Covariant Quantum Error-Correcting Codes
with Metrological Entanglement Advantage

Cheng-Ju Lin,[1, 2] Zi-Wen Liu,[3, 4] Victor V. Albert,[1] and Alexey V. Gorshkov[1, 2]

[1]*Joint Center for Quantum Information and Computer Science,*
NIST/University of Maryland, College Park, Maryland 20742, USA
[2]*Joint Quantum Institute, NIST/University of Maryland, College Park, Maryland 20742, USA*
[3]*Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, China*
[4]*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*
(Dated: October 1, 2024)

We show that a subset of the basis for the irreducible representations of the total $SU(2)$ rotation forms a covariant approximate quantum error-correcting code with transversal $U(1)$ logical gates. Using only properties of the angular momentum algebra, we obtain bounds on the code inaccuracy against generic noise on any known d sites and against heralded d-local erasures, generalizing and improving previous works on the "thermodynamic code" to general local spin and different irreducible representations. We demonstrate that this family of codes can host and protect a probe state with quantum Fisher information surpassing the standard quantum limit when the sensing parameter couples to the generator of the $U(1)$ logical gate.

Quantum error-correcting codes have found applications beyond fault-tolerant quantum computation to a wide array of quantum technologies and physical scenarios, each requiring unique features. For quantum metrology, protected encodings allow for a metrological *entanglement advantage*, where the ability to measure a signal, quantified by the quantum Fisher information (QFI), surpasses the standard quantum limit [1–12]. Another example is code covariance with continuous transversal (i.e., tensor-product) gates, which is relevant in scenarios like reference-frame error correction [13, 14]. While these gates are incompatible with exact finite-dimensional codes due to the Eastin-Knill theorem [15], they are achievable with approximate quantum error-correcting codes (AQECCs) [16–19].

One can then consider the task of measuring the variable parametrizing a continuous transversal gate for metrological purposes. Various studies have established bounds on the inaccuracy of covariant AQECCs from this perspective [20–25]. Conversely, codes with continuous gates can be candidates for QFI protection, though code covariance is not necessary for this application.

Motivated by these ideas, we introduce a family of AQECCs with transversal continuous gates, based on simple symmetry considerations. Additionally, these codes provide a metrological entanglement advantage for certain code parameters. More specifically, we consider N spin-s degrees of freedom and the irreducible representations (irreps) of their total $SU(2)$ rotation. The irrep basis is $|J,M\rangle$, where $J \leq sN$ is the total spin quantum number, and M is the magnetic quantum number. We show that the states $|J,M\rangle$ with certain choices of M form an AQECC with transversal $U(1)$ logical gates, providing bounds on its code inaccuracy against generic noise on any known d sites for $J \leq sN$ and against heralded d-local erasures for $J = sN$.

Our results stem from generalizing and improving upon results in Refs. [10, 20, 25, 26]. A special case of our

AQECCs is $s = \frac{1}{2}$ and $J = sN = \frac{N}{2}$, in which case $|J,M\rangle$ are the Dicke states [27], and the corresponding code is the "thermodynamic code". Its code-inaccuracy bound against generic noise on any known d sites was first analyzed in Ref. [26], which also reported a coincidental identical bound for the $(s, J) = (1, N)$ code. We generalize this code to general s and $J \sim N^a$ ($a \leq 1$) and tighten the bound against the same noise. We obtain our bound using the $SU(2)$ ladder algebra without explicit expressions for the code words, therefore producing a general s- and J-dependent bound for this family of codes and explaining the coincidence from Ref. [26].

In addition, consider the state $|J = \frac{N}{2} - 1, M = J\rangle = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} e^{i2\pi j/N} \hat{q}_j^- |\uparrow\rangle^{\otimes N}$ along with $|J = \frac{N}{2} - 1, M\rangle$ for the rest of M, where $\hat{q}_j^z |\uparrow\rangle = \frac{1}{2} |\uparrow\rangle$. In Ref. [28], a subset of the states with $M \sim J - N^\kappa$ and $\kappa < 1$ was shown to form an approximate error detection code ("magnon code"). Complementing Ref. [28], we consider code words $|J = \frac{N}{2} - 1, M\rangle$ with $|M| = O(N^b)$ for some $b < 1$.

The inaccuracy bound of the "thermodynamic code" with code words $|M| = O(1)$ in the large-N limit against heralded d-local erasure errors for a constant d is reported in Refs. [20, 25]. We generalize it to $s \geq \frac{1}{2}$ and $J = sN$, but with $|M| = O(N^b)$ and $d \sim N^c$ for some exponents b and c. We obtain a general s-dependent code-inaccuracy bound in this case using the Clebsch-Gordan coefficients without writing out the code words.

Our results extend the codes to a regime containing probe states with a metrological entanglement advantage. The code-inaccuracy bounds also let us bound the loss of QFI. Accordingly, we rigorously demonstrate (and numerically confirm) that a noisy probe state in our codes can retain a metrological entanglement advantage under certain conditions.

Our $SU(2)$-symmetry-irrep-based approach offers a general route for constructing covariant AQECCs with continuous transversal gates, complementing randomized constructions [29–31], reference-frame-assisted construc-

tions [14, 32, 33], and an approach which identifies exact codes from certain discrete-group irreps [34–39]. The protection of a metrological entanglement advantage is an additional feature.

Code words from $SU(2)$ irrep states.—Consider a system of N spin-s degrees of freedom (i.e. qudits of dimension $2s+1$), where s can be a half integer or an integer. The total $SU(2)$-rotation unitary is a tensor product of the $SU(2)$ rotations on the individual spins, generated by the $su(2)$ algebra

$$\hat{Q}^z = \sum_{j=1}^{N} \hat{q}_j^z \ , \quad \hat{Q}^+ = \sum_{j=1}^{N} \hat{q}_j^+ \ , \quad \hat{Q}^- = \sum_{j=1}^{N} \hat{q}_j^- \ , \quad (1)$$

where j labels the spin and $[\hat{q}_j^+, \hat{q}_j^-] = 2\hat{q}_j^z$, $[\hat{q}_j^z, \hat{q}_j^\pm] = \pm\hat{q}_j^\pm$. This yields $[\hat{Q}^+, \hat{Q}^-] = 2\hat{Q}^z$ and $[\hat{Q}^z, \hat{Q}^\pm] = \pm\hat{Q}^\pm$.

The irreps of the total $SU(2)$-rotation unitary are labeled by $J \leq sN$, with a possible multiplicity, and each irrep has dimension $2J+1$ with the basis labeled by $|J,M\rangle$, $M \in \{-J, \cdots, J\}$. The state $|J,M\rangle$ has the following properties:

$$Q^z|J,M\rangle = M|J,M\rangle \ , \quad Q^\pm|J,M\rangle = c_M^\pm|J,M\pm1\rangle \ , \quad (2)$$

where $c_M^\pm = \sqrt{(J \mp M)(J \pm M + 1)}$.

We will show that the code space $\mathfrak{C} = \text{span}\{|J,M\rangle; M = M_{\min}, M_{\min} + \Delta, M_{\min} + 2\Delta, \ldots, M_{\max}\}$, equipped with transversal $U(1)$ logical gates generated by $\hat{Q}^z$, forms an AQECC against generic noise on any known d sites and against heralded d-local erasures for some parameters. For the maximum $J = sN$, $|J,M\rangle$ is totally permutation symmetric, and the irrep multiplicity is one. For a general $J < sN$, one can have multiple irreps with the same J, and we consider $|J,M\rangle$ in any irrep or from any linear superposition of different irreps with the same J, as long as the ladder algebra in Eq. (2) holds among the resulting $|J,M\rangle$. As such, we suppress the multiplicity index.

Code inaccuracy against generic errors.—We use the notation $((N, k))$ and $\epsilon(\mathcal{N})$ for an AQECC, where N denotes the number of qudits, k denotes the number of logical qubits, and $\epsilon(\mathcal{N})$ denotes the optimal recovery inaccuracy against noise channel $\mathcal{N}$ (with encoding):

$$\epsilon(\mathcal{N}) = \min_{\mathcal{R}} D_P(\mathcal{R} \circ \mathcal{N}, \text{id}) \ , \quad (3)$$

where $D_P(\mathcal{E}, \mathcal{F})$ is the purified distance between channels $\mathcal{E}$ and $\mathcal{F}$ (see App. S1 for a list of definitions) and $\mathcal{R}$ is the recovery channel. As we only consider noise on at most d sites, we use $\epsilon(d)$ to stress its dependence on d, while suppressing $\mathcal{N}$ whose specific choices should be clear from the context.

We bound the code inaccuracy $\epsilon(\mathcal{N})$ using a generalization [18] of the Knill-Laflamme conditions [40]. If we let Π be the projector onto the code space and $\mathcal{N}(\bullet) = \sum_j K_j(\bullet)K_j^\dagger$ be the noise channel, then $\epsilon(\mathcal{N}) \leq \delta$ if and only if

$$\Pi K_i^\dagger K_j \Pi = \lambda_{ij}\Pi + \Pi B_{ij}\Pi \ , \quad (4)$$

where λ_{ij} are the components of a density operator, and $D_P(\Lambda + \mathcal{B}, \Lambda) \leq \delta$, where $\Lambda(\rho) = \sum_{i,j} \lambda_{ij}\text{Tr}(\rho)|i\rangle\langle j|$ and $(\Lambda + \mathcal{B})(\rho) = \Lambda(\rho) + \sum_{i,j} \text{Tr}(\rho B_{ij})|i\rangle\langle j|$.

To this end, we start with the following two lemmas.

Lemma 1. *Assume $\hat{F}$ is a d-local operator, then $\langle J,n|\hat{F}|J,m\rangle = 0$ if $|n - m| \geq 2sd + 1$.*

Physically, this lemma holds since a d-local operator can change the magnetic quantum number M by at most $2sd$. It ensures the "off-diagonal" part of the Knill-Laflamme conditions to be exactly zero if the code words have enough spacing in M. The proof is simple and given in App. S2.

To bound the "diagonal part" of the Knill-Laflamme conditions, we have the following lemma.

Lemma 2. *Assume $\hat{F}$ is a d-local operator, then $|\langle J,n|\hat{F}|J,n\rangle - \langle J,m|\hat{F}|J,m\rangle| \leq dq_0\|\hat{F}\|_{op}C(n,m)$, where $C(n,m) := \sum_{M=n}^{m-1}[c_M^+]^{-1}$ and q_0 is a constant. In particular, for $0 \leq b < a \leq 1$, if we take $J \sim N^a$ and $|m - n| \sim N^b$, then we have $|\langle J,n|\hat{F}|J,n\rangle - \langle J,m|\hat{F}|J,m\rangle| = \|\hat{F}\|_{op}O(dN^{b-a})$.*

Relegating the proof to App. S3, we comment that this bound is obtained by bounding $|\langle J, n+1|\hat{F}|J, n+1\rangle - \langle J,n|\hat{F}|J,n\rangle|$ and using the triangle inequality. We emphasize that the proof only uses the ladder algebra Eq. (2) and the locality of the operator $Q^\pm$ without the explicit expression for $|J,M\rangle$.

Using these two lemmas, we obtain the following upper bound on the code inaccuracy for generic noise on any known d sites $\mathcal{N}(\rho) = \sum_{j=1}^{n_d} K_j(\rho)K_j^\dagger$ whose Kraus operators K_j are at most d-local, acting on a fixed set of d qudits, and $n_d = (2s+1)^{2d}$ is the number of independent Kraus operators. The proof is presented in App. S4.

Theorem 1. *Consider the code space $\mathfrak{C} = \text{span}\{|J,M\rangle; M = M_{\min}, M_{\min}+\Delta, M_{\min}+2\Delta, \ldots, M_{\max}\}$, where $J \sim N^a$, $M_{\max} - M_{\min} \sim N^b$, and $b < a \leq 1$. $\mathfrak{C}$ forms an $((N, O(b\log_2 N)))$ AQECC against a d-local noise channel with $\epsilon(d) = O(\sqrt{d}(2s+1)^d N^{(b-a)/2})$ if we take $\Delta \geq 2sd + 1$.*

We focus on parameters where $\epsilon \to 0$ in the thermodynamic limit $N \to \infty$. For size $d \sim c\log_2 N$, we have $\epsilon = O(\sqrt{\log_2 N}N^{(2c\log_2(2s+1)+b-a)/2})$, requiring $c < (a - b)/(4\log_2(2s+1))$ and $b < a$ to have $\epsilon \to 0$ when $N \to \infty$ as shown in Fig. 1(a) as the shaded area.

Our code-inaccuracy bounds generalize and improve the results in Ref. [26], which gives $\epsilon' = \tilde{O}(n_d N^{(5b-1)/2})$ for special cases $(s, J) = (\frac{1}{2}, \frac{N}{2})$ and $(s, J) = (1, N)$, while we achieve $\epsilon = \tilde{O}(\sqrt{n_d}N^{(b-1)/2})$. Our improvements from N^{5b} to N^b and from n_d to $\sqrt{n_d}$ come from setting the "off-diagonal" Knill-Laflamme conditions to zero and a different proof technique.

