensional quantum codes. Here we remark on possible extensions to non-compact Lie groups and infinite-dimensional codes.

First, note that our definitions of δ_{local} and δ_{charge} can be naturally extended to the situations of non-compact groups where H_S and H_L are arbitrary finite-dimensional Hermitian operators but the group transformations are not periodic, because their definitions only depend on the local geometry of the symmetry group. For the global measure δ_{group} , we need to assume compact Lie groups, i.e., the physical and logical group transformations are both periodic with a common period.

Moreover, when the physical system S is infinite-dimensional, one may naturally consider some finite-dimensional truncation \tilde{H}_S of H_S . The trade-off relations we derive below hold for \tilde{H}_S and H_L , so when the truncation is suitably chosen we can still obtain nontrivial results that well indicate the behaviors of H_S . For example, when $\|(\mathcal{E}_{S \leftarrow L})^\dagger(H_S) - (\mathcal{E}_{S \leftarrow L})^\dagger(\tilde{H}_S)\|$ ($\|\cdot\|$ is the spectral norm) is small, \tilde{H}_S is a good substitute for H_S in terms of the charge conservation violation [12].

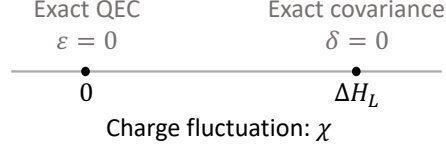
III. SUPPLEMENTARY NOTE 3: TRADE-OFF BETWEEN QEC AND GLOBAL COVARIANCE: CHARGE FLUCTUATION APPROACH

In this section, we derive trade-off relations between the QEC inaccuracy ε and the global covariance violation δ_{group} by connecting them to a quantity which we call the charge fluctuation χ (note that this notion is distinct from the charge conservation violation although they are in some way related as will be discussed). Our approach essentially proceed in two steps. First, we connect δ_{group} and χ by providing a lower bound on δ_{group} which depends on χ . Then we prove upper bounds on $|\chi|$ in terms of the QEC inaccuracy ε using two different methods. The first one is based on analyzing the deviation of the approximate QEC code from the the KL conditions, which we call the KL-based method, and the second one is based on treating the problem as a channel estimation problem and using quantum metrology techniques. These methods eventually lead to two types of trade-off bounds between QEC and global covariance. Note that we assume quantum codes are isometric throughout this section unless stated otherwise.

A. Bounding global covariance violation by charge fluctuation

Consider a code defined by encoding isometry $\mathcal{E}_{S \leftarrow L}$. We start by considering the situation where the code achieves exact QEC under the noise channel $\mathcal{N}_S(\cdot) = \sum_i K_{S,i}(\cdot)K_{S,i}^\dagger$. According to the KL conditions,

$$\Pi K_{S,i}^\dagger K_{S,j} \Pi \propto \Pi,\tag{27}$$



Supplementary Figure 1. For exact QEC codes ($\varepsilon = 0$) which satisfy the HKS condition, the charge fluctuation $\chi = 0$. For exactly covariant codes ($\delta_{\text{group}} = 0$), the charge fluctuation $\chi = \Delta H_L$. The trade-off can be derived by investigating the relations between the distances of χ from ΔH_L and 0, and the symmetry and QEC measures.

where Π is the projection onto the code subspace. In particular, let $|0_L\rangle$ and $|1_L\rangle$ be eigenstates corresponding to the largest and the smallest eigenvalues of H_L . (We do not specify the exact choices of $|0_L\rangle$ and $|1_L\rangle$ even when H_L is degenerate, as long as they correspond to the largest and smallest eigenvalues, respectively.) Using Eq. (27), we have

$$\langle 0_L | (\mathcal{E}_{S \leftarrow L})^\dagger (K_{S,i}^\dagger K_{S,j}) | 0_L \rangle = \langle 1_L | (\mathcal{E}_{S \leftarrow L})^\dagger (K_{S,i}^\dagger K_{S,j}) | 1_L \rangle. \quad (28)$$

Using the HKS condition $H_S \in \text{span}\{K_{S,i}^\dagger K_{S,j}, \forall i, j\}$, we must also have

$$\langle 0_L | (\mathcal{E}_{S \leftarrow L})^\dagger (H_S) | 0_L \rangle = \langle 1_L | (\mathcal{E}_{S \leftarrow L})^\dagger (H_S) | 1_L \rangle. \quad (29)$$

The incompatibility between QEC and symmetry could be understood through the incompatibility between Eq. (29) and Eq. (17). Eq. (29) implies that $\langle 0_L | H_L | 0_L \rangle - \langle 1_L | H_L | 1_L \rangle = 0$ when the code achieves exact QEC, which contradicts with $\langle 0_L | H_L | 0_L \rangle - \langle 1_L | H_L | 1_L \rangle = \Delta H_L$ for exactly covariant codes from Eq. (17), implying the non-existence of exact QEC codes with exact covariance.

For general codes, we define the *charge fluctuation*:

$$\chi := \langle 0_L | (\mathcal{E}_{S \leftarrow L})^\dagger (H_S) | 0_L \rangle - \langle 1_L | (\mathcal{E}_{S \leftarrow L})^\dagger (H_S) | 1_L \rangle, \quad (30)$$

Based on the discussion above, one can see that χ embodies the transition between exact QEC and exact symmetry quantitatively—when a code is close to being an exactly covariant code, χ cannot be too far away from ΔH_L , and when a code is close to being an exact QEC code, χ cannot be too far away from 0 (see an illustration in Fig. 1). Thus the trade-off relation between ε and δ_{group} can be derived by connecting χ to each of them separately.

We now derive the following lower bound on the global covariance violation δ_{group} in terms of the charge fluctuation χ , which directly connects δ_{group} with χ . Note that, in this Supplementary Information, “ \gtrsim ”, “ \lesssim ”, and “ \simeq ” mean “ \geq ”, “ \leq ”, and “ $=$ ”, respectively, up to the leading order.

Proposition 2. *Consider an isometric quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. Then when $|\Delta H_L - \chi| \leq \Delta H_S$, it holds that*

$$\delta_{\text{group}} \geq \min \left\{ \frac{\sqrt{|\Delta H_L - \chi| (\Delta H_S - \frac{1}{2} |\Delta H_L - \chi|)}}{\Delta H_S}, \sqrt{\frac{3}{8}} \right\}, \quad (31)$$

and when $|\Delta H_L - \chi| > \Delta H_S$, $\delta_{\text{group}} \geq \sqrt{3/8}$. In particular, when $|\Delta H_L - \chi| \ll \Delta H_S$, we have

$$\delta_{\text{group}} \gtrsim \sqrt{\frac{|\Delta H_L - \chi|}{\Delta H_S}}. \quad (32)$$

Proof. Since $U_{S,\theta}$ and $U_{L,\theta}$ are both periodic with a common period, we assume H_S and H_L both have integer eigenvalues. We also assume the smallest eigenvalue of H_S is zero because constant shifts do not affect the definitions of symmetry measures. When $\mathcal{E}_{S \leftarrow L}(\cdot) = W(\cdot)W^\dagger$ is isometric, let $|\mathbf{c}_0\rangle = W|0_L\rangle$, $|\mathbf{c}_1\rangle = W|1_L\rangle$, and write

$$|\mathbf{c}_0\rangle = \sum_{\eta=0}^{\Delta H_S} c_\eta^0 |\eta^0\rangle, \quad |\mathbf{c}_1\rangle = \sum_{\eta=0}^{\Delta H_S} c_\eta^1 |\eta^1\rangle, \quad (33)$$

where $\sum_{\eta} |c_{\eta}^0|^2 = \sum_{\eta} |c_{\eta}^1|^2 = 1$ and $|\eta^0\rangle$ and $|\eta^1\rangle$ are eigenstates of H_S with eigenvalue η . $|\eta^0\rangle$ and $|\eta^1\rangle$ may not be the same when H_S is degenerate. Note that when η is not an eigenvalue of H_S , we simply take $c_{\eta}^0 = 0$ (or $c_{\eta}^1 = 0$) so that c_{η}^0 (or c_{η}^1) is well-defined for any integer η . Let $|\psi\rangle = \frac{1}{\sqrt{2}}(|0_L\rangle|0_R\rangle + |1_L\rangle|1_R\rangle)$, where R is a reference system. Then the channel fidelity

$$\begin{aligned}
f_{\theta} &:= f(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^{\dagger}, \mathcal{E}_{S \leftarrow L}) \\
&= \min_{\rho} f((\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^{\dagger} \otimes \mathbb{1}_R)(\rho), (\mathcal{E}_{S \leftarrow L} \otimes \mathbb{1}_R)(\rho)) \\
&\leq f((\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^{\dagger} \otimes \mathbb{1}_R)(|\psi\rangle), (\mathcal{E}_{S \leftarrow L} \otimes \mathbb{1}_R)(|\psi\rangle)) \\
&= \left| \langle \psi | W^{\dagger} U_{S,\theta} W U_{L,\theta}^{\dagger} | \psi \rangle \right| \\
&= \left| \frac{1}{2} \sum_{\eta=0}^{\Delta H_S} |c_{\eta}^0|^2 e^{-i\eta\theta + i\Delta H_L\theta} + \frac{1}{2} \sum_{\eta=0}^{\Delta H_S} |c_{\eta}^1|^2 e^{-i\eta\theta} \right| \\
&= \left| \sum_{\eta=-\Delta H_L}^{\Delta H_S} c_{\eta} e^{-i\eta\theta} \right| = \left| c_{\eta_*} + \sum_{\eta \neq \eta_*} c_{\eta} e^{-i(\eta - \eta_*)\theta} \right|, \tag{34}
\end{aligned}$$

where we define $c_{\eta} := \frac{1}{2} |c_{\eta+\Delta H_L}^0|^2 + \frac{1}{2} |c_{\eta}^1|^2$ for $\eta \in [-\Delta H_L, \Delta H_S]$ and choose η_* such that $c_{\eta_*} \geq c_{\eta}$ for all η . Note that there is always a θ such that $\sum_{\eta \neq \eta_*} c_{\eta} \cos((\eta - \eta_*)\theta) = 0$ (because the integration of it from 0 to 2π is zero) and that $\sum_{\eta \neq \eta_*} c_{\eta} e^{-i(\eta - \eta_*)\theta}$ is imaginary. Then we must have

$$\min_{\theta} f_{\theta} \leq \sqrt{c_{\eta_*}^2 + (1 - c_{\eta_*})^2}. \tag{35}$$

To arrive at a non-trivial lower bound on $\delta_{\text{group}} = \sqrt{1 - \min_{\theta} f_{\theta}^2}$, we need an upper bound of $\min_{\theta} f_{\theta}$ which is smaller than 1. To this end, we analyze c_{η_*} in detail. In particular, we consider two situations:

(1) $c_{\eta_*} \leq 1/2$ and a constant upper bound on $\min_{\theta} f_{\theta}$ exists. We can always find a subset of $\{\eta\}$ denoted by \mathfrak{s} such that $1/4 \leq \sum_{\eta \in \mathfrak{s}} c_{\eta} \leq 1/2$. To find such a set, we first include η_* in \mathfrak{s} and add new elements into \mathfrak{s} one by one until their sum is at least $1/4$. Then there is always a θ such that $(\sum_{\eta \in \mathfrak{s}} c_{\eta} e^{-i\eta\theta}) \cdot (\sum_{\eta \notin \mathfrak{s}} c_{\eta} e^{-i\eta\theta})$ is imaginary, in which case $\min_{\theta} f_{\theta} \leq \sqrt{(1/4)^2 + (3/4)^2} = \sqrt{5/8}$ and we have

$$\delta_{\text{group}} \geq \sqrt{1 - \min_{\theta} f_{\theta}^2} \geq \sqrt{3/8}. \tag{36}$$

(2) $c_{\eta_*} > 1/2$. Then $\sqrt{c_{\eta_*}^2 + (1 - c_{\eta_*})^2}$ is a monotonically increasing function of c_{η_*} and we only need to find an upper bound on c_{η_*} . To see such an upper bound exists, we first consider the special case where $\varepsilon = 0$ and, according to the KL conditions and the HKS condition, $\langle \mathbf{c}_0 | H_S | \mathbf{c}_0 \rangle = \langle \mathbf{c}_1 | H_S | \mathbf{c}_1 \rangle$ holds true. On the other hand, if $c_{\eta_*} = \frac{1}{2} |c_{\eta_*+\Delta H_L}^0|^2 + \frac{1}{2} |c_{\eta_*}^1|^2 = 1$, we must have $|c_{\eta_*+\Delta H_L}^0|^2 = |c_{\eta_*}^1|^2 = 1$ and $\langle \mathbf{c}_0 | H_S | \mathbf{c}_0 \rangle - \langle \mathbf{c}_1 | H_S | \mathbf{c}_1 \rangle = \Delta H_L > 0$, leading to a contradiction.

In general, to derive a non-trivial upper bound on c_{η_*} , we first note that $0 \leq \eta^* \leq \Delta H_S - \Delta H_L$ because otherwise either $c_{\eta_*+\Delta H_L}^0 = 0$ or $c_{\eta_*}^1 = 0$ which contradicts with $c_{\eta_*} > 1/2$. We have $\chi = \langle \mathbf{c}_0 | H_S | \mathbf{c}_0 \rangle - \langle \mathbf{c}_1 | H_S | \mathbf{c}_1 \rangle$ and

$$\begin{aligned}
\sum_{\eta} |c_{\eta}^0|^2 \eta &= \sum_{\eta} |c_{\eta}^1|^2 \eta + \chi, \\
\Leftrightarrow \Delta H_L - \chi &= -\left(1 - |c_{\eta_*}^1|^2\right) \left(\eta_* - \frac{\sum_{\eta \neq \eta_*} |c_{\eta}^1|^2 \eta}{\sum_{\eta \neq \eta_*} |c_{\eta}^1|^2} \right) + \left(1 - |c_{\eta_*+\Delta H_L}^0|^2\right) \left(\eta_* + \Delta H_L - \frac{\sum_{\eta \neq \eta_*+\Delta H_L} |c_{\eta}^0|^2 \eta}{\sum_{\eta \neq \eta_*+\Delta H_L} |c_{\eta}^0|^2} \right).
\end{aligned}$$

Note that both $\left| \eta_* - \frac{\sum_{\eta \neq \eta_*} |c_{\eta}^1|^2 \eta}{\sum_{\eta \neq \eta_*} |c_{\eta}^1|^2} \right|$ and $\left| \eta_* + \Delta H_L - \frac{\sum_{\eta \neq \eta_*+\Delta H_L} |c_{\eta}^0|^2 \eta}{\sum_{\eta \neq \eta_*+\Delta H_L} |c_{\eta}^0|^2} \right|$ are at most ΔH_S . Therefore, $c_{\eta_*} \leq 1 - |\Delta H_L - \chi| / (2\Delta H_S)$ and

$$\delta_{\text{group}} \geq \sqrt{1 - \min_{\theta} f_{\theta}^2} \geq \sqrt{2c_{\eta_*}(1 - c_{\eta_*})} \geq \frac{\sqrt{|\Delta H_L - \chi| (\Delta H_S - \frac{1}{2} |\Delta H_L - \chi|)}}{\Delta H_S}. \tag{37}$$

Proposition 2 then follows from combining Eq. (36) and Eq. (37). \square

B. Bounding charge fluctuation by QEC inaccuracy

We now need to establish connections between χ and the QEC inaccuracy in order to link the global covariance violation to the QEC inaccuracy. We discuss two different methods that achieve this.

1. Knill–Laflamme-based method

Intuitively, a non-zero charge fluctuation leads to a violation of the KL conditions (Eq. (27)), which indicates a non-zero QEC inaccuracy. Therefore, we may bound the QEC inaccuracy through analyzing the deviation from the KL condition. We call this method the KL-based method. Specifically, we have

Proposition 3. *Consider an isometric quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. Then it holds that*

$$|\chi| \leq 2\varepsilon\mathfrak{J}, \quad (38)$$

where \mathfrak{J} is a function of H_S and \mathcal{N}_S defined by

$$\mathfrak{J}(\mathcal{N}_S, H_S) := \min_{h: H_S = \sum_{ij} h_{ij} K_{S,i}^\dagger K_{S,j}} \Delta h, \quad (39)$$

where h is Hermitian.

One can verify that $\mathfrak{J}(\mathcal{N}_S, H_S)$ is efficiently computable using the following semidefinite program [74]:

$$\min_{\substack{x, \nu \in \mathbb{R} \\ h \text{ is Hermitian}}} 2x, \quad \text{s.t.} \quad \begin{pmatrix} x\mathbb{1} & h - \nu\mathbb{1} \\ h - \nu\mathbb{1} & x\mathbb{1} \end{pmatrix} \geq 0, \quad H_S = \sum_{ij} h_{ij} K_{S,i}^\dagger K_{S,j}. \quad (40)$$

The proof of Proposition 3 is partly based on a useful lemma from Ref. [38] which connects the QEC inaccuracy ε to the deviation from the KL conditions:

Lemma 4 ([38]). *Let Π be the projection onto the code subspace of an isometric quantum code $\mathcal{E}_{S \leftarrow L}$ and the noise channel is $\mathcal{N}_S = \sum_{i=1}^r K_i(\cdot) K_i^\dagger$. Then*

$$\varepsilon(\mathcal{N}_S, \mathcal{E}_{S \leftarrow L}) = \min_{\Lambda} \sqrt{1 - f^2(\Lambda, \Lambda + \mathcal{B})}, \quad (41)$$

where $\Lambda(\rho) = \sum_{ij} \lambda_{ij} \text{Tr}(\rho) |i\rangle \langle j|$, $(\Lambda + \mathcal{B})(\rho) = \Lambda(\rho) + \sum_{ij} \text{Tr}(B_{ij} \mathcal{E}_{S \leftarrow L}(\rho)) |i\rangle \langle j|$, and λ_{ij} and B_{ij} are constant numbers and operators satisfying $\Pi K_i^\dagger K_j \Pi = \lambda_{ij} \Pi + \Pi B_{ij} \Pi$.

Proposition 3 then follows by connecting the deviation from the KL conditions to the charge fluctuation. The proof goes as follows.

Proof of Proposition 3. Let Π be the projection onto the code subspace under consideration, $|\mathbf{c}_0\rangle = W|0_L\rangle$, $|\mathbf{c}_1\rangle = W|1_L\rangle$ where $\mathcal{E}_{S \leftarrow L}(\cdot) = W(\cdot)W^\dagger$, and the simplified notations $\langle K_{S,i}^\dagger K_{S,j} \rangle_0 := \langle \mathbf{c}_0 | K_{S,i}^\dagger K_{S,j} | \mathbf{c}_0 \rangle$ and $\langle K_{S,i}^\dagger K_{S,j} \rangle_1 := \langle \mathbf{c}_1 | K_{S,i}^\dagger K_{S,j} | \mathbf{c}_1 \rangle$. Assume λ_{ij} and B_{ij} satisfies $\Pi K_{S,i}^\dagger K_{S,j} \Pi = \lambda_{ij} \Pi + \Pi B_{ij} \Pi$. Let $|\psi\rangle = \psi_0 |0_L\rangle |0_R\rangle + \psi_1 |1_L\rangle |1_R\rangle$ where $|\psi_0|^2 + |\psi_1|^2 = 1$, we have

$$f(\Lambda + \mathcal{B}, \Lambda) \leq \min_{\psi_0, \psi_1} f((\Lambda + \mathcal{B}) \otimes \mathbb{1}_R)(|\psi\rangle), (\Lambda \otimes \mathbb{1}_R)(|\psi\rangle)) = \min_{\psi_0} f(\lambda + \tilde{B}, \lambda), \quad (42)$$

where $\tilde{B}_{ij} = |\psi_0|^2 \langle K_{S,i}^\dagger K_{S,j} \rangle_0 + |\psi_1|^2 \langle K_{S,i}^\dagger K_{S,j} \rangle_1 - \lambda_{ij}$. According to the Fuchs–van de Graaf inequality $f(\rho, \sigma) \leq \sqrt{1 - \frac{1}{4} \|\rho - \sigma\|_1^2}$ [64],

$$\varepsilon = \min_{\Lambda, \mathcal{B}} \sqrt{1 - f(\Lambda + \mathcal{B}, \Lambda)^2} \geq \min_{\lambda} \max_{\psi_0} \sqrt{1 - f(\lambda + \tilde{B}, \lambda)^2} \geq \frac{1}{2} \min_{\lambda} \max_{\psi_0} \|\tilde{B}\|_1. \quad (43)$$

According to the HKS condition, $H_S = \sum_{ij} h_{ij} K_{S,i}^\dagger K_{S,j}$ for some Hermitian matrix h . Without loss of generality, we assume h is diagonal and $H_S = \sum_i h_{ii} K_{S,i}^\dagger K_{S,i}$ because if not, we can always choose another set of Kraus operators

that diagonalizes h . We can also assume $\max_i h_{ii} = -\min_i h_{ii} = \frac{\Delta h}{2}$ because we can replace H_S with $H_S - \nu \mathbb{1}$ for any $\nu \in \mathbb{R}$. Then we have

$$\begin{aligned}
\varepsilon &\geq \frac{1}{2} \min_{\lambda} \max_{\psi_0} \left\| \tilde{B} \right\|_1 \geq \frac{1}{2} \min_{\lambda} \max_{\psi_0} \sum_i \left| \tilde{B}_{ii} \right| \\
&= \frac{1}{2} \min_{\lambda} \max_{\psi_0} \sum_i \left| |\psi_0|^2 \langle K_{S,i}^\dagger K_{S,i} \rangle_0 + |\psi_1|^2 \langle K_{S,i}^\dagger K_{S,i} \rangle_1 - \lambda_{ii} \right| \\
&\geq \frac{1}{4} \min_{\lambda} \sum_i \left| \langle K_{S,i}^\dagger K_{S,i} \rangle_0 - \lambda_{ii} \right| + \left| \langle K_{S,i}^\dagger K_{S,i} \rangle_1 - \lambda_{ii} \right| \\
&\geq \frac{1}{4} \sum_i \left| \langle K_{S,i}^\dagger K_{S,i} \rangle_0 - \langle K_{S,i}^\dagger K_{S,i} \rangle_1 \right| \\
&\geq \frac{1}{4} \frac{1}{\max_i h_{ii}} \sum_i \left| h_{ii} \left(\langle K_{S,i}^\dagger K_{S,i} \rangle_0 - \langle K_{S,i}^\dagger K_{S,i} \rangle_1 \right) \right| \geq \frac{|\chi|}{2\Delta h}.
\end{aligned} \tag{44}$$

Note that there might be many different choices of h such that $H_S = \sum_{ij} h_{ij} K_{S,i}^\dagger K_{S,j}$ holds true. In order to obtain the tightest lower bound on ε , we can minimize Δh over all possible h such that $H_S = \sum_{ij} h_{ij} K_{S,i}^\dagger K_{S,j}$, leading to $\varepsilon \geq |\chi|/2\mathfrak{J}$, where $\mathfrak{J} := \min_{h: H_S = \sum_{ij} h_{ij} K_{S,i}^\dagger K_{S,j}} \Delta h$. \square

2. Quantum metrology method

Besides the KL-based method, the relationship between the charge fluctuation and the QEC inaccuracy could be understood through the lens of quantum metrology, which results in another inequality concerning χ and ε , as shown in the following. A detailed comparison between the two bounds obtained from the KL-based method and the quantum metrology method ([Proposition 3](#) and [Proposition 5](#)) will later be given in [Sec. III C](#) and [Sec. III D](#).

Proposition 5. *Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. Then it holds that*

$$|\chi| \leq 2\varepsilon \left(\sqrt{(1-\varepsilon^2)\mathfrak{F}} + \mathfrak{B} \right). \tag{45}$$

Here

$$\mathfrak{B} := \max_{|\psi\rangle} \sqrt{8\mathbb{V}_{H_S}(\mathcal{E}_{S \leftarrow L}(|\psi\rangle))} \leq \sqrt{2}\Delta H_S, \tag{46}$$

where the variance $\mathbb{V}_H(\rho) := \text{Tr}(H^2\rho) - (\text{Tr}(H\rho))^2$, and \mathfrak{F} is a function of \mathcal{N}_S and H_S defined by

$$\mathfrak{F}(\mathcal{N}_S, H_S) := 4 \min_{h: H_S = \sum_{ij} h_{ij} K_{S,i}^\dagger K_{S,j}} \left\| \sum_{ij} (h^2)_{ij} K_{S,i}^\dagger K_{S,j} - H_S^2 \right\|, \tag{47}$$

where h is Hermitian. In particular, when $\varepsilon \ll 1$ and $\mathfrak{B} \ll \sqrt{\mathfrak{F}}$, we have

$$|\chi| \lesssim 2\varepsilon \sqrt{\mathfrak{F}}. \tag{48}$$

Unlike $\mathfrak{J}(\mathcal{N}_S, H_S)$ introduced in [Sec. III B 1](#), $\mathfrak{F}(\mathcal{N}_S, H_S)$ appearing in [Proposition 5](#) has a clear operational meaning:

$$\mathfrak{F}(\mathcal{N}_S, H_S) \equiv F^\infty(\mathcal{N}_{S,\theta}), \quad \forall \theta \in \mathbb{R}. \tag{49}$$

Here $F^\infty(\mathcal{N}_{S,\theta})$ is the regularized QFI [[40](#), [75](#)] of the quantum channel $\mathcal{N}_{S,\theta} := \mathcal{N}_S \circ \mathcal{U}_{S,\theta}$ where θ is the unknown parameter to be estimated. (Generally, the regularized QFI of quantum channel Φ_θ is defined by $F^\infty(\Phi_\theta) := \lim_{N \rightarrow \infty} F(\Phi_\theta^{\otimes N})/N$.) Note that $F^\infty(\mathcal{N}_{S,\theta})$ is independent of θ and computable using semidefinite programming [[76](#)]. Also, $F^\infty(\mathcal{N}_{S,\theta}) \leq \infty$ if and only if the HKS condition is satisfied. The channel QFI inherits many nice properties from the QFI of quantum states. For example, here F^∞ obeys the monotonicity property, i.e. $F^\infty(\mathcal{R} \circ \Phi_\theta \circ \mathcal{E}) \leq F^\infty(\Phi_\theta)$ for arbitrary parameter-independent channels \mathcal{R} and \mathcal{E} .