Moreover, the derivation of the bounds in Ref. [26] relies on the $b \leq \frac{1}{2}$ assumption; our approach allows $1 > b > \frac{1}{2}$, which enables a probe state with a metrological
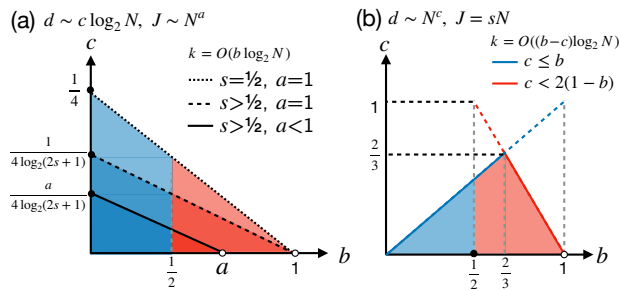
FIG. 1. (a) For a d-local noise channel, the regimes of b and c with $\epsilon \to 0$ (blue and red shaded area) and for metrology (red shaded area). (b) For heralded d-local erasure errors, the regimes of b and c with $\epsilon \to 0$ (blue and red shaded area) and for metrology (red shaded area), all independent of s.

entanglement advantage, as we will discuss below. Notably, using the ladder algebra of $SU(2)$, we obtain a general s- and J-dependent bound on the code inaccuracy, explaining the coincidence of code-inaccuracy bounds for $(s, J) = (\frac{1}{2}, \frac{N}{2})$ and $(s, J) = (1, N)$ in Ref. [26].

The reason to consider the noise on a fixed set of d sites is to control the number n_d. For generic d-local noise $\mathcal{N}(\rho) = \sum_{j=1}^{n_d} K_j(\rho)K_j^\dagger$, where K_i can act on *different* sets of d qudits, $\mathfrak{C}$ forms an $((N, O(b\log_2 N)))$ AQECC with $\epsilon(d) = O(\sqrt{d}n_d N^{(b-a)/2})$ if we take $\Delta \geq 4sd + 1$. (See App. S4 for this corollary and its proof.)

Code inaccuracy against heralded erasures.—The inaccuracy of the "thermodynamic code" against heralded d-local erasures for $d = O(1)$ and $|M| = O(1)$ was found in Refs. [20, 25] to be $\epsilon = O(N^{-1})$. The heralded d-local erasure channel is $\mathcal{N}_\alpha(\rho) := \text{Tr}_\alpha[\rho]$, where α denotes the set of d erased qudits.

We now generalize the code-inaccuracy bound against heralded erasures to $s, d \sim N^c$ and $|M| = O(N^b)$. In this case, we consider codes with $J = sN$ only, where $|J,M\rangle$ is totally permutation symmetric. To bound the code inaccuracy, we follow Ref. [18] by bounding the channel distance via the complementary channel $\hat{\mathcal{N}}_\alpha(\bullet) = \text{Tr}_{\bar{\alpha}}(\bullet)$, where $\bar{\alpha}$ is the complement of α. This can be achieved by the following lemma.

Lemma 3. *Consider the irrep of $SU(2)$ formed by $|J,M\rangle$, where $J = sN$, and the reduced density matrix $\rho_M := \text{Tr}_{\bar{\alpha}}[|J,M\rangle\langle J,M|]$, where $\bar{\alpha}$ is the complement of the set of erased qudits α. We then have the fidelity $f(\rho_M, \rho_{M=0}) = 1 - O(\frac{dM^2}{sN^2}) = 1 - O(N^{2b+c-2})$ asymptotically for $M \sim N^b$, $d \sim N^c$, and $1 > b \geq c$.*

Physically, the above lemma states that the environment can barely tell the code words apart from each other. In the proof (presented in App. S5), we use Clebsch-Gordan coefficients without writing the code words out explicitly, enabling a general s-dependent bound on the fidelity for $J = sN$. In particular, we can rewrite $|J,M\rangle = \sum_{m_1=-j_1}^{j_1} C_{j_1,m_1,j_2,m_2}^{JM} |j_1,m_1\rangle_\alpha \otimes |j_2,m_2\rangle_{\bar{\alpha}}$, where $m_2 = M - m_1$, $j_1 = sd$, $j_2 = J - j_1$, and

$C_{j_1,m_1,j_2,m_2}^{JM} = (\langle j_1, m_1|\langle j_2, m_2|)|J,M\rangle$ is the Clebsch-Gordan coefficient. The fidelity between ρ_M and $\rho_{M=0}$ can therefore be calculated accordingly.

Our proof for a general s has essentially the same asymptotic scaling as the proof of Lemma 13 in Ref. [20] for $s = \frac{1}{2}$, which reported $f(\rho_M, \rho_{M=0}) = 1 - D_{d,M}/N^2 + O(N^{-3})$, where $D_{d,M} = d^2/8 + [d(1 + M^2 + 2M) + 2d^3 + d^2(2M - 1)]/4$. This indeed becomes our result in the asymptotic regime $M \sim N^b$, $d \sim N^c$, and $1 > b \geq c$, making the leading-order term $\sim dM^2$.

The following theorem bounds the code inaccuracy against a heralded d-local erasure error.

Theorem 2. *Consider the code space $\mathfrak{C} = \text{span}\{|J,M\rangle; M = M_{\min}, M_{\min}+\Delta, M_{\min}+2\Delta, \ldots, M_{\max}\}$, where $J = sN$, $M_{\max} - M_{\min} \sim N^b$ and $b < 1$. $\mathfrak{C}$ forms an $((N, O((b - c)\log_2 N)))$ AQECC against heralded d-local erasures of size $d \sim N^c$, where $0 \leq c \leq b < 1$, with $\epsilon(d) = O(N^{b+c/2-1})$ if we take $\Delta \geq 2sd + 1$.*

Theorem 2 can be obtained by combining our Lemma 1, our Lemma 3, and Theorem 3 of Ref. [20]. To make our presentation self-contained, we present the proof of Theorem 2 tailored to our case in App. S6. Refs. [20, 25] showed that $\epsilon = \Omega(N^{-1})$ for any $U(1)$ covariant code. Thus, this family of $SU(2)$ AQECCs is asymptotically optimal for general s since $\epsilon = O(N^{-1})$ when $b = c = 0$.

For general b and c, to have $\epsilon \to 0$ when $N \to \infty$, we need $c < 2(1 - b)$. Additionally, we need $\Delta = \Omega(N^c)$ for Lemma 1 and $c \leq b$ to ensure at least one logical qubit. This parameter regime is shaded in Fig. 1(b), which is independent of s and k. The code words with $M \sim N^b$ for $b > \frac{1}{2}$ [red shaded area in Fig. 1(b)] are crucial for a metrological entanglement advantage.

Quantum metrological advantage.—We now apply our codes to quantum metrology and show that they can have an entanglement advantage with noise resistance. In the local parameter estimation problem [10], we would like to estimate the parameter θ imprinted on the probe state $|\psi_0\rangle \in \mathfrak{C}$ as $|\psi_\theta\rangle = e^{-i\hat{Q}^z\theta}|\psi_0\rangle \in \mathfrak{C}$. Abbreviating $\psi_\theta := |\psi_\theta\rangle\langle\psi_\theta|$, the noisy probe state is then $\rho_\theta := \mathcal{N}(\psi_\theta)$ for the noise channel $\mathcal{N}$.

In quantum metrology, QFI is a central figure of merit, which characterizes how well a parameter can be estimated via the quantum Cramér-Rao bound [41]. For the states ρ_θ parameterized by θ, the QFI is defined as

$$\mathfrak{F}(\rho_\theta; \partial_\theta \rho_\theta) := \text{Tr}(\rho_\theta R^2) , \qquad (5)$$

where R satisfies $\rho_\theta R + R\rho_\theta = 2\partial_\theta \rho_\theta$.

If θ is generated from $e^{-i\theta\hat{Q}^z}$, then without entanglement, the QFI is upper bounded by the standard quantum limit, scaling as $\mathfrak{F} \sim N$. Achieving $\mathfrak{F} \sim N^\gamma$ for $1 < \gamma \leq 2$ indicates a metrological entanglement advantage, with $\gamma = 2$ as the Heisenberg limit. Since $\mathfrak{F}(\psi_\theta; \partial_\theta \psi_\theta) = 4(\langle\psi_\theta|(\hat{Q}^z)^2|\psi_\theta\rangle - \langle\psi_\theta|\hat{Q}^z|\psi_\theta\rangle^2)$ for a pure state, our code can have an entanglement advantage if it contains a logical state whose variance of $\hat{Q}^z$ scales

super-linearly with N. This is possible with the probe state $|\psi_0\rangle \propto |J,M\rangle + |J,-M\rangle$, whose QFI is $\mathfrak{F}(\psi_\theta) = 4M^2 \sim N^{2b}$ for $M \sim N^b$ and $b > \frac{1}{2}$.

While a GHZ(-like) state $|\psi'\rangle = (|J,J\rangle + |J,-J\rangle)/\sqrt{2}$ can achieve the Heisenberg limit [42–44], its QFI becomes zero even if only one qudit is erased (see App. S7). Various other probe states also provide an entanglement advantage, e.g., spin-squeezed [45–49], Dicke [44, 50], and scrambled [51] states. Additionally, various studies have employed different methods to assess the the effect noise in metrology [52–58]. Differently, we use AQECC inaccuracy to rigorously bound the QFI loss under certain noise, showing that entanglement advantage persists. Notably, the QFI protection here is achieved without repeated quantum error correction [10], differing from the typical error-corrected metrology scenarios [1–8, 11]. In a recent work [59], a non-maximally symmetric state also achieves metrological entanglement advantage without error correction.

To this end, we use the following theorem adapted from Ref. [10] to bound the loss of QFI.

Proposition 27 (Adapted from Ref. [10]). *Let ψ_θ be a pure state and $\partial_\theta \psi_\theta = -i[\hat{Q}^z, \psi_\theta]$. If, for $\epsilon > 0$, $\mathcal{M}$ is a channel with a diamond distance from identity satisfying $\|\mathcal{M} - \mathrm{id}\|_\diamond \leq 2\epsilon$, then*

$$\mathfrak{F}(\psi_\theta; \partial_\theta \psi_\theta) - \mathfrak{F}(\mathcal{M}(\psi_\theta); \mathcal{M}(\partial_\theta \psi)) \leq 16\epsilon \|\partial_\theta \psi_\theta\|_1 \|\partial_\theta \psi_\theta\|_{op}. \tag{6}$$

Since $\|\mathcal{E} - \mathcal{F}\|_\diamond \leq 2D_P(\mathcal{E}, \mathcal{F})$ (see App. S1), taking $\mathcal{M} = \mathcal{R} \circ \mathcal{N}$, we have $\|\mathcal{M} - \mathrm{id}\|_\diamond \leq 2\epsilon$ bounded by the AQECC inaccuracy. In fact, from the data processing inequality [10, 60], we have $\mathfrak{F}(\mathcal{R} \circ \mathcal{N}(\rho_\theta); \mathcal{R} \circ \mathcal{N}(\partial_\theta \rho_\theta)) \leq \mathfrak{F}(\mathcal{N}(\rho_\theta); \mathcal{N}(\partial_\theta \rho_\theta))$. That is, the protection of QFI can already be achieved by encoding alone.

For metrology purposes, $\mathfrak{C} = \mathrm{span}\{|J,M\rangle, |J,-M\rangle\}$ suffices to host the probe state $|\psi_0\rangle = (|J,M\rangle + |J,-M\rangle)/\sqrt{2}$, which has $\|\partial_\theta \psi_\theta\|_1 = \|\partial_\theta \psi_\theta\|_{op} = M$. Therefore, we have $\mathfrak{F}(\mathcal{N}(\psi_\theta); \mathcal{N}(\partial_\theta \psi_\theta)) \geq (1 - 4\epsilon)4M^2$, and a protected entanglement advantage is possible if $\epsilon \to 0$ when $N \to \infty$ while having $M \sim N^b$ with $b > 1/2$.

Now considering the d-local noise, for $J \sim N^a$, $M \sim N^b$, and $d \sim c\log_2(N)$, Theorem 1 gives $\epsilon = \tilde{O}(N^{(2c\log_2(2s+1)+b-a)/2})$. So for $a > b > 1/2$, the noisy probe state can maintain an entanglement advantage $\mathfrak{F}(\mathcal{N}(\psi_\theta)) = \Theta(N^{2b})$ if $c < (a-b)/(4\log_2(2s+1))$ (illustrated in Fig. 1(a)). For heralded d-local erasure errors with $d \sim N^c$, Theorem 2 gives $\epsilon = O(N^{b+c/2-1})$. The partially-erased probe state still has an entanglement advantage $\mathfrak{F}(\mathcal{N}(\psi_\theta)) = \Theta(N^{2b})$ with $1 > b > 1/2$ and $\epsilon \to 0$ provided $c < 2(1-b)$ and $c \leq b$ (illustrated in Fig. 1(b)).