The operational meaning of the quantity \mathfrak{B} is not immediately clear for general encoding channels, but when $\mathcal{E}_{S \leftarrow L}$ is isometric we have that

$$\mathfrak{B} = \sqrt{2F(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L})|_{\theta=0}}, \quad (50)$$

which satisfies

$$|\mathfrak{B} - \sqrt{2}\delta_{\text{local}}| \leq \sqrt{2}\Delta H_L, \quad (51)$$

due to the chain rule of the square root of the channel QFI [72]: $\sqrt{F(\Phi_{1,\theta} \circ \Phi_{2,\theta})} \leq \sqrt{F(\Phi_{1,\theta})} + \sqrt{F(\Phi_{2,\theta})}$ for any $\Phi_{1,\theta}$ and $\Phi_{2,\theta}$, and $F(\mathcal{U}_{L,\theta}) = F(\mathcal{U}_{L,\theta}^\dagger) = (\Delta H_L)^2$. In general, \mathfrak{B} depends on specific encodings and in order to obtain an code-independent bound we should replace \mathfrak{B} with its upper bound ΔH_S so that Eq. (45) becomes

$$|\chi| \leq 2\varepsilon \left(\sqrt{(1-\varepsilon^2)\mathfrak{F}} + \sqrt{2}\Delta H_S \right), \quad (52)$$

which still leads to useful bounds, e.g., for single-erasure noise as discussed in Sec. III D. However, in many cases \mathfrak{B} is negligible, i.e., $\mathfrak{B} \ll \sqrt{\mathfrak{F}}$ (or $\delta_{\text{local}} \ll \sqrt{\mathfrak{F}}$ for isometric codes) in the examples we study later in Sec. VII.

The monotonicity of the regularized QFI is a key ingredient in the proof of Proposition 5. Specifically, we introduce a two-level system C and an ancillary qubit A and consider the channel estimation of the error-corrected noise channel $\mathcal{N}_{C,\theta} = \mathcal{R}_{C \leftarrow SA} \circ (\mathcal{N}_{S,\theta} \otimes \mathbb{1}_A) \circ \mathcal{E}_{SA \leftarrow C}$ (see Fig. 2). $\mathcal{R}_{C \leftarrow SA}$ and $\mathcal{E}_{SA \leftarrow C}$ is carefully chosen such that $F^\infty(\mathcal{N}_{C,\theta})$ is roughly $\Theta((\chi/\varepsilon)^2)$ around $\theta = 0$. Intuitively, one might interpret $F^\infty(\mathcal{N}_{C,\theta})$ handwavyly as a quantity proportional to the square of the “signal-to-noise ratio” where the QEC inaccuracy ε is roughly the noise rate of $\mathcal{N}_{C,\theta}$ and the charge fluctuation $|\chi|$ is roughly the signal strength. Proposition 5 then follows from the monotonicity of QFI:

$$F^\infty(\mathcal{N}_{C,\theta}) \leq F^\infty(\mathcal{N}_{S,\theta}). \quad (53)$$

We now explain the error-corrected metrology protocol in detail. We first introduce an ancilla-assisted two-level encoding. Consider a two-level system C spanned by $|0_C\rangle$ and $|1_C\rangle$ and a Hamiltonian

$$H_C = \frac{\Delta H_L}{2} \cdot Z_C, \quad \mathcal{U}_{C,\theta}(\cdot) = e^{-iH_C\theta}(\cdot)e^{iH_C\theta}, \quad (54)$$

where Z_C is the Pauli-Z operator. We define a repetition encoding from C to LA ,

$$\mathcal{E}_{LA \leftarrow C}^{\text{rep}}(|0_C\rangle) := |0_L 0_A\rangle, \quad \mathcal{E}_{LA \leftarrow C}^{\text{rep}}(|1_C\rangle) := |1_L 1_A\rangle, \quad (55)$$

where A is the ancillary qubit. The corresponding repetition recovery channel is

$$\mathcal{R}_{C \leftarrow LA}^{\text{rep}}(\cdot) := \sum_{i=0}^{d_L-1} R_i^{\text{rep}}(\cdot) R_i^{\text{rep}\dagger}, \quad (56)$$

where $R_0^{\text{rep}} = |0_C\rangle\langle 0_L 0_A| + |1_C\rangle\langle 1_L 1_A|$ and $R_1^{\text{rep}} = |0_C\rangle\langle 1_L 0_A| + |1_C\rangle\langle 0_L 1_A|$ and $R_{i>1}^{\text{rep}} = |0_C\rangle\langle i_L 0_A| + |1_C\rangle\langle i_L 1_A|$. Clearly, $\mathcal{R}_{C \leftarrow LA}^{\text{rep}} \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}} = \mathbb{1}_C$. The repetition encoding $\mathcal{E}_{LA \leftarrow C}^{\text{rep}}$ is covariant with respect to H_C and H_L , i.e., $\mathcal{E}_{LA \leftarrow C}^{\text{rep}} \circ \mathcal{U}_{C,\theta} = (\mathcal{U}_{L,\theta} \otimes \mathbb{1}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}$. Moreover, the repetition code corrects all bit-flip noise. When concatenated with $\mathcal{E}_{S \leftarrow L}$ and $\mathcal{R}_{L \leftarrow S}$, the error-corrected noisy channel $\mathcal{N}_{C,\theta} = \mathcal{R}_{C \leftarrow SA}^{\text{rep}} \circ (\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \otimes \mathbb{1}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}$ becomes a *rotated dephasing channel*, namely, a single-qubit channel which is a composition of dephasing channel $(1-p)(\cdot) + pZ(\cdot)Z$ and a Pauli-Z phase rotation $e^{-i\phi_Z}$ [22] (see Fig. 2). The regularized QFI of any rotated dephasing channel Φ_θ is [40]

$$F^\infty(\Phi_\theta) = \frac{|\partial_\theta x_\theta|^2}{1 - |x_\theta|^2}, \quad (57)$$

where the complex number $x_\theta = \langle 0|\Phi_\theta(|0\rangle\langle 1|)|1\rangle$. We consider the estimation around $\theta = 0$ for $\mathcal{N}_{C,\theta}$. The monotonicity of the regularized QFI guarantees that

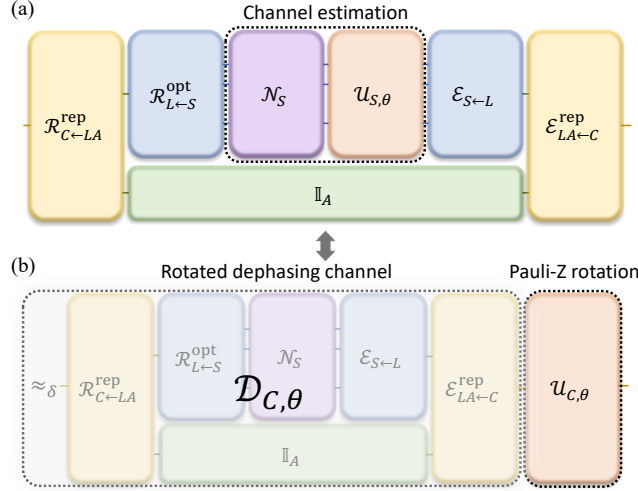
$$F^\infty(\mathcal{N}_{C,\theta})|_{\theta=0} = \frac{|\partial_\theta \xi_\theta|^2|_{\theta=0}}{1 - |\xi_{\theta=0}|^2} \leq F^\infty(\mathcal{N}_{S,\theta}), \quad (58)$$

where

$$\xi_\theta := \langle 0_C|\mathcal{N}_{C,\theta}(|0_C\rangle\langle 1_C|)|1_C\rangle. \quad (59)$$

Proposition 5 can then be proven, connecting ξ_θ with ε and χ .

Now we are ready to present the formal proof of Proposition 5.



Supplementary Figure 2. A two-level encoding scheme for the estimation of θ . (a) Definition of the encoded channel $\mathcal{N}_{C,\theta}$ in the system C . $\mathcal{N}_{C,\theta} = \mathcal{R}_{C \leftarrow SA} \circ (\mathcal{N}_{S,\theta} \otimes \mathbb{I}_A) \circ \mathcal{E}_{SA \leftarrow C}$ where $\mathcal{R}_{C \leftarrow SA} = \mathcal{R}_{LA \leftarrow C}^{\text{rep}} \circ (\mathcal{R}_{L \leftarrow S}^{\text{opt}} \otimes \mathbb{I}_A)$ and $\mathcal{E}_{SA \leftarrow C} = (\mathcal{E}_{S \leftarrow L} \otimes \mathbb{I}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}$. It includes a concatenation of the repetition encoding from C to LA and the encoding under consideration $\mathcal{E}_{S \leftarrow L}$ from L to S with the optimal recovery channels chosen accordingly. (b) $\mathcal{N}_{C,\theta}$ is the composition of $\mathcal{D}_{C,\theta}$ and $\mathcal{U}_{C,\theta}$ where $\mathcal{U}_{C,\theta}$ is the unitary rotation with respect to a Pauli-Z Hamiltonian H_C and $\mathcal{D}_{C,\theta}$ is a rotated dephasing channel which is at most ε -far from identity at $\theta = 0$. When $\delta_{\text{group}} \approx 0$, $\mathcal{D}_{C,\theta} \approx \mathcal{R}_{C \leftarrow SA} \circ (\mathcal{N}_S \otimes \mathbb{I}_A) \circ \mathcal{E}_{SA \leftarrow C}$ is almost θ -independent. Note that the gates are applied from right to left.

Proof of Proposition 5. Let $\mathcal{R}_{L \leftarrow S}^{\text{opt}}$ be the recovery channel such that $\varepsilon = P(\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{I}_L)$. Let $\mathcal{R}_{C \leftarrow SA} = \mathcal{R}_{C \leftarrow SA}^{\text{rep}} \circ (\mathcal{R}_{L \leftarrow S}^{\text{opt}} \otimes \mathbb{I}_A)$ and $\mathcal{E}_{SA \leftarrow C} = (\mathcal{E}_{S \leftarrow L}^{\text{opt}} \otimes \mathbb{I}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}$, we have two rotated dephasing channels (see Fig. 2):

$$\mathcal{N}_{C,\theta} = \mathcal{R}_{C \leftarrow SA} \circ (\mathcal{N}_{S,\theta} \otimes \mathbb{I}_A) \circ \mathcal{E}_{SA \leftarrow C} =: \mathcal{D}_{C,\theta} \circ \mathcal{U}_{C,\theta}, \quad (60)$$

where $\mathcal{D}_{C,\theta}$ and $\mathcal{N}_{C,\theta}$ are rotated dephasing channels of the following forms:

$$\begin{aligned} \mathcal{D}_{C,\theta}(\cdot) &= (1 - p_\theta) e^{-i \frac{\phi_\theta}{2} Z_C} (\cdot) e^{i \frac{\phi_\theta}{2} Z_C} + p_\theta Z_C e^{-i \frac{\phi_\theta}{2} Z_C} (\cdot) e^{i \frac{\phi_\theta}{2} Z_C} Z_C, \\ \mathcal{N}_{C,\theta}(\cdot) &= (1 - p_\theta) e^{-i \frac{\phi_\theta + \Delta H_L \theta}{2} Z_C} (\cdot) e^{i \frac{\phi_\theta + \Delta H_L \theta}{2} Z_C} + p_\theta Z_C e^{-i \frac{\phi_\theta + \Delta H_L \theta}{2} Z_C} (\cdot) e^{i \frac{\phi_\theta + \Delta H_L \theta}{2} Z_C} Z_C, \end{aligned}$$

and $\xi_\theta = \langle 0_C | \mathcal{D}_{C,\theta} \circ \mathcal{U}_{C,\theta} | 0_C \rangle \langle 1_C | 1_C \rangle = (1 - 2p_\theta) e^{-i(\phi_\theta + \Delta H_L \theta)}$. $\mathcal{D}_{C,\theta}$ is identity when the code is both exactly covariant and exactly error-correcting; it is θ -independent when the code is exactly covariant (see also Ref. [22]). Note that we will not use the channel $\mathcal{D}_{C,\theta}$ in this proof (it will be used later in Sec. IV), but we introduce the notation here to clarify its physical meaning.

Consider the parameter estimation of θ in the neighborhood of $\theta = 0$. On one hand, for rotated dephasing channels (see Sec. X for the purified distance of rotated dephasing channels from identity), we have

$$\sqrt{p_{\theta=0}} \leq P(\mathcal{D}_{C,\theta=0}, \mathbb{I}_C) = P(\mathcal{N}_{C,\theta=0}, \mathbb{I}_C). \quad (61)$$

On the other hand, we have

$$P(\mathcal{N}_{C,\theta=0}, \mathbb{I}_C) = P(\mathcal{R}_{C \leftarrow SA} \circ (\mathcal{N}_S \otimes \mathbb{I}_A) \circ \mathcal{E}_{SA \leftarrow C}, \mathbb{I}_C) \leq P(\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{I}_L) \leq \varepsilon, \quad (62)$$

where we use the monotonicity of the purified distance [54] and the definition of ε . Combining Eq. (61) and Eq. (62), we have

$$|\xi_{\theta=0}| \geq 1 - 2\varepsilon^2. \quad (63)$$

As shown in Sec. XI, we also have

$$|\partial_\theta \xi_\theta|^2 \big|_{\theta=0} \geq (|\chi| - 2\varepsilon\mathfrak{B})^2, \quad (64)$$

when $|\chi| \geq 2\varepsilon\mathfrak{B}$. Hence, when $|\chi| \geq 2\varepsilon\mathfrak{B}$, we must have

$$F^\infty(\mathcal{N}_{S,\theta}) = F^\infty(\mathcal{N}_{S,\theta}) \big|_{\theta=0} \geq F^\infty(\mathcal{N}_{C,\theta}) \big|_{\theta=0} \geq \frac{(|\chi| - 2\varepsilon\mathfrak{B})^2}{4\varepsilon^2(1 - \varepsilon^2)}, \quad (65)$$

completing the proof. \square

C. Consequent bounds on the trade-off between QEC and global covariance

In [Sec. III A](#) and [Sec. III B](#), we derived bounds on δ_{group} and ε separately, using the notion of charge fluctuation χ . Combining these results, we immediately obtain the trade-off relations between δ_{group} and ε .

Theorem 6. *Consider an isometric quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. It holds that, when $0 \leq G(\varepsilon) \leq \Delta H_S$,*

$$\delta_{\text{group}} \geq \min \left\{ \frac{\sqrt{G(\varepsilon) (\Delta H_S - \frac{1}{2} G(\varepsilon))}}{\Delta H_S}, \sqrt{\frac{3}{8}} \right\}, \quad (66)$$

and when $G(\varepsilon) > \Delta H_S$, $\delta_{\text{group}} \geq \sqrt{3/8}$, where we could take either

$$G(\varepsilon) = \Delta H_L - 2\varepsilon\mathfrak{J}, \quad (67)$$

or

$$G(\varepsilon) = \Delta H_L - 2\varepsilon(\sqrt{1 - \varepsilon^2}\mathfrak{F} + \mathfrak{B}), \quad (68)$$

where \mathfrak{J} , \mathfrak{F} , and \mathfrak{B} are given by [Eq. \(39\)](#), [Eq. \(47\)](#), and [Eq. \(46\)](#), respectively.

For the extreme cases of exactly covariant codes and exactly error-correcting codes, we have the following corollaries:

Corollary 7. *Consider an isometric quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. When $\varepsilon = 0$, i.e., the code is exactly error-correcting, it holds that when $\Delta H_L \leq \Delta H_S$,*

$$\delta_{\text{group}} \geq \frac{\sqrt{\Delta H_L (\Delta H_S - \frac{1}{2} \Delta H_L)}}{\Delta H_S}, \quad (69)$$

and when $\Delta H_L > \Delta H_S$, $\delta_{\text{group}} \geq \sqrt{3/8}$.

Corollary 8. *Consider an isometric quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. When $\delta_{\text{group}} = 0$, i.e., the code is exactly covariant, it holds that*

$$\varepsilon \geq \frac{\Delta H_L}{2\mathfrak{J}}, \quad (70)$$

where \mathfrak{J} is given by [Eq. \(39\)](#).

Corollary 9. *Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. When $\delta_{\text{group}} = 0$, i.e., the code is exactly covariant, we must have either $2\varepsilon\mathfrak{B} \geq \Delta H_L$ or*

$$\varepsilon \frac{\sqrt{1 - \varepsilon^2}}{1 - 2\varepsilon\mathfrak{B}/\Delta H_L} \geq \frac{\Delta H_L}{2\sqrt{\mathfrak{F}}}, \quad (71)$$

where \mathfrak{F} and \mathfrak{B} are given by [Eq. \(47\)](#) and [Eq. \(46\)](#), respectively.

We make a few remarks on the scope of application of these results. Although [Proposition 2](#) and [Proposition 3](#) need the isometric encoding assumption, [Proposition 5](#) (and thus [Corollary 9](#)) holds for arbitrary codes. Also, [Proposition 2](#) only holds when H_L and H_S share a common period, but [Proposition 3](#) and [Proposition 5](#) hold true for arbitrary Hamiltonians without the $U(1)$ assumption. Finally, a keen reader might have already noticed that the choice of the pair of orthonormal states $\{|0_L\rangle, |1_L\rangle\}$ in the proofs of [Proposition 2](#), [Proposition 3](#) and [Proposition 5](#) is quite arbitrary (chosen only for the purpose of proving [Theorem 6](#)) and we can in principle replace it with any other pair and the proofs will still hold, leading to refinements of these propositions. We present these refinements in detail in [Sec. XII](#). In particular, [Proposition 2](#) leads to an inequality between δ_{group} and δ_{charge} .

To compare the results from the KL-based method and the quantum metrology method, we first consider the limiting situation where $\delta_{\text{group}} \ll 1$ and $\varepsilon \ll 1$. Then we have

$$\delta_{\text{group}} \gtrsim \sqrt{\frac{\Delta H_L - 2\varepsilon\mathfrak{J}}{\Delta H_S}} \Leftrightarrow \varepsilon + \delta_{\text{group}}^2 \frac{\Delta H_S}{2\mathfrak{J}} \gtrsim \frac{\Delta H_L}{2\mathfrak{J}}, \quad (72)$$

$$\delta_{\text{group}} \gtrsim \sqrt{\frac{\Delta H_L - 2\varepsilon(\sqrt{\mathfrak{F}} + \mathfrak{B})}{\Delta H_S}} \Leftrightarrow \varepsilon + \delta_{\text{group}}^2 \frac{\Delta H_S}{2(\sqrt{\mathfrak{F}} + \mathfrak{B})} \gtrsim \frac{\Delta H_L}{2(\sqrt{\mathfrak{F}} + \mathfrak{B})}. \quad (73)$$

When $\mathfrak{B} \ll \sqrt{\mathfrak{F}}$, the metrology bound performs no worse than the KL-based bound because we always have $\mathfrak{J}^2 \geq \mathfrak{F}$ (proof in [Sec. XIII](#)). For the examples we study later in [Sec. VII](#), we find that \mathfrak{B} is negligible, but in practice one may need to bound the parameter \mathfrak{B} a priori using properties of specific codes to obtain desired trade-off relations. It still open in general under which conditions $\mathfrak{B} \ll \sqrt{\mathfrak{F}}$ holds, and whether [Proposition 5](#) might be further improved with \mathfrak{B} removed.

A byproduct of our results are lower bounds on ε ([Eq. \(70\)](#) and [Eq. \(71\)](#)) for exactly covariant codes, a special case which has been extensively studied in previous works [[12](#), [13](#), [21–23](#)]. As discussed below, the bound [Eq. \(70\)](#) for random local erasure noise behaves almost the same as the one in [Ref. \[12\]](#) and our [Proposition 3](#) provides an alternative method to obtain this result. However, compared to [Proposition 5](#), the bound in [Ref. \[22\]](#)

$$\varepsilon \frac{\sqrt{1 - \varepsilon^2}}{1 - 2\varepsilon^2} \geq \frac{\Delta H_L}{2\sqrt{\mathfrak{F}}} \quad (74)$$

does not involve the parameter \mathfrak{B} , implying that the proof of our [Proposition 5](#) might be further improved.

D. Noise models and explicit behaviors of the bounds

Now we explicitly discuss how the bounds in [Theorem 6](#) behave under difference types of noise in an n -partite system. We consider 1-local Hamiltonians $H_S = \sum_{l=1}^n H_{S_l}$, so $\Delta H_S = O(n)$. In this case we have $\delta_{\text{group}} = \Omega(1/\sqrt{\Delta H_S}) = \Omega(1/\sqrt{n})$ as long as $G(\varepsilon) = \Omega(1)$. On the other hand, when $G(\varepsilon) = o(1)$, the scaling of ε must be lower bounded by $\Omega(1/\mathfrak{J})$ (or $\Omega(1/(\sqrt{\mathfrak{F}} + \mathfrak{B}))$) so it is important to understand the scalings of \mathfrak{J} , \mathfrak{F} and \mathfrak{B} . When n is large, the values of \mathfrak{J} and \mathfrak{F} may be not efficiently computable. However, in [Proposition 3](#), [Proposition 5](#) and [Theorem 6](#), we could always replace them with their efficiently computable upper bounds and the trade-off relations still hold then. We discuss the following two general noise models [[12](#), [13](#), [21–23](#)] (there are still other types of noises that we will not cover, e.g., random long-range phase errors [[13](#)]):

(1) Random local noise, where different local noise channels acting on a constant number of subsystems randomly. Specifically, $\mathcal{N}_S = \sum_i q_i \mathcal{N}_S^{(i)}$ and $H_S = \sum_i H_S^{(i)}$, where $\mathcal{N}_S^{(i)}$ represent the local noise channels acting on a constant number of subsystems, q_i represent their probabilities ($q_i > 0$ and $\sum_i q_i = 1$), and the HKS condition is satisfied for each pair of $(\mathcal{N}_S^{(i)}, H_S^{(i)})$. Then we have

$$\mathfrak{J}(\mathcal{N}_S, H_S) \leq \max_i \frac{1}{q_i} \mathfrak{J}(\mathcal{N}_S^{(i)}, H_S^{(i)}), \quad (75)$$

$$\mathfrak{F}(\mathcal{N}_S, H_S) \leq \sum_i \frac{1}{q_i} \mathfrak{F}(\mathcal{N}_S^{(i)}, H_S^{(i)}), \quad (76)$$

where

$$\tilde{\mathfrak{F}}(\mathcal{N}_S, H_S) := 4 \min_{h: H_S = \sum_{ij} h_{ij} K_{S,i}^\dagger K_{S,j}} \left\| \sum_{ij} (h^2)_{ij} K_{S,i}^\dagger K_{S,j} \right\|. \quad (77)$$

We prove [Eq. \(75\)](#) in [Sec. XIV](#) and [Eq. \(76\)](#) was previously known in [Ref. \[22\]](#). Note that $\tilde{\mathfrak{F}}$ might be different when we replace H_{S_l} with $H_{S_l} - \nu \mathbb{1}$ for some $\nu \in \mathbb{R}$, one need to minimize over ν to find the optimal $\tilde{\mathfrak{F}}$ [[22](#)]. For example, consider single-erasure noise in an n -partite system and let the erasure channel of the ℓ -th subsystem be $\mathcal{N}_{S_\ell}(\cdot) = |\emptyset\rangle\langle\emptyset|_{S_\ell} \otimes \text{Tr}_{S_\ell}(\cdot)$ (we use $|\emptyset\rangle$ to represent the vacuum state after erasure). When $\mathcal{N}_S = \sum_{l=1}^n \frac{1}{n} \mathcal{N}_{S_l}$ and the Hamiltonian takes the 1-local form $H_S = \sum_{l=1}^n H_{S_l}$, we have $\mathfrak{J}(\mathcal{N}_{S_l}, H_{S_l}) = \Delta H_{S_l}$ and $\tilde{\mathfrak{F}}(\mathcal{N}_{S_l}, H_{S_l}) = (\Delta H_{S_l})^2$. Then

we have

$$\mathfrak{J}(\mathcal{N}_S, H_S) \leq n \max_l \Delta H_{S_l}, \quad (78)$$

$$\mathfrak{F}(\mathcal{N}_S, H_S) \leq n \sum_{l=1}^n (\Delta H_{S_l})^2, \quad (79)$$

$$\sqrt{\mathfrak{F}(\mathcal{N}_S, H_S)} + \mathfrak{B}(\mathcal{N}_S, H_S) \leq n \left(\sqrt{\frac{\sum_{l=1}^n (\Delta H_{S_l})^2}{n}} + \sqrt{2} \frac{\Delta H_S}{n} \right). \quad (80)$$

Note that using Eq. (78) and Eq. (70), we obtain $\varepsilon \geq \Delta H_L / (2n \max_l \Delta H_{S_l})$ which is identical to Theorem 1 in Ref. [12]. In Eq. (80), we use

$$\mathfrak{B} \leq \sqrt{2} \Delta H_S \quad (81)$$

Comparing Eq. (80) with Eq. (78), we find that when \mathfrak{B} is not negligible, the quantum metrology method can still outperform the KL-based method in some cases (e.g., when one of ΔH_{S_l} is extremely large). For other types of random local noise acting on each subsystem uniformly randomly, we also have $\mathfrak{J} = O(n)$ and $\sqrt{\mathfrak{F}} + \mathfrak{B} = O(n)$ and the behaviors of the trade-off relations from the KL-based method and the quantum metrology method are similar. From Eq. (72) and Eq. (73), we have $\varepsilon + \Theta(1) \cdot \delta_{\text{group}}^2 = \Omega(1/n)$, meaning that when both ε and δ_{group} are sufficiently small, their optimal scalings are $\Theta(1/n)$ and $\Theta(1/\sqrt{n})$, respectively.

(2) Independent noise, where noise channels act on each subsystem independently. Note that independent noise is considered a “stronger” noise model than local noise because the noise actions are no longer guaranteed to be local. Specifically, $\mathcal{N}_S = \bigotimes_{l=1}^n \mathcal{N}_{S_l}$ and $H_S = \sum_{l=1}^n H_{S_l}$ where \mathcal{N}_{S_l} represent independent noise channels acting on each subsystem and H_{S_l} acts only non-trivially on the subsystem S_l , and the HKS condition is satisfied for each pair of $(\mathcal{N}_{S_l}, H_{S_l})$. Here we have

$$\mathfrak{J}(\mathcal{N}_S, H_S) \leq \sum_{l=1}^n \mathfrak{J}(\mathcal{N}_{S_l}, H_{S_l}), \quad (82)$$

$$\mathfrak{F}(\mathcal{N}_S, H_S) = \sum_{l=1}^n \mathfrak{F}(\mathcal{N}_{S_l}, H_{S_l}). \quad (83)$$

The proof of Eq. (82) is provided in Sec. XIV, and Eq. (83) follows directly from the additivity of the regularized QFI $\mathfrak{F} = F^\infty(\bigotimes_{l=1}^n \mathcal{N}_{S_l, \theta}) = \sum_{l=1}^n F^\infty(\mathcal{N}_{S_l, \theta})$ [22]. Therefore we have $\mathfrak{J} = O(n)$ and $\sqrt{\mathfrak{F}} = O(\sqrt{n})$ and there is now a quadratic gap between them. If \mathfrak{B} can be upper bounded by $O(\sqrt{n})$ (e.g., in Sec. VII), from Eq. (73), we have $\varepsilon + \Theta(\sqrt{n}) \cdot \delta_{\text{group}}^2 = \Omega(1/\sqrt{n})$, meaning that when both ε and δ_{group} are sufficiently small, their optimal scalings are both $\Theta(1/\sqrt{n})$. From Eq. (72), we only have $\varepsilon + \Theta(1) \cdot \delta_{\text{group}}^2 = \Omega(1/n)$ and a worse lower bound $\varepsilon = \Omega(1/n)$ for small δ_{group} . In general, for independent noise, the trade-off bound from the KL-based method is asymptotically weaker than the one from the quantum metrology method as long as $\mathfrak{B} = o(n)$.