To assess our QFI-loss bound, we numerically calculate $\Delta\mathfrak{F} := \mathfrak{F}(\psi_\theta) - \mathfrak{F}(\mathcal{N}(\psi_\theta))$ for $J = sN$, where the channel is $\mathcal{N}(\psi_\theta) = \mathrm{Tr}_\alpha[\psi_\theta]$ and α contains d spins (so $j_1 = sd$). In Fig. 2, we plot $\Delta\mathfrak{F}/(4M^2)$ for various b and c in the red-shaded region of Fig. 1(b). The numerical calculations are detailed in App. S7. Applying the QFI-loss bound Eq. (6), we have $\Delta\mathfrak{F}/(4M^2) \leq 4\epsilon$. We ob-
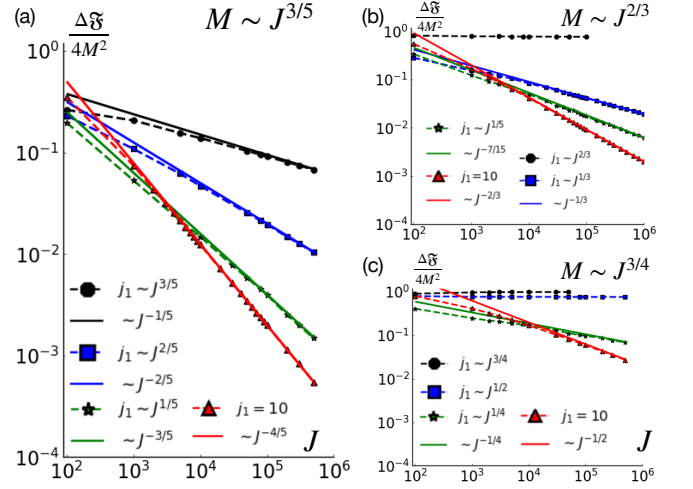

FIG. 2. QFI loss $\Delta\mathfrak{F}$ of the probe state $|\psi\rangle \propto |J,M\rangle + |J,-M\rangle$ ($J = sN$) for (a) $M \sim J^{\frac{3}{5}}$, (b) $M \sim J^{\frac{2}{3}}$, and (c) $M \sim J^{\frac{3}{4}}$, after erasing $d = j_1/s$ qudits. We calculate $\Delta\mathfrak{F}$ by picking values of J and j_1, giving us results for any spin-s and corresponding $N = J/s$ and $d = j_1/s$. Numerical results show $\Delta\mathfrak{F}/(4M^2) \sim J^{2b+c-2}$ for the parameters in the red-shaded region of Fig. 1(b).

serve $\Delta\mathfrak{F}/M^2 \sim J^{2b+c-2}$, which is indeed asymptotically upper-bounded by $J^{b+c/2-1} \sim N^{b+c/2-1}$. Though the bound is not tight, it could be due to the Fuchs-van de Graaf inequality [61] not being saturated asymptotically. However, within the red region in Fig. 1(b), $\mathfrak{F}(\psi_\theta)$ and $\mathfrak{F}(\mathcal{N}(\psi_\theta))$ both scale as N^{2b}, both indicating a metrological entanglement advantage.

In App. S7, we present a local measurement that saturates the quantum Cramér-Rao bound for the ideal probe state $|\psi_\theta\rangle$. For the partially-erased probe state, a global measurement can achieve the QFI scaling, while a local measurement can surpass the standard limit in certain parameter regimes, despite not saturating the quantum Cramér-Rao bound.

Discussion.—We establish code-inaccuracy bounds for a family of covariant approximate quantum error-correcting codes whose codewords house irreducible representations of $SU(2)$. We show that a particular (entangled) codeword can be used for quantum sensing in a way that outperforms classical strategies, even in the presence of certain noise.

Ref. [62] shows that an N-qubit Dicke state $|J = \frac{N}{2}, M\rangle$ can be prepared using a quantum circuit with a number of gates polynomial in N and M. One can then prepare the probe state $|\psi_0\rangle$ using an ancilla qubit and a controlled version of this circuit. Exploring whether this method can be extended to qudits would be valuable.

We have proposed some measurement schemes for the ideal and partially-erased probe states in App. S7. It is worth investigating whether simpler measurements can saturate the Cramér-Rao bound for our noisy state [63]. While the recovery channel [18, 64] cannot increase quantum Fisher information [10, 60], it may aid in devising a

measurement protocol for any noisy probe.

With modifications, our approximate quantum error-correcting code results can be applied to other $SU(2)$ code words, like the quantum many-body scar states in the spin-1 XY-Dzyaloshinskii-Moriya model [65–67] (see App. S8 for details). Accordingly, the $SU(2)$-scar probe states could be noise-resilient, complementing the results using scar states for sensing [67–69]. For states with approximate $SU(2)$ symmetry [24, 25], like the scars in a deformed PXP model [70–74], some exact codes [24, 25], or approximately prepared states [75], further study is warranted to determine if they also provide a noise-resilient entanglement advantage.

Our result suggests a new way of finding approximate quantum error-correcting codes with continuous transversal logical gates by using symmetry irreps, complementing existing methods [14, 29, 31–33]. Here, one selects a subset of the $SU(2)$-irrep basis to form an approximate quantum error-correcting code, reducing the symmetry to $U(1)$ and giving rise to continuous transversal logical gates. This raises the question of whether one can also devise approximate quantum error-correcting codes by using irreps from other continuous symmetries (say, $SU(3)$), which could potentially yield transversal gates generated by multiple (possibly non-commuting) generators and an entanglement advantage for multi-parameter sensing.

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

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

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

[4] X.-M. Lu, S. Yu, and C. H. Oh, Robust quantum metrological schemes based on protection of quantum Fisher information, Nat Commun **6**, 7282 (2015).

[5] Y. Matsuzaki and S. Benjamin, Magnetic-field sensing with quantum error detection under the effect of energy relaxation, Phys. Rev. A **95**, 032303 (2017).

[6] P. Sekatski, M. Skotiniotis, J. Kołodyński, and W. Dür, Quantum metrology with full and fast quantum control, Quantum **1**, 27 (2017).

[7] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Achieving the Heisenberg limit in quantum metrology using quantum error correction, Nat Commun **9**, 78 (2018).

[8] S. Zhou and L. Jiang, Optimal approximate quantum error correction for quantum metrology, Phys. Rev. Research **2**, 013235 (2020).

[9] Y. Ouyang, N. Shettell, and D. Markham, Robust quantum metrology with explicit symmetric states, IEEE Trans. Inform. Theory **68**, 1809 (2022), arXiv:1908.02378 [quant-ph].

[10] P. Faist, M. P. Woods, V. V. Albert, J. M. Renes, J. Eisert, and J. Preskill, Time-Energy Uncertainty Relation for Noisy Quantum Metrology, PRX Quantum **4**, 040336 (2023).

[11] I. Rojkov, D. Layden, P. Cappellaro, J. Home, and F. Reiter, Bias in error-corrected quantum sensing, Phys. Rev. Lett. **128**, 140503 (2022).

[12] Y. Ouyang and G. K. Brennen, Finite-round quantum error correction on symmetric quantum sensors (2024), arXiv:2212.06285 [quant-ph].

[13] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Reference frames, superselection rules, and quantum information, Rev. Mod. Phys. **79**, 555 (2007).

[14] P. Hayden, S. Nezami, S. Popescu, and G. Salton, Error Correction of Quantum Reference Frame Information, PRX Quantum **2**, 010326 (2021).

[15] B. Eastin and E. Knill, Restrictions on Transversal Encoded Quantum Gate Sets, Phys. Rev. Lett. **102**, 110502 (2009).

[16] D. W. Leung, M. A. Nielsen, I. L. Chuang, and Y. Yamamoto, Approximate quantum error correction can lead to better codes, Phys. Rev. A **56**, 2567 (1997).

[17] C. Crépeau, D. Gottesman, and A. Smith, Approximate quantum error-correcting codes and secret sharing schemes, in *Annual International Conference on the Theory and Applications of Cryptographic Techniques* (Springer, 2005) pp. 285–301.

[18] C. Bény and O. Oreshkov, General conditions for approximate quantum error correction and near-optimal recovery channels, Phys. Rev. Lett. **104**, 120501 (2010).

[19] J. Yi, W. Ye, D. Gottesman, and Z.-W. Liu, Complexity and order in approximate quantum error-correcting codes, Nat. Phys. , 1 (2024).

[20] P. Faist, S. Nezami, V. V. Albert, G. Salton, F. Pastawski, P. Hayden, and J. Preskill, Continuous symmetries and approximate quantum error correction,

Phys. Rev. X **10**, 041018 (2020).

[21] S. Zhou, Z.-W. Liu, and L. Jiang, New perspectives on covariant quantum error correction, Quantum **5**, 521 (2021).

[22] A. Kubica and R. Demkowicz-Dobrzański, Using Quantum Metrological Bounds in Quantum Error Correction: A Simple Proof of the Approximate Eastin-Knill Theorem, Phys. Rev. Lett. **126**, 150503 (2021).

[23] K. Fang and Z.-W. Liu, No-go theorems for quantum resource purification: New approach and channel theory, PRX Quantum **3**, 010337 (2022).

[24] Z.-W. Liu and S. Zhou, Quantum error correction meets continuous symmetries: Fundamental trade-offs and case studies (2022), arXiv:2111.06360 [quant-ph].

[25] Z.-W. Liu and S. Zhou, Approximate symmetries and quantum error correction, npj Quantum Inf **9**, 119 (2023).

[26] F. G. S. L. Brandao, E. Crosson, M. B. Şahinoğlu, and J. Bowen, Quantum Error Correcting Codes in Eigenstates of Translation-Invariant Spin Chains, Phys. Rev. Lett. **123**, 110502 (2019).

[27] R. H. Dicke, Coherence in Spontaneous Radiation Processes, Phys. Rev. **93**, 99 (1954).

[28] M. Gschwendtner, R. Koenig, B. Şahinoğlu, and E. Tang, Quantum Error-Detection at Low Energies, J. High Energ. Phys. **2019** (9), 21.

[29] L. Kong and Z.-W. Liu, Near-Optimal Covariant Quantum Error-Correcting Codes from Random Unitaries with Symmetries, PRX Quantum **3**, 020314 (2022).

[30] Z. Li, H. Zheng, J. Liu, L. Jiang, and Z.-W. Liu, Designs from local random quantum circuits with SU(d) symmetry (2024), arXiv:2309.08155 [quant-ph].

[31] Z. Li, H. Zheng, Y. Wang, L. Jiang, Z.-W. Liu, and J. Liu, SU(d)-Symmetric Random Unitaries: Quantum Scrambling, Error Correction, and Machine Learning (2023), arXiv:2309.16556 [quant-ph].

[32] M. P. Woods and Á. M. Alhambra, Continuous groups of transversal gates for quantum error correcting codes from finite clock reference frames, Quantum **4**, 245 (2020).

[33] Y. Yang, Y. Mo, J. M. Renes, G. Chiribella, and M. P. Woods, Optimal Universal Quantum Error Correction via Bounded Reference Frames, Phys. Rev. Research **4**, 023107 (2022).

[34] J. A. Gross, Designing Codes around Interactions: The Case of a Spin, Phys. Rev. Lett. **127**, 010504 (2021).

[35] S. P. Jain, E. R. Hudson, W. C. Campbell, and V. V. Albert, Æ codes, arXiv arXiv:2311.12324.

[36] E. Kubischta and I. Teixeira, The Not-So-Secret Fourth Parameter of Quantum Codes (2024), arXiv:2310.17652 [quant-ph].

[37] E. Kubischta and I. Teixeira, Quantum Codes from Twisted Unitary t -Groups, Phys. Rev. Lett. **133**, 030602 (2024).

[38] E. Kubischta and I. Teixeira, Quantum Codes and Irreducible Products of Characters (2024), arXiv:2403.08999 [quant-ph].

[39] A. Denys and A. Leverrier, Quantum error-correcting codes with a covariant encoding (2024), arXiv:2306.11621 [quant-ph].

[40] E. Knill, R. Laflamme, and L. Viola, Theory of Quantum Error Correction for General Noise, Phys. Rev. Lett. **84**, 2525 (2000).

[41] S. L. Braunstein and C. M. Caves, Statistical distance and the geometry of quantum states, Phys. Rev. Lett. **72**, 3439 (1994).

[42] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nature Photon **5**, 222 (2011).

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

[44] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Quantum metrology with nonclassical states of atomic ensembles, Rev. Mod. Phys. **90**, 035005 (2018).

[45] J. Ma, X. Wang, C. Sun, and F. Nori, Quantum spin squeezing, Physics Reports **509**, 89 (2011).

[46] J. Franke, S. R. Muleady, R. Kaubruegger, F. Kranzl, R. Blatt, A. M. Rey, M. K. Joshi, and C. F. Roos, Quantum-enhanced sensing on optical transitions through finite-range interactions, Nature **621**, 740 (2023).

[47] G. Bornet, G. Emperauger, C. Chen, B. Ye, M. Block, M. Bintz, J. A. Boyd, D. Barredo, T. Comparin, F. Mezzacapo, T. Roscilde, T. Lahaye, N. Y. Yao, and A. Browaeys, Scalable spin squeezing in a dipolar Rydberg atom array, Nature **621**, 728 (2023).

[48] J. A. Hines, S. V. Rajagopal, G. L. Moreau, M. D. Wahrman, N. A. Lewis, O. Marković, and M. Schleier-Smith, Spin Squeezing by Rydberg Dressing in an Array of Atomic Ensembles, Phys. Rev. Lett. **131**, 063401 (2023).

[49] W. J. Eckner, N. Darkwah Oppong, A. Cao, A. W. Young, W. R. Milner, J. M. Robinson, J. Ye, and A. M. Kaufman, Realizing spin squeezing with Rydberg interactions in an optical clock, Nature **621**, 734 (2023).

[50] Z. H. Saleem, M. Perlin, A. Shaji, and S. K. Gray, Achieving the Heisenberg limit with Dicke states in noisy quantum metrology, Phys. Rev. A **109**, 052615 (2024).

[51] Z. Li, S. Colombo, C. Shu, G. Velez, S. Pilatowsky-Cameo, R. Schmied, S. Choi, M. Lukin, E. Pedrozo-Peñafiel, and V. Vuletić, Improving metrology with quantum scrambling, Science **380**, 1381 (2023).

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

[53] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, The elusive Heisenberg limit in quantum-enhanced metrology, Nat Commun **3**, 1063 (2012).

[54] J. Kołodyński and R. Demkowicz-Dobrzański, Efficient tools for quantum metrology with uncorrelated noise, New J. Phys. **15**, 073043 (2013).

[55] R. Chaves, J. B. Brask, M. Markiewicz, J. Kolodynski, and A. Acin, Noisy metrology beyond the standard quantum limit, Phys. Rev. Lett. **111**, 120401 (2013).

[56] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Adaptive Quantum Metrology under General Markovian Noise, Phys. Rev. X **7**, 041009 (2017).

[57] S. Zhou and L. Jiang, Asymptotic theory of quantum channel estimation, PRX Quantum **2**, 010343 (2021).

[58] C. Yin and A. Lucas, Heisenberg-limited metrology with perturbing interactions (2023), arXiv:2308.10929 [math-ph, physics:quant-ph].

[59] C. Yin, V. V. Albert, and S. Zhou, Small correlation is sufficient for optimal noisy quantum metrology (2024), arXiv:2408.00079 [quant-ph].

[60] C. Ferrie, Data-processing inequalities for quantum metrology, Phys. Rev. A **90**, 014101 (2014).

[61] C. Fuchs and J. van de Graaf, Cryptographic distinguishability measures for quantum-mechanical states, IEEE Transactions on Information Theory **45**, 1216 (1999).

[62] A. Bärtschi and S. Eidenbenz, Deterministic preparation of Dicke states, in *International Symposium on Fundamentals of Computation Theory* (Springer, 2019) pp. 126–139.

[63] S. Zhou, C.-L. Zou, and L. Jiang, Saturating the quantum Cram\'er-Rao bound using LOCC, Quantum Sci. Technol. **5**, 025005 (2020).

[64] H. K. Ng and P. Mandayam, A simple approach to approximate quantum error correction based on the transpose channel, Phys. Rev. A **81**, 062342 (2010).

[65] M. Schecter and T. Iadecola, Weak Ergodicity Breaking and Quantum Many-Body Scars in Spin-1 XY Magnets, Phys. Rev. Lett. **123**, 147201 (2019).

[66] D. K. Mark, C.-J. Lin, and O. I. Motrunich, Unified structure for exact towers of scar states in the Affleck-Kennedy-Lieb-Tasaki and other models, Phys. Rev. B **101**, 195131 (2020).

[67] S. Dooley, Robust Quantum Sensing in Strongly Interacting Systems with Many-Body Scars, PRX Quantum **2**, 020330 (2021).

[68] S. Dooley, S. Pappalardi, and J. Goold, Entanglement enhanced metrology with quantum many-body scars, Phys. Rev. B **107**, 035123 (2023).

[69] A. Yoshinaga, Y. Matsuzaki, and R. Hamazaki, Quantum Metrology Protected by Hilbert Space Fragmentation (2022), arXiv:2211.09567 [cond-mat, physics:quant-ph].

[70] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, Probing many-body dynamics on a 51-atom quantum simulator, Nature **551**, 579 (2017).

[71] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, Weak ergodicity breaking from quantum many-body scars, Nat. Phys. **14**, 745 (2018).

[72] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, Quantum scarred eigenstates in a Rydberg atom chain: Entanglement, breakdown of thermalization, and stability to perturbations, Phys. Rev. B **98**, 155134 (2018).

[73] C.-J. Lin and O. I. Motrunich, Exact Quantum Many-Body Scar States in the Rydberg-Blockaded Atom Chain, Phys. Rev. Lett. **122**, 173401 (2019).

[74] S. Choi, C. J. Turner, H. Pichler, W. W. Ho, A. A. Michailidis, Z. Papić, M. Serbyn, M. D. Lukin, and D. A. Abanin, Emergent SU(2) dynamics and perfect quantum many-body scars, Phys. Rev. Lett. **122**, 220603 (2019).

[75] L. Piroli, G. Styliaris, and J. I. Cirac, Approximating many-body quantum states with quantum circuits and measurements (2024), arXiv:2403.07604 [cond-mat, physics:quant-ph].

Supplemental Material for "Covariant Quantum Error-correcting Codes with Metrological Entanglement Advantage"

Cheng-Ju Lin[1,2], Zi-Wen Liu[3,4], Victor V. Albert[1], Alexey V. Gorshkov[1,2]

[1] *Joint Center for Quantum Information and Computer Science,*
NIST/University of Maryland, College Park, Maryland 20742, USA
[2] *Joint Quantum Institute, NIST/University of Maryland, College Park, Maryland 20742, USA*
[3] *Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, China*
[4] *Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*

This Supplemental Material is organized as follows. In Sec. S1, we list the definitions of various quantities used throughout the paper. We then present the proofs of Lemma 1 (Sec. S2), Lemma 2 (Sec. S3), Theorem 1 & Corollary 1 (Sec. S4), Lemma 3 (Sec. S5), and Theorem 2 (Sec. S6). In particular, Lemma 1 pertains to the "off-diagonal" part of the Knill-Laflamme conditions, while Lemma 2 bounds the "diagonal" part. Theorem 1 and Corollary 1 give a bound on the code inaccuracy against a generic d-local noise channel. Lemma 3 pertains to the fidelity between the reduced density matrices from code words of different magnetic quantum numbers. Theorem 2 describes an upper bound on the code inaccuracy against a heralded d-local erasure error. In Sec. S7, we provide the details behind the numerical calculation of the quantum Fisher information and behind the measurement protocols described in the main text. Finally, in Sec. S8, we show that the quantum many-body scar states in the spin-1 XY-Dzyaloshinskii-Moriya model can also form an covariant AQECC.

S1. A LIST OF DEFINITIONS

In this section, we list the definitions used throughout the paper.

Definition 1. *(Trace norm)* $\|\rho\|_1 := Tr\sqrt{\rho^\dagger \rho}$,

which induces the trace distance between two density matrices ρ and σ as $\|\rho - \sigma\|_1$. The trace distance then induces a distance measure between quantum channels $\mathcal{E}$ and $\mathcal{F}$:

Definition 2. *(Diamond distance)* $\|\mathcal{E} - \mathcal{F}\|_\diamond := \max_{|\psi\rangle} \|\mathcal{E} \otimes \mathrm{id}(|\psi\rangle\langle\psi|) - \mathcal{F} \otimes \mathrm{id}(|\psi\rangle\langle\psi|)\|_1$,

where $|\psi\rangle$ is a purification of ρ and id is the identity channel on the auxiliary space whose input has the same Hilbert space dimension as the input of $\mathcal{E}$ and $\mathcal{F}$.

Definition 3. *(Fidelity)* $f(\rho, \sigma) := \|\sqrt{\sigma}\sqrt{\rho}\|_1 = Tr\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$,

which induces the entanglement fidelity between channels $\mathcal{E}$ and $\mathcal{F}$:

Definition 4. *(Entanglement fidelity)* $F(\mathcal{E}, \mathcal{F}) := \min_{|\psi\rangle} f(\mathcal{E} \otimes \mathrm{id}(|\psi\rangle\langle\psi|), \mathcal{F} \otimes \mathrm{id}(|\psi\rangle\langle\psi|))$,

where again $|\psi\rangle$ is a purification of ρ and id is the identity channel on the auxiliary space whose input has the same Hilbert space dimension as the input of $\mathcal{E}$ and $\mathcal{F}$. The purified channel distance is defined as follows:

Definition 5. *(Purified channel distance)* $D_P(\mathcal{E}, \mathcal{F}) := \sqrt{1 - F(\mathcal{E}, \mathcal{F})^2}$.

Note that one can also define another channel distance measure, called Bures distance as follows:

Definition 6. *(Bures channel distance)* $D_B(\mathcal{E}, \mathcal{F}) := \sqrt{1 - F(\mathcal{E}, \mathcal{F})}$.

In Ref. [18], the code inaccuracy is quantified by the Bures distance. On the other hand, in this work, we opt for using purified distance. These channel distances are related, however, via the following Fuchs-van de Graaf inequality [61]:

$$1 - f(\rho, \sigma) \le \frac{1}{2}\|\rho - \sigma\|_1 \le \sqrt{1 - f(\rho, \sigma)^2} \ . \tag{S1}$$

This immediately gives us

$$D_B(\mathcal{E}, \mathcal{F})^2 \le \frac{1}{2}\|\mathcal{E} - \mathcal{F}\|_\diamond \le D_P(\mathcal{E}, \mathcal{F}) \ . \tag{S2}$$

S2. PROOF OF LEMMA 1

In this section, we present the proof of Lemma 1, which pertains to the "off-diagonal" part of the Knill-Laflamme conditions.

Lemma 1: Assume $\hat{F}$ is a d-local operator, then $\langle J, n|\hat{F}|J, m\rangle = 0$ if $|n - m| \geq 2sd + 1$.

Proof. Recall that $\hat{Q}^z|J, n\rangle = n|J, n\rangle$, where $\hat{Q}^z = \sum_{j=1}^{N} \hat{q}_j^z$. Since we consider a system of spin-s degrees of freedom, the charge $\hat{q}^z$ can have eigenvalues $q^z = -s, -s + 1, \cdots, s$. Therefore, a d-local operator can change the total charge by at most $2sd$. So we have $\langle J, n|\hat{F}|J, m\rangle = 0$, if $|n - m| \geq 2sd + 1$.

$\square$

S3. PROOF OF LEMMA 2

In this section, we present the proof of Lemma 2, which bounds the "diagonal" part of the Knill-Laflamme conditions.

Lemma 2: Assume $\hat{F}$ is a d-local operator, then $|\langle J, n|\hat{F}|J, n\rangle - \langle J, m|\hat{F}|J, m\rangle| \leq dq_0\|\hat{F}\|_{op}C(n, m)$, where $C(n, m) := \sum_{M=n}^{m-1}[c_M^+]^{-1}$ and q_0 is a constant. In particular, for $0 \leq b < a \leq 1$, if we take $J \sim N^a$ and $|m - n| \sim N^b$, then we have $|\langle J, n|\hat{F}|J, n\rangle - \langle J, m|\hat{F}|J, m\rangle| = \|\hat{F}\|_{op}O(dN^{b-a})$.

Proof. Assume I is the set of qudits $\hat{F}$ acts on. We then have

$$\|[\hat{F}, \hat{Q}^+]\|_{op} \leq \|[\hat{F}, \sum_{j\in I} \hat{q}_j^+]\|_{op} \leq dq_0\|\hat{F}\|_{op} , \tag{S3}$$

where we denoted $q_0 = 2\|\hat{q}_j^+\|_{op}$ for convenience, which is a constant. To bound $|\langle J, n + 1|\hat{F}|J, n + 1\rangle - \langle J, n|\hat{F}|J, n\rangle|$, note that

$$\langle J, n + 1|\hat{F}|J, n + 1\rangle = \frac{1}{(c_n^+)^2}\langle J, n|\hat{Q}^-\hat{F}\hat{Q}^+|J, n\rangle = \frac{1}{(c_n^+)^2}\langle J, n|\hat{Q}^-\hat{Q}^+\hat{F}|J, n\rangle + \frac{1}{c_n^+}\langle J, n + 1|[\hat{F}, \hat{Q}^+]|J, n\rangle . \tag{S4}$$

However, recall $c_n^\pm = \sqrt{(J \mp n)(J \pm n + 1)}$, we also have

$$\frac{1}{(c_n^+)^2}\langle J, n|\hat{Q}^-\hat{Q}^+\hat{F}|J, n\rangle = \frac{c_{n+1}^-}{c_n^+}\langle J, n|\hat{F}|J, n\rangle = \langle J, n|\hat{F}|J, n\rangle , \tag{S5}$$

which gives us

$$|\langle J, n + 1|\hat{F}|J, n + 1\rangle - \langle J, n|\hat{F}|J, n\rangle| \leq dq_0\|\hat{F}\|_{op} \cdot (c_n^+)^{-1} , \tag{S6}$$

or

$$\begin{aligned}|\langle J, m|\hat{F}|J, m\rangle - \langle J, n|\hat{F}|J, n\rangle| &= \left|\sum_{\ell=n}^{m-1}\langle J, \ell + 1|\hat{F}|J, \ell + 1\rangle - \langle J, \ell|\hat{F}|J, \ell\rangle\right| \\ &\leq \sum_{\ell=n}^{m-1}|\langle J, \ell + 1|\hat{F}|J, \ell + 1\rangle - \langle J, \ell|\hat{F}|J, \ell\rangle| \\ &\leq dq_0\|\hat{F}\|_{op} \cdot C(m, n) , \end{aligned} \tag{S7}$$

where $C(m, n) = \sum_{\ell=n}^{m-1}[c_\ell^+]^{-1} = \sum_{\ell=n}^{m-1}[(J - \ell)(J + \ell + 1)]^{-1/2}$ and we used the triangle inequality in the first inequality.