Finally, we remark here that the exact values of $\mathfrak{J}(\mathcal{N}_{S_l}, H_{S_l})$, $\mathfrak{F}(\mathcal{N}_{S_l}, H_{S_l})$ and $\tilde{\mathfrak{F}}(\mathcal{N}_{S_l}, H_{S_l})$ can also be analytically calculated (or upper bounded) for not only erasure noise, but also other types of practically relevant noise, e.g., depolarizing noise. In principle, to derive an upper bound on $\mathfrak{J}(\mathcal{N}_{S_l}, H_{S_l})$, $\mathfrak{F}(\mathcal{N}_{S_l}, H_{S_l})$ or $\tilde{\mathfrak{F}}(\mathcal{N}_{S_l}, H_{S_l})$, one only need to find a Hermitian matrix h that satisfies $H_{S_l} = \sum_{ij} h_{ij} K_{S_l, i}^\dagger K_{S_l, j}$ and use the target functions Δh , $\left\| \sum_{ij} (h^2)_{ij} K_{S_l, i}^\dagger K_{S_l, j} - H_{S_l}^2 \right\|$ or $\left\| \sum_{ij} (h^2)_{ij} K_{S_l, i}^\dagger K_{S_l, j} \right\|$ as the upper bound. One can further tighten the bound by minimizing the target functions over all possible h . We give a few examples below.

First, we note that for an erasure noise channel $\mathcal{N}(\cdot) = (1-p)(\cdot) + p|\emptyset\rangle\langle\emptyset| \text{Tr}(\cdot)$, we have

$$\mathfrak{J}(\mathcal{N}, H) = \frac{\Delta H}{p}, \quad \mathfrak{F}(\mathcal{N}, H) = (\Delta H)^2 \frac{1-p}{p}, \quad \tilde{\mathfrak{F}}(\mathcal{N}, H) = \frac{(\Delta H)^2}{p}. \quad (84)$$

To derive these, we assume the system is d -dimensional and let $K_1 = \sqrt{1-p} \mathbb{1}$, $K_{i+1} = \sqrt{p} |\emptyset\rangle\langle i|$, for $i = 1, \dots, d$. Then the d -dimensional Hermitian matrix h such that $H = \sum_{ij} h_{ij} K_i^\dagger K_j$ must be $h = \begin{pmatrix} \frac{h_{11}}{1-p} & 0 \\ 0 & \frac{H - h_{00} \mathbb{1}}{p} \end{pmatrix}$ for some $h_{11} \in \mathbb{R}$. The above equations follow straightforwardly by minimizing the target functions over h_{11} (see also Ref. [22]).

Similarly, for single-qubit depolarizing noise $\mathcal{N}(\cdot) = (1-p)(\cdot) + p\frac{\mathbb{1}}{2}$, we have

$$\mathfrak{J}(\mathcal{N}, H) \leq \frac{2\Delta H}{p(1-p/2)}, \quad \mathfrak{F}(\mathcal{N}, H) = (\Delta H)^2 \frac{2(1-p)^2}{p(3-2p)}, \quad \tilde{\mathfrak{F}}(\mathcal{N}, H) = (\Delta H)^2 \frac{2-p}{p(3-2p)}. \quad (85)$$

To derive these, we let $K_1 = \sqrt{1 - \frac{3}{4}p}\mathbb{1}$, $K_2 = \sqrt{\frac{p}{4}}X$, $K_3 = \sqrt{\frac{p}{4}}Y$ and $K_4 = \sqrt{\frac{p}{4}}Z$. Then the 4-dimensional Hermitian matrix such that $H = \sum_{ij} h_{ij} K_i^\dagger K_j$ must be $h = \begin{pmatrix} 0 & 0 & 0 & h_{14} \\ 0 & 0 & ih_{23} & 0 \\ 0 & -ih_{23} & 0 & 0 \\ h_{14} & 0 & 0 & 0 \end{pmatrix}$ for some $h_{14}, h_{23} \in \mathbb{R}$ when $\left\| \sum_{ij} (h^2)_{ij} K_{S,i}^\dagger K_{S,j} - H_S^2 \right\|$ in \mathfrak{F} and $\left\| \sum_{ij} (h^2)_{ij} K_{S,i}^\dagger K_{S,j} \right\|$ in \mathfrak{J} are minimized (see Ref. [40]). The above equations follow straightforwardly by minimizing the target functions over h_{14} and h_{23} . Note that for \mathfrak{J} , there is no guarantee that the anti-diagonal form of h is optimal, so it only provides an upper bound on \mathfrak{J} .

Finally, for depolarizing noise on qudits: $\mathcal{N}(\cdot) = (1-p)(\cdot) + p\frac{\mathbb{1}}{d}$, we have from Ref. [22] that

$$\mathfrak{F}(\mathcal{N}, H) \leq (\Delta H)^2 \frac{1-p}{p}, \quad \tilde{\mathfrak{F}}(\mathcal{N}, H) \leq (\Delta H)^2 \frac{1}{p}, \quad (86)$$

and how to find a simple upper bound on \mathfrak{J} is still open.

IV. SUPPLEMENTARY NOTE 4: TRADE-OFF BETWEEN QEC AND GLOBAL COVARIANCE: GATE IMPLEMENTATION ERROR APPROACH

In this section, we introduce another framework that also enables us to derive the trade-off between the QEC inaccuracy ε and the global covariance violation δ_{group} and could be interesting in its own right. Here the idea is to analyze a key notion which we call the gate implementation error γ that allow us to treat ε and δ_{group} on the same footing. More specifically, we first formally define γ in Sec. IV A and show that $\varepsilon + \delta_{\text{group}} \geq \gamma$. Then we derive two lower bounds on γ using two different methods from quantum metrology and quantum resource theory, which automatically induce two trade-off relations between the QEC inaccuracy and the global covariance violation. We will compare the gate implementation error approach to the charge fluctuation approach at the end of this section.

A. Gate implementation error as a unification of QEC inaccuracy and global covariance violation

Consider a practical quantum computing scenario where we want to implement a set of logical gates $U_{L,\theta} = e^{-iH_L\theta}$ for $\theta \in \mathbb{R}$ using physical gates $U_{S,\theta} = e^{-iH_S\theta}$ under noise \mathcal{N}_S . We would like to design an encoding and a recovery channel such that $\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}$ simulate $\mathcal{U}_{L,\theta}$. We call the error in such simulations the *gate implementation error* and the *Choi gate implementation error*, defined by

$$\gamma := \min_{\mathcal{R}_{L \leftarrow S}} \max_{\theta} P(\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{U}_{L,\theta}), \quad (87)$$

$$\bar{\gamma} := \min_{\mathcal{R}_{L \leftarrow S}} \max_{\theta} \bar{P}(\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{U}_{L,\theta}). \quad (88)$$

Both the QEC inaccuracy and the covariance violation contribute to the gate implementation error (see Fig. 3). Clearly, $\gamma = 0$ when the quantum code is exactly error-correcting and covariant. In general, γ is upper bounded by the sum of ε and δ_{group} , as shown in the following proposition.

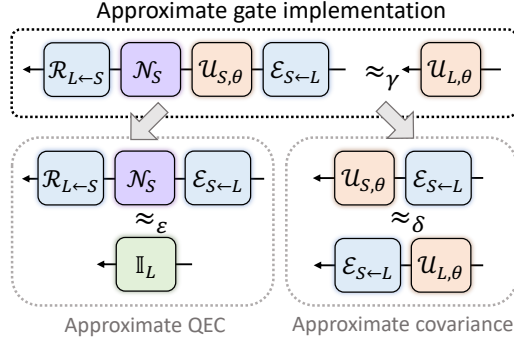
Proposition 10. *Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . It holds that*

$$\varepsilon + \delta_{\text{group}} \geq \gamma, \quad (89)$$

$$\bar{\varepsilon} + \bar{\delta}_{\text{group}} \geq \bar{\gamma}. \quad (90)$$

Proof. Using the triangular inequality of the purified distance [54], we have

$$P(\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{U}_{L,\theta}) \leq P(\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}) + P(\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}, \mathcal{U}_{L,\theta}). \quad (91)$$



Supplementary Figure 3. Both the QEC inaccuracy ε and the covariance violation δ_{group} contribute to the error in approximate gate implementation. Specifically, the (Choi) gate implementation error γ ($\bar{\gamma}$) is upper bounded by $\delta_{\text{group}} + \varepsilon$ ($\bar{\delta}_{\text{group}} + \bar{\varepsilon}$). Note that the gates are applied from right to left.

The first term is upper bounded by $P(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta})$ using the monotonicity of the purified distance and the second term is equal to $P(\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{I}_L)$ by definition. Then $\gamma \leq \varepsilon + \delta_{\text{group}}$ follows by taking the maximization over θ and the minimization over $\mathcal{R}_{L \leftarrow S}$ on both sides. The above discussion also holds when we replace the purified distance $P(\cdot, \cdot)$ with the Choi purified distance $\bar{P}(\cdot, \cdot)$, implying that $\bar{\gamma} \leq \bar{\varepsilon} + \bar{\delta}_{\text{group}}$. \square

B. Bounding gate implementation error

1. Quantum metrology method

Now we derive a lower bound on the gate implementation error γ , where we consider the approximate gate implementation of $\mathcal{U}_{L,\theta}$ using noisy gates $\mathcal{N}_{S,\theta}$ as an error-corrected metrology protocol where θ is an unknown parameter to be estimated.

Again, we use the ancilla-assisted two-level encoding, as introduced in Sec. III B 2. We choose the repetition code concatenated with the quantum code under study, so that the error-corrected noise channel $\mathcal{N}_{C,\theta} = \mathcal{R}_{C \leftarrow SA}^{\text{rep}} \circ (\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \otimes \mathbb{I}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}$ becomes a rotated dephasing channel. The main difference between the error-corrected metrology protocol we use here and the one in Sec. III B 2 is that now we choose the recovery channel $\mathcal{R}_{L \leftarrow S}$ to be the optimal recovery channel which minimizes the gate implementation error (instead of the QEC inaccuracy) and guarantees a lower noise rate over the entire group of θ (instead of just around $\theta = 0$). In this case, we show that there always exists some θ_* such that $F^\infty(\mathcal{N}_{C,\theta_*}) = \Theta((\Delta H_L/\gamma)^2)$, which then provide us a lower bound on γ using the monotonicity of the regularized QFI. Now we state and prove Theorem 11 which provides a lower bound on γ (and thus on $\varepsilon + \delta_{\text{group}}$).

Theorem 11. *Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. Then it holds that*

$$\varepsilon + \delta_{\text{group}} \geq \gamma \geq \ell_1 \left(\frac{\Delta H_L}{2\sqrt{\mathfrak{F}}} \right), \quad (92)$$

where \mathfrak{F} is given by Eq. (47), $\ell_1(x) = x + O(x^2)$ is the inverse function of the monotonic function $x = y \frac{\sqrt{1-y^2}}{1-2y^2}$ on $[0, 1/\sqrt{2})$.

In particular, for exact QEC codes, we have the following corollary:

Corollary 12. *Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. When $\varepsilon = 0$, i.e., when the code is exactly error-correcting, it holds that $\delta_{\text{group}} \geq \ell_1(\Delta H_L/2\sqrt{\mathfrak{F}})$, where \mathfrak{F} is given by Eq. (47).*

Proof of Theorem 11. By definition, there exists a $\mathcal{R}_{L \leftarrow S}^{\text{opt}(G)}$ such that $\gamma = \max_\theta P(\mathcal{R}_{L \leftarrow S}^{\text{opt}(G)} \circ \mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{U}_{L,\theta})$. Let $\mathcal{R}_{C \leftarrow SA} = \mathcal{R}_{C \leftarrow SA}^{\text{rep}} \circ (\mathcal{R}_{L \leftarrow S}^{\text{opt}(G)} \otimes \mathbb{I}_A)$ and $\mathcal{E}_{SA \leftarrow C} = (\mathcal{E}_{S \leftarrow L}^{\text{opt}(G)} \otimes \mathbb{I}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}$, we have $\mathcal{N}_{C,\theta} = \mathcal{R}_{C \leftarrow SA} \circ (\mathcal{N}_{S,\theta} \otimes \mathbb{I}_A) \circ$

$\mathcal{E}_{SA \leftarrow C} = \mathcal{D}_{C,\theta} \circ \mathcal{U}_{C,\theta}$, and $\mathcal{D}_{C,\theta}$ and $\mathcal{N}_{C,\theta}$ are rotated dephasing channels of the following forms:

$$\mathcal{D}_{C,\theta}(\cdot) = (1 - p_\theta) e^{-i \frac{\phi_\theta}{2} Z_C} (\cdot) e^{i \frac{\phi_\theta}{2} Z_C} + p_\theta Z_C e^{-i \frac{\phi_\theta}{2} Z_C} (\cdot) e^{i \frac{\phi_\theta}{2} Z_C} Z_C, \quad (93)$$

$$\mathcal{N}_{C,\theta}(\cdot) = (1 - p_\theta) e^{-i \frac{\phi_\theta + \Delta H_L \theta}{2} Z_C} (\cdot) e^{i \frac{\phi_\theta + \Delta H_L \theta}{2} Z_C} + p_\theta Z_C e^{-i \frac{\phi_\theta + \Delta H_L \theta}{2} Z_C} (\cdot) e^{i \frac{\phi_\theta + \Delta H_L \theta}{2} Z_C} Z_C. \quad (94)$$

where $p_\theta \in (0, 1)$ and $\phi_\theta \in [0, 2\pi)$. Let $\xi_\theta = \langle 0_C | \mathcal{N}_{C,\theta}(|0_C\rangle \langle 1_C|) | 1_C \rangle = (1 - 2p_\theta) e^{-i(\phi_\theta + \Delta H_L \theta)}$. The regularized channel QFI of rotated dephasing channels is

$$F^\infty(\mathcal{N}_{C,\theta}) = \frac{|\partial_\theta \xi_\theta|^2}{1 - |\xi_\theta|^2} = \frac{(1 - 2p_\theta)^2 (\partial_\theta \phi_\theta + \Delta H_L)^2}{4p_\theta(1 - p_\theta)} + \frac{(\partial_\theta p_\theta)^2}{4p_\theta(1 - p_\theta)}. \quad (95)$$

In order to get a lower bound on $F^\infty(\mathcal{N}_{C,\theta})$ as a function of γ . We note that the purified distance between $\mathcal{D}_{C,\theta}$ and $\mathbb{1}_C$ (see Sec. X) is upper bounded by γ :

$$P(\mathcal{D}_{C,\theta}, \mathbb{1}_C) = \sqrt{\frac{1 - (1 - 2p_\theta) \cos \phi_\theta}{2}} \leq \gamma, \quad (96)$$

because

$$\begin{aligned} P(\mathcal{D}_{C,\theta}, \mathbb{1}_C) &= P(\mathcal{R}_{C \leftarrow SA} \circ (\mathcal{N}_{S,\theta} \otimes \mathbb{1}_A) \circ \mathcal{E}_{SA \leftarrow C} \circ \mathcal{U}_{C,\theta}^\dagger, \mathbb{1}_C) \\ &= P(\mathcal{R}_{C \leftarrow SA} \circ (\mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^\dagger \otimes \mathbb{1}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}, \mathbb{1}_C) \\ &\leq P(\mathcal{R}_{L \leftarrow S}^{\text{opt(G)}} \circ \mathcal{N}_S \circ \mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^\dagger, \mathbb{1}_L) \leq \gamma, \end{aligned}$$

where we use the monotonicity of the purified distance and the definition of γ . Eq. (96) implies $\sqrt{p_\theta} \leq P(\mathcal{D}_{C,\theta}, \mathbb{1}_C) \leq \gamma$ for all $\theta \in \mathbb{R}$. Since $U_{L,\theta}$ and $U_{S,\theta}$ are periodic with a common period τ , we must have $\phi_\theta = \phi_{\theta+\tau}$. Therefore, there must exists a θ_* such that $\partial_\theta \phi_\theta|_{\theta=\theta_*} = 0$. Then using Eq. (95) and the monotonicity of the regularized QFI, we see that

$$F^\infty(\mathcal{N}_{S,\theta}) = F^\infty(\mathcal{N}_{S,\theta_*}) \geq F^\infty(\mathcal{N}_{C,\theta_*}) \geq \frac{(1 - 2p_{\theta_*})^2 (\Delta H_L)^2}{4p_{\theta_*}(1 - p_{\theta_*})} \geq \frac{(1 - 2\gamma^2)^2 (\Delta H_L)^2}{4\gamma^2(1 - \gamma^2)}. \quad (97)$$

Theorem 11 then follows from Proposition 10. □

Note that Theorem 11 coincides with Theorem 1 in Ref. [22] in the special case where $\delta_{\text{group}} = 0$.

2. Quantum resource theory method

Now we present another derivation based on quantum resource theory, which allows us to derive not only a lower bound on the worst-case gate implementation error, but also on the Choi gate implementation error.

We work with a resource theory of coherence [42] where the free (incoherent) states are those whose density operators commute with the Hamiltonian and the free (covariant) operations are those that commute with the Hamiltonian evolution, e.g., a covariant operation $\mathcal{C}_{L \leftarrow S}$ from S to L satisfies $\mathcal{C}_{L \leftarrow S} \circ \mathcal{U}_{S,\theta} = \mathcal{U}_{L,\theta} \circ \mathcal{C}_{L \leftarrow S}$ for all $\theta \in \mathbb{R}$. Assuming that the recovery operations $\mathcal{R}_{L \leftarrow S}$ and the noise channel \mathcal{N}_S are covariant, we can formulate the covariant QEC as a resource conversion task from noisy physical states to error-corrected logical states and the noise rate of the latter is upper bounded by γ , illustrated by the following lemma:

Proposition 13. *Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose \mathcal{N}_S commutes with $\mathcal{U}_{S,\theta}$. Then the QEC inaccuracy measures under the restriction that the recovery channel is covariant satisfy*

$$\varepsilon_{\text{cov}} = \min_{\mathcal{R}_{L \leftarrow S}^{\text{cov}}} P(\mathcal{R}_{L \leftarrow S}^{\text{cov}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L) \leq \gamma, \quad (98)$$

$$\bar{\varepsilon}_{\text{cov}} = \min_{\mathcal{R}_{L \leftarrow S}^{\text{cov}}} \bar{P}(\mathcal{R}_{L \leftarrow S}^{\text{cov}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L) \leq \bar{\gamma}, \quad (99)$$

where $\mathcal{R}_{L \leftarrow S}^{\text{cov}}$ is a recovery channel satisfying $\mathcal{U}_{L,\theta} \circ \mathcal{R}_{L \leftarrow S}^{\text{cov}} = \mathcal{R}_{L \leftarrow S}^{\text{cov}} \circ \mathcal{U}_{S,\theta}$.

Proof. Let $\mathcal{R}_{L \leftarrow S}^{\text{opt}(G)}$ be a recovery channel such that $\gamma = \max_{\theta} P(\mathcal{R}_{L \leftarrow S}^{\text{opt}(G)} \circ \mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathbb{U}_{L,\theta})$. Suppose $U_{S,\theta}$ and $U_{L,\theta}$ share a common period τ . Consider the following recovery channel:

$$\mathcal{R}_{L \leftarrow S}^{\text{cov}(G)} = \frac{1}{\tau} \int_0^{\tau} d\theta \mathcal{U}_{L,\theta}^{\dagger} \circ \mathcal{R}_{L \leftarrow S}^{\text{opt}(G)} \circ \mathcal{U}_{S,\theta}. \quad (100)$$

It can be verified that $\mathcal{R}_{L \leftarrow S}^{\text{cov}(G)}$ must be covariant and

$$\begin{aligned} P(\mathcal{R}_{L \leftarrow S}^{\text{cov}(G)} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L) &= P\left(\frac{1}{\tau} \int_0^{\tau} d\theta \mathcal{U}_{L,\theta}^{\dagger} \circ \mathcal{R}_{L \leftarrow S}^{\text{opt}(G)} \circ \mathcal{N}_S \circ \mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L\right) \\ &\leq \max_{\theta} P\left(\mathcal{U}_{L,\theta}^{\dagger} \circ \mathcal{R}_{L \leftarrow S}^{\text{opt}(G)} \circ \mathcal{N}_S \circ \mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L\right) = \gamma, \end{aligned}$$

where we used the concavity of $f^2(\Phi, \mathbb{1})$ with respect to Φ [77], leading to Eq. (98).

Similarly, let $\mathcal{R}_{L \leftarrow S}^{\text{opt}(C)}$ be a recovery channel such that $\bar{\gamma} = \max_{\theta} \bar{P}(\mathcal{R}_{L \leftarrow S}^{\text{opt}(C)} \circ \mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathbb{U}_{L,\theta})$. We can define $\mathcal{R}_{L \leftarrow S}^{\text{cov}(C)} = \frac{1}{\tau} \int_0^{\tau} d\theta \mathcal{U}_{L,\theta}^{\dagger} \circ \mathcal{R}_{L \leftarrow S}^{\text{opt}(C)} \circ \mathcal{U}_{S,\theta}$ and verify that $\bar{P}(\mathcal{R}_{L \leftarrow S}^{\text{cov}(C)} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L) \leq \bar{\gamma}$, leading to Eq. (99). \square

In order to derive a concrete lower bound on γ and $\bar{\gamma}$ using Proposition 13, we choose a resource monotone based on another type of QFI of quantum states called the RLD QFI [70] defined by $F^{\mathcal{R}}(\rho_{\theta}) = \text{Tr}((\partial_{\theta} \rho_{\theta})^2 \rho_{\theta}^{-1})$ when the support of $\partial_{\theta} \rho_{\theta}$ is contained in ρ_{θ} and $= +\infty$ otherwise. The resource monotone satisfies

$$F^{\mathcal{R}}(\mathcal{C}_{L \leftarrow S}(\rho_S), H_L) \leq F^{\mathcal{R}}(\rho_S, H_S), \quad (101)$$

for all ρ_S and covariant operations $\mathcal{C}_{L \leftarrow S}$, where

$$F^{\mathcal{R}}(\rho, H) := F^{\mathcal{R}}(e^{-iH\theta} \rho e^{iH\theta}) = \begin{cases} \text{Tr}(H \rho^2 H \rho^{-1}) - \text{Tr}(\rho H^2) & \text{supp}(H \rho H) \subseteq \text{supp}(\rho), \\ +\infty & \text{otherwise.} \end{cases} \quad (102)$$

Consider an error-corrected logical state $\mathcal{R}_{L \leftarrow S}^{\text{cov}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}(|\psi_L\rangle)$ using covariant recovery operations. On one hand, its RLD QFI is lower bounded by $\Theta(1/\gamma^2)$ when ρ is γ -close to a coherent pure state in terms of purified distance [22, 42]. On the other hand, its RLD QFI is upper bounded by the RLD QFI of the noisy physical state $F^{\mathcal{R}}(\mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}(|\psi_L\rangle), H_S)$ is no less than the channel RLD QFI $F^{\mathcal{R}}(\mathcal{N}_{S,\theta}) = \max_{\rho} F^{\mathcal{R}}((\mathcal{N}_{S,\theta} \otimes \mathbb{1})(\rho))$ [71, 72]. Specifically,

$$F^{\mathcal{R}}(\mathcal{N}_{S,\theta}) = \|\text{Tr}_S(\Psi_{SR}^{\mathcal{N}_{S,H_S}} (\Psi_{SR}^{\mathcal{N}_S})^{-1} \Psi_{SR}^{\mathcal{N}_{S,H_S}})\|, \quad (103)$$

where $\Gamma_{SR}^{\mathcal{N}_S} = (\mathcal{N}_S \otimes \mathbb{1}_R)(\Gamma_{SR})$, $\Gamma_{SR}^{\mathcal{N}_{S,H_S}} = (\mathcal{N}_S \otimes \mathbb{1}_R)((H_S \otimes \mathbb{1}_R)\Gamma_{SR}) - (\mathcal{N}_S \otimes \mathbb{1}_R)(\Gamma_{SR}(H_S \otimes \mathbb{1}_R))$, $\Gamma_{SR} = |\Gamma\rangle_{SR} \langle \Gamma|_{SR}$ and $|\Gamma\rangle_{SR} = \sum_i |i\rangle_S |i\rangle_R$.

We now state and prove Theorem 14 which provides lower bounds on γ and $\bar{\gamma}$ by considering different input logical states $|\psi_L\rangle$.

Theorem 14. Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose \mathcal{N}_S commutes with $\mathcal{U}_{S,\theta}$. Then it holds that

$$\varepsilon + \delta_{\text{group}} \geq \gamma \geq \ell_2 \left(\frac{\Delta H_L}{\sqrt{4F^{\mathcal{R}}(\mathcal{N}_{S,\theta})}} \right), \quad (104)$$

$$\bar{\varepsilon} + \bar{\delta}_{\text{group}} \geq \bar{\gamma} \geq \ell_3 \left(\sqrt{\frac{\frac{1}{d_L} \text{Tr}(H_L^2) - \frac{1}{d_L^2} \text{Tr}(H_L)^2}{F^{\mathcal{R}}(\mathcal{N}_{S,\theta})}} \right), \quad (105)$$

where $\ell_2(x) = x + O(x^2)$ is the inverse function of the monotonic increasing function $x = \frac{y}{\sqrt{(1-3y^2+y^4)(1-6\sqrt{2}y^2)}}$ for $y \in [0, 1/(6\sqrt{2})]$ and $\ell_3(x) = x + O(x^2)$ is the inverse function of the monotonic increasing function $x = \frac{y}{\sqrt{(1-3y^2+y^4)(1-(3(\Delta H_L)^2 y)/\sqrt{2}(\frac{1}{d_L} \text{Tr}(H_L^2) - \frac{1}{d_L^2} \text{Tr}(H_L)^2))}}$ for $y \in \left[0, \frac{\sqrt{2}(\frac{1}{d_L} \text{Tr}(H_L^2) - \frac{1}{d_L^2} \text{Tr}(H_L)^2)}{3(\Delta H_L)^2}\right]$.