We can obtain a bound on $C(n, m)$ as follows. Consider the function $g(x) := [(J - x)(J + x + 1)]^{-1/2}$ in the domain $x \in (-J - 1, J)$. It is easy to see that the minimum of $g(x)$ is at $x^* = -1/2$, and $g(x)$ is a decreasing function from $x = -J - 1$ to $x = x^*$ and an increasing function from $x = x^*$ to $x = J$. It is also easy to check that $g(x) = g(2x^* - x)$. We first consider the case $n < 0 < m$. We separate $C(n, m)$ into two pieces, $C(n, m) = C_1(n) + C_2(m)$, where $C_1(n) := \sum_{\ell=n}^{\ell=-1}[c_\ell^+]^{-1}$ and $C_2(m) := \sum_{\ell=0}^{m-1}[c_\ell^+]^{-1}$. These two pieces can be bounded as

$$\int_n^0 g(x)dx < C_1(n) < \int_n^0 g(x - 1)dx ,$$

$$\int_0^m g(x - 1)dx < C_2(m) < \int_0^m g(x)dx . \tag{S8}$$

The integrals of the upper bounds are given by $\int_0^m g(x)dx = 2\arcsin\left(\sqrt{\frac{J}{2J+1}}\right) - 2\arcsin\left(\sqrt{\frac{J-m}{2J+1}}\right)$ and $\int_n^0 dx\, g(x-1) = \int_n^0 dx\, g(-x) = \int_0^{-n} dy\, g(y) = 2\arcsin\left(\sqrt{\frac{J}{2J+1}}\right) - 2\arcsin\left(\sqrt{\frac{J+n}{2J+1}}\right)$.

We are interested in the regime where $n/J \to 0$, $m/J \to 0$, and $J \sim N^a$ ($a \le 1$) when $N \to \infty$. Consider the function

$$f(\tilde{x}) = 2\arcsin\left(\sqrt{\frac{1+\tilde{x}}{2+J^{-1}}}\right) , \tag{S9}$$

where $\tilde{x} = x/J$. Its Taylor series is $f(\tilde{x}) - f(0) = f'(0)\tilde{x} + O(f''(0)\tilde{x}^2)$, where $f(0) = 2\arcsin\left(\sqrt{\frac{J}{2J+1}}\right)$, $f'(0) = \sqrt{\frac{J}{1+J}}$ and $f''(0) = -\frac{J^{\frac{1}{2}}}{2(1+J)^{\frac{3}{2}}}$. Therefore, we have

$$\begin{aligned}
2\arcsin\left(\sqrt{\frac{J}{2J+1}}\right) - 2\arcsin\left(\sqrt{\frac{J+x}{2J+1}}\right) &= -\frac{x}{\sqrt{J(J+1)}} + O\left(\frac{x^2}{[J(1+J)]^{\frac{3}{2}}}\right) \\
&= -\frac{x}{J}(1+J^{-1})^{-\frac{1}{2}} + O\left(\frac{x^2}{J^3}(1+J^{-1})^{-\frac{3}{2}}\right) \\
&= -\frac{x}{J} + O(xJ^{-2}) . \tag{S10}
\end{aligned}$$

Therefore, for $n < 0 < m$, we have

$$C(m,n) = \frac{|m-n|}{J} + O(|m-n|J^{-2}) . \tag{S11}$$

If $0 < n < m$, then we only need to bound $\sum_{\ell=n}^{m-1}[c_\ell^+]^{-1} < \int_n^m g(x)dx = \int_0^m g(x)dx - \int_0^n g(x)dx = |m-n|J^{-1} + O(|m-n|J^{-2})$; if $n < m < 0$, we bound $\sum_{\ell=n}^{m-1}[c_\ell^+]^{-1} < \int_n^m g(x-1)dx = \int_{-m}^{-n} g(x)dx = |m-n|J^{-1} + O(|m-n|J^{-2})$.

Therefore, for $J \sim N^a$ and $|m-n| \sim N^b$, where $b < a \le 1$, we have

$$|\langle J,m|\hat{F}|J,m\rangle - \langle J,n|\hat{F}|J,n\rangle| = \|\hat{F}\|_{op} O(dN^{b-a}) . \tag{S12}$$

$\square$

S4. PROOF OF THEOREM 1 AND COROLLARY 1

In this section, we present the proof of Theorem 1, which gives an upper bound on the code inaccuracy against a generic d-local noise channel.

Theorem 1: Consider the code space $\mathfrak{C} = \mathrm{span}\{|J,M\rangle; M = M_{\min}, M_{\min}+\Delta, M_{\min}+2\Delta, \ldots, M_{\max}\}$, where $J \sim N^a$, $M_{\max} - M_{\min} \sim N^b$, and $b < a \le 1$. $\mathfrak{C}$ forms an $((N, O(b\log_2 N)))$ AQECC against a d-local noise channel with $\epsilon(d) = O(\sqrt{d}(2s+1)^d N^{(b-a)/2})$ if we take $\Delta \ge 2sd + 1$.

Proof. The proof of this theorem uses the Knill-Laflamme condition for the AQECC developed by Bény and Oreshkov in Ref. [18], which we now state.

(Bény and Oreshkov) Consider a noise channel with the Kraus representation $\mathcal{N}(\bullet) = \sum_j K_j(\bullet)K_j^\dagger$, and let Π be the projector onto the code space. Then the code is ϵ-correctable (i.e., there exists a recovery channel $\mathcal{R}$ such that $D_P(\mathcal{R}\mathcal{N}, id) \le \epsilon$) if and only if

$$\Pi K_i^\dagger K_j \Pi = \lambda_{ij}\Pi + \Pi B_{ij}\Pi , \tag{S13}$$

where λ_{ij} are the components of a density operator, and $D_P(\Lambda + \mathcal{B}, \Lambda) \le \epsilon$, where $\Lambda(\rho) = \sum_{i,j} \lambda_{ij} Tr(\rho)|i\rangle\langle j|$ and $(\Lambda + \mathcal{B})(\rho) = \Lambda(\rho) + \sum_{i,j} Tr(\rho B_{ij})|i\rangle\langle j|$.

This theorem in fact has an interpretation from the complementary channel point of view, as described in Ref. [18]. More specifically, consider the noise channel $\mathcal{N}(\rho) = \sum_i K_i(\rho)K_i^\dagger$, whose complementary channel is $\hat{\mathcal{N}}(\rho) = \sum_{i,j} Tr(K_j\rho K_i^\dagger)|i\rangle\langle j| = (\Lambda + \mathcal{B})(\rho)$, where $|i\rangle\langle j|$ acts on the environment Hilbert space. If we have perfect Knill-Laflamme conditions, the complementary channel of the noise becomes $\hat{\mathcal{N}}(\rho) = \sum_{i,j} \lambda_{ij}|i\rangle\langle j| := \lambda_0 = \Lambda(\rho)$,

which is a fixed density matrix λ_0 independent of the input ρ. In other words, the condition for having a perfect quantum error-correcting code is equivalent to the condition that the environment gains no information from the code space. And we indeed see that the AQECC generalization of the Knill-Laflamme conditions is a perturbation from the complementary channel point of view, where the environment can gain some ϵ-small information from the code space. Note that, in the original theorem in Ref. [18], Bures distance $D_B(\mathcal{N}, \mathcal{M})$ is used as the measure of channel distance, while here we opt for using the purified distance $D_P(\mathcal{N}, \mathcal{M})$, where $D_B(\mathcal{N}, \mathcal{M})^2 \leq D_P(\mathcal{N}, \mathcal{M})$ can be obtained using Eq. (S2).

Here we consider the noise channel $\mathcal{N}(\rho) = \sum_{i=1}^{n_d} K_i(\rho) K_i^\dagger$, where $n_d = (2s+1)^{2d}$ is the number of independent Kraus operators. We assume K_j is at most d-local, and that all K_j act on the same set of d qudits, so $K_i^\dagger K_j$ is at most d-local as well. We are also interested in the regime $J \sim N^a$ and $(M_{\max} - M_{\min}) \sim N^b$. From *Lemma 1*, for $|J,n\rangle, |J,m\rangle \in \mathfrak{C}$ and $n \neq m$, we have $\langle J,n|K_i^\dagger K_j|J,m\rangle = 0$ if $\Delta \geq 2sd + 1$. Now assuming we take $\lambda_{ij} = \langle J,M_{\min}|K_i^\dagger K_j|J,M_{\min}\rangle$, we then have $\Lambda(\rho) = \sum_{i,j=1}^{N_d} \lambda_{ij} \text{Tr}[\rho]|i\rangle\langle j|$ and $(\Lambda + \mathcal{B})(\rho) = \sum_{i,j=1}^{N_d} \text{Tr}[K_j \rho K_i^\dagger]|i\rangle\langle j|$, where the inputs of both channels are from the logical space.

Our goal now is to bound the entanglement fidelity, $F(\Lambda, \Lambda + \mathcal{B}) := \min_{|\psi\rangle} f((\Lambda + \mathcal{B}) \otimes \text{id}(|\psi\rangle\langle\psi|), \Lambda \otimes \text{id}(|\psi\rangle\langle\psi|))$. Recall that the inputs of Λ and $\Lambda + \mathcal{B}$ are from the logical space, and the auxiliary space is isomorphic to the input space. Let us use $|x\rangle := V^\dagger|J,M_x\rangle$ for $x = 1, \cdots, 2^k$ to denote the basis of the logical space for convenience, where $|J,M_x\rangle \in \mathfrak{C}$, and V is the isometric encoding such that $VV^\dagger = \Pi$ is the projection onto the code space.

For any state $|\psi\rangle = \sum_{x,y} A_{xy}|x\rangle \otimes |y\rangle$, if we consider the operators $A = \sum_{x,y} A_{xy}|x\rangle\langle y|$ and $B := A^T = \sum_{x,y} A_{xy}|y\rangle\langle x|$ and the state $|\Phi\rangle = \sum_{x=1}^{2^k}|x\rangle \otimes |x\rangle$, then we can express $|\psi\rangle = (A \otimes \text{id})|\Phi\rangle = (\text{id} \otimes B)|\Phi\rangle$. We can therefore rewrite $\mathcal{E} \otimes \text{id}(|\psi\rangle\langle\psi|) = (\text{id} \otimes B)(\mathcal{E} \otimes \text{id})(|\Phi\rangle\langle\Phi|)(\text{id} \otimes B^\dagger)$ for $\mathcal{E} = \Lambda$ and $\mathcal{E} = \Lambda + \mathcal{B}$. Moreover,

$$(\Lambda + \mathcal{B}) \otimes \text{id}(|\Phi\rangle\langle\Phi|) = \sum_{x=1}^{2^k} \sum_{i,j=1}^{n_d} \langle J,M_x|K_i^\dagger K_j|J,M_x\rangle \cdot |i\rangle\langle j| \otimes |x\rangle\langle x| := \sum_{x=1}^{2^k} \sum_{i,j=1}^{n_d} \Sigma_{ij}^x \cdot |i\rangle\langle j| \otimes |x\rangle\langle x|, \tag{S14}$$

$$\Lambda \otimes \text{id}(|\Phi\rangle\langle\Phi|) = \sum_{x=1}^{2^k} \sum_{i,j=1}^{n_d} \langle J,M_{\min}|K_i^\dagger K_j|J,M_{\min}\rangle \cdot |i\rangle\langle j| \otimes |x\rangle\langle x| = \sum_{x=1}^{2^k} \sum_{i,j=1}^{n_d} \lambda_{ij} \cdot |i\rangle\langle j| \otimes |x\rangle\langle x|, \tag{S15}$$

where we have used Lemma 1 and denoted $\Sigma_{ij}^x = \langle J,M_x|K_i^\dagger K_j|J,M_x\rangle$.

Note that the normalization condition $\langle\psi|\psi\rangle = \sum_x \langle x|B^\dagger B|x\rangle = 1$, so we can interpret $\langle x|B^\dagger B|x\rangle$ as a probability. The fidelity can be rewritten as

$$f((\Lambda + \mathcal{B}) \otimes \text{id}(|\psi\rangle\langle\psi|), \Lambda \otimes \text{id}(|\psi\rangle\langle\psi|))$$

$$= f\left(\sum_{x=1}^{2^k}\langle x|B^\dagger B|x\rangle \cdot \sum_{i,j=1}^{n_d} \Sigma_{ij}^x|i\rangle\langle j| \otimes \frac{B|x\rangle\langle x|B^\dagger}{\langle x|B^\dagger B|x\rangle}, \sum_{x=1}^{2^k}\langle x|B^\dagger B|x\rangle \cdot \sum_{i,j=1}^{n_d} \lambda_{ij}|i\rangle\langle j| \otimes \frac{B|x\rangle\langle x|B^\dagger}{\langle x|B^\dagger B|x\rangle}\right)$$

$$\geq \sum_{x=1}^{2^k}\langle x|B^\dagger B|x\rangle \cdot f\left(\sum_{i,j=1}^{n_d} \Sigma_{ij}^x|i\rangle\langle j| \otimes \frac{B|x\rangle\langle x|B^\dagger}{\langle x|B^\dagger B|x\rangle}, \sum_{i,j=1}^{n_d} \lambda_{ij}|i\rangle\langle j| \otimes \frac{B|x\rangle\langle x|B^\dagger}{\langle x|B^\dagger B|x\rangle}\right)$$

$$= \sum_{x=1}^{2^k}\langle x|B^\dagger B|x\rangle \cdot f\left(\sum_{i,j=1}^{n_d} \Sigma_{ij}^x|i\rangle\langle j|, \sum_{i,j=1}^{n_d} \lambda_{ij}|i\rangle\langle j|\right), \tag{S16}$$

where we have used the joint concavity of the fidelity in the first inequality.