In particular, when $\varepsilon = 0$, i.e., when the code is exactly error-correcting, we have the following corollary:

Corollary 15. Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose \mathcal{N}_S commutes with $\mathcal{U}_{S,\theta}$. When $\varepsilon = 0$, i.e., when the code is exactly error-correcting, it holds that $\delta_{\text{group}} \geq \ell_2(\Delta H_L / \sqrt{4F^{\mathcal{R}}(\mathcal{N}_{S,\theta})})$ and $\bar{\delta}_{\text{group}} \geq \ell_3(\Delta H_L / \sqrt{4F^{\mathcal{R}}(\mathcal{N}_{S,\theta})})$.

Proof of Theorem 14. Let $|+L\rangle = \frac{|0_L\rangle + |1_L\rangle}{\sqrt{2}}$. Then according to Proposition 13, there exists a covariant recovery channel $\mathcal{R}_{L \leftarrow S}^{\text{cov}}$ such that

$$1 - \langle +L | \rho_L | +L \rangle \leq \varepsilon_{\text{cov}}^2 \leq \gamma^2, \quad (106)$$

where $\rho_L = (\mathcal{R}_{L \leftarrow S}^{\text{cov}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L})(|+L\rangle \langle +L|)$. According to Lemma 3 in Ref. [40],

$$F^{\mathcal{R}}(\rho_L, H_L) \geq \frac{1 - 3\gamma^2 + \gamma^4}{\gamma^2} \cdot \left(\mathbb{V}_{H_L}(|+L\rangle) - \frac{3\sqrt{2}\gamma(\Delta H_L)^2}{2} \right), \quad (107)$$

where the variance $\mathbb{V}_{H_L}(|+L\rangle) = \langle +L | H_L^2 | +L \rangle - \langle +L | H_L | +L \rangle^2 = \frac{(\Delta H_L)^2}{4}$. $\gamma < 1/(6\sqrt{2})$ guarantees the right-hand side is positive. On the other hand, using Eq. (101),

$$F^{\mathcal{R}}(\rho_L, H_L) \leq F^{\mathcal{R}}(\rho_S, H_S) = F^{\mathcal{R}}(\mathcal{N}_{S,\theta}(\mathcal{E}_{S \leftarrow L}(|+L\rangle \langle +L|))) \leq F^{\mathcal{R}}(\mathcal{N}_{S,\theta}), \quad (108)$$

where $\rho_S = (\mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L})(|+L\rangle \langle +L|)$. Using Eq. (107) and Eq. (108), we have

$$\gamma^2 \cdot \frac{1}{(1 - 3\gamma^2 + \gamma^4)(1 - 6\sqrt{2}\gamma)} \geq \frac{(\Delta H_L)^2}{4F^{\mathcal{R}}(\mathcal{N}_{S,\theta})}, \quad (109)$$

proving Eq. (104).

Similarly, let $|\Psi_{LR}\rangle = \frac{1}{\sqrt{d_L}} \sum_{i=1}^{d_L} |i\rangle_L |i\rangle_R$. Then according to Proposition 13, there exists a covariant recovery channel $\mathcal{R}_{L \leftarrow S}^{\text{opt}(C)}$ such that

$$1 - \langle \Psi_{LR} | \rho_{LR} | \Psi_{LR} \rangle \leq \bar{\varepsilon}_{\text{cov}} \leq \bar{\gamma}, \quad (110)$$

where $\rho_{LR} = (\mathcal{R}_{L \leftarrow S}^{\text{cov}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L})(|\Psi_{LR}\rangle \langle \Psi_{LR}|)$. According to Lemma 3 in Ref. [40],

$$\begin{aligned} F^{\mathcal{R}}(\rho_{LR}, H_L \otimes \mathbb{1}_R) &\geq \frac{1 - 3\bar{\gamma}^2 + \bar{\gamma}^4}{\bar{\gamma}^2} \left(\mathbb{V}_{H_L \otimes \mathbb{1}_R}(|\Psi_{LR}\rangle) - \frac{3\sqrt{2}\bar{\gamma}(\Delta H_L)^2}{2} \right) \\ &\geq \frac{1 - 3\bar{\gamma}^2 + \bar{\gamma}^4}{\bar{\gamma}^2} \left(\frac{\text{Tr}(H_L^2)}{d_L} - \frac{\text{Tr}(H_L)^2}{d_L^2} - \frac{3\sqrt{2}\bar{\gamma}(\Delta H_L)^2}{2} \right). \end{aligned} \quad (111)$$

The rest of the proof is exactly the same as in the proof of the lower bound on the worst-case gate implementation error. \square

In fact, the proof of Theorem 14 follows almost exactly from the proof of Theorem 2 in Ref. [22] and our new contribution here is Proposition 13.

To compare Theorem 14 to Theorem 11, we first note that $F^{\mathcal{R}}(\Phi_\theta) \geq F^\infty(\Phi_\theta)$ for any Φ_θ because $F^{\mathcal{R}}(\rho_\theta) \geq F(\rho_\theta)$ [78] for any ρ_θ and $F^{\mathcal{R}}(\Phi_\theta) = \lim_{N \rightarrow \infty} F^{\mathcal{R}}(\Phi_\theta^{\otimes N})/N$ [71, 72]. Moreover, Theorem 14 requires the commutativity between the noise and the Hamiltonian and $\ell_1(x) \geq \ell_2(x)$, so Theorem 14 provides a weaker bound on the (worst-case) gate implementation error than Theorem 11. Note that $F^{\mathcal{R}}(\mathcal{N}_{S,\theta}) < +\infty$ only when $\text{span}\{K_{S,i}H_S, \forall i\} \subseteq \text{span}\{K_{S,i}, \forall i\}$ which is also a stronger condition than the HKS condition. The resource theory method leads to a bound on the Choi gate implementation error, however, which is not available using the quantum metrology method.

Also note recent works Refs. [79, 80] (results implied by Ref. [25]) which considered the coherence cost of implementing unitary gates based on relevant insights.

C. Explicit behaviors of the bounds and comparison with the charge fluctuation approach

We first make a general comparison between the trade-off relations derived using the gate implementation error approach (Theorem 11 and Theorem 14) and the charge fluctuation approach (Theorem 6) in Sec. III. Two clear advantages of the gate implementation error approach are that 1) it applies to general quantum codes (e.g., the

non-isometric encodings in [13, 23]) while the charge fluctuation approach only holds for isometric codes; 2) it leads to a trade-off relation for the Choi measures. Additionally, for the special case of $\delta_{\text{group}} = 0$, the results based on the gate implementation error approach directly reduce to the previous known result for exactly covariant codes in Ref. [22], while there is still some discrepancy with previous results using the charge fluctuation approach (see more discussion in Sec. III C). For the special case of $\varepsilon = 0$ which was not previously studied, we have two lower bounds on δ from Corollary 12 and Corollary 7 which behave as follows:

$$\delta_{\text{group}} \gtrsim \frac{\Delta H_L}{2\sqrt{\mathfrak{F}}}, \quad \delta_{\text{group}} \gtrsim \sqrt{\frac{\Delta H_L}{\Delta H_S}}. \quad (112)$$

It is interesting to observe that the first bound depends on the noise channel while the second one does not (as long as the HKS condition is satisfied).

We now remark on the explicit scalings of our bounds for different noise models as in Sec. III D. Again, consider a n -partite system and a local physical Hamiltonian with $\Delta H_S = O(n)$. For random local noise which acts uniformly randomly on each subsystem, $\mathfrak{F} = O(n^2)$ and the two bounds give $\delta_{\text{group}} = \Omega(1/n)$ and $\delta_{\text{group}} = \Omega(1/\sqrt{n})$, respectively. That is, the charge fluctuation approach outperforms the gate implementation error approach in this case. For noise acting independently on each subsystems, we have $\mathfrak{F} = O(n)$ which gives $\delta_{\text{group}} = \Omega(1/\sqrt{n})$ using the gate implementation error approach. In this situation, the bounds based on the two approaches are comparable. Note that in situations where $\mathfrak{F} = o(n)$, i.e., the noise is even stronger than independent noise so that the regularized QFI is sublinear, the bound based the gate implementation error approach should outperform the bound based on the charge fluctuation approach. In general situations where both the QEC inaccuracy and the global covariance violation are non-vanishing, we expect a similar behavior, i.e., the gate implementation error approach performs better in the extremely strong noise regime, while the charge fluctuation approach performs better in weaker noise regimes.

V. SUPPLEMENTARY NOTE 5: LIMITATIONS ON TRANSVERSAL LOGICAL GATES

Note that a key implication of our results is symmetry constraints on QEC codes that achieve a given accuracy, which extends the scope of previous knowledge on the incompatibility between symmetries and QEC to general codes, especially exact QEC codes which are most commonly studied. As we shall discuss in this section, such constraints actually allow us to derive restrictions on the transversally implementable gates for general QEC codes, advancing our understanding of fault tolerance. A key intuition is that the precision of gate implementation is associated with the degree of symmetry. Recall that there are no QEC codes which admit a continuous symmetry acting transversally on physical qubits and thus there are no transversal universal gate sets, according to the Eastin–Knill theorem. For stabilizer codes, the incompatibility between QEC and symmetry are reflected in the classification of transversally logical gates in finite levels of the Clifford hierarchy [45–48]. Here we present new restrictions of transversal gates for arbitrary QEC codes from the perspective of global covariance violations.

The following corollary puts a restriction on the logical transversal gates using Corollary 7. Namely, the non-trivial logical gates cannot be too close to the identity operators when implemented by transversal physical gates in the vicinity of identity operators because δ_{group} has a lower bound of $\Theta(1/\sqrt{n})$. Note that here we implicitly consider exact QEC codes under single-erasure noise (so that the HKS condition is satisfied for 1-local Hamiltonians) and in this case Corollary 7 outperforms Corollary 12, so we will only use Corollary 7 in this section.

Corollary 16. *Suppose an n -qudit QEC code with distance at least 2 admits a transversal implementation $V_S = \bigotimes_{i=1}^n e^{-i2\pi T_{S_i}/D}$ of the logical gate $V_L = e^{-i2\pi T_L/D}$ where D is a positive integer and $T_{L,S}$ have integer eigenvalues. Then it holds that*

$$D \leq \max \left\{ 4\pi \sqrt{\frac{2}{3}} \left(\Delta T_L + \sum_{i=1}^n \Delta T_{S_i} \right), 2\sqrt{2}\pi \sqrt{\frac{\sum_{i=1}^n \Delta T_{S_i}}{\Delta T_L}} \left(\Delta T_L + \sum_{i=1}^n \Delta T_{S_i} \right) \right\}. \quad (113)$$

In particular, $D = O(\text{poly}(n))$ when $\Delta T_S = O(\text{poly}(n))$, $\Delta T_L = O(\text{poly}(n))$ and $\Delta T_S/\Delta T_L = O(\text{poly}(n))$, where $\Delta T_S = \sum_{\ell=1}^n \Delta T_{S_\ell}$.

Proof. Any codes with distance at least 2 can correct single-erasure noise. Let $H_L = T_L$ and $H_S = \sum_i T_{S_i}$. They have integer eigenvalues implies that $\mathcal{U}_{S,\theta}$ and $\mathcal{U}_{L,\theta}$ share a common period 2π . According to Corollary 7, the code

⁴ The conditions are satisfied in common settings; see, e.g., the proof of Corollary 17.

must satisfy

$$\delta_{\text{group}} \geq \min \left\{ \frac{\sqrt{\Delta H_L (\Delta H_S - \frac{1}{2} \Delta H_L)}}{\Delta H_S}, \sqrt{\frac{3}{8}} \right\}. \quad (114)$$

We can always write $\theta = \frac{2j\pi}{D} + \theta_1$ for some $j \in \mathbb{N}$ and $\theta_1 \in [0, 2\pi/D)$. Then we have

$$\begin{aligned} P(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^\dagger, \mathcal{E}_{S \leftarrow L}) &= P(\mathcal{U}_{S,\theta_1} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta_1}^\dagger, \mathcal{E}_{S \leftarrow L}) \\ &\leq P(\mathcal{U}_{S,\theta_1} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta_1}^\dagger, \mathcal{U}_{S,\theta_1} \circ \mathcal{E}_{S \leftarrow L}) + P(\mathcal{U}_{S,\theta_1} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{E}_{S \leftarrow L}) \\ &\leq P(\mathcal{U}_{L,\theta_1}^\dagger, \mathbb{1}_L) + P(\mathcal{U}_{S,\theta_1}, \mathbb{1}_S), \end{aligned} \quad (115)$$

where we use the monotonicity and the triangular inequality of the purified distance. Without loss of generality, assume $H_L = \sum_{i=1}^{d_L} \lambda_i |i\rangle \langle i|$ and let $|\psi\rangle = \sum_{i=1}^{d_L} \psi_i |i\rangle |i\rangle$. Consider first the situation where $D > 2 \max\{\Delta H_L, \Delta H_S\}$, then we have $\theta_1 \max\{\Delta H_L, \Delta H_S\} < \pi$ and

$$\begin{aligned} P(\mathcal{U}_{L,\theta_1}^\dagger, \mathbb{1}_L) &= \max_{\{\psi_i\}} P((\mathcal{U}_{L,\theta_1}^\dagger \otimes \mathbb{1})(|\psi\rangle \langle \psi|), |\psi\rangle \langle \psi|) \\ &= \max_{\{\psi_i\}} \sqrt{1 - \left| \sum_i |\psi_i|^2 e^{i\theta_1 \lambda_i} \right|^2} = \max_{i,i'} \left| \sin \left(\frac{1}{2} \theta_1 (\lambda_i - \lambda_{i'}) \right) \right| \leq \theta_1 \Delta H_L. \end{aligned} \quad (116)$$

Similarly, $P(\mathcal{U}_{S,\theta_1}, \mathbb{1}_S) \leq \theta_1 \Delta H_S = \theta_1 \sum_{l=1}^n \Delta H_{S_l}$. Since $\delta_{\text{group}} = \max_{\theta} P(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^\dagger, \mathcal{E}_{S \leftarrow L}) = \max_{\theta_1} P(\mathcal{U}_{S,\theta_1} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta_1}^\dagger, \mathcal{E}_{S \leftarrow L})$, we obtain

$$\delta_{\text{group}} \leq \frac{2\pi}{D} \left(\Delta H_L + \sum_{l=1}^n \Delta H_{S_l} \right), \quad (117)$$

Otherwise,

$$D \leq 2 \max\{\Delta H_L, \Delta H_S\}. \quad (118)$$

The result then follows by combining Eq. (114), Eq. (117) and Eq. (118). \square

Corollary 16 shows that the precision of transversal logical gates under certain restrictions only increases polynomially in the number of qubits. For the important case of stabilizer codes, this implies that the levels of the Clifford hierarchy that can be reached only increase polynomially in the number of qubits. Specifically, consider an n -qubit stabilizer code with distance at least 2. The following corollary describes the limitation on the transversally implementable logical gates for stabilizer codes:

Corollary 17. *Let $\tilde{V}_S = V_1 (\bigotimes_{l=1}^n e^{-i2\pi a_l Z_l/D}) V_2$ be a transversal logical gate for an n -qubit stabilizer code with distance at least 2, where D is a power of two and a_l is an integer and $V_{1,2}$ are transversal Clifford operators. (This describes the most general form of transversal logical gates for stabilizer codes [44, 47]). When $a_l = O(\text{poly}(n))$, we must have $D = O(\text{poly}(n))$ and \tilde{V}_S implements a logical gate \tilde{V}_L in the $O(\log n)$ -th level of the Clifford hierarchy.*

Proof. Let Π be the projection onto the stabilizer code under consideration and $Q_S = \bigotimes_{l=1}^n e^{-i2\pi a_l Z_l/D}$. Then

$$\tilde{V}_S \Pi = \Pi \tilde{V}_S \Pi, \Rightarrow Q_S \Pi_2 = \Pi_1 Q_S \Pi_2, \quad (119)$$

where $\Pi_1 = V_1^\dagger \Pi V_1$ and $\Pi_2 = V_2 \Pi V_2^\dagger$. Both Π_1 and Π_2 are stabilizer codes with the same code distance as Π . Without loss of generality, we assume Π , Π_1 and Π_2 are two-dimensional stabilizer codes (by considering subcodes of the original codes).

As proven in Proposition 4 in Ref. [47], $V_S = (Q_S)^4$ must be a logical gate on Π_2 , satisfying

$$(Q_S)^4 \Pi_2 = \Pi_2 (Q_S)^4 \Pi_2, \quad (120)$$

and the logical gate V_L has the form $V_L = e^{-i2\pi a Z_L/D}$ where a is an integer. First consider the situation where $a = 0$ (for any choice of two-dimensional codes), i.e., $(Q_S)^4 \Pi_2 = \Pi_2$. By writing down the stabilizer code Π_2 in its

computational basis, it is easy to observe that either $Q_S \Pi_2 = \Pi_2$, then \tilde{V}_S implements a Clifford logical gate and the Corollary holds, or $Q_S \Pi_2 \neq \Pi_2$, then $\sum_{l=1}^n a_l/D$ must be a positive constant and $D = O(\sum_{l=1}^n a_l) = O(\text{poly}(n))$.

Now we consider the situation where $a \neq 0$. By writing down the stabilizer code Π_2 in its computational basis, we observe $1 \leq a \leq \sum_{l=1}^n 4a_l = O(\text{poly}(n))$. Let $H_L = aZ_L$ and $H_{S_l} = 4a_l Z_l$. Then we must have

$$\frac{\sqrt{\Delta H_L (\Delta H_S - \frac{1}{2} \Delta H_L)}}{\Delta H_S} \geq \frac{1}{O(\text{poly}(n))}. \quad (121)$$

Using Corollary 16, we have $D = O(\text{poly}(n))$. Since D is a power of 2, for all l , $e^{-i2\pi a_l Z_l/D}$ (see Proposition 1 in Ref. [47]) and thus $Q_S = \bigotimes_{l=1}^n e^{-i2\pi a_l Z_l/D}$ must be in the $(\log D)$ -th level of the Clifford hierarchy. Corollary 17 then follows from the fact that Clifford operators V_1, V_2 preserve the level of the Clifford hierarchy and any physical gate in the j -th level of the Clifford hierarchy implements a logical gate in the j -th level of the Clifford hierarchy (because logical Pauli operators can be implemented by physical Pauli operators for stabilizer codes). \square

Corollary 17 provides a simple proof on the limitations of transversal logical gates for stabilizer codes from the perspective of continuous symmetries. Note that the relevant results previously known for stabilizer codes [45, 46, 48] were obtained using very different techniques.

VI. SUPPLEMENTARY NOTE 6: TRADE-OFF BETWEEN QEC AND LOCAL SYMMETRY MEASURES

In this section, we study relations between QEC and local symmetry measures, that is, the local covariance violation and the charge conservation violation. We will first prove a lemma which links the charge conservation to the charge fluctuation and then derive trade-off relations using Proposition 3 and Proposition 5. We will also derive a lower bound on the local covariance violation using the quantum metrology method.

Note that the results in this section (Theorem 19, Theorem 21 and Theorem 23) hold true for arbitrary Hermitian operators H_L and H_S , which do not necessarily share a common period as generators of $U(1)$ representations.

A. Bounds via charge fluctuation

We first observe a simple connection between the charge fluctuation χ and the charge conservation violation δ_{charge} :

Lemma 18. *Consider a quantum code $\mathcal{E}_{S \leftarrow L}$, a physical Hamiltonian H_S and a logical Hamiltonian H_L . Then*

$$|\chi| \geq \Delta H_L - \delta_{\text{charge}}. \quad (122)$$

Proof. By definition, $|\chi| = |\langle 0_L | (\mathcal{E}_{S \leftarrow L})^\dagger (H_S) | 0_L \rangle - \langle 1_L | (\mathcal{E}_{S \leftarrow L})^\dagger (H_S) | 1_L \rangle|$. Then we must have $|\chi| \geq \Delta H_L - \delta_{\text{charge}}$, because $\delta_{\text{charge}} \geq |\langle 0_L | ((\mathcal{E}_{S \leftarrow L})^\dagger (H_S) - H_L) | 0_L \rangle - \langle 1_L | ((\mathcal{E}_{S \leftarrow L})^\dagger (H_S) - H_L) | 1_L \rangle|$, and $\Delta H_L = |\langle 0_L | H_L | 0_L \rangle - \langle 1_L | H_L | 1_L \rangle|$. \square

Using the KL-based method (Proposition 3) and Proposition 1, we immediately have the following trade-off relations:

Theorem 19. *Consider an isometric quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. It holds that*

$$\delta_{\text{local}} + 2\varepsilon \mathfrak{J} \geq \Delta H_L, \quad (123)$$

$$\delta_{\text{charge}} + 2\varepsilon \mathfrak{J} \geq \Delta H_L, \quad (124)$$

where \mathfrak{J} is given by Eq. (39).

Note that Eq. (124) reduces to Corollary 3 in Ref. [12] for random local erasure noise.

In particular, for exact QEC codes, we have the following corollary:

Corollary 20. *Consider an isometric quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. When $\varepsilon = 0$, i.e., when the code is exactly error-correcting, we must have $\delta_{\text{local}} \geq \Delta H_L$ and $\delta_{\text{charge}} \geq \Delta H_L$.*

Similarly, using the quantum metrology method (Proposition 5) and Proposition 1, we have the following trade-off relations:

Theorem 21. *Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. It holds that*

$$\delta_{\text{charge}} + 2\varepsilon \left(\sqrt{(1 - \varepsilon^2)\mathfrak{F}} + \mathfrak{B} \right) \geq \Delta H_L. \quad (125)$$

In particular, when $\varepsilon \ll 1$ and $\mathfrak{B} \ll \sqrt{\mathfrak{F}}$, we have

$$\delta_{\text{charge}} + 2\varepsilon \sqrt{\mathfrak{F}} \gtrsim \Delta H_L, \quad (126)$$

where \mathfrak{F} and \mathfrak{B} are given by Eq. (47) and Eq. (46), respectively. Furthermore, when the code is isometric, it holds that

$$\delta_{\text{local}} + 2\varepsilon \left(\sqrt{(1 - \varepsilon^2)\mathfrak{F}} + \mathfrak{B} \right) \geq \Delta H_L. \quad (127)$$

Corollary 22. *Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. When $\varepsilon = 0$, i.e., when the code is exactly error-correcting, we must have $\delta_{\text{charge}} \geq \Delta H_L$.*

Note that Corollary 22 is slightly more general than Corollary 20 as the former covers the situation where the encoding is non-isometric.

B. Bounding local covariance violation using quantum metrology

The trade-off relation between ε and δ_{local} in Theorem 21 requires the code to be isometric. In fact, we can show a cleaner version of the trade-off between the QEC inaccuracy and the local covariance violation using the quantum metrology method which does not contain \mathfrak{B} and also covers the non-isometric scenario, as shown below.

Theorem 23. *Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. It holds that*

$$\delta_{\text{local}} + 2\varepsilon \left(\sqrt{(1 - \varepsilon^2)\mathfrak{F}} + \varepsilon \Delta H_L \right) \geq \Delta H_L. \quad (128)$$

When $\varepsilon \ll 1$, we have

$$\delta_{\text{local}} + 2\varepsilon \sqrt{\mathfrak{F}} \gtrsim \Delta H_L, \quad (129)$$

where \mathfrak{F} and \mathfrak{B} are given by Eq. (47) and Eq. (46), respectively.

Corollary 24. *Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. When $\varepsilon = 0$, i.e., when the code is exactly error-correcting, it holds that $\delta_{\text{local}} \geq \Delta H_L$.*

Proof. Let $\mathcal{R}_{L \leftarrow S}^{\text{opt}}$ be the recovery channel such that $\varepsilon = P(\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L)$. Let (see Fig. 2)

$$\begin{aligned} \mathcal{N}_{C, \theta} &= \mathcal{R}_{C \leftarrow SA}^{\text{rep}} \circ (\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \otimes \mathbb{1}_A) \circ \\ &(\mathcal{U}_{S, \theta} \circ \mathcal{E}_{S \leftarrow L} \otimes \mathbb{1}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}. \end{aligned} \quad (130)$$

Consider the parameter estimation of θ in the neighborhood of $\theta = 0$ and let $\xi_\theta = \langle 0_C | \mathcal{N}_{C, \theta} (|0_C\rangle \langle 1_C|) | 1_C \rangle$. Following the proof of Proposition 5, we have from Eq. (63) that

$$|\xi_{\theta=0}| \geq 1 - 2\varepsilon^2. \quad (131)$$

As shown in Sec. XI, we also have

$$|\partial_\theta \xi_\theta|^2 \big|_{\theta=0} \geq ((1 - 2\varepsilon^2)\Delta H_L - \delta_{\text{local}})^2, \quad (132)$$

when $(1 - 2\varepsilon^2)\Delta H_L \geq \delta_{\text{local}}$. Hence, when $(1 - 2\varepsilon^2)\Delta H_L \geq \delta_{\text{local}}$, we must have

$$F^\infty(\mathcal{N}_{S, \theta}) = F^\infty(\mathcal{N}_{S, \theta}) \big|_{\theta=0} \geq F^\infty(\mathcal{N}_{C, \theta}) \big|_{\theta=0} \geq \frac{((1 - 2\varepsilon^2)\Delta H_L - \delta_{\text{local}})^2}{4\varepsilon^2(1 - \varepsilon^2)}, \quad (133)$$

completing the proof. \square

C. Remarks on the behaviors of the bounds

From [Theorem 19](#), [Theorem 21](#) and [Theorem 23](#), we observe that in n -partite systems, the local covariance violation δ_{local} and the charge conservation violation δ_{charge} are usually lower bounded by constants for small ε which does not vanish as $n \rightarrow \infty$ like the global covariance violation δ_{group} . However, also note that δ_{local} and δ_{charge} may naturally be superconstant (for example, for the trivial encoding $\mathcal{E}_{S \leftarrow L} = 1$ we usually have $\delta_{\text{local}} = \delta_{\text{charge}} = \Delta(H_S - H_L) = \Theta(n)$), indicating that the constant or even sublinear scaling of δ_{local} and δ_{charge} requires non-trivial code structures.

Also note that the bounds on δ_{local} in both [Theorem 19](#) and [Theorem 21](#) rely on the fact that $\delta_{\text{local}} \geq \delta_{\text{charge}}$, indicating that these bounds may not be tight when there is a gap between δ_{local} and a function of δ_{charge} . Such a gap does exist as we shown later in examples (see [Sec. VII](#)) and we provide a possible explanation of the existence of the gap in [Sec. XVI](#).