We will use the Fuchs-van de Graaf inequality Eq. (S1) to bound the fidelity via the trace norm. To this end, we would like to bound $\|\Sigma_x - \lambda_0\|_1$, where $\Sigma_x := \sum_{i,j=1}^{N_d} \Sigma_{ij}^x|i\rangle\langle j|$ and $\lambda_0 := \sum_{i,j=1}^{N_d} \lambda_{ij}|i\rangle\langle j|$. Note that since $\Sigma_x^\dagger = \Sigma_x$ and $\lambda_0^\dagger = \lambda_0$, for each x, we can find a unitary transformation U_x that diagonalizes $\Sigma_x - \lambda_0$, giving us $\gamma_i^x \delta_{i\ell} = \sum_{j,k} [U_x^\dagger]_{ij}(\Sigma_{jk}^x - \lambda_{jk})[U_x]_{k\ell}$. Furthermore,

$$\gamma_i^x = \sum_{j,k} [U_x^\dagger]_{ij}(\langle J,M|K_j^\dagger K_k|J,M\rangle - \langle J,M_{\min}|K_j^\dagger K_k|J,M_{\min}\rangle)[U_x]_{ki}$$

$$:= \langle J,M|F_i^\dagger F_i|J,M\rangle - \langle J,M_{\min}|F_i^\dagger F_i|J,M_{\min}\rangle, \tag{S17}$$

where $F_i := \sum_j K_j[U_x]_{ji}$ is at most d-local and $\sum_i F_i^\dagger F_i = I$. (Note that F_i depends on x in general, though here we suppress it to simplify the notation.) We therefore have $|\gamma_i^x| = \|F_i\|_{op} \cdot O(dN^{b-a}) = O(dN^{b-a})$ from Lemma 2 and the

fact that $\|F_i\|_{op} \leq 1$. This enables us to bound $\|\Sigma_x - \rho_0\|_1 = \|\sum_{i=1}^{n_d} \gamma_i^x |i\rangle\langle i|\|_1 = O(dn_d N^{b-a})$, which is independent of x. Using the Fuchs-van de Graaf inequality Eq. (S1), or $f(\Sigma_x, \lambda_0) \geq 1 - \frac{1}{2}\|\Sigma_x - \lambda_0\|_1 = 1 - O(n_d d N^{b-a})$, we have

$$
\begin{aligned}
f((\Lambda + \mathcal{B}) \otimes \mathrm{id}(|\psi\rangle\langle\psi|), \Lambda \otimes \mathrm{id}(|\psi\rangle\langle\psi|)) &\geq \sum_{x=1}^{2^k} \langle x|B^\dagger B|x\rangle \cdot f\left(\sum_{i,j=1}^{n_d} \Sigma_{ij}^x |i\rangle\langle j|, \sum_{i,j=1}^{n_d} \lambda_{ij} |i\rangle\langle j|\right) \\
&\geq \sum_{x=1}^{2^k} \langle x|B^\dagger B|x\rangle \cdot [1 - O(dn_d N^{b-a})] \\
&= 1 - O(dn_d N^{b-a}) \ .
\end{aligned}
\tag{S18}
$$

Since the above inequality is independent of $|\psi\rangle$, we have $F(\Lambda, \Lambda + \mathcal{B}) = 1 - O(dn_d N^{b-a})$. We therefore have

$$
\epsilon = D_P(\Lambda, \Lambda + \mathcal{B}) = \sqrt{1 - F(\Lambda, \Lambda + \mathcal{B})^2} = O(\sqrt{dn_d} N^{(b-a)/2}) = O(\sqrt{d}(2s+1)^d N^{(b-a)/2}) \ ,
\tag{S19}
$$

where we have used $n_d = (2s+1)^{2d}$. Since $k = \log_2[\frac{M_{\max} - M_{\min}}{\Delta} + 1]$, and $\Delta \geq 2sd + 1$, we have $k = O(b \log_2 N)$ for $|M_{\max} - M_{\min}| \sim N^b$.

$\square$

We see that the reason to assume the Kraus operators K_i all act on the same set of qudits is so that the number of Kraus operator $n_d = (2s+1)^{2d}$ can be controlled. In general, if we consider the noise channel $\mathcal{N}(\rho) = \sum_{i=1}^{n_d} K_i \rho K_i^\dagger$, where K_i can act on different sets of at most d qudits, then we have the following corollary.

Corollary 1. *Consider the code space* $\mathfrak{C} = span\{|J,M\rangle; M = M_{\min}, M_{\min}+\Delta, M_{\min}+2\Delta, \ldots, M_{\max}\}$, *where* $J \sim N^a$, $M_{\max} - M_{\min} \sim N^b$, *and* $b < a \leq 1$. $\mathfrak{C}$ *forms an* $((N, O(b \log_2 N)))$ *AQECC against a d-local noise channel with* $\epsilon(d) = O(\sqrt{d}n_d N^{(b-a)/2})$ *if we take* $\Delta \geq 4sd + 1$.

Proof. The proof follows from the proof of Theorem 1. The difference is that, here, $K_i^\dagger K_j$ is at most 2d-local instead of d-local, so we need to take $\Delta \geq 4sd + 1$. Moreover, since $F_i = \sum_j K_j [U_x]_{ji}$ is not necessary d-local any more, we can only bound

$$
\|\Sigma_x - \lambda_0\|_1 = O(dn_d^2 N^{(b-a)}) \ ,
\tag{S20}
$$

since there are n_d^2 terms. This gives us

$$
D_P(\Lambda, \Lambda + \mathcal{B}) = O(\sqrt{d}n_d N^{(b-a)/2}) \ .
\tag{S21}
$$

$\square$

S5. PROOF OF LEMMA 3

In this section, we present the proof of Lemma 3, which bounds the fidelity between the reduced density matrices from code words of different magnetic quantum numbers. This proof is in fact mostly parallel to the proof of Lemma 14 in Ref. [20] but with a different and generalized physical interpretation, enabled by the usage of Clebsch-Gordan coefficients.

Lemma 3: Consider the irrep of $SU(2)$ *formed by* $|J,M\rangle$, *where* $J = sN$, *and the reduced density matrix* $\rho_M := Tr_{\bar\alpha}[|J,M\rangle\langle J,M|]$, *where* $\bar\alpha$ *is the complement of the set of erased qudits* α. *We then have the fidelity* $f(\rho_M, \rho_{M=0}) = 1 - O(\frac{dM^2}{sN^2}) = 1 - O(N^{2b+c-2})$ *asymptotically for* $M \sim N^b$, $d \sim N^c$, *and* $1 > b \geq c$.

Proof. Note that, since $|J,M\rangle = \sum_{m_1=-j_1}^{j_1} C_{j_1,m_1,j_2,m_2}^{JM} |j_1,m_1\rangle \otimes |j_2,m_2\rangle$, where $C_{j_1,m_1,j_2,m_2}^{JM} = (\langle j_1, m_1| \otimes \langle j_2, m_2|)|J,M\rangle$ is the Clebsch-Gordan coefficient, we have $\rho_M = \sum_{m_1=-j_1}^{j_1} (C_{j_1,m_1,j_2,M-m_1}^{JM})^2 |j_1 m_1\rangle\langle j_1 m_1|$, where $j_1 = sd$ and $j_2 = J - j_1$, since $J = sN$ is the irrep with the highest possible total spin quantum number with multiplicity one. The fidelity between ρ_M and $\rho_{M=0}$ is therefore $f(\rho_M, \rho_{M=0}) = \sum_{m_1=-j_1}^{j_1} C_{j_1,m_1,j_2,M-m_1}^{JM} C_{j_1,m_1,j_2,-m_1}^{JM=0}$.

For $J = j_1 + j_2$, the Clebsch-Gordan coefficient is

$$
C_{j_1,m_1,j_2,m_2}^{JM} = \delta_{M=m_1+m_2} \left[\frac{\binom{2j_1}{j_1+m_1}\binom{2J-2j_1}{J-j_1+M-m_1}}{\binom{2J}{J+M}}\right]^{\frac{1}{2}} \ ,
\tag{S22}
$$

where $\binom{N}{k} = \frac{N!}{k!(N-k)!}$ is the binomial coefficient. The fidelity is therefore

$$f(\rho_M, \rho_{M=0}) = \left[\binom{2J}{J+M}\binom{2J}{J}\right]^{-\frac{1}{2}} \sum_{m_1=-j_1}^{j_1} \binom{2j_1}{j_1+m_1}\left[\binom{2J-2j_1}{J-j_1+M-m_1}\binom{2J-2j_1}{J-j_1-m_1}\right]^{\frac{1}{2}}. \tag{S23}$$

We are after the asymptotic behavior of the fidelity in the powers of J (equivalently N), expecting $M \sim J^b$ and $j_1 \sim J^c$. Using Stirling's approximation $\ln(N!) = N \ln N - N + \frac{1}{2}\ln(2\pi N) + \frac{1}{12N} + O(N^{-3})$, we have

$$\ln\binom{2R}{R+\Delta} = 2R\ln(2R) - 2R + \frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln(2R) + \frac{1}{24R} + O(R^{-3})$$

$$- (R+\Delta)\ln(R+\Delta) + (R+\Delta) - \frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(R+\Delta) - \frac{1}{12(R+\Delta)} + O((R+\Delta)^{-3})$$

$$- (R-\Delta)\ln(R-\Delta) + (R-\Delta) - \frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(R-\Delta) - \frac{1}{12(R-\Delta)} + O((R-\Delta)^{-3}). \tag{S24}$$

Expecting $\Delta/R \to 0$ when $R \to \infty$, but also the possibility that $R \sim N$ and $\Delta \sim N^b$ $(b < 1)$, we use

$$\ln(R+\Delta) = \ln R + \ln(1 + \frac{\Delta}{R}) = \ln R + \frac{\Delta}{R} - \frac{\Delta^2}{2R^2} + \frac{\Delta^3}{3R^3} + O(\Delta^4 R^{-4}) \tag{S25}$$

$$\ln(R-\Delta) = \ln R + \ln(1 - \frac{\Delta}{R}) = \ln R - \frac{\Delta}{R} - \frac{\Delta^2}{2R^2} - \frac{\Delta^3}{3R^3} + O(\Delta^4 R^{-4}) \tag{S26}$$

$$(R\pm\Delta)^{-1} = \frac{1}{R} \mp \frac{\Delta}{R^2} + O(\Delta^2 R^{-3}), \tag{S27}$$

and obtain

$$\ln\binom{2R}{R+\Delta} = (2\ln 2)R - \frac{1}{2}\ln R - \frac{1}{2}\ln\pi - \frac{1}{8R} - \frac{\Delta^2}{R} + \frac{\Delta^2}{2R^2} + O(\Delta^4 R^{-3}), \tag{S28}$$

or

$$\binom{2R}{R+\Delta} = \frac{2^{2R}}{\sqrt{\pi R}}\exp\left(-\frac{1}{8R} - \frac{\Delta^2}{R} + \frac{\Delta^2}{2R^2} + O(\Delta^4 R^{-3})\right). \tag{S29}$$

Using Eq. (S29), we have

$$\left[\binom{2J}{J+M}\binom{2J}{J}\right]^{-\frac{1}{2}} = \frac{2^{2J}}{\sqrt{\pi J}}\exp\left(-\frac{1}{8J} - \frac{M^2}{J} + \frac{M^2}{J^2} + O(M^4 J^{-3})\right). \tag{S30}$$

Applying Eq. (S28) to $\binom{2J-2j_1}{J-j_1+M-m_1}$ and $\binom{2J-2j_1}{J-j_1-m_1}$, we have

$$\ln\binom{2J-2j_1}{J-j_1+M-m_1} = (2\ln 2)(J-j_1) - \frac{1}{2}\ln(J-j_1) + \frac{1}{2}\ln\pi - \frac{1}{8(J-j_1)} - \frac{(M-m_1)^2}{(J-j_1)} + \frac{(M-m_1)^2}{2(J-j_1)^2} + O(J^{-3}), \tag{S31}$$

$$\ln\binom{2J-2j_1}{J-j_1-m_1} = (2\ln 2)(J-j_1) - \frac{1}{2}\ln(J-j_1) + \frac{1}{2}\ln\pi - \frac{1}{8(J-j_1)} - \frac{m_1^2}{(J-j_1)} + \frac{m_1^2}{2(J-j_1)^2} + O(J^{-3}). \tag{S32}$$

Using

$$(J-j_1)^{-1} = J^{-1}\left(1 + \frac{j_1}{J} + O(J^{-2})\right),$$

$$(J-j_1)^{-2} = J^{-2} + O(J^{-3}), \tag{S33}$$

we have

$$\left[\binom{2J-2j_1}{J-j_1+M-m_1}\binom{2J-2j_1}{J-j_1-m_1}\right]^{\frac{1}{2}} = \frac{2^{2(J-j_1)}}{\sqrt{\pi(J-j_1)}}\exp\left(-\frac{1}{8J} - \frac{j_1}{8J^2} - \frac{(M-m_1)^2+m_1^2}{2J}\right.$$

$$\left. + \frac{(M-m_1)^2+m_1^2}{4J^2}(1-2j_1) + O(J^{-3})\right). \tag{S34}$$

Combining the above equations together, using $(1 - j_1/J)^{-1/2} = (1 + j_1/2J + 3j_1^2/8J^2 + O(J^{-3}))$, and expanding the equation up to $O(J^{-3})$, we have

$$f(\rho_M, \rho_{M=0}) = \frac{2^{-2j_1}}{\sqrt{1 - j_1/J}} \exp\left(-\frac{j_1}{8J^2} - \frac{j_1 M^2}{2J^2} + O(J^{-3})\right)$$

$$\times \sum_{m_1=-j_1}^{j_1} \binom{2j_1}{j_1 + m_1} \exp\left(\frac{Mm_1 - m_1^2}{J} - \frac{Mm_1 - m_1^2}{2J^2}(1 - 2j_1)\right)$$

$$= 2^{-2j_1}\left(1 + \frac{j_1}{2J} + \frac{3j_1^2}{8J^2} + O(J^{-3})\right)\left(1 - \frac{j_1(1 + 4M^2)}{8J^2} + O(J^{-3})\right)$$

$$\times \sum_{m_1=-j_1}^{j_1} \binom{2j_1}{j_1 + m_1}\left(1 + \frac{Mm_1 - m_1^2}{J} + \frac{(Mm_1 - m_1^2)^2}{2J^2} - \frac{(Mm_1 - m_1^2)}{2J^2}(1 - 2j_1) + O(J^{-3})\right)$$

$$= 2^{-2j_1} \sum_{m_1=-j_1}^{j_1} \binom{2j_1}{j_1 + m_1}(1 + J^{-1}A_{M,j_1,m_1} + J^{-2}B_{M,j_1,m_1}) , \tag{S35}$$

where

$$A_{M,j_1,m_1} = \frac{1}{2}j_1 + Mm_1 - m_1^2 ,$$

$$B_{M,j_1,m_1} = \frac{3j_1^2}{8} - \frac{j_1}{8} - \frac{1}{2}M^2 j_1 + \frac{1}{2}j_1(Mm_1 - m_1^2) + \frac{1}{2}(Mm_1 - m_1)^2 + (j_1 - \frac{1}{2})(Mm_1 - m_1^2) . \tag{S36}$$

Upon the change of variables $r = j_1 + m_1$, using equations

$$\sum_{r=0}^{2j_1} \binom{2j_1}{r} = 2^{2j_1} ,$$

$$\sum_{r=0}^{2j_1} \binom{2j_1}{r}r = j_1 2^{2j_1} ,$$

$$\sum_{r=0}^{2j_1} \binom{2j_1}{r}r^2 = (\frac{1}{2}j_1 + j_1^2)2^{2j_1} , \tag{S37}$$

we have $\sum_{r=0}^{2j_1} \binom{2j_1}{r}A_{M,j_1,m_1} = 0$.

In addition, using equations

$$\sum_{r=0}^{2j_1} \binom{2j_1}{r}r^3 = \frac{1}{2}j_1^2(2j_1 + 3)2^{2j_1} ,$$

$$\sum_{r=0}^{2j_1} \binom{2j_1}{r}r^4 = \frac{1}{4}j_1(1 - 3j_1 + 16j_1^2 + 4j_1^3)2^{2j_1} , \tag{S38}$$

we have

$$2^{-2j_1} \sum_{r=0}^{2j_1} \binom{2j_1}{r}B_{M,j_1,m_1} = -\frac{1}{4}j_1 M^2 - \frac{7}{8}j_1^2 + 2j_1^2 M^2 + 4j_1^3 M + \frac{1}{2}j_1^3 + \frac{3j_1}{8} . \tag{S39}$$

So we have

$$f(\rho_M, \rho_{M=0}) = 1 - J^{-2}\left(\frac{1}{4}j_1 M^2 + \frac{7}{8}j_1^2 - 2j_1^2 M^2 - 4j_1^3 M - \frac{1}{2}j_1^3 - \frac{3j_1}{8}\right) + O(J^{-3})$$

$$= 1 - O\left(\frac{j_1 M^2}{J^2}\right) = 1 - O\left(\frac{dM^2}{sN^2}\right) = 1 - O(N^{2b+c-2}) , \tag{S40}$$

where we have used $J = sN$, $M \sim J^b$ and $j_1 = sd \sim J^c$ with $1 > b \geq c$. $\qquad\square$

S6. PROOF OF THEOREM 2

In this section, we present the proof of Theorem 2, which bounds the code inaccuracy against the d-local erasure error. This theorem in fact can be obtained by combining our Lemma 1, our Lemma 3, and Theorem 3 of Ref. [20]. Here we present the proof tailored to our case to make our presentation self-contained.

Theorem 2: Consider the code space $\mathfrak{C} = span\{|J,M\rangle; M = M_{\min}, M_{\min}+\Delta, M_{\min}+2\Delta, \ldots, M_{\max}\}$, where $J = sN$, $M_{\max} - M_{\min} \sim N^b$ and $b < 1$. $\mathfrak{C}$ forms an $((N, O((b-c)\log_2 N)))$ AQECC against d-local erasures of size $d \sim N^c$, where $0 \le c \le b < 1$, with $\epsilon(d) = O(N^{b+c/2-1})$ if we take $\Delta \ge 2sd + 1$.

Proof. The bound of the code inaccuracy against erasure errors can be obtained using Lemma 3 and Theorem 3 in Ref. [20]. (See also Ref. [25].) A potential concern is the applicability of the aforementioned theorem in Ref. [20] when $d \sim N^c$ for some power $c < 1$, as the theorem in Ref. [20] is only addressing the case with $d = O(1)$. To alleviate the above concern and to make our presentation self-contained, we prove our code-inaccuracy bound against erasure errors at known locations.

We consider the erasure errors at known locations expressed as $\mathcal{N}(\bullet) = \sum_\alpha p_\alpha |\alpha\rangle\langle\alpha|_C \otimes \mathcal{N}^\alpha(\bullet)$, where α denotes the set of qudits that are erased whose cardinal number is at most d, C is the classical registry that records the set of lost qudits, and $\mathcal{N}^\alpha(\bullet) := \text{Tr}_\alpha(\bullet)$. To bound the code inaccuracy, following Bény and Oreshkov [18], we bound the entanglement fidelity between the complementary error channel $\widehat{N}(\rho) = (\Lambda + \mathcal{B})(\rho)$ and some channel $\Lambda(\rho)$, which needs to be independent of the input ρ. Note that the complementary channel of $\mathcal{N}^\alpha$ is $\widehat{\mathcal{N}}^\alpha(\bullet) = \text{Tr}_{\bar{\alpha}}(\bullet)$, where $\bar{\alpha}$ is the complement of α. The complementary channel $\widehat{\mathcal{N}}$ to the full erasure channel is just the original channel, but with each $\mathcal{N}^\alpha \to \widehat{\mathcal{N}}^\alpha$ and with a different reference system C'. We consider the channels $\Lambda(\rho) := \sum_\alpha p_\alpha |\alpha\rangle\langle\alpha|_{C'} \otimes \mathcal{T}^\alpha(\rho)$ and $\mathcal{T}^\alpha(\rho) := \text{Tr}[\rho] \cdot \tau_\alpha^0$, both independent of the input ρ and $\tau_\alpha^0 := \text{Tr}_{\bar{\alpha}}(|J, M = 0\rangle\langle J, M = 0|)$. The goal now is to bound the entanglement fidelity $F(\widehat{\mathcal{N}} = \Lambda + \mathcal{B}, \Lambda)$.

Recall that the code space is $\mathfrak{C} = span\{|J,M\rangle; M = M_{\min}, M_{\min}+\Delta, M_{\min}+2\Delta, \ldots, M_{\max}\}$, and we use $|x\rangle_L = V^\dagger |J, M_x\rangle_A$, $x = 1, \cdots, 2^k$, to denote the basis in the logical space for $|J, M_x\rangle \in \mathfrak{C}$, where V is the isometric encoding such that $VV^\dagger = \Pi$ is the projection onto the code space. Consider $\rho_\alpha^{x,x'} = \text{Tr}_{\bar{\alpha}}[|J, M_x\rangle\langle J, M_{x'}|]$, which is an at-most d-local operator. If we pick $\Delta \ge 2sd + 1$, from Lemma 1, we should have $\rho_\alpha^{x,x'} = 0$ if $x \ne x'$, assuming the size of α is at most d. We therefore will only consider the case $x = x'$ and denote $\rho_\alpha^x := \text{Tr}_{\bar{\alpha}}[|J, M_x\rangle\langle J, M_x|]$ for later convenience.

First, using the joint concavity of fidelity for the probability p_α, we have

$$f(\widehat{\mathcal{N}} \otimes \text{id}(|\psi\rangle\langle\psi|), \Lambda \otimes \text{id}(|\psi\rangle\langle\psi|)) = f\left(\sum_\alpha p_\alpha |\alpha\rangle\langle\alpha|'_C \otimes (\widehat{\mathcal{N}}^\alpha \otimes \text{id})(|\psi\rangle\langle\psi|), \sum_\alpha p_\alpha |\alpha\rangle\langle\alpha|'_C \otimes (\mathcal{T}^\alpha \otimes \text{id})(|\psi\rangle\langle\psi|)\right)$$

$$\ge \sum_\alpha p_\alpha f[(\widehat{\mathcal{N}}^\alpha \otimes \text{id})(|\psi\rangle\langle\psi|), (\mathcal{T}^\alpha \otimes \text{id})(|\psi\rangle\langle\psi|)] . \tag{S41}$$

To bound $f[(\widehat{\mathcal{N}}^\alpha \otimes \text{id})(|\psi\rangle\langle\psi|), (\mathcal{T}^\alpha \otimes \text{id})(|\psi\rangle\langle\psi|)]$, recall that the auxiliary space of $[\widehat{\mathcal{N}}^\alpha \otimes \text{id}]$ and $[\mathcal{T}^\alpha \otimes \text{id}]$ is isomorphic to the input space, which is the logical space. For any state $|\psi\rangle = \sum_{x,y} A_{xy}|x\rangle \otimes |y\rangle$, consider the operators $A = \sum_{x,y} A_{xy}|x\rangle\langle y|$ and $B := A^T = \sum_{x,y} A_{xy}|y\rangle\langle x|$ and the state $|\Phi\rangle = \sum_{x=1}^{2^k} |x\rangle \otimes |x\rangle$, we have $|\psi\rangle = A \otimes \text{id}|\Phi\rangle = \text{id} \otimes B|\Phi\rangle$. Therefore, we can write

$$(\widehat{\mathcal{N}}^\alpha \otimes \text{id})(|\psi\rangle\langle\psi|) = (\text{id} \otimes B)\widehat{\mathcal{N}}^\alpha(|\Phi\rangle\langle\Phi|)(\text{id} \otimes B^\dagger) = \sum_{x=1}^{2^k} \rho_\alpha^x \otimes B|x\rangle\langle x|B^\dagger$$

$$(\mathcal{T}^\alpha \otimes \text{id})(|\psi\rangle\langle\psi|) = (\text{id} \otimes B)\widehat{\mathcal{N}}^\alpha(|\Phi\rangle\langle\Phi|)(\text{id} \otimes B^\dagger) = \sum_{x=1}^{2^k} \tau_\alpha^0 \otimes B|x\rangle\langle x|B^\dagger . \tag{S42}$$

Recall that the normalization condition $\langle\psi|\psi\rangle = \sum_x \langle x|B^\dagger B|x\rangle = 1$, so we can interpret $\langle x|B^\dagger B|x\rangle$ as a probability.

Using the joint concavity of fidelity again, we have

$$f[(\widehat{\mathcal{N}}^\alpha \otimes \mathrm{id})(|\psi\rangle\langle\psi|), (\mathcal{T}^\alpha \otimes \mathrm{id})(|\psi\rangle\langle\psi|)]$$

$$= f\left(\sum_{x=1}^{2^k} \langle x|B^\dagger B|x\rangle \cdot \rho_\alpha^x \otimes \frac{B|x\rangle\langle x|B^\dagger}{\langle x|B^\dagger B|x\rangle}, \sum_{x=1}^{2^k} \langle x|B^\dagger B|x\rangle \cdot \tau_\alpha^0 \otimes \frac{B|x\rangle\langle x|B^\dagger}{\langle x|B^\dagger B|x\rangle}\right)$$

$$\geq \sum_{x=1}^{2^k} \langle x|B^\dagger B|x\rangle \cdot f\left(\rho_\alpha^x \otimes \frac{B|x\rangle\langle x|B^\dagger}{\langle x|B^\dagger B|x\rangle}, \tau_\alpha^0 \otimes \frac{B|x\rangle\langle x|B^\dagger}{\langle x|B^\dagger B|x\rangle}\right)$$

$$= \sum_{x=1}^{2^k} \langle x|B^\dagger B|x\rangle \cdot f(\rho_\alpha^x, \tau_\alpha^0)$$

$$\geq \sum_x \langle x|B^\dagger B|x\rangle \left(1 - O\left(\frac{j_1 M_{\max}^2}{4J^2}\right)\right)$$

$$= 1 - O\left(\frac{j_1 M_{\max}^2}{4J^2}\right) = 1 - O\left(\frac{|\alpha| M_{\max}^2}{4sN^2}\right)$$

$$= 1 - O\left(\frac{d M_{\max}^2}{4sN^2}\right), \tag{S43}$$

where we have used Lemma 3, $j_1 = s|\alpha|$, $J = sN$, $|\alpha| \leq d$, and assumed $|M_{\max}| \geq |M_{\min}|$, without loss of generality. We therefore have

$$f(\widehat{\mathcal{N}} \otimes \mathrm{id}(|\psi\rangle\langle\psi|), \Lambda \otimes \mathrm{id}(|\psi\rangle\langle\psi|)) = 1 - O\left(\frac{d M_{\max}^2}{4sN^2}\right). \tag{S44}$$

Since it holds for arbitrary $|\psi\rangle$, we have $F(\widehat{N}, \Lambda) = 1 - O\left(\frac{d M_{\max}^2}{4sN^2}\right)$, or $\epsilon = D_P(\widehat{\mathcal{N}}, \Lambda) = O\left(\sqrt{\frac{d}{2s}}\frac{M_{\max}}{N}\right)$. In the scaling regime $M_{\max} - M_{\min} \sim N^b$ and $d \sim N^c$, we have $\epsilon = O(N^{b+c/2-1})$. The number of logical qubits is $k = \log_2\left(\frac{M_{\max}-M_{\min}}{\Delta} + 1\right)$. Since we need to take $\Delta = \Omega(N^c)$, we have $k = O((b-c)\log_2 N)$ as claimed.