VII. SUPPLEMENTARY NOTE 7: CASE STUDIES OF EXPLICIT CODES

In the above, we have derived several forms of fundamental limits on the QEC accuracy and degree of symmetry or charge conservation that a quantum code can possibly admit. Then a natural question is to what extent these limits can be attained by certain codes. Furthermore, explicit constructions of approximately covariant codes would be important for our understanding of the QEC-symmetry trade-off and may find broad applications. In this section, we introduce and analyze two code examples with interesting approximate covariance features to address these needs. In the first example, we generalize a covariant code called the thermodynamic code [[9](#), [12](#)] to a class of general quantum codes which exhibits a full trade-off between symmetry and QEC via a smooth transition from exact covariance to exact QEC. The second one involves a well-known QEC code called the quantum Reed–Muller codes [[81](#), [82](#)], which can be seen as a prominent example of approximately covariant exact QEC codes. In particular, we explicitly compute their QEC and symmetry measures, and compare them to the fundamental limits. Remarkably, the scalings of the global covariance violation and the charge conservation violation for both examples match well with the optimal scalings from our bounds.

A. Modified thermodynamic codes

Thermodynamic codes [[9](#), [12](#)] are n -qubit quantum codes given by certain Dicke states with different magnetic charges which become approximately quantum error-correcting for large n . Specifically, a two-dimensional thermodynamic code have codewords

$$|c_0\rangle = |m_n\rangle, \quad |c_1\rangle = |(-m)_n\rangle, \quad (134)$$

where $|m_n\rangle$ for $m \in [-n, n]$ is the Dicke state defined by

$$|m_n\rangle = \frac{1}{\sqrt{\binom{n}{\frac{n+m}{2}}}} \sum_{\substack{\mathbf{j} \in \{0,1\}^n: \\ \sum_i j_i = \frac{n+m}{2}}} |\mathbf{j}\rangle, \quad (135)$$

satisfying $\sum_{l=1}^n Z_l |m_n\rangle = -m |m_n\rangle$. Note that $m+n$ must be an even number and we assume $2 \leq m \ll n$. It is easy to verify that the thermodynamic code is exactly covariant with respect to $H_L = \frac{m}{2} Z_L$ and $H_S = -\frac{1}{2} \sum_{l=1}^n Z_l$ and it was proven that for single-erasure noise [[40](#)] $\varepsilon = m/2n + O(m^2/n^2)$ which is infinitely small when $m/n \rightarrow 0$. Here we extend the thermodynamic code in such a way that it smoothly transitions from an exactly covariant code to an exact QEC code as tuned by a continuous parameter $0 \leq q \leq 1$. Specifically, our modified thermodynamic code is defined by

$$|c_0^q\rangle = \sqrt{\frac{n}{n+qm}} |m_n\rangle + \sqrt{\frac{qm}{n+qm}} |(-n)_n\rangle, \quad (136)$$

$$|c_1^q\rangle = \sqrt{\frac{n}{n+qm}} |(-m)_n\rangle + \sqrt{\frac{qm}{n+qm}} |n_n\rangle. \quad (137)$$

In particular, when $q = 0$, we have the original thermodynamic code, and when $q = 1$, we obtain an modified code which is exactly error-correcting under single-erasure noise. We shall compute the QEC inaccuracy and the different covariance violation measures, and compare them with our trade-off bounds.

1. QEC inaccuracy

Here we compute the QEC inaccuracy of modified thermodynamic codes $\varepsilon(\mathcal{N}_S, \mathcal{E}_{S \leftarrow L})$ where $\mathcal{E}_{S \leftarrow L}(\cdot) = W(\cdot)W^\dagger$ with $W = |\mathfrak{c}_0^q\rangle\langle 0_L| + |\mathfrak{c}_1^q\rangle\langle 1_L|$ and \mathcal{N}_S is the single-erasure noise channel $\mathcal{N}_S = \sum_{l=1}^n \frac{1}{n} \mathcal{N}_{S_l}$, where $\mathcal{N}_{S_l}(\cdot) = |\emptyset\rangle\langle\emptyset|_{S_l} \otimes \text{Tr}_{S_l}(\cdot)$.

We need to use the following lemma which compute the purified distance between error-corrected channels, employing the formalism of complementary channels [38]. Let $\hat{\Phi}_{B \leftarrow A}(\cdot) = \text{Tr}_A(V_{AB \leftarrow A}(\cdot)V_{AB \leftarrow A}^\dagger)$ be the complementary channel of channel $\Phi_A(\cdot) = \text{Tr}_B(V_{AB \leftarrow A}(\cdot)V_{AB \leftarrow A}^\dagger)$, where $V_{AB \leftarrow A}$ is a Stinespring dilation of Φ_A . Then we have

Lemma 25 ([38]).

$$\min_{\mathcal{R}} P(\mathcal{R} \circ \Phi_1, \Phi_2) = \min_{\mathcal{R}'} P(\hat{\Phi}_1, \mathcal{R}' \circ \hat{\Phi}_2), \quad (138)$$

for arbitrary $\Phi_{1,2}$, where the minimizations are taken over all channels with the appropriate input and output spaces.

Choosing $\Phi_1 = \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}$ and $\Phi_2 = \mathbb{1}_L$, we have

$$\varepsilon = \min_{\mathcal{R}'_{B \leftarrow L}} P(\hat{\mathcal{N}}_{B \leftarrow S} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{R}'_{B \leftarrow L} \circ \hat{\mathbb{1}}_L), \quad (139)$$

As detailed in Sec. XV, we have that

$$\varepsilon = \frac{(1-q)m}{2n} + O\left(\frac{m^2}{n^2}\right), \quad (140)$$

and furthermore, explicitly construct a recovery channel $\mathcal{R}_{L \leftarrow S}^{\text{opt}}$ which achieves the optimal QEC inaccuracy up to the lowest order of m/n :

$$\tilde{\varepsilon} = P(\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L) = \sqrt{\frac{1}{2} - \sqrt{\frac{(n+m)(n+(2q-1)m)}{4(n+qm)^2}}} \approx \frac{(1-q)m}{2n}. \quad (141)$$

2. Symmetry violation measures

We now compute all the approximate symmetry measures associated with our modified thermodynamic codes. Note that we let $H_L = \frac{m}{2} Z_L$ and $H_S = -\frac{1}{2} \sum_{l=1}^n Z_l$, which guarantees that the code tends to be covariant as $n \rightarrow \infty$.

We first compute δ_{group} and δ_{local} . Let $|\psi\rangle = |0_L\rangle|\psi_R^0\rangle + |1_L\rangle|\psi_R^1\rangle$ be an arbitrary pure state on $L \otimes R$. Then

$$\begin{aligned} |\psi_\theta\rangle &:= U_{S,\theta} W U_{L,\theta}^\dagger |\psi\rangle \\ &= \left(\sqrt{\frac{n}{n+qm}} |m_n\rangle + e^{\frac{i(m+n)\theta}{2}} \sqrt{\frac{qm}{n+qm}} |(-n)_n\rangle \right) |\psi_R^0\rangle + \left(\sqrt{\frac{n}{n+qm}} |(-m)_n\rangle + e^{\frac{-i(m+n)\theta}{2}} \sqrt{\frac{qm}{n+qm}} |n_n\rangle \right) |\psi_R^1\rangle, \end{aligned} \quad (142)$$

and

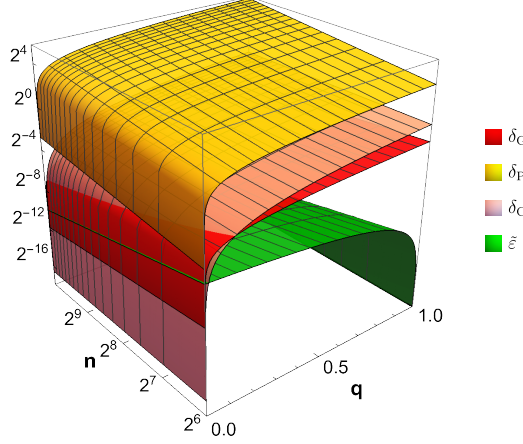
$$P(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}) = \max_{\psi} P(|\psi_\theta\rangle, |\psi\rangle) = \sqrt{1 - \left| \frac{n + qm \cos((m+n)\theta/2)}{n + qm} \right|^2}, \quad (143)$$

where the maximum of $P(|\psi_\theta\rangle, |\psi\rangle)$ is attained at $|\psi\rangle = (|0_L\rangle + |1_L\rangle)/\sqrt{2}$ (here the reference system can be one-dimensional, namely $|\psi_R^0\rangle = |\psi_R^1\rangle = 1/\sqrt{2}$, because $U_{S,\theta} W U_{L,\theta}^\dagger |0_L\rangle$ does not overlap with $U_{S,\theta} W U_{L,\theta}^\dagger |1_L\rangle$ for all θ). Therefore, the global covariance violation is given by

$$\delta_{\text{group}} = \max_{\theta} P(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}) = \sqrt{1 - \left(\frac{n - qm}{n + qm} \right)^2} = \sqrt{\frac{4qm}{n}} + O\left(\left(\frac{m}{n}\right)^{3/2}\right). \quad (144)$$

The code is exactly covariant when $\theta = \frac{4k\pi}{m+n}$ and k is an integer, and the corresponding local covariance violation is given by

$$\delta_{\text{local}} = \sqrt{2\partial_\theta^2 P(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta})^2} \Big|_{\theta = \frac{4k\pi}{m+n}, \forall k \in \mathbb{Z}} = \sqrt{\frac{qm(m+n)^2}{n + qm}} = \sqrt{qm n} + O\left(m \sqrt{\frac{m}{n}}\right), \quad (145)$$



Supplementary Figure 4. Various types of symmetry measures and QEC accuracy of thermodynamic codes, where $m = 2$, $n \in [2^6, 2^{10}]$ and $q \in [10^{-5}, 1 - 10^{-5}]$. Here we plot δ_{group} , δ_{local} , δ_C and $\tilde{\varepsilon}$ (which is approximately equal to ε for large n). As q increases from 0 to 1, $\tilde{\varepsilon}$ decreases while the symmetry measures increase. We can also see that fixing q , the slopes of δ_{group} , δ_{local} , δ_C , and $\tilde{\varepsilon}$ with respect to n are $-1/2$, $1/2$, 0 , and -1 , respectively, matching our calculations.

To compute the charge conservation violation δ_{charge} , note that $\mathcal{E}_{L \leftarrow S}^\dagger(H_S) = \frac{mn(1-q)}{2(n+qm)}Z_L$ and $\mathcal{E}_{L \leftarrow S}^\dagger(H_S^2) = \frac{mn(m+qn)}{4(n+qm)}\mathbb{1}_L$ so we have

$$\delta_{\text{charge}} = \Delta(H_L - (\mathcal{E}_{S \leftarrow L})^\dagger(H_S)) = \frac{qm(n+m)}{(n+qm)} = qm + O\left(m \cdot \frac{m}{n}\right). \quad (146)$$

Also note that the parameter \mathfrak{B} which shows up in [Theorem 6](#) and [Theorem 21](#) is given by

$$\mathfrak{B} = 2\sqrt{2} \max_{|\psi\rangle} \sqrt{|\langle\psi|\mathcal{E}^\dagger(H_S^2)|\psi\rangle - |\langle\psi|\mathcal{E}^\dagger(H_S)|\psi\rangle|^2} = \sqrt{\frac{2mn(m+qn)}{n+qm}} = \sqrt{2qnm} + O\left(m\sqrt{\frac{m}{n}}\right), \quad (147)$$

when $qn \gg m$ and $\mathfrak{B} = O(m)$ otherwise.

3. Trade-off between QEC and symmetry, and explicit comparisons with lower bounds

Let us first overview the behavior of modified thermodynamic codes. Our calculations above indicate that up to the leading order, $\varepsilon \simeq (1-q)m/2n$, while $\delta_{\text{group}} \simeq \sqrt{4qm/n}$, $\delta_{\text{local}} \simeq \sqrt{qmn}$, and $\delta_{\text{charge}} \simeq qm$ (see [Fig. 4](#)). That is, as q varies from 0 to 1, the symmetry violation (in terms of different measures) and the QEC inaccuracy exhibit trade-off behaviors—the former increases from 0 while the latter decreases to 0.

We now discuss the comparison with our lower bounds, focusing on the large n asymptotics. Note that $H_L = \frac{m}{2}Z_L$ and $H_S = -\frac{1}{2}\sum_{l=1}^n Z_l$, so we have $\Delta H_L = m$ and $\Delta H_{S_l} = 1$ for each l . For single-erasure noise channels (as shown in [Sec. III D](#)), we have $\mathfrak{J} \leq n \max_l \Delta H_{S_l} = n$ and $\mathfrak{F} \leq n \sum_{l=1}^n (\Delta H_{S_l})^2 = n^2$, and [Theorem 6](#) then gives:

$$\delta_{\text{group}} \geq \sqrt{\frac{m-2n\varepsilon}{n}} + O\left(\left(\frac{m}{n}\right)^{3/2}\right). \quad (148)$$

Plugging in the QEC inaccuracy $\varepsilon \simeq \sqrt{(1-q)m/n}$, we have

$$\delta_{\text{group}} \geq \sqrt{\frac{qm}{n}} + O\left(\left(\frac{m}{n}\right)^{3/2}\right). \quad (149)$$

Recall that for the modified thermodynamic code we have $\delta_{\text{group}} \simeq \sqrt{4qm/n}$, which saturates this lower bound on δ_{group} up to a constant factor of 2 in the leading order of m/n . Similarly, we could also plug the actual value

$\delta_{\text{group}} \simeq \sqrt{4qm/n}$ into Eq. (148) and obtain the lower bound

$$\varepsilon \geq \frac{(1-4q)m}{2n} + O\left(\frac{m^2}{n^2}\right), \quad (150)$$

which shows that the actual value $\varepsilon \simeq (1-q)m/2n$ of the modified thermodynamic code saturates this lower bound up to a constant factor in the leading order of m/n for $q < 1/4$.

For the local symmetry measures, we first note that for the modified thermodynamic code with $q > 0$ we have $\delta_{\text{local}}/m \simeq \sqrt{qn/m}$ which becomes larger than 1 as $m/n \rightarrow 0$, thus Theorem 23 is not saturated. We provide one possible explanation of this gap between δ_{local} and its lower bound in Sec. XVI, where we show a refinement of Theorem 23 by replacing δ_{local} in Theorem 23 with δ_{local}^* ($\leq \delta_{\text{local}}$) which is defined using the QFI of the error-corrected channel $\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_{S, \theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L, \theta}^\dagger$ instead of the QFI of $\mathcal{U}_{S, \theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L, \theta}^\dagger$ at $\theta = 0$. We show that the gap between δ_{local} and its lower bound could be explained by its gap with δ_{local}^* , explaining the looseness of Theorem 23.

Recall that for the modified thermodynamic code the charge conservation violation is $\delta_{\text{charge}} \simeq qm$. From Theorem 19, we have

$$\delta_{\text{charge}} \geq \Delta H_L - 2\varepsilon\mathfrak{J} = qm + O\left(\frac{m^2}{n}\right). \quad (151)$$

Namely, δ_{charge} exactly saturates this lower bound in the leading order of m/n .

Note that $\mathfrak{B} = O(\sqrt{n})$ and $\mathfrak{F} = O(n^2)$ satisfies $\mathfrak{B} \ll \sqrt{\mathfrak{F}}$ in Theorem 6 and Theorem 21 so that \mathfrak{B} is negligible in the trade-off relations from Theorem 6 and Theorem 21 for modified thermodynamic codes in the large n asymptotics.

Finally, note that the trade-off relation given by the diamond distance,

$$\delta_{\text{group}, \diamond}^2 + 2\sqrt{\varepsilon_\diamond} \geq \delta_{\text{group}}^2 + 2\varepsilon \geq \frac{m}{n}, \quad (152)$$

which follows from Theorem 6 and the discussion in Sec. II, is also saturated up to a constant factor because $\delta_{\text{group}, \diamond} = \delta_{\text{group}}$ and $\varepsilon_\diamond = \varepsilon^2 + O(m^3/n^3)$ (see Sec. XV for details).

B. Quantum Reed–Muller codes

Reed–Muller codes constitute a family of error-correcting codes of great theoretical and technological interest. The classical Reed–Muller code $R(s, t)$ [82] is a $[2^t, \sum_{i=0}^s \binom{t}{i}, 2^{t-s}]$ code whose codewords correspond to Boolean functions of t variables of degree at most s . Then the shortened Reed–Muller codes $\overline{R}(s, t) = [2^t - 1, \sum_{i=1}^s \binom{t}{i}, 2^{t-s}]$ are obtained by selecting the codewords of $R(s, t)$ whose first digits are 0 and deleting their first digits.

The generalization to the quantum regime based on the stabilizer formalism and CSS construction, which leads to the quantum Reed–Muller codes, are also an important type of QEC codes [81]. Given the nice structures and features of quantum Reed–Muller codes, they provide a natural platform for understanding code properties. For example, quantum Reed–Muller codes were widely applied in magic state distillation and implementing transversal non-Clifford operations [49–52]. Quantum Reed–Muller codes were also known to reach the highest level of the Clifford hierarchy possible under the disjointness restriction [48]. Here we consider the $[[n = 2^t - 1, 1, 3]]$ quantum Reed–Muller code, which is a CSS code [2] whose X stabilizers correspond to $\overline{R}(1, t)$ and Z stabilizers correspond to $\overline{R}(t-2, t)$. It is exactly error-correcting under single-erasure noise and admits a transversal implementation $\bigotimes_l (e^{i\pi Z_l/2^{t-1}})$ of the logical operator $e^{-i\pi Z_L/2^{t-1}}$. We now compute its symmetry violation measures. This code has the following form in the computation basis:

$$|\mathbf{c}_0\rangle = \frac{1}{\sqrt{2^t}} \left(|\mathbf{0}\rangle + \sum_{\mathbf{x} \in \overline{R}(1, t) \setminus \{\mathbf{0}\}} |\mathbf{x}\rangle \right), \quad (153)$$

$$|\mathbf{c}_1\rangle = \frac{1}{\sqrt{2^t}} \left(|\mathbf{1}\rangle + \sum_{\mathbf{x} \in \overline{R}(1, t) \setminus \{\mathbf{0}\}} |\mathbf{1} + \mathbf{x}\rangle \right), \quad (154)$$

where we use \mathbf{x} to denote n -bit strings ($\mathbf{0}$ and $\mathbf{1}$ are all-zero and all-one strings, respectively). All strings in $\overline{R}(1, t) \setminus \{\mathbf{0}\}$ have weight 2^{t-1} . Let W be the encoding isometry $W = |\mathbf{c}_0\rangle\langle 0_L| + |\mathbf{c}_1\rangle\langle 1_L|$. Consider $H_L = \frac{1}{2}Z_L$ and $H_S =$

$-\frac{1}{2} \sum_{l=1}^n Z_l$, which again guarantees that the code tends to be covariant as $n \rightarrow \infty$. Let $|\psi\rangle = |0_L\rangle |\psi_R^0\rangle + |1_L\rangle |\psi_R^1\rangle$ be an arbitrary pure state on $L \otimes R$. Then

$$\begin{aligned} |\psi_\theta\rangle &= U_{S,\theta} W U_{L,\theta}^\dagger |\psi\rangle \\ &= \left(\frac{1}{\sqrt{2^t}} \left(e^{i2^{t-1}\theta} |0\rangle + \sum_{\mathbf{x} \in \overline{R(1,t)} \setminus \{0\}} |\mathbf{x}\rangle \right) \right) |\psi_R^0\rangle + \left(\frac{1}{\sqrt{2^t}} \left(e^{-i2^{t-1}\theta} |1\rangle + \sum_{\mathbf{x} \in \overline{R(1,t)} \setminus \{0\}} |1+\mathbf{x}\rangle \right) \right) |\psi_R^1\rangle, \end{aligned} \quad (155)$$

and thus

$$P(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}) = \max_{\psi} P(|\psi_\theta\rangle, |\psi\rangle) = \sqrt{1 - \left| \frac{n + \cos((n+1)\theta/2)}{n+1} \right|^2}. \quad (156)$$

Therefore, we have

$$\delta_{\text{group}} = \max_{\theta} P(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^\dagger, \mathcal{E}_{S \leftarrow L}) = \sqrt{1 - \left(\frac{n-1}{n+1} \right)^2} = \sqrt{\frac{4}{n}} + O\left(\left(\frac{1}{n}\right)^{3/2}\right). \quad (157)$$

The lower bound from [Theorem 6](#) gives

$$\delta_{\text{group}} \geq \sqrt{\frac{\Delta H_L (\Delta H_S - \frac{1}{2} \Delta H_L)}{\Delta H_S}} = \sqrt{\frac{n - \frac{1}{2}}{n^2}} = \sqrt{\frac{1}{n}} + O\left(\left(\frac{1}{n}\right)^{3/2}\right). \quad (158)$$

Similar to the modified thermodynamic code, δ_{group} saturates its lower bound up to a constant factor of 2 in the leading order of $1/n$. Also note that $\delta_{\text{group}, \diamond} = \delta_{\text{group}}$ in this case according to the discussion in [Sec. II](#), indicating the saturation of the lower bound when we consider the diamond distance. The code is exactly covariant when $\theta = \frac{4k\pi}{n+1}$ and k is an integer, and the corresponding local covariance violation can also be easily computed from [Eq. \(156\)](#):

$$\delta_{\text{local}} = \sqrt{2\partial_\theta^2 P(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}, \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta})^2 \big|_{\theta = \frac{4k\pi}{n+1}, \forall k \in \mathbb{Z}}} = \sqrt{n+1}, \quad (159)$$

which has a quadratic gap with its lower bound $\Delta H_L = 1$. To compute the charge conservation violation δ_{charge} , we note that $(\mathcal{E}_{S \leftarrow L})^\dagger(H_S) = 0$ and $(\mathcal{E}_{S \leftarrow L})^\dagger(H_S^2) = n/4$, thus

$$\delta_{\text{charge}} = \Delta(H_L - (\mathcal{E}_{S \leftarrow L})^\dagger(H_S)) = 1, \quad (160)$$

matching our lower bound ΔH_L . Also note that $\mathfrak{B} = \sqrt{2n}$, so it is negligible in the trade-off relations from [Theorem 6](#) and [Theorem 21](#).

VIII. SUPPLEMENTARY DISCUSSION

In this work, we devised and explored various approaches that enable us to quantitatively understand the fundamental trade-off between the QEC capability and several different characterizations of the degree of continuous symmetries associated with general QEC codes, including the violation of covariance conditions in both global and local senses as well as the violation of charge conservation (see [Sec. IX](#) for a summary). In particular, we introduced two intuitive and powerful frameworks based on the notions of charge fluctuation and gate implementation error respectively, and employed several different methods from approximate QEC, quantum metrology, and quantum resource theory, to derive various forms of the trade-off relations in terms of distance metrics that address both worst-case and average-case inputs. Our results and techniques are expected to have numerous interesting applications to quantum computation as well as physics (see the main text). We specifically discussed the consequent restrictions on the transversal logical gates for general QEC codes, which could be of interest for fault tolerance. We also provided detailed analysis of two interesting examples of approximately covariant codes—a parametrized extension of the (covariant) thermodynamic code, which gives a code family that continuously interpolates between exactly covariant and error-correcting, and the quantum Reed–Muller codes. We showed that both codes can saturate the lower bounds asymptotically up to constant factors, indicating that the bounds are quite tight.

We would like to point out a few issues arising from our technical analysis that are not yet satisfactorily understood and could be worth further investigation:

- For both of the code examples we studied, the global covariance violation and charge conservation violation attain the optimal asymptotic scaling as indicated by the bounds based on the charge fluctuation approach, but the local covariance violation does not (there is a $\Theta(\sqrt{n})$ vs. $O(1)$ gap). Note the observation (discussed above and in [Sec. XVI](#)) that if we additionally consider a recovery step in the definition of local covariance violation then it exhibits a tight scaling. We would hope to close this gap by further understanding both sides of it. This is potentially key to a complete understanding of the behavior and practical meanings of the local symmetry measures.
- The gate implementation error approach provides bounds that behave worse than the corresponding bounds from the charge fluctuation approach under uniformly random local noise. It would be interesting to further understand whether this gap stems from the looseness of [Proposition 10](#). On the other hand, the discussion in [Sec. IV C](#) also indicates that for extremely strong noise (so that the regularized QFI is sublinear), the gate implementation error approach outperforms the charge fluctuation approach. The question remains whether there is a universal bound which exhibits optimal scalings under any noise models.

There are also several important directions for future study:

- Gate implementation error. We introduced the gate implementation error as a notion that nicely unifies QEC inaccuracy and global covariance violation, and in turn serves as a tool for deriving the trade-off between them. We believe that this quantity is interesting in its own right and expect it to find broader applications in the analysis of QEC, distillation etc.
- General continuous symmetry groups. Here we mainly carried out the discussion in terms of $U(1)$ which corresponds to a single conserved quantity, but obviously the symmetry groups are often more complicated in quantum computation and physical scenarios. It would be useful to extend our study to other important continuous symmetry groups such as $SU(d)$, for which we expect that our analysis for $U(1)$ provides a basis and serves as a sub-theory but it is useful to invoke corresponding representation theory machinery (like in Refs. [\[12, 27\]](#)).
- Discrete symmetries. Given the incompatibility results for continuous symmetries, it is natural to ask whether discrete symmetries, which are also broadly important, place restrictions on QEC. It is known that for discrete symmetry groups one can in principle construct a covariant code which also achieves exact QEC [\[11\]](#), indicating that the incompatibility is not as fundamental as continuous symmetries. However, we do know interesting cases where exact QEC is forbidden even in the presence of discrete symmetries under simple additional constraints (e.g., AdS/CFT codes—see Refs. [\[12, 18, 28\]](#)). It would be interesting to further explore both the possible limitations as well as good code constructions for QEC with discrete symmetries in more general terms.

Furthermore, we expect the study of how QEC interacts symmetries to be relevant in wide-ranging physical scenarios. In the main text, we discussed potential applications of our theory and techniques to several topics of great interest in physics, including AdS/CFT, black hole radiation, and many-body physics. It would be interesting to further consolidate these ideas. To this end, an important task is to bridge the language of quantum information used here and those commonly used in high energy and condensed matter physics. To conclude, our study enriches the “physical” understanding of QEC using a wide variety of approaches in quantum information. We hope it will stimulate further interest into exploring the interaction between QEC, quantum information, and physics.

Appendices

IX. SUPPLEMENTARY METHOD 1: SUMMARY AND COMPARISON OF THE DIFFERENT TRADE-OFF RESULTS

Trade-off Relations	Summary of Strengths and Weaknesses
Theorem 6 (δ_{group} and ε)	For single-erasure errors, the bound is near-optimal (as shown in Sec. VII). The bound only applies to isometric encodings.
Theorem 11 (δ_{group} and ε)	The bound is suboptimal for single-erasure errors, but it outperforms Theorem 6 for strong noise when $\mathfrak{F} = o(n)$. The bound applies to general encoding channels.
Theorem 14 (δ_{group} and ε ; $\bar{\delta}_{\text{group}}$ and $\bar{\varepsilon}$)	The bounds apply to general encoding channels, and can characterize the average behavior based on the Choi purified distance. The worst-case bound is weaker than Theorem 11 .
Theorem 19 (δ_{charge} and ε ; δ_{local} and ε)	The trade-off between δ_{charge} and ε is near-optimal for single-erasure errors (as shown in Sec. VII). The bounds only apply to isometric encodings.
Theorem 21 (δ_{charge} and ε ; δ_{local} and ε)	The trade-off between δ_{charge} and ε is near-optimal for single-erasure errors (as shown in Sec. VII). It contains the parameter \mathfrak{B} , the subtlety related to which is discussed in Sec. III B 2 . The bound on δ_{charge} and ε applies to general encoding channels. The bound on δ_{local} and ε only applies to isometric encodings.
Theorem 23 (δ_{local} and ε)	The bound does not contain parameter \mathfrak{B} and also applies to general encoding channels. The bound is suboptimal in some cases, see further discussions in Sec. XVI .

Table I. Summary of the strengths and weaknesses of the various trade-off relations we derived using different approaches.

X. SUPPLEMENTARY METHOD 2: PURIFIED DISTANCE AND DIAMOND DISTANCE BETWEEN ROTATED DEPHASING CHANNELS AND THE IDENTITY

Here, we calculate the purified distance between rotated dephasing channels and the identity for completeness. The same calculations could be found, e.g., in Ref. [\[22\]](#).

Lemma 26. *Consider rotated dephasing channels $\mathcal{D}(\cdot) = (1-p)e^{-i\frac{\phi}{2}Z}\rho e^{i\frac{\phi}{2}Z} + pe^{-i\frac{\phi}{2}Z}Z\rho Ze^{i\frac{\phi}{2}Z}$, where Z is the Pauli- Z operator. It holds that $P(\mathcal{D}, \mathbb{1}) = \sqrt{\frac{1}{2}(1 - (1-2p)\cos\phi)}$ and $D_{\diamond}(\mathcal{D}, \mathbb{1}) = \frac{1}{2}\sqrt{1 - 2(1-2p)\cos\phi + (1-2p)^2}$.*

Proof. Let the input state be $|\psi\rangle = \psi_{00}|00\rangle + \psi_{01}|01\rangle + \psi_{10}|10\rangle + \psi_{11}|11\rangle$, then

$$(\mathcal{D} \otimes \mathbb{1})(|\psi\rangle\langle\psi|) = \begin{pmatrix} \psi_{00}\psi_{00}^* & \psi_{00}\psi_{01}^* & (1-2p)e^{-i\phi}\psi_{00}\psi_{10}^* & (1-2p)e^{-i\phi}\psi_{00}\psi_{11}^* \\ \psi_{00}\psi_{01}^* & \psi_{01}\psi_{01}^* & (1-2p)e^{-i\phi}\psi_{01}\psi_{10}^* & (1-2p)e^{-i\phi}\psi_{01}\psi_{11}^* \\ (1-2p)e^{i\phi}\psi_{10}\psi_{00}^* & (1-2p)e^{i\phi}\psi_{10}\psi_{01}^* & \psi_{10}\psi_{10}^* & \psi_{10}\psi_{11}^* \\ (1-2p)e^{i\phi}\psi_{11}\psi_{00}^* & (1-2p)e^{i\phi}\psi_{11}\psi_{01}^* & \psi_{11}\psi_{10}^* & \psi_{11}\psi_{11}^* \end{pmatrix}. \quad (161)$$

Then

$$P(\mathcal{D}, \mathbb{1}) = \sqrt{1 - f^2(\mathcal{D}, \mathbb{1})} \quad (162)$$

$$= \max_{\psi_{00,01,10,11}} (2\text{Re}[(1-2p)e^{-i\phi}](|\psi_{00}|^2 + |\psi_{01}|^2)(|\psi_{10}|^2 + |\psi_{11}|^2))^{1/2} = \sqrt{\frac{1}{2}(1 - (1-2p)\cos\phi)}, \quad (163)$$

and

$$D_{\diamond}(\mathcal{D}, \mathbb{1}) = \frac{1}{2} \|\mathcal{D} - \mathbb{1}\|_{\diamond} \quad (164)$$

$$= 2 \left| 1 - (1-2p)e^{-i\phi} \right| \max_{\psi_{00,01,10,11}} \left\| \begin{pmatrix} \psi_{00} \\ \psi_{01} \end{pmatrix} \begin{pmatrix} \psi_{10}^* & \psi_{11}^* \end{pmatrix} \right\|_1 = \frac{1}{2} \sqrt{1 - 2(1-2p)\cos\phi + (1-2p)^2}. \quad (165)$$

□

XI. SUPPLEMENTARY METHOD 3: TWO LOWER BOUNDS ON $|\partial_\theta \xi_\theta|^2|_{\theta=0}$.

Here, we present the proofs of the two lower bounds on $|\partial_\theta \xi_\theta|^2|_{\theta=0}$ used in [Proposition 5](#) and [Theorem 23](#).

Lemma 27. *Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. Let $\mathcal{R}_{L \leftarrow S}^{\text{opt}}$ be the recovery channel such that $\varepsilon = P(\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L)$ and $\mathcal{N}_{C,\theta} = \mathcal{R}_{C \leftarrow SA}^{\text{rep}} \circ (\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \otimes \mathbb{1}_A) \circ (\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \otimes \mathbb{1}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}$. Then $\xi_\theta = \langle 0_C | \mathcal{N}_{C,\theta}(|0_C\rangle \langle 1_C|) | 1_C \rangle$ satisfies*

$$|\partial_\theta \xi_\theta|^2|_{\theta=0} \geq ((1 - 2\varepsilon^2)\Delta H_L - \delta_{\text{local}})^2, \quad (166)$$

when $(1 - 2\varepsilon^2)\Delta H_L \geq \delta_{\text{local}}$.

$$|\partial_\theta \xi_\theta|^2|_{\theta=0} \geq (|\chi| - 2\varepsilon\mathfrak{B})^2, \quad (167)$$

when $|\chi| \geq 2\varepsilon\mathfrak{B}$. Here $\mathfrak{B} = \max_{|\psi\rangle} \sqrt{8\mathbb{V}_{H_S}(\mathcal{E}_{S \leftarrow L}(|\psi\rangle))}$.

Proof. We have two rotated dephasing channels:

$$\mathcal{D}_{C,\theta} = \mathcal{R}_{C \leftarrow SA}^{\text{rep}} \circ (\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \otimes \mathbb{1}_A) \circ (\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \otimes \mathcal{U}_{L,\theta}^\dagger \otimes \mathbb{1}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}, \quad (168)$$

$$\mathcal{N}_{C,\theta} = \mathcal{R}_{C \leftarrow SA}^{\text{rep}} \circ (\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \otimes \mathbb{1}_A) \circ (\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \otimes \mathbb{1}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}, \quad (169)$$

of the following forms:

$$\mathcal{D}_{C,\theta}(\cdot) = (1 - p_\theta)e^{-i\frac{\phi_\theta}{2}Z_C}(\cdot)e^{i\frac{\phi_\theta}{2}Z_C} + p_\theta Z_C e^{-i\frac{\phi_\theta}{2}Z_C}(\cdot)e^{i\frac{\phi_\theta}{2}Z_C} Z_C, \quad (170)$$

$$\mathcal{N}_{C,\theta}(\cdot) = (1 - p_\theta)e^{-i\frac{\phi_\theta + \Delta H_L \theta}{2}Z_C}(\cdot)e^{i\frac{\phi_\theta + \Delta H_L \theta}{2}Z_C} + p_\theta Z_C e^{-i\frac{\phi_\theta + \Delta H_L \theta}{2}Z_C}(\cdot)e^{i\frac{\phi_\theta + \Delta H_L \theta}{2}Z_C} Z_C. \quad (171)$$

Consider the parameter estimation of θ in the neighborhood of $\theta = 0$. Then let $\mathcal{R}_{C \leftarrow SA} = \mathcal{R}_{C \leftarrow SA}^{\text{rep}} \circ (\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathbb{1}_A)$ and $\mathcal{E}_{SA \leftarrow C} = (\mathcal{E}_{S \leftarrow L} \circ \mathbb{1}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}$, we have

$$\begin{aligned} \sqrt{p_{\theta=0}} &\leq P(\mathcal{D}_{C,\theta=0}, \mathbb{1}_C) = P(\mathcal{R}_{C \leftarrow SA} \circ (\mathcal{N}_S \otimes \mathbb{1}_A) \circ \mathcal{E}_{SA \leftarrow C}, \mathbb{1}_C) \\ &\leq P(\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L) \leq \varepsilon, \end{aligned} \quad (172)$$

where we use [Lemma 26](#), the monotonicity of the purified distance and the definition of ε .

(1) We first prove [Eq. \(166\)](#). The channel QFI [\[40, 68, 76\]](#) of rotated dephasing channel $\mathcal{D}_{C,\theta}$ is

$$\begin{aligned} F(\mathcal{D}_{\theta,C}) &= 4 \min_h \|(\partial_\theta \mathbf{K}_C - ih\mathbf{K}_C)^\dagger (\partial_\theta \mathbf{K} - ih\mathbf{K}_C)\| \\ &= (1 - 2p_\theta)^2 (\partial_\theta \phi_\theta)^2 + \frac{(\partial_\theta p_\theta)^2}{(1 - p_\theta)p_\theta}, \end{aligned} \quad (173)$$

where h is an arbitrary Hermitian matrix and $\mathbf{K}_C = \left(\frac{\sqrt{1 - p_\theta} e^{-i\phi_\theta Z_C}}{\sqrt{p_\theta} e^{-i\phi_\theta Z_C} Z_C} \right)$. Moreover, using the monotonicity of the channel QFI,

$$F(\mathcal{D}_{\theta,C})|_{\theta=0} \leq F(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^\dagger)|_{\theta=0} = (\delta_{\text{local}})^2. \quad (174)$$

Then

$$|\partial_\theta \xi_\theta|^2|_{\theta=0} = |\langle 0_C | \partial_\theta \mathcal{D}_{C,\theta} \circ \mathcal{U}_{C,\theta}(|0_C\rangle \langle 1_C|) | 1_C \rangle + \langle 0_C | \mathcal{D}_{C,\theta} \circ \partial_\theta \mathcal{U}_{C,\theta}(|0_C\rangle \langle 1_C|) | 1_C \rangle|^2|_{\theta=0} \quad (175)$$

$$= |\langle 0_C | \partial_\theta \mathcal{D}_{C,\theta} \circ \mathcal{U}_{C,\theta}(|0_C\rangle \langle 1_C|) | 1_C \rangle + (-i\Delta H_L) \langle 0_C | \mathcal{D}_{C,\theta} \circ \mathcal{U}_{C,\theta}(|0_C\rangle \langle 1_C|) | 1_C \rangle|^2|_{\theta=0} \quad (176)$$

$$\geq |(1 - 2p_\theta)i(-\partial_\theta \phi_\theta)e^{-i\phi_\theta} + (-i\Delta H_L)(1 - 2p_\theta)e^{-i\phi_\theta}|^2|_{\theta=0} \geq ((1 - 2\varepsilon^2)\Delta H_L - \delta_{\text{local}})^2. \quad (177)$$

(2) Next we prove [Eq. \(167\)](#). First, note that

$$\partial_\theta (\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L})(\cdot) = -i[H_S, \mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L}(\cdot)]. \quad (178)$$

Then

$$\partial_\theta \xi_\theta|_{\theta=0} = \langle 0_C | \partial_\theta \mathcal{N}_{C,\theta}(|0_C\rangle \langle 1_C|) | 1_C \rangle|_{\theta=0} = \langle 0_C | \mathcal{R}_{C \leftarrow SA} \circ \mathcal{N}_{SA}(-i[H_{SA}, \mathcal{E}_{SA \leftarrow C}(|0_C\rangle \langle 1_C|)]) | 1_C \rangle, \quad (179)$$

where $\mathcal{N}_{SA} := \mathcal{N}_S \otimes \mathbb{1}_A$ and $H_{SA} := H_S \otimes \mathbb{1}_A$.

We claim that

$$\left| \langle 0_C | \partial_\theta \mathcal{N}_{C,\theta}(|0_C\rangle \langle 1_C|) | 1_C \rangle|_{\theta=0} - \langle 0_L | \partial_\theta \tilde{\mathcal{U}}_{L,\theta}(|0_L\rangle \langle 1_L|) | 1_L \rangle|_{\theta=0} \right| \leq 2\varepsilon \mathfrak{B}, \quad (180)$$

where $\tilde{\mathcal{U}}_{L,\theta} = e^{-i(\mathcal{E}_{S \leftarrow L})^\dagger(H_S)\theta}(\cdot)e^{i(\mathcal{E}_{S \leftarrow L})^\dagger(H_S)\theta}$. Clearly,

$$\langle 0_L | \partial_\theta \tilde{\mathcal{U}}_{L,\theta}(|0_L\rangle \langle 1_L|) | 1_L \rangle|_{\theta=0} = -i \langle 0_L | [(\mathcal{E}_{S \leftarrow L})^\dagger(H_S), |0_L\rangle \langle 1_L|] | 1_L \rangle = -i\chi. \quad (181)$$

Eq. (167) is then proven combining Eq. (180) and Eq. (181).

Now we prove Eq. (180). First, let $\mathcal{E}_{SA \leftarrow C}(\cdot) = \sum_j E_j(\cdot)E_j^\dagger$, $\mathcal{R}_{C \leftarrow SA} \circ (\mathcal{N}_S \otimes \mathbb{1}_A)(\cdot) = \sum_i R_i(\cdot)R_i^\dagger$. We have

$$f^2 \left(\sum_{ij} R_i E_j \rho E_j^\dagger R_i^\dagger, \rho \right) = f^2(\mathcal{R}_{C \leftarrow SA} \circ (\mathcal{N}_S \otimes \mathbb{1}_A) \circ \mathcal{E}_{SA \leftarrow C}(\rho), \rho) \geq f^2(\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L) = 1 - \varepsilon^2, \quad (182)$$

for any ρ in system $C \otimes R$ where R is a two-dimensional reference system. In particular, choose ρ to be the maximally entangled state in $C \otimes R$, we have

$$2 \left(1 - f^2 \left(\sum_{ij} R_i E_j \rho E_j^\dagger R_i^\dagger, \rho \right) \right) = \sum_{ij} \left(\text{Tr}(E_j^\dagger R_i^\dagger R_i E_j) - \frac{|\text{Tr}(R_i E_j)|^2}{2} \right) = \sum_{ij} \|\check{A}_{ij}\|_{HS}^2 \leq 2\varepsilon^2, \quad (183)$$

where $A_{ij} = R_i E_j$ and $\check{(\cdot)} = (\cdot) - \frac{\text{Tr}(\cdot)\mathbb{1}}{2}$, and $\|\cdot\|_{HS}$ denotes the Hilbert-Schmidt norm. Note that

$$\begin{aligned} & \langle 0_C | \partial_\theta \mathcal{N}_{C,\theta}(|0_C\rangle \langle 1_C|) | 1_C \rangle|_{\theta=0} - \langle 0_L | \partial_\theta \tilde{\mathcal{U}}_{L,\theta}(|0_L\rangle \langle 1_L|) | 1_L \rangle|_{\theta=0} \\ &= \langle 0_C | \mathcal{R}_{C \leftarrow SA} \circ \mathcal{N}_{SA}(-i[H_{SA}, \mathcal{E}_{SA \leftarrow C}(|0_C\rangle \langle 1_C|)]) | 1_C \rangle + i \langle 0_C | [\mathcal{E}_{C \leftarrow SA}^\dagger(H_{SA}), |0_C\rangle \langle 1_C|] | 1_C \rangle \\ &= -i \langle 0_C | \left(\sum_{ij} R_i H_{SA} E_j | 0_C \rangle \langle 1_C | E_j^\dagger R_i^\dagger - \sum_j E_j^\dagger H_{SA} E_j | 0_C \rangle \langle 1_C | \right) | 1_C \rangle \\ &+ i \langle 0_C | \left(\sum_{ij} R_i E_j | 0_C \rangle \langle 1_C | E_j^\dagger H_{SA} R_i^\dagger - \sum_j | 0_C \rangle \langle 1_C | E_j^\dagger H_{SA} E_j \right) | 1_C \rangle. \end{aligned} \quad (184)$$

Then

$$\begin{aligned} & \left\| \sum_{ij} R_i H_{SA} E_j | 0_C \rangle \langle 1_C | E_j^\dagger R_i^\dagger - \sum_j E_j^\dagger H_{SA} E_j | 0_C \rangle \langle 1_C | \right\| \\ & \leq \left\| \sum_{ij} R_i H_{SA} E_j | 0_C \rangle \langle 1_C | E_j^\dagger R_i^\dagger - \sum_j E_j^\dagger H_{SA} E_j | 0_C \rangle \langle 1_C | \right\|_{HS} \end{aligned} \quad (185)$$

$$= \left\| \sum_{ij} R_i H_{SA} E_j | 0_C \rangle \langle 1_C | A_{ij}^\dagger - A_{ij}^\dagger R_i H_{SA} E_j | 0_C \rangle \langle 1_C | \right\|_{HS} \quad (186)$$

$$= \left\| \sum_{ij} R_i H_{SA} E_j | 0_C \rangle \langle 1_C | \check{A}_{ij}^\dagger - \check{A}_{ij}^\dagger R_i H_{SA} E_j | 0_C \rangle \langle 1_C | \right\|_{HS} \quad (187)$$

$$\leq 2 \sum_{ij} \|\check{A}_{ij}\|_{HS} \|R_i H_{SA} E_j | 0_C \rangle \langle 1_C |\|_{HS} \quad (188)$$

$$\leq 2 \sqrt{\sum_{ij} \|\check{A}_{ij}\|_{HS}^2 \sum_{ij} \|R_i H_{SA} E_j | 0_C \rangle \langle 1_C |\|_{HS}^2} \quad (189)$$

$$\leq 2\sqrt{2\varepsilon^2 \langle 0_C | \mathcal{E}_{C \leftarrow SA}^\dagger(H_{SA}^2) | 0_C \rangle} \quad (190)$$

$$\leq 2\sqrt{2\varepsilon} \sqrt{\|(\mathcal{E}_{S \leftarrow L})^\dagger(H_S^2)\|}, \quad (191)$$

where we used the triangular inequality, the submultiplicity of $\|\cdot\|_{H_S}$ and the Cauchy–Schwartz inequalities. Similarly,

$$\left\| \sum_{ij} R_i E_j |0_C\rangle \langle 1_C| E_j^\dagger H_{SA} R_i^\dagger - \sum_j (|0_C\rangle \langle 1_C|) E_j^\dagger H_{SA} E_j \right\| \leq 2\sqrt{2}\varepsilon \sqrt{\|(\mathcal{E}_{S \leftarrow L})^\dagger(H_S^2)\|}. \quad (192)$$

We have

$$\left| \langle 0_C | \partial_\theta \mathcal{N}_{C,\theta} (|0_C\rangle \langle 1_C|) | 1_C \rangle \Big|_{\theta=0} - \langle 0_L | \partial_\theta \tilde{\mathcal{U}}_{L,\theta} (|0_L\rangle \langle 1_L|) | 1_L \rangle \Big|_{\theta=0} \right| \leq 4\sqrt{2}\varepsilon \sqrt{\|(\mathcal{E}_{S \leftarrow L})^\dagger(H_S^2)\|}. \quad (193)$$

Note that the equation above still holds when replacing H_S with $H_S - \nu \mathbb{1}$, and that

$$\min_{\nu \in \mathbb{R}} 2\sqrt{2} \sqrt{\|(\mathcal{E}_{S \leftarrow L})^\dagger((H_S - \nu \mathbb{1})^2)\|} = 2\sqrt{2} \sqrt{\min_{\nu} \max_{|\psi\rangle} \langle \psi | (\mathcal{E}_{S \leftarrow L})^\dagger(H_S^2) | \psi \rangle - 2\nu \langle \psi | (\mathcal{E}_{S \leftarrow L})^\dagger(H_S) | \psi \rangle + \nu^2} \quad (194)$$

$$= 2\sqrt{2} \sqrt{\max_{|\psi\rangle} \min_{\nu} \langle \psi | (\mathcal{E}_{S \leftarrow L})^\dagger(H_S^2) | \psi \rangle - 2\nu \langle \psi | (\mathcal{E}_{S \leftarrow L})^\dagger(H_S) | \psi \rangle + \nu^2} \quad (195)$$

$$= \max_{|\psi\rangle} 2\sqrt{2} \sqrt{\langle \psi | (\mathcal{E}_{S \leftarrow L})^\dagger(H_S^2) | \psi \rangle - (\langle \psi | (\mathcal{E}_{S \leftarrow L})^\dagger(H_S) | \psi \rangle)^2} = \mathfrak{B}, \quad (196)$$

where in the second step we used the Sion's minimax theorem [83]. Eq. (180) is then proven. \square

XII. SUPPLEMENTARY METHOD 4: REFINEMENTS OF PROPOSITION 2, PROPOSITION 3 AND PROPOSITION 5

Here, we present the refinements of Proposition 2, Proposition 3 and Proposition 5. First, we can slightly modify the proof of Proposition 2 to obtain the following:

Proposition 28. *Consider an isometric quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. Then when $\delta_{\text{charge}} \leq \Delta H_S$, the following trade-off relation holds:*

$$\delta_{\text{group}} \geq \min \left\{ \frac{\sqrt{\delta_{\text{charge}} (\Delta H_S - \frac{1}{2} \delta_{\text{charge}})}}{\Delta H_S}, \sqrt{\frac{3}{8}} \right\}, \quad (197)$$

and when $\delta_{\text{charge}} > \Delta H_S$, $\delta_{\text{group}} \geq \sqrt{3/8}$. In particular, when $\delta_{\text{charge}} \ll \Delta H_S$, we have

$$\delta_{\text{group}} \gtrsim \sqrt{\frac{\delta_{\text{charge}}}{\Delta H_S}}. \quad (198)$$

Proof. Since $U_{S,\theta}$ and $U_{L,\theta}$ are both periodic with a common period, we assume H_S and H_L both have integer eigenvalues. We also assume the smallest eigenvalue of H_S is zero because constant shifts do not affect the definitions of symmetry measures. Choose orthonormal $|\tilde{0}_L\rangle$ and $|\tilde{1}_L\rangle$ such that

$$\tilde{\delta}_{\text{charge}} = |\langle \tilde{0}_L | (H_L - (\mathcal{E}_{S \leftarrow L})^\dagger(H_S)) | \tilde{0}_L \rangle - \langle \tilde{1}_L | (H_L - (\mathcal{E}_{S \leftarrow L})^\dagger(H_S)) | \tilde{1}_L \rangle| \quad (199)$$

is arbitrarily close to δ_{charge} and that $\widetilde{\Delta H_L} = \langle \tilde{0}_L | H_L | \tilde{0}_L \rangle - \langle \tilde{1}_L | H_L | \tilde{1}_L \rangle \geq 0$ is a rational number. We can always multiply both H_S and H_L by a large integer (e.g., the denominator of $\widetilde{\Delta H_L}$) such that $\widetilde{\Delta H_L}$ becomes an integer. This rescaling does not change the value of δ_{group} and the right-hand side of Eq. (197). Therefore, without loss of generality, we assume $\widetilde{\Delta H_L}$ is an integer.