$\square$

S7. CALCULATION OF QFI LOSS AND MEASUREMENTS

In this section, we present the details of the numerical calculation for the QFI loss shown in the main text. In addition, we propose a simple θ-independent local measurement for the noiseless probe state and a θ-independent global measurement for the partially-erased probe state that can saturate the quantum Cramér-Rao bound. We also propose a local measurement for the partially-erased probe state, which can give a scaling better than the standard quantum limit in some parameter regime, though it does not saturate the quantum Cramér-Rao bound.

We consider a system with N spin-s degrees of freedom and the ideal probe state $|\psi\rangle = \frac{1}{\sqrt{2}}(|J, M\rangle + |J, -M\rangle)$, where $J = sN$, and the quantum metrology problem where the sensing parameter θ couples to $\hat{Q}_z$, $|\psi_\theta\rangle = e^{-i\theta\hat{Q}_z}|\psi\rangle$, or $\psi_\theta = \frac{1}{2}(|J, M\rangle\langle J, M| + |J, -M\rangle\langle J, -M| + e^{-i2M\theta}|J, M\rangle\langle J, -M| + e^{i2M\theta}|J, -M\rangle\langle J, M|)$. Since $\hat{Q}_z|J, M\rangle = M|J, M\rangle$, we see that the QFI of this probe state is $\mathfrak{F}(\psi_\theta) = 4(\langle\hat{Q}_z^2\rangle - \langle\hat{Q}_z\rangle^2) = 4M^2$.

Consider the probe state going through an erasure channel $\rho_\theta = \mathcal{N}(\psi_\theta) := \mathrm{Tr}_d[\psi_\theta]$, where $\mathrm{Tr}_d[...]$ means tracing out d sites. Note that, since the $J = sN$ irrep is totally permutationally symmetric, we can just assume that we trace out the first d sites. To calculate the QFI of ρ_θ, we need to find its eigendecomposition. This can be achieved as follows. First, note that we can rewrite

$$|J, M\rangle = \sum_{m_1=-j_1}^{j_1} C_{j_1,m_1,j_2,M-m_1}^{J,M} |j_1, m_1\rangle_d \otimes |j_2, M - m_1\rangle_{\bar{d}}, \tag{S45}$$

where $C_{j_1,m_1,j_2,m_2}^{J,M}$ are the Clebsch-Gordan coefficients, $j_1 = sd$, $j_2 = J - j_1$, and $\bar{d}$ is the complement of the d sites

to be traced out. Defining $a_m = C^{J,M}_{j_1,m,j_2,M-m}$ and $b_m = C^{J,-M}_{j_1,m,j_2,-M-m}$, we have

$$\mathcal{N}(|J,M\rangle\langle J,M|) = \sum_{m=-j_1}^{j_1} a_m^2 |j_2, M-m\rangle\langle j_2, M-m|,$$

$$\mathcal{N}(|J,-M\rangle\langle J,-M|) = \sum_{m=-j_1}^{j_1} b_m^2 |j_2, -M-m\rangle\langle j_2, -M-m|,$$

$$\mathcal{N}(|J,M\rangle\langle J,-M|) = \sum_{m=-j_1}^{j_1} a_m b_m |j_2, M-m\rangle\langle j_2, -M-m|,$$

$$\mathcal{N}(|J,-M\rangle\langle J,M|) = \sum_{m=-j_1}^{j_1} a_m b_m |j_2, -M-m\rangle\langle j_2, M-m| . \tag{S46}$$

Note that we consider the case $M > j_1$ so that $|j_2, M-m\rangle$ and $|j_2, -M-m'\rangle$ are always distinct states for $m, m' = -j_1 \ldots j_1$. Therefore, we obtain

$$\rho_\theta = \frac{1}{2} \bigoplus_{m=-j_1}^{j_1} \begin{pmatrix} a_m^2 & e^{-i2M\theta} a_m b_m \\ e^{i2M\theta} a_m b_m & b_m^2 \end{pmatrix} := \bigoplus_{m=-j_1}^{j_1} B_m , \tag{S47}$$

where the matrix B_m is represented by the basis $|j_2, M-m\rangle$ and $|j_2, -M-m\rangle$. We also see that ρ_θ can be interpreted as an ensemble labeled by $m = -j_1, \cdots, j_1$, where the θ information is encoded in the coherence between the states $|j_2, M-m\rangle$ and $|j_2, -M-m\rangle$, which have a magnetic quantum number separation $2M$. The eigenvalues of the 2×2 matrix B_m are $\frac{1}{2}(a_m^2 + b_m^2)$ and 0, and the corresponding eigenvectors are

$$|m\rangle = \frac{1}{\sqrt{2\lambda_m}} (e^{-iM\theta} a_m |j_2, M-m\rangle + e^{iM\theta} b_m |j_2, -M-m\rangle) , \tag{S48}$$

$$|\bar{m}\rangle = \frac{1}{\sqrt{2\lambda_m}} (e^{-iM\theta} b_m |j_2, M-m\rangle - e^{iM\theta} a_m |j_2, -M-m\rangle) , \tag{S49}$$

respectively. Note that $\partial_\theta \rho_\theta = -iM \sum_{m=-j_1}^{j_1} a_m b_m (e^{-2iM\theta}|j_2, M-m\rangle\langle j_2, -M-m| - e^{2iM\theta}|j_2, -M-m\rangle\langle j_2, M-m|)$. Therefore, we obtain $\langle m|\partial_\theta \rho_\theta |m'\rangle = 0$ for both $m \neq m'$ and $m = m'$. On the other hand, $(\partial_\theta \rho_\theta)^2 = \bigoplus_{m=-j_1}^{j_1} M^2 a_m^2 b_m^2 I_m$, where I_m is the identity matrix in the $|J, M-m\rangle, |J, -M-m\rangle$ subspace. Denoting the set of m such that $\lambda_m \neq 0$ as Λ (it is possible that some of λ_m are zero), we have

$$\begin{aligned}
\mathfrak{F}(\rho_\theta; \partial_\theta \rho_\theta) &= \sum_{m,m':\lambda_m+\lambda_{m'}>0} \frac{2|\langle m|\partial_\theta \rho_\theta |m'\rangle|^2}{\lambda_m + \lambda_{m'}} \\
&= \sum_{m,m'\in\Lambda} \frac{2|\langle m|\partial_\theta \rho_\theta |m'\rangle|^2}{\lambda_m + \lambda_{m'}} + \sum_{m\in\Lambda} \frac{4}{\lambda_m} \langle m|\partial_\theta \rho_\theta (I - \sum_{m'\in\Lambda} |m'\rangle\langle m'|)\partial_\theta \rho_\theta |m\rangle \\
&= \sum_{m\in\Lambda} \frac{1}{\lambda_m} |\langle m|\partial_\theta \rho_\theta |m\rangle|^2 + \sum_{m\in\Lambda} \frac{4}{\lambda_m} \langle m|\partial_\theta \rho_\theta (I - \sum_{m'\in\Lambda} |m'\rangle\langle m'|)\partial_\theta \rho_\theta |m\rangle \\
&= \sum_{m\in\Lambda} \frac{4a_m^2 b_m^2}{\lambda_m} M^2 = \sum_{m\in\Lambda} \left(\frac{2a_m^2 b_m^2}{a_m^2 + b_m^2}\right) 4M^2 . \tag{S50}
\end{aligned}$$

Finally, we note that $\mathfrak{F}(\psi_\theta; \partial_\theta \psi_\theta) = 4M^2$ and $\Delta\mathfrak{F} = \mathfrak{F}(\psi_\theta; \partial_\theta \psi_\theta) - \mathfrak{F}(\rho_\theta; \partial_\theta \rho_\theta)$, and the Clebsch-Gordan coefficients can be calculated efficiently via a recursive formula or an explicit formula, which gives us an efficient method to calculate the QFI loss. We calculate the QFI by picking values of J and j_1. So the result is applicable to different s with the corresponding $N = J/s$ and $d = j_1/s$.

It is interesting to examine the QFI loss if the probe state is a GHZ(-like) state $|\psi'\rangle = \frac{1}{\sqrt{2}}(|J,J\rangle + |J,-J\rangle)$, which would achieve the Heisenberg scaling in the absence of noise. We can express $|J,J\rangle = |j_1, j_1\rangle \otimes |j_2, j_2\rangle$ and $|J,-J\rangle = |j_1, -j_1\rangle \otimes |j_2, -j_2\rangle$, which gives us $\rho_\theta = \frac{1}{2}(|j_2, j_2\rangle\langle j_2, j_2| + |j_2, -j_2\rangle\langle j_2, -j_2|)$ independent of θ. This implies that $\mathfrak{F}(\rho_\theta; \partial_\theta \rho_\theta) = 0$. This can also be obtained from Eq. (S50). Note that, in this case, the only nonzero λ_m is $m = j_1$ or $m = -j_1$, and we have $(a_m, b_m) = (1, 0)$ if $m = j_1$ and $(a_m, b_m) = (0, 1)$ if $m = -j_1$, respectively, so $\mathfrak{F}(\rho_\theta; \partial_\theta \rho_\theta) = 0$. As expected, noise obstructs us from using GHZ-like states for optimal sensing.

We propose a simple local measurement for the ideal probe state $|\psi\rangle = \frac{1}{\sqrt{2}}(|J,M\rangle + |J,-M\rangle)$ which saturates the quantum Cramér-Rao bound. Specifically, we consider the observable $\mathbf{D} := \bigotimes_{j=1}^{N} D_j$, where D_j satisfies $D_j \hat{q}_j^z D_j = -\hat{q}_j^z$. This is a local measurement since one can measure $\mathbf{D}$ by measuring D_j locally and multiplying the outcomes. For example, in the basis $\hat{q}_j^z |\sigma\rangle = \sigma|\sigma\rangle$, where $\sigma = -s, \cdots s$, one can pick $D_j = \sum_{\sigma=-s}^{s} |s\rangle\langle -s|$, which would give $\mathbf{D}^2 = I$. The unknown parameter θ can therefore be obtained by measuring $\langle \mathbf{D} \rangle = \langle \psi_\theta | \mathbf{D} | \psi_\theta \rangle = \cos(2M\theta)$, and we have $\theta = \arccos(\langle \mathbf{D} \rangle)/(2M)$. The variance is $\sigma_{\mathbf{D}}^2 = \langle \mathbf{D}^2 \rangle - \langle \mathbf{D} \rangle^2 = 1 - \cos^2(2M\theta)$, which gives the standard error of the measurement as $\Delta\mathbf{D} = \sigma_{\mathbf{D}}/\sqrt{\nu}$, where ν is the number of samples. So we have $\Delta\theta = \frac{\Delta\mathbf{D}}{2M}\frac{\arccos(x)}{dx}|_{x=\langle \mathbf{D} \rangle} = \frac{1}{2M\sqrt{\nu}}$, which indeed saturates the quantum Cramér-Rao bound since QFI $\mathfrak{F}(\psi_\theta) = 4M^2$.

For the partially-erased probe state $\rho_\theta = \sum_{m=-j_1}^{j_1} \lambda_m |m\rangle\langle m|$, recall that the θ information is encoded in the coherence between the states $|j_2, M-m\rangle$ and $|j_2, -M-m\rangle$. A global measurement such as $\mathbf{D}' = \sum_{m=-j_1}^{j_1} |j_2, M - m\rangle\langle j_2, -M - m| + \text{h.c.}$ would be a candidate to estimate the unknown θ. Indeed, we see $\langle \mathbf{D}' \rangle = \text{Tr}[\rho_\theta \mathbf{D}'] = \sum_{m=-j_1}^{j_1} a_m b_m \cos(2M\theta) = A\cos(2M\theta)$, where $A := \sum_{m=-j_1}^{j_1} a_m b_m$. Since $\mathbf{D}'^2 = \sum_{m=-j_1}^{j_1} I_m$, we have $\sigma_{\mathbf{D}'}^2 = 1 - A^2\cos^2(2M\theta)$, and

$$\Delta\theta = \frac{\sigma_{\mathbf{D}'}}{\sqrt{\nu}2MA\sin(2M\theta)} \sim \frac{1}{2MA} \ . \tag{S51}$$

Note that

$$A = \sum_{m=-j_1}^{j_1} a_m b_m = \