When $\mathcal{E}_{S \leftarrow L}(\cdot) = W(\cdot)W^\dagger$ is isometric, let $|\tilde{c}_0\rangle = W|\tilde{0}_L\rangle$, $|\tilde{c}_1\rangle = W|\tilde{1}_L\rangle$, and write

$$|\tilde{c}_0\rangle = \sum_{\eta=0}^{\Delta H_S} c_\eta^0 |\eta^0\rangle, \quad |\tilde{c}_1\rangle = \sum_{\eta=0}^{\Delta H_S} c_\eta^1 |\eta^1\rangle, \quad (200)$$

where $\sum_{\eta} |c_{\eta}^0|^2 = \sum_{\eta} |c_{\eta}^1|^2 = 1$ and $|\eta^0\rangle$ and $|\eta^1\rangle$ are eigenstates of H_S with eigenvalue η . $|\eta^0\rangle$ and $|\eta^1\rangle$ may not be the same when H_S is degenerate. Note that when η is not an eigenvalue of H_S , we simply take $c_{\eta}^0 = 0$ (or $c_{\eta}^1 = 0$) so that c_{η}^0 (or c_{η}^1) is well-defined for any integer η . Let $|\psi\rangle = \frac{1}{\sqrt{2}}(|\tilde{0}_L\rangle|0_R\rangle + |\tilde{1}_L\rangle|1_R\rangle)$. Then the channel fidelity

$$\begin{aligned}
f_{\theta} &:= f(\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^{\dagger}, \mathcal{E}_{S \leftarrow L}) \\
&= \min_{\rho} f((\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^{\dagger} \otimes \mathbb{1}_R)(\rho), (\mathcal{E}_{S \leftarrow L} \otimes \mathbb{1}_R)(\rho)) \\
&\leq f((\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^{\dagger} \otimes \mathbb{1}_R)(|\psi\rangle), (\mathcal{E}_{S \leftarrow L} \otimes \mathbb{1}_R)(|\psi\rangle)) \\
&= |\langle \psi | W^{\dagger} U_{S,\theta} W U_{L,\theta}^{\dagger} | \psi \rangle| \\
&= \left| \frac{1}{2} \sum_{\eta=0}^{\Delta H_S} |c_{\eta}^0|^2 e^{-i\eta\theta + i\widetilde{\Delta H_L}\theta} + \frac{1}{2} \sum_{\eta=0}^{\Delta H_S} |c_{\eta}^1|^2 e^{-i\eta\theta} \right| \\
&= \left| \sum_{\eta=-\widetilde{\Delta H_L}}^{\Delta H_S} c_{\eta} e^{-i\eta\theta} \right| = \left| c_{\eta_*} + \sum_{\eta \neq \eta_*} c_{\eta} e^{-i(\eta-\eta_*)\theta} \right|, \tag{201}
\end{aligned}$$

where we define $c_{\eta} := \frac{1}{2} |c_{\eta+\widetilde{\Delta H_L}}^0|^2 + \frac{1}{2} |c_{\eta}^1|^2$ for $\eta \in [-\widetilde{\Delta H_L}, \Delta H_S]$ and choose η_* such that $c_{\eta_*} \geq c_{\eta}$ for all η . Note that there is always a θ such that $\sum_{\eta \neq \eta_*} c_{\eta} \cos((\eta - \eta_*)\theta) = 0$ (because the integration of it from 0 to 2π is zero) and that $\sum_{\eta \neq \eta_*} c_{\eta} e^{-i(\eta-\eta_*)\theta}$ is imaginary. Then we must have

$$\min_{\theta} f_{\theta} \leq \sqrt{c_{\eta_*}^2 + (1 - c_{\eta_*})^2}. \tag{202}$$

To arrive at a non-trivial lower bound on $\delta_{\text{group}} = \sqrt{1 - \min_{\theta} f_{\theta}^2}$, we need an upper bound of $\min_{\theta} f_{\theta}$ which is smaller than 1. To this end, we analyze c_{η_*} in detail. In particular, we consider two situations:

(1) $c_{\eta_*} \leq 1/2$ and a constant upper bound on $\min_{\theta} f_{\theta}$ exists. We can always find a subset of $\{\eta\}$ denoted by \mathfrak{s} such that $1/4 \leq \sum_{\eta \in \mathfrak{s}} c_{\eta} \leq 1/2$. To find such a set, we first include η_* in \mathfrak{s} and add new elements into \mathfrak{s} one by one until their sum is at least $1/4$. Then there is always a θ such that $(\sum_{\eta \in \mathfrak{s}} c_{\eta} e^{-i\eta\theta}) \cdot (\sum_{\eta \notin \mathfrak{s}} c_{\eta} e^{-i\eta\theta})$ is imaginary, in which case $\min_{\theta} f_{\theta} \leq \sqrt{(1/4)^2 + (3/4)^2} = \sqrt{5/8}$ and we have

$$\delta_{\text{group}} \geq \sqrt{1 - \min_{\theta} f_{\theta}^2} \geq \sqrt{3/8}. \tag{203}$$

(2) $c_{\eta_*} > 1/2$. Then $\sqrt{c_{\eta_*}^2 + (1 - c_{\eta_*})^2}$ is a monotonically increasing function of c_{η_*} and we only need to find an upper bound on c_{η_*} . We first note that $0 \leq \eta_* \leq \Delta H_S - \widetilde{\Delta H_L}$ because otherwise either $c_{\eta_* + \widetilde{\Delta H_L}}^0 = 0$ or $c_{\eta_*}^1 = 0$ which contradicts with $c_{\eta_*} > 1/2$. We have $\tilde{\chi} = \langle \tilde{c}_0 | H_S | \tilde{c}_0 \rangle - \langle \tilde{c}_1 | H_S | \tilde{c}_1 \rangle$, $|\widetilde{\Delta H_L} - \tilde{\chi}| = \tilde{\delta}_{\text{charge}}$ and

$$\begin{aligned}
&\sum_{\eta} |c_{\eta}^0|^2 \eta = \sum_{\eta} |c_{\eta}^1|^2 \eta + \tilde{\chi}, \\
&\Leftrightarrow \widetilde{\Delta H_L} - \tilde{\chi} = - \left(1 - |c_{\eta_*}^1|^2\right) \left(\eta_* - \frac{\sum_{\eta \neq \eta_*} |c_{\eta}^1|^2 \eta}{\sum_{\eta \neq \eta_*} |c_{\eta}^1|^2}\right) + \left(1 - |c_{\eta_* + \widetilde{\Delta H_L}}^0|^2\right) \left(\eta_* + \widetilde{\Delta H_L} - \frac{\sum_{\eta \neq \eta_* + \widetilde{\Delta H_L}} |c_{\eta}^0|^2 \eta}{\sum_{\eta \neq \eta_* + \widetilde{\Delta H_L}} |c_{\eta}^0|^2}\right).
\end{aligned}$$

Note that both $\left|\eta_* - \frac{\sum_{\eta \neq \eta_*} |c_{\eta}^1|^2 \eta}{\sum_{\eta \neq \eta_*} |c_{\eta}^1|^2}\right|$ and $\left|\eta_* + \widetilde{\Delta H_L} - \frac{\sum_{\eta \neq \eta_* + \widetilde{\Delta H_L}} |c_{\eta}^0|^2 \eta}{\sum_{\eta \neq \eta_* + \widetilde{\Delta H_L}} |c_{\eta}^0|^2}\right|$ are at most ΔH_S . Therefore, $c_{\eta_*} \leq 1 - \tilde{\delta}_{\text{charge}}/(2\Delta H_S)$ and

$$\delta_{\text{group}} \geq \sqrt{1 - \min_{\theta} f_{\theta}^2} \geq \sqrt{2c_{\eta_*}(1 - c_{\eta_*})} \geq \frac{\sqrt{\tilde{\delta}_{\text{charge}} \left(\Delta H_S - \frac{1}{2}\tilde{\delta}_{\text{charge}}\right)}}{\Delta H_S}. \tag{204}$$

Proposition 28 then follows from [Eq. \(203\)](#), [Eq. \(204\)](#) and the fact that $\tilde{\delta}_{\text{charge}}$ can be arbitrarily close to δ_{charge} . \square

Proposition 28 shows that δ_{group} is always lower bounded by a monotonic function of δ_{charge} . We also immediately see that such a relation cannot hold true for general non-isometric encodings because there are situations when $\delta_{\text{charge}} > 0$ and $\delta_{\text{group}} = 0$ [73]. The proof of **Proposition 28** is essentially the same as the proof of **Proposition 2** except that we replace the pair of logical states $|0_L\rangle$ and $|1_L\rangle$ with another pair of logical states $|\tilde{0}_L\rangle$ and $|\tilde{1}_L\rangle$. Similar tricks can be used to obtain the refinements of **Proposition 3** and **Proposition 5**.

Proposition 29. *Consider an isometric quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. Then the following inequality holds:*

$$\Delta((\mathcal{E}_{S \leftarrow L})^\dagger(H_S)) \leq 2\varepsilon\mathfrak{J}. \quad (205)$$

Proposition 30. *Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. Then the following inequality holds:*

$$\Delta((\mathcal{E}_{S \leftarrow L})^\dagger(H_S)) \leq 2\varepsilon \left(\sqrt{(1 - \varepsilon^2)\mathfrak{F}} + \mathfrak{B} \right). \quad (206)$$

In particular, when $\varepsilon \ll 1$ and $\mathfrak{B} \ll \sqrt{\mathfrak{F}}$, we have

$$\Delta((\mathcal{E}_{S \leftarrow L})^\dagger(H_S)) \lesssim 2\varepsilon\sqrt{\mathfrak{F}}. \quad (207)$$

We omit the proofs of **Proposition 29** and **Proposition 30** here, as they follow directly from the proofs of **Proposition 3** and **Proposition 5** where we replace $|0_L\rangle$ and $|1_L\rangle$ with orthonormal $|\tilde{0}_L\rangle$ and $|\tilde{1}_L\rangle$ which satisfy

$$\Delta((\mathcal{E}_{S \leftarrow L})^\dagger(H_S)) = \langle \tilde{0}_L | (\mathcal{E}_{S \leftarrow L})^\dagger(H_S) | \tilde{0}_L \rangle - \langle \tilde{1}_L | (\mathcal{E}_{S \leftarrow L})^\dagger(H_S) | \tilde{1}_L \rangle. \quad (208)$$

XIII. SUPPLEMENTARY METHOD 5: AN INEQUALITY BETWEEN $\mathfrak{J}(\mathcal{N}_S, H_S)$ AND $\mathfrak{F}(\mathcal{N}_S, H_S)$

Here, we prove a simple inequality of \mathfrak{J} and \mathfrak{F} .

Lemma 31. *Suppose the HKS condition is satisfied. Then it holds that $\mathfrak{J}(\mathcal{N}_S, H_S)^2 \geq \mathfrak{F}(\mathcal{N}_S, H_S)$.*

Proof. Let $\mathbf{K}_S = \begin{pmatrix} K_{S,1} \\ \vdots \\ K_{S,r} \end{pmatrix}$. According to the definitions Eq. (39) and Eq. (47),

$$\mathfrak{J}(\mathcal{N}_S, H_S)^2 = \min_{h: \mathbf{K}_S^\dagger h \mathbf{K}_S = H_S} (\Delta h)^2 = 4 \min_{h, \nu: \mathbf{K}_S^\dagger h \mathbf{K}_S = H_S - \nu \mathbb{1}} \|h^2\|, \quad (209)$$

and

$$\mathfrak{F}(\mathcal{N}_S, H_S) = F^\infty(\mathcal{N}_S, \theta) = 4 \min_{h: \mathbf{K}_S^\dagger h \mathbf{K}_S = H_S} \left\| \mathbf{K}_S^\dagger h^2 \mathbf{K}_S - H_S^2 \right\| = 4 \min_{h, \nu: \mathbf{K}_S^\dagger h \mathbf{K}_S = H_S - \nu \mathbb{1}} \left\| \mathbf{K}_S^\dagger h^2 \mathbf{K}_S - (H_S - \nu \mathbb{1})^2 \right\|. \quad (210)$$

Then it is clear that $\mathfrak{F}(\mathcal{N}_S, H_S) \leq \mathfrak{J}(\mathcal{N}_S, H_S)^2$ because

$$\left\| \mathbf{K}_S^\dagger h^2 \mathbf{K}_S - (H_S - \nu \mathbb{1})^2 \right\| \leq \left\| \mathbf{K}_S^\dagger h^2 \mathbf{K}_S \right\| \leq \|h^2\| \left\| \mathbf{K}_S \mathbf{K}_S^\dagger \right\| = \|h^2\|. \quad (211)$$

□

XIV. SUPPLEMENTARY METHOD 6: UPPER BOUNDS ON $\mathfrak{J}(\mathcal{N}_S, H_S)$ FOR RANDOM LOCAL NOISE AND INDEPENDENT NOISE

In this section, we prove two useful upper bounds (Eq. (75) and Eq. (82)) on \mathfrak{J} for random local noise and independent noise, respectively.

Lemma 32. *Let $\mathcal{N}_S = \sum_{i=1}^m q_i \mathcal{N}_S^{(i)}$ and $H_S = \sum_i H_S^{(i)}$ where $q_i > 0$ and $\sum_i q_i = 1$. Suppose the HKS condition is satisfied for each pair of $\mathcal{N}_S^{(i)}$ and $H_S^{(i)}$. Then Eq. (75) holds, i.e.,*

$$\mathfrak{J}(\mathcal{N}_S, H_S) \leq \max_i \frac{1}{q_i} \mathfrak{J}(\mathcal{N}_S^{(i)}, H_S^{(i)}). \quad (212)$$

Proof. Let $\mathcal{N}_S^{(i)}(\cdot) = \sum_{j=1}^{r_i} K_{S,j}^{(i)}(\cdot) K_{S,j}^{(i)\dagger}$. Then we could take $\mathbf{K}_S = \begin{pmatrix} \sqrt{q_1} \mathbf{K}^{(1)} \\ \sqrt{q_2} \mathbf{K}^{(2)} \\ \vdots \\ \sqrt{q_m} \mathbf{K}^{(m)} \end{pmatrix}$, where $\mathbf{K}^{(i)} = \begin{pmatrix} K_{S,1}^{(i)} \\ \vdots \\ K_{S,r_i}^{(i)} \end{pmatrix}$. Recall that

$$\mathfrak{J}(\mathcal{N}_S, H_S) = \min_{h: \mathbf{K}_S^\dagger h \mathbf{K}_S = H_S} (\Delta h) = 4 \min_{h, \nu: \mathbf{K}_S^\dagger h \mathbf{K}_S = H_S - \nu \mathbb{1}} \|h\|, \quad (213)$$

$$\mathfrak{J}(\mathcal{N}_S^{(i)}, H_S^{(i)}) = \min_{h^{(i)}: \mathbf{K}^{(i)\dagger} h^{(i)} \mathbf{K}^{(i)} = H_S^{(i)}} (\Delta h^{(i)}) = 4 \min_{h^{(i)}, \nu^{(i)}: \mathbf{K}^{(i)\dagger} h^{(i)} \mathbf{K}^{(i)} = H_S^{(i)} - \nu^{(i)} \mathbb{1}} \|h^{(i)}\|. \quad (214)$$

Suppose $(h_*^{(i)}, \nu_*^{(i)})$ is optimal for $\mathfrak{J}(\mathcal{N}_S^{(i)}, H_S^{(i)})$, then let $\nu_* = \sum_i \nu_*^{(i)}$ and

$$h = \begin{pmatrix} \frac{h_*^{(1)}}{q_1} & & & \\ & \frac{h_*^{(2)}}{q_2} & & \\ & & \ddots & \\ & & & \frac{h_*^{(m)}}{q_m} \end{pmatrix}, \quad (215)$$

We must have $\mathbf{K}_S^\dagger h_* \mathbf{K}_S = \sum_i \mathbf{K}_S^{(i)\dagger} h_* \mathbf{K}_S^{(i)} = H_S - \nu_* \mathbb{1}$. Therefore,

$$\mathfrak{J}(\mathcal{N}_S, H_S) = 4 \min_{h, \nu: \mathbf{K}_S^\dagger h \mathbf{K}_S = H_S - \nu \mathbb{1}} \|h\| \leq 4 \|h_*\| = 4 \max_i \frac{1}{q_i} \|h_*^{(i)}\| = \max_i \frac{1}{q_i} \mathfrak{J}(\mathcal{N}_S^{(i)}, H_S^{(i)}). \quad (216)$$

□

Lemma 33. Let $\mathcal{N}_S = \bigotimes_{l=1}^n \mathcal{N}_{S_l}$ and $H_S = \sum_{l=1}^n H_{S_l}$. Suppose the HKS condition is satisfied for each pair of \mathcal{N}_{S_l} and H_{S_l} . Then Eq. (82) holds, i.e.,

$$\mathfrak{J}(\mathcal{N}_S, H_S) \leq \sum_{l=1}^n \mathfrak{J}(\mathcal{N}_{S_l}, H_{S_l}). \quad (217)$$

Proof. Let $\mathcal{N}_{S_l}(\cdot) = \sum_{j=1}^{r_l} K_{S_l,j}(\cdot) K_{S_l,j}^\dagger$ and $\mathbf{K}_{S_l} = \begin{pmatrix} K_{S_l,1} \\ \vdots \\ K_{S_l,r_l} \end{pmatrix}$. Then we could take $\mathbf{K}_S = \bigotimes_{l=1}^n \mathbf{K}_{S_l}$.

Recall that

$$\mathfrak{J}(\mathcal{N}_S, H_S) = \min_{h: \mathbf{K}_S^\dagger h \mathbf{K}_S = H_S} (\Delta h) = 4 \min_{h, \nu: \mathbf{K}_S^\dagger h \mathbf{K}_S = H_S - \nu \mathbb{1}} \|h\|, \quad (218)$$

$$\mathfrak{J}(\mathcal{N}_{S_l}, H_{S_l}) = \min_{h_l: \mathbf{K}_{S_l}^\dagger h_l \mathbf{K}_{S_l} = H_{S_l}} (\Delta h_l) = 4 \min_{h_l, \nu_l: \mathbf{K}_{S_l}^\dagger h_l \mathbf{K}_{S_l} = H_{S_l} - \nu_l \mathbb{1}} \|h_l\|. \quad (219)$$

Suppose $(h_{l,*}, \nu_{l,*})$ is optimal for $\mathfrak{J}(\mathcal{N}_{S_l}, H_{S_l})$, then let $\nu_* = \sum_l \nu_{l,*}$ and

$$h_* = \sum_{l=1}^n \mathbb{1} \otimes \cdots \underbrace{h_l}_{l\text{-th}} \otimes \cdots \otimes \mathbb{1}. \quad (220)$$

We must have $\mathbf{K}_S^\dagger h_* \mathbf{K}_S = \sum_l H_{S_l} - \nu_{l,*} \mathbb{1} = H_S - \nu_* \mathbb{1}$. Therefore,

$$\mathfrak{J}(\mathcal{N}_S, H_S) = 4 \min_{h, \nu: \mathbf{K}_S^\dagger h \mathbf{K}_S = H_S - \nu \mathbb{1}} \|h\| \leq 4 \|h_*\| = 4 \sum_l \|h_{l,*}\| = \sum_{l=1}^n \mathfrak{J}(\mathcal{N}_{S_l}, H_{S_l}). \quad (221)$$

□

XV. SUPPLEMENTARY METHOD 7: COMPUTING THE QEC INACCURACY OF MODIFIED THERMODYNAMIC CODES

Let $\mathcal{N}_S = \sum_{i=1}^n \frac{1}{n} \mathcal{N}_{S_i}$, where $\mathcal{N}_{S_i}(\cdot) = |\emptyset\rangle \langle \emptyset|_{S_i} \otimes \text{Tr}_{S_i}(\cdot)$. Since for erasure channels the noise locations are effectively known (simply by measuring $|\emptyset\rangle$), to compute $\varepsilon = \min_{\mathcal{R}_{L \leftarrow S}} P(\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L)$, we could equivalently replace \mathcal{N}_S with the completely erasure noise on the first qubit \mathcal{N}_{S_1} and then $\widehat{\mathcal{N}}_{S \rightarrow B}(\cdot) = \sum_{i,j=0}^1 \text{Tr}(\langle i_{S_1} | (\cdot) | j_{S_1} \rangle) |i_B\rangle \langle j_B|$. Note that

$$|0_L\rangle = \sqrt{\frac{n+m}{2(n+qm)}} |1\rangle | (m-1)_{n-1} \rangle + |0\rangle \left(\sqrt{\frac{n-m}{2(n+qm)}} | (m+1)_{n-1} \rangle + \sqrt{\frac{qm}{n+qm}} | 0^{\otimes n-1} \rangle \right) \quad (222)$$

$$|1_L\rangle = |1\rangle \left(\sqrt{\frac{n-m}{2(n+qm)}} | (-m-1)_{n-1} \rangle + \sqrt{\frac{qm}{n+qm}} | 1^{\otimes n-1} \rangle \right) + \sqrt{\frac{n+m}{2(n+qm)}} |0\rangle | (-m+1)_{n-1} \rangle \quad (223)$$

Then (for simplicity, we sometimes omit the subscripts $L, S, S \leftarrow L$ and $L \leftarrow S$)

$$\widehat{\mathcal{N} \circ \mathcal{E}}(|0_L\rangle \langle 0_L|) = \widehat{\mathcal{N}}(|\mathbf{c}_0^q\rangle \langle \mathbf{c}_0^q|) = \frac{n+(2q-1)m}{2(n+qm)} |0_B\rangle \langle 0_B| + \frac{n+m}{2(n+qm)} |1_B\rangle \langle 1_B| =: \rho_0, \quad (224)$$

$$\widehat{\mathcal{N} \circ \mathcal{E}}(|1_L\rangle \langle 1_L|) = \widehat{\mathcal{N}}(|\mathbf{c}_1^q\rangle \langle \mathbf{c}_1^q|) = \frac{n+(2q-1)m}{2(n+qm)} |1_B\rangle \langle 1_B| + \frac{n+m}{2(n+qm)} |0_B\rangle \langle 0_B| =: \rho_1, \quad (225)$$

$$\widehat{\mathcal{N} \circ \mathcal{E}}(|0_L\rangle \langle 1_L|) = \widehat{\mathcal{N} \circ \mathcal{E}}(|1_L\rangle \langle 0_L|) = 0. \quad (226)$$

Then let $|\psi\rangle = \psi_0 |0_L\rangle |\psi_R^0\rangle + \psi_1 |1_L\rangle |\psi_R^1\rangle$ be an arbitrary pure state on $L \otimes R$. We have

$$\varepsilon = \min_{\mathcal{R}'} P(\widehat{\mathcal{N} \circ \mathcal{E}}, \mathcal{R}' \circ \widehat{\mathbb{1}}) \quad (227)$$

$$= \min_{\zeta} \max_{|\psi\rangle} P(|\psi_0|^2 \rho_0 \otimes |\psi_R^0\rangle \langle \psi_R^0| + |\psi_1|^2 \rho_1 \otimes |\psi_R^1\rangle \langle \psi_R^1|, \zeta \otimes (|\psi_0|^2 |\psi_R^0\rangle \langle \psi_R^0| + |\psi_1|^2 |\psi_R^1\rangle \langle \psi_R^1|)) \quad (228)$$

$$\geq \min_{\zeta} \max_{|\psi_0|^2 + |\psi_1|^2 = 1} P(|\psi_0|^2 \rho_0 + |\psi_1|^2 \rho_1, \zeta) \quad (229)$$

$$\geq \min_{\zeta} \frac{1}{2} (P(\rho_0, \zeta) + P(\rho_1, \zeta)) \quad (230)$$

$$\geq \frac{1}{2} P(\rho_0, \rho_1) \quad (231)$$

$$= \frac{(1-q)m}{2(n+qm)} = \frac{(1-q)m}{2n} + O\left(\frac{m^2}{n^2}\right), \quad (232)$$

where we used the monotonicity of purified distance for the third line and the triangular inequality for the fifth line. On the other hand, we have the following upper bound:

$$\varepsilon = \min_{\mathcal{R}'} P(\widehat{\mathcal{N} \circ \mathcal{E}}, \mathcal{R}' \circ \widehat{\mathbb{1}}) \quad (233)$$

$$= \min_{\zeta} \max_{|\psi\rangle} P(|\psi_0|^2 \rho_0 \otimes |\psi_R^0\rangle \langle \psi_R^0| + |\psi_1|^2 \rho_1 \otimes |\psi_R^1\rangle \langle \psi_R^1|, \zeta \otimes (|\psi_0|^2 |\psi_R^0\rangle \langle \psi_R^0| + |\psi_1|^2 |\psi_R^1\rangle \langle \psi_R^1|)) \quad (234)$$

$$\leq \max_{|\psi\rangle} P(|\psi_0|^2 \rho_0 \otimes |\psi_R^0\rangle \langle \psi_R^0| + |\psi_1|^2 \rho_1 \otimes |\psi_R^1\rangle \langle \psi_R^1|, \zeta \otimes (|\psi_0|^2 |\psi_R^0\rangle \langle \psi_R^0| + |\psi_1|^2 |\psi_R^1\rangle \langle \psi_R^1|)) \Big|_{\zeta = \frac{\rho_0 + \rho_1}{2}} \quad (235)$$

$$\leq \max_{|\psi_0|^2 + |\psi_1|^2 = 1} \sqrt{1 - \left(|\psi_0|^2 f(\rho_0, \zeta) + |\psi_1|^2 f(\rho_1, \zeta) \right)^2} \Big|_{\zeta = \frac{\rho_0 + \rho_1}{2}} \quad (236)$$

$$= P\left(\rho_1, \frac{1}{2}(\rho_0 + \rho_1)\right) \quad (237)$$

$$= \sqrt{\frac{1}{2} - \frac{\sqrt{(n+m)(n+(2q-1)m)}}{2(n+qm)}} = \frac{(1-q)m}{2n} + O\left(\frac{m^2}{n^2}\right), \quad (238)$$

where we used the joint concavity of the fidelity [2] for the fourth line and the fact that $P(\rho_0, \zeta) = P(\rho_1, \zeta)$ when $\zeta = \frac{\rho_0 + \rho_1}{2}$ for the fifth line. Therefore, we conclude that

$$\varepsilon(\mathcal{N}_S, \mathcal{E}_{S \leftarrow L}) = \frac{(1-q)m}{2n} + O\left(\frac{m^2}{n^2}\right). \quad (239)$$

We claim that an optimal recovery channel (up to the leading order in m/n) achieving the smallest ε is

$$\begin{aligned} \mathcal{R}_{L \leftarrow S}^{\text{opt}} = & \sum_{l=1}^n \left(|0_L\rangle \langle (m-1)_{n-1}|_{S \setminus S_l} + |1_L\rangle \langle \varphi_1|_{S \setminus S_l} \right) \langle \emptyset|_{S_l} (\cdot) |\emptyset\rangle_{S_l} \left(|(m-1)_{n-1}\rangle_{S \setminus S_l} \langle 0_L| + |\varphi_1\rangle_{S \setminus S_l} \langle 1_L| \right) \\ & + \left(|0_L\rangle \langle \varphi_0|_{S \setminus S_l} + |1_L\rangle \langle (-m+1)_{n-1}|_{S \setminus S_l} \right) \langle \emptyset|_{S_l} (\cdot) |\emptyset\rangle_{S_l} \left(|\varphi_0\rangle_{S \setminus S_l} \langle 0_L| + |(-m+1)_{n-1}\rangle_{S \setminus S_l} \langle 1_L| \right) \\ & + |0_L\rangle \langle 0_L| \text{Tr} \left((\cdot) \tilde{\Pi}^\perp \right), \end{aligned} \quad (240)$$

where

$$|\varphi_1\rangle = \sqrt{\frac{n-m}{n+(2q-1)m}} |(-m-1)_{n-1}\rangle + \sqrt{\frac{2qm}{n+(2q-1)m}} |1^{\otimes n-1}\rangle, \quad (241)$$

$$|\varphi_0\rangle = \sqrt{\frac{n-m}{n+(2q-1)m}} |(m+1)_{n-1}\rangle + \sqrt{\frac{2qm}{n+(2q-1)m}} |0^{\otimes n-1}\rangle, \quad (242)$$

$\tilde{\Pi}$ is the projection onto $\text{span}\{|\emptyset\rangle_{S_l} |(m-1)_{n-1}\rangle_{S \setminus S_l}, |\emptyset\rangle_{S_l} |(-m+1)_{n-1}\rangle_{S \setminus S_l}, |\emptyset\rangle_{S_l} |\varphi_1\rangle_{S \setminus S_l}, |\emptyset\rangle_{S_l} |\varphi_0\rangle_{S \setminus S_l}, \forall l\}$ and $\tilde{\Pi}^\perp$ is its orthogonal projector. We have

$$\mathcal{R}^{\text{opt}} \circ \mathcal{N} \circ \mathcal{E}(|0_L\rangle \langle 0_L|) = |0_L\rangle \langle 0_L|, \quad (243)$$

$$\mathcal{R}^{\text{opt}} \circ \mathcal{N} \circ \mathcal{E}(|1_L\rangle \langle 1_L|) = |1_L\rangle \langle 1_L|, \quad (244)$$

$$\mathcal{R}^{\text{opt}} \circ \mathcal{N} \circ \mathcal{E}(|0_L\rangle \langle 1_L|) = \sqrt{\frac{(n+m)(n+(2q-1)m)}{(n+qm)^2}} |0_L\rangle \langle 1_L|. \quad (245)$$

That is, $\mathcal{R}^{\text{opt}} \circ \mathcal{N} \circ \mathcal{E}$ is a dephasing channel. Using [Lemma 26](#), we obtain

$$P(\mathcal{R}^{\text{opt}} \circ \mathcal{N} \circ \mathcal{E}, \mathbb{1}) = \sqrt{\frac{1}{2} - \frac{\sqrt{(n+m)(n+(2q-1)m)}}{2(n+qm)}} = \frac{(1-q)m}{2n} + O\left(\frac{m^2}{n^2}\right), \quad (246)$$

and

$$D_\diamond(\mathcal{R}^{\text{opt}} \circ \mathcal{N} \circ \mathcal{E}, \mathbb{1}) = \frac{1}{2} - \frac{\sqrt{(n+m)(n+(2q-1)m)}}{2(n+qm)} = \left(\frac{(1-q)m}{2n}\right)^2 + O\left(\frac{m^3}{n^3}\right). \quad (247)$$

Since

$$\varepsilon^2 \leq \varepsilon_\diamond \leq D_\diamond(\mathcal{R}^{\text{opt}} \circ \mathcal{N} \circ \mathcal{E}, \mathbb{1}), \quad (248)$$

we have $\varepsilon_\diamond = \left(\frac{(1-q)m}{2n}\right)^2 + O\left(\frac{m^3}{n^3}\right)$.

XVI. SUPPLEMENTARY METHOD 8: REFINEMENT OF THEOREM 23

Here, we provide a refinement of [Theorem 23](#) by replacing the local covariance violation δ_{local} with a quantity δ_{local}^* which characterizes the local covariance violation under noise and recovery. Then we show that δ_{local}^* could be much more smaller than δ_{local} , which provides one explanation of the looseness of [Theorem 23](#). Specifically,

$$\delta_{\text{local}}^* := \min_{\substack{\mathcal{R}_{L \leftarrow S}: \\ P(\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L) = \varepsilon}} \sqrt{F(\mathcal{R}_{L \leftarrow S} \circ \mathcal{N}_S \circ \mathcal{U}_{S, \theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L, \theta}^\dagger) \big|_{\theta=0}}. \quad (249)$$

Clearly, $\delta_{\text{local}}^* \leq \delta_{\text{local}}$ because of the monotonicity of the QFI. And we have the following refinement of [Theorem 23](#):

Theorem 34. *Consider a quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. When $1 - 2\varepsilon^2 > \delta_{\text{local}}^*/\Delta H_L$, it holds that*

$$\frac{\varepsilon\sqrt{1-\varepsilon^2}}{1-2\varepsilon^2-\delta_{\text{local}}^*/(\Delta H_L)} \geq \frac{\Delta H_L}{\sqrt{4\mathfrak{F}}}. \quad (250)$$

When $\varepsilon \ll 1$, we have

$$\delta_{\text{local}}^* + 2\varepsilon\sqrt{\mathfrak{F}} \gtrsim \Delta H_L. \quad (251)$$

Proof. In order to prove [Theorem 34](#), we only need to prove a refinement of [Lemma 27](#) where δ_{local} is replaced by δ_{local}^* , i.e., $\xi_\theta = \langle 0_C | \mathcal{N}_{C,\theta}(|0_C\rangle \langle 1_C|) | 1_C \rangle$ satisfies

$$|\partial_\theta \xi_\theta|^2|_{\theta=0} \geq ((1 - 2\varepsilon^2)\Delta H_L - \delta_{\text{local}}^*)^2, \quad (252)$$

when $(1 - 2\varepsilon^2)\Delta H_L \geq \delta_{\text{local}}^*$. The rest of the proof follows exactly the same from the proof of [Theorem 23](#). Here

$$\mathcal{N}_{C,\theta} = \mathcal{R}_{C \leftarrow SA}^{\text{rep}} \circ (\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \otimes \mathbb{1}_A) \circ (\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \otimes \mathbb{1}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}, \quad (253)$$

where $\mathcal{R}_{L \leftarrow S}^{\text{opt}}$ is a recovery channel such that $\varepsilon = P(\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L}, \mathbb{1}_L)$. To see that [Eq. \(252\)](#) must be true, we only need to revisit the proof of [Lemma 27](#) in [Sec. XI](#) and note that δ_{local} in [Eq. \(174\)](#) and then in [Eq. \(177\)](#) can be replaced by δ_{local}^* because

$$F(\mathcal{D}_{\theta,C})|_{\theta=0} \leq F(\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^\dagger)|_{\theta=0} = (\delta_{\text{local}}^*)^2, \quad (254)$$

where $\mathcal{D}_{C,\theta} = \mathcal{R}_{C \leftarrow SA}^{\text{rep}} \circ (\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \otimes \mathbb{1}_A) \circ (\mathcal{U}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^\dagger \otimes \mathbb{1}_A) \circ \mathcal{E}_{LA \leftarrow C}^{\text{rep}}$. □

Now we show that δ_{local}^* could be much more smaller than δ_{local} . In [Sec. VII](#), we had two exact QEC code examples (the modified thermodynamic code at $q = 1$, and the $[[n = 2^t - 1, 1, 3]]$ quantum Reed–Muller code) where $\delta_{\text{local}} = \Theta(\sqrt{n})$ and $\delta_{\text{charge}} = O(1)$. The following proposition, combined with [Proposition 1](#), indicates that $\delta_{\text{local}}^* = O(1)$ and has at least a quadratic gap to δ_{local} in these examples.

Proposition 35. *Consider an isometric quantum code defined by $\mathcal{E}_{S \leftarrow L}$. Consider physical Hamiltonian H_S , logical Hamiltonian H_L , and noise channel \mathcal{N}_S . Suppose the HKS condition is satisfied. When $\varepsilon = 0$, i.e., the code is exactly error-correcting, it holds that $\delta_{\text{local}}^* \leq \delta_{\text{charge}}$.*

Proof. Suppose $\mathcal{E}_{S \leftarrow L}(\cdot) = W(\cdot)W^\dagger$ where W is isometric. Then $\delta_{\text{charge}} = \Delta(H_L - W^\dagger H_S W)$. When $\varepsilon = 0$, we must have $\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S \circ \mathcal{E}_{S \leftarrow L} = \mathbb{1}_L$. Let Π be the projection onto the code subspace and $\Pi^\perp = \mathbb{1} - \Pi$. Then according to Theorem 10.1 in Ref. [\[2\]](#), we see that there exists a recovery channel $\mathcal{R}_{L \leftarrow S}^{\text{opt}}$ such that $\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S$ have the following form:

$$\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_S(\cdot) = W^\dagger(\cdot)W + \mathcal{R}^\perp(\Pi^\perp(\cdot)\Pi^\perp), \quad (255)$$

for some CPTP map \mathcal{R}^\perp . Let $|\psi_\theta\rangle = U_{S,\theta} W U_{L,\theta}^\dagger |\psi\rangle = e^{-iH_S\theta} W e^{iH_L\theta} |\psi\rangle$, then $|\partial_\theta \psi_\theta\rangle = e^{-iH_S\theta} W e^{iH_L\theta} iH_L |\psi\rangle - iH_S e^{-iH_S\theta} W e^{iH_L\theta} |\psi\rangle$. Then

$$\rho_\theta = \mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^\dagger(|\psi\rangle \langle \psi|) = W^\dagger |\psi_\theta\rangle \langle \psi_\theta| W + \mathcal{R}^\perp(\Pi^\perp(|\psi_\theta\rangle \langle \psi_\theta|)\Pi^\perp), \quad (256)$$

$$\rho_\theta|_{\theta=0} = W^\dagger |\psi_\theta\rangle \langle \psi_\theta| W = |\psi\rangle \langle \psi|, \quad (257)$$

$$\partial_\theta \rho_\theta|_{\theta=0} = W^\dagger |\partial_\theta \psi_\theta\rangle \langle \psi_\theta| W + W^\dagger |\psi_\theta\rangle \langle \partial_\theta \psi_\theta| W = |\psi'\rangle \langle \psi| + |\psi\rangle \langle \psi'|, \quad (258)$$

where $|\psi'\rangle = i(H_L - W^\dagger H_S W) |\psi\rangle$. Clearly, $\langle \psi | \psi' \rangle + \langle \psi' | \psi \rangle = 0$. Then

$$(\delta_{\text{local}}^*)^2 \leq F(\mathcal{R}_{L \leftarrow S}^{\text{opt}} \circ \mathcal{N}_{S,\theta} \circ \mathcal{E}_{S \leftarrow L} \circ \mathcal{U}_{L,\theta}^\dagger)|_{\theta=0} \quad (259)$$

$$= \max_{\psi} F(\rho_\theta)|_{\theta=0} = \max_{\psi} 4(|\langle \psi' | \psi' \rangle|^2 - |\langle \psi | \psi' \rangle|^2) = (\delta_{\text{charge}})^2. \quad (260)$$

□

-
- [1] Shor, P. W. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A* **52**, R2493–R2496 (1995). URL <https://link.aps.org/doi/10.1103/PhysRevA.52.R2493>.
- [2] Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, 2010).
- [3] Gottesman, D. An introduction to quantum error correction and fault-tolerant quantum computation. In *Quantum information science and its contributions to mathematics, Proceedings of Symposia in Applied Mathematics*, vol. 68, 13–58 (2010).

- [4] Lidar, D. A. & Brun, T. A. *Quantum error correction* (Cambridge university press, 2013).
- [5] Almheiri, A., Dong, X. & Harlow, D. Bulk locality and quantum error correction in ads/cft. *J. High Energy Phys.* **2015**, 163 (2015). URL [https://doi.org/10.1007/JHEP04\(2015\)163](https://doi.org/10.1007/JHEP04(2015)163).
- [6] Pastawski, F., Yoshida, B., Harlow, D. & Preskill, J. Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence. *J. High Energy Phys.* **2015**, 149 (2015). URL [https://doi.org/10.1007/JHEP06\(2015\)149](https://doi.org/10.1007/JHEP06(2015)149).
- [7] Kitaev, A. Fault-tolerant quantum computation by anyons. *Ann. Phys.* **303**, 2–30 (2003). URL [https://doi.org/10.1016/s0003-4916\(02\)00018-0](https://doi.org/10.1016/s0003-4916(02)00018-0).
- [8] Zeng, B., Chen, X., Zhou, D.-L. & Wen, X.-G. *Quantum information meets quantum matter: From quantum entanglement to topological phases of many-body systems* (Springer, 2019).
- [9] Brandão, F. G. S. L., Crosson, E., Şahinoğlu, M. B. & Bowen, J. Quantum error correcting codes in eigenstates of translation-invariant spin chains. *Phys. Rev. Lett.* **123**, 110502 (2019). URL <https://link.aps.org/doi/10.1103/PhysRevLett.123.110502>.
- [10] Wang, D.-S., Affleck, I. & Raussendorf, R. Topological qubits from valence bond solids. *Phys. Rev. Lett.* **120**, 200503 (2018). URL <https://link.aps.org/doi/10.1103/PhysRevLett.120.200503>.
- [11] Hayden, P., Nezami, S., Popescu, S. & Salton, G. Error correction of quantum reference frame information. *PRX Quantum* **2**, 010326 (2021). URL <https://link.aps.org/doi/10.1103/PRXQuantum.2.010326>.
- [12] Faist, P. *et al.* Continuous symmetries and approximate quantum error correction. *Phys. Rev. X* **10**, 041018 (2020). URL <https://link.aps.org/doi/10.1103/PhysRevX.10.041018>.
- [13] Woods, M. P. & Alhambra, Á. M. Continuous groups of transversal gates for quantum error correcting codes from finite clock reference frames. *Quantum* **4**, 245 (2020). URL <https://doi.org/10.22331/q-2020-03-23-245>.
- [14] Wang, D.-S., Zhu, G., Okay, C. & Laflamme, R. Quasi-exact quantum computation. *Phys. Rev. Res.* **2**, 033116 (2020). URL <https://doi.org/10.1103/PhysRevResearch.2.033116>.
- [15] Eastin, B. & Knill, E. Restrictions on transversal encoded quantum gate sets. *Phys. Rev. Lett.* **102**, 110502 (2009). URL <https://link.aps.org/doi/10.1103/PhysRevLett.102.110502>.
- [16] Shor, P. W. Fault-tolerant quantum computation. In *Proceedings of 37th Conference on Foundations of Computer Science*, 56–65 (IEEE, 1996). URL <https://doi.org/10.1109/SFCS.1996.548464>.
- [17] Wen, X.-G. Classifying gauge anomalies through symmetry-protected trivial orders and classifying gravitational anomalies through topological orders. *Phys. Rev. D* **88**, 045013 (2013). URL <https://link.aps.org/doi/10.1103/PhysRevD.88.045013>.
- [18] Harlow, D. & Ooguri, H. Symmetries in Quantum Field Theory and Quantum Gravity. *Commun. Math. Phys.* **383**, 1669–1804 (2021). URL <https://doi.org/10.1007/s00220-021-04040-y>.
- [19] Dong, X., Harlow, D. & Wall, A. C. Reconstruction of bulk operators within the entanglement wedge in gauge-gravity duality. *Phys. Rev. Lett.* **117**, 021601 (2016). URL <https://link.aps.org/doi/10.1103/PhysRevLett.117.021601>.
- [20] May, A., Sorce, J. & Yoshida, B. The connected wedge theorem and its consequences. *J. High Energy Phys.* **2022**, 153 (2022). URL [https://doi.org/10.1007/JHEP11\(2022\)153](https://doi.org/10.1007/JHEP11(2022)153).
- [21] Kubica, A. & Demkowicz-Dobrzański, R. Using quantum metrological bounds in quantum error correction: a simple proof of the approximate eastin-knill theorem. *Phys. Rev. Lett.* **126**, 150503 (2021). URL <https://doi.org/10.1103/physrevlett.126.150503>.
- [22] Zhou, S., Liu, Z.-W. & Jiang, L. New perspectives on covariant quantum error correction. *Quantum* **5**, 521 (2021). URL <https://doi.org/10.22331/q-2021-08-09-521>.
- [23] Yang, Y., Mo, Y., Renes, J. M., Chiribella, G. & Woods, M. P. Optimal universal quantum error correction via bounded reference frames. *Phys. Rev. Res.* **4**, 023107 (2022).
- [24] Fang, K. & Liu, Z.-W. No-go theorems for quantum resource purification: New approach and channel theory. *PRX Quantum* **3**, 010337 (2022). URL <https://link.aps.org/doi/10.1103/PRXQuantum.3.010337>.
- [25] Tajima, H. & Saito, K. Universal limitation of quantum information recovery: symmetry versus coherence. *arXiv:2103.01876* (2021). URL <https://arxiv.org/abs/2103.01876>.
- [26] Wang, D.-S., Wang, Y.-J., Cao, N., Zeng, B. & Laflamme, R. Theory of quasi-exact fault-tolerant quantum computing and valence-bond-solid codes. *New J. Phys.* **24**, 023019 (2022).
- [27] Kong, L. & Liu, Z.-W. Near-optimal covariant quantum error-correcting codes from random unitaries with symmetries. *PRX Quantum* **3**, 020314 (2022). URL <https://link.aps.org/doi/10.1103/PRXQuantum.3.020314>.
- [28] Harlow, D. & Ooguri, H. Constraints on symmetries from holography. *Phys. Rev. Lett.* **122**, 191601 (2019). URL <https://link.aps.org/doi/10.1103/PhysRevLett.122.191601>.
- [29] Kohler, T. & Cubitt, T. Toy models of holographic duality between local hamiltonians. *J. High Energy Phys.* **2019**, 17 (2019). URL [https://doi.org/10.1007/JHEP08\(2019\)017](https://doi.org/10.1007/JHEP08(2019)017).
- [30] Brading, K., Castellani, E. & Teh, N. Symmetry and Symmetry Breaking. In Zalta, E. N. (ed.) *The Stanford Encyclopedia of Philosophy* (Metaphysics Research Lab, Stanford University, 2021), Fall 2021 edn.
- [31] Witten, E. Symmetry and emergence. *Nat. Phys.* **14**, 116–119 (2018). URL <https://doi.org/10.1038/nphys4348>.
- [32] Misner, C. W. & Wheeler, J. A. Classical physics as geometry. *Ann. Phys.* **2**, 525–603 (1957). URL [https://doi.org/10.1016/0003-4916\(57\)90049-0](https://doi.org/10.1016/0003-4916(57)90049-0).
- [33] Giddings, S. B. & Strominger, A. Loss of incoherence and determination of coupling constants in quantum gravity. *Nucl. Phys. B* **307**, 854–866 (1988). URL [https://doi.org/10.1016/0550-3213\(88\)90109-5](https://doi.org/10.1016/0550-3213(88)90109-5).
- [34] Kallosh, R., Linde, A., Linde, D. & Susskind, L. Gravity and global symmetries. *Phys. Rev. D* **52**, 912–935 (1995). URL <https://link.aps.org/doi/10.1103/PhysRevD.52.912>.

- [35] Arkani-Hamed, N., Motl, L., Nicolis, A. & Vafa, C. The string landscape, black holes and gravity as the weakest force. *J. High Energy Phys.* **2007**, 060–060 (2007). URL <https://doi.org/10.1088/1126-6708/2007/06/060>.
- [36] Banks, T. & Seiberg, N. Symmetries and strings in field theory and gravity. *Phys. Rev. D* **83**, 084019 (2011). URL <https://link.aps.org/doi/10.1103/PhysRevD.83.084019>.
- [37] Knill, E. & Laflamme, R. Theory of quantum error-correcting codes. *Phys. Rev. A* **55**, 900–911 (1997). URL <https://link.aps.org/doi/10.1103/PhysRevA.55.900>.
- [38] Bény, C. & Oreshkov, O. General conditions for approximate quantum error correction and near-optimal recovery channels. *Phys. Rev. Lett.* **104**, 120501 (2010). URL <https://link.aps.org/doi/10.1103/PhysRevLett.104.120501>.
- [39] Giovannetti, V., Lloyd, S. & Maccone, L. Advances in quantum metrology. *Nat. Photonics* **5**, 222 (2011). URL <https://doi.org/10.1038/nphoton.2011.35>.
- [40] Zhou, S. & Jiang, L. Asymptotic theory of quantum channel estimation. *PRX Quantum* **2**, 010343 (2021). URL <https://link.aps.org/doi/10.1103/PRXQuantum.2.010343>.
- [41] Chitambar, E. & Gour, G. Quantum resource theories. *Rev. Mod. Phys.* **91**, 025001 (2019). URL <https://doi.org/10.1103/RevModPhys.91.025001>.
- [42] Marvian, I. Coherence distillation machines are impossible in quantum thermodynamics. *Nat. Commun.* **11**, 25 (2020). URL <https://doi.org/10.1038/s41467-019-13846-3>.
- [43] Fang, K. & Liu, Z.-W. No-go theorems for quantum resource purification. *Phys. Rev. Lett.* **125**, 060405 (2020). URL <https://link.aps.org/doi/10.1103/PhysRevLett.125.060405>.
- [44] Zeng, B., Cross, A. & Chuang, I. L. Transversality versus universality for additive quantum codes. *IEEE T. Inform. Theory* **57**, 6272–6284 (2011). URL <https://doi.org/10.1109/tit.2011.2161917>.
- [45] Bravyi, S. & König, R. Classification of topologically protected gates for local stabilizer codes. *Phys. Rev. Lett.* **110**, 170503 (2013). URL <https://link.aps.org/doi/10.1103/PhysRevLett.110.170503>.
- [46] Pastawski, F. & Yoshida, B. Fault-tolerant logical gates in quantum error-correcting codes. *Phys. Rev. A* **91**, 012305 (2015). URL <https://link.aps.org/doi/10.1103/PhysRevA.91.012305>.
- [47] Anderson, J. T. & Jochym-O'Connor, T. Classification of transversal gates in qubit stabilizer codes. *Quantum Information and Computation* **16**, 771–802 (2016). URL <https://doi.org/10.26421/qic16.9-10-3>.
- [48] Jochym-O'Connor, T., Kubica, A. & Yoder, T. J. Disjointness of stabilizer codes and limitations on fault-tolerant logical gates. *Phys. Rev. X* **8**, 021047 (2018). URL <https://link.aps.org/doi/10.1103/PhysRevX.8.021047>.
- [49] Bravyi, S. & Haah, J. Magic-state distillation with low overhead. *Phys. Rev. A* **86**, 052329 (2012). URL <https://doi.org/10.1103/PhysRevA.86.052329>.
- [50] Anderson, J. T., Duclos-Cianci, G. & Poulin, D. Fault-tolerant conversion between the steane and reed-muller quantum codes. *Phys. Rev. Lett.* **113**, 080501 (2014). URL <https://doi.org/10.1103/PhysRevLett.113.080501>.
- [51] Haah, J. & Hastings, M. B. Codes and protocols for distilling t , controlled- s , and toffoli gates. *Quantum* **2**, 71 (2018). URL <https://doi.org/10.22331/q-2018-06-07-71>.
- [52] Hastings, M. B. & Haah, J. Distillation with sublogarithmic overhead. *Phys. Rev. Lett.* **120**, 050504 (2018). URL <https://doi.org/10.1103/PhysRevLett.120.050504>.
- [53] Gilchrist, A., Langford, N. K. & Nielsen, M. A. Distance measures to compare real and ideal quantum processes. *Phys. Rev. A* **71**, 062310 (2005). URL <https://link.aps.org/doi/10.1103/PhysRevA.71.062310>.
- [54] Tomamichel, M. *Quantum Inf. Process. with finite resources: mathematical foundations*, vol. 5 (Springer, 2015).
- [55] Helstrom, C. W. *Quantum detection and estimation theory* (Academic press New York, 1976).
- [56] Holevo, A. S. *Probabilistic and statistical aspects of quantum theory* (North Holland, 1982).
- [57] Hübner, M. Explicit computation of the bures distance for density matrices. *Phys. Lett. A* **163**, 239–242 (1992). URL [https://doi.org/10.1016/0375-9601\(92\)91004-B](https://doi.org/10.1016/0375-9601(92)91004-B).
- [58] Sommers, H.-J. & Zyczkowski, K. Bures volume of the set of mixed quantum states. *J. Phys. A-math. Gen.* **36**, 10083–10100 (2003). URL <https://doi.org/10.1088/0305-4470/36/39/308>.
- [59] Watrous, J. *The theory of quantum information* (Cambridge University Press, 2018).
- [60] Liu, Z.-W. & Winter, A. Resource theories of quantum channels and the universal role of resource erasure. *arXiv:1904.04201* (2019). URL <https://arxiv.org/abs/1904.04201>.
- [61] Horodecki, M., Horodecki, P. & Horodecki, R. General teleportation channel, singlet fraction, and quasidistillation. *Phys. Rev. A* **60**, 1888 (1999). URL <https://doi.org/10.1103/PhysRevA.60.1888>.
- [62] Nielsen, M. A. A simple formula for the average gate fidelity of a quantum dynamical operation. *Phys. Lett. A* **303**, 249–252 (2002). URL [https://doi.org/10.1016/S0375-9601\(02\)01272-0](https://doi.org/10.1016/S0375-9601(02)01272-0).
- [63] Aharonov, D., Kitaev, A. & Nisan, N. Quantum circuits with mixed states. In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98, 20–30 (Association for Computing Machinery, New York, NY, USA, 1998). URL <https://doi.org/10.1145/276698.276708>.
- [64] Fuchs, C. A. & Van De Graaf, J. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE T. Inform. Theory* **45**, 1216–1227 (1999). URL <https://doi.org/10.1109/18.761271>.
- [65] Kessler, E. M., Lovchinsky, I., Sushkov, A. O. & Lukin, M. D. Quantum error correction for metrology. *Phys. Rev. Lett.* **112**, 150802 (2014). URL <https://link.aps.org/doi/10.1103/PhysRevLett.112.150802>.
- [66] Arrad, G., Vinkler, Y., Aharonov, D. & Retzker, A. Increasing sensing resolution with error correction. *Phys. Rev. Lett.* **112**, 150801 (2014). URL <https://link.aps.org/doi/10.1103/PhysRevLett.112.150801>.
- [67] Gottesman, D. Quantum fault tolerance in small experiments. *arXiv:1610.03507* (2016). URL <https://arxiv.org/abs/1610.03507>.
- [68] Fujiwara, A. & Imai, H. A fibre bundle over manifolds of quantum channels and its application to quantum statistics. *J.*

- Phys. A-math. Theor.* **41**, 255304 (2008). URL <https://doi.org/10.1088%2F1751-8113%2F41%2F25%2F255304>.
- [69] Braunstein, S. L. & Caves, C. M. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.* **72**, 3439–3443 (1994). URL <https://link.aps.org/doi/10.1103/PhysRevLett.72.3439>.
- [70] Yuen, H. & Lax, M. Multiple-parameter quantum estimation and measurement of nonselfadjoint observables. *IEEE T. Inform. Theory* **19**, 740–750 (1973). URL <https://doi.org/10.1109/tit.1973.1055103>.
- [71] Hayashi, M. Comparison between the cramer-rao and the mini-max approaches in quantum channel estimation. *Commun. Math. Phys.* **304**, 689–709 (2011). URL <https://doi.org/10.1007/s00220-011-1239-4>.
- [72] Katariya, V. & Wilde, M. M. Geometric distinguishability measures limit quantum channel estimation and discrimination. *Quantum Inf. Process.* **20**, 78 (2021). URL <https://doi.org/10.1007/s11128-021-02992-7>.
- [73] Cirstoiu, C., Korzekwa, K. & Jennings, D. Robustness of noether’s principle: Maximal disconnects between conservation laws and symmetries in quantum theory. *Phys. Rev. X* **10**, 041035 (2020). URL <https://link.aps.org/doi/10.1103/PhysRevX.10.041035>.
- [74] Boyd, S. & Vandenberghe, L. *Convex optimization* (Cambridge university press, 2004).
- [75] Kołodyński, J. & Demkowicz-Dobrzański, R. Efficient tools for quantum metrology with uncorrelated noise. *New J. Phys.* **15**, 073043 (2013). URL <https://doi.org/10.1088%2F1367-2630%2F15%2F7%2F073043>.
- [76] Demkowicz-Dobrzański, R., Kołodyński, J. & Guţă, M. The elusive heisenberg limit in quantum-enhanced metrology. *Nat. Commun.* **3**, 1063 (2012). URL <https://doi.org/10.1038/ncomms2067>.
- [77] Schumacher, B. Sending entanglement through noisy quantum channels. *Phys. Rev. A* **54**, 2614–2628 (1996). URL <https://link.aps.org/doi/10.1103/PhysRevA.54.2614>.
- [78] Petz, D. & Ghinea, C. Introduction to quantum fisher information. In *Quantum probability and related topics*, 261–281 (World Scientific, 2011). URL https://doi.org/10.1142/9789814338745_0015.
- [79] Tajima, H., Shiraishi, N. & Saito, K. Uncertainty relations in implementation of unitary operations. *Phys. Rev. Lett.* **121**, 110403 (2018). URL <https://doi.org/10.1103/PhysRevLett.121.110403>.
- [80] Tajima, H., Shiraishi, N. & Saito, K. Coherence cost for violating conservation laws. *Phys. Rev. Res.* **2**, 043374 (2020). URL <https://doi.org/10.1103/PhysRevResearch.2.043374>.
- [81] Steane, A. Quantum reed-muller codes. *IEEE T. Inform. Theory* **45**, 1701–1703 (1999). URL <https://doi.org/10.1109/18.771249>.
- [82] MacWilliams, F. J. & Sloane, N. J. A. *The theory of error correcting codes*, vol. 16 (Elsevier, 1977).
- [83] Komiya, H. Elementary proof for sion’s minimax theorem. *Kodai Math. J.* **11**, 5–7 (1988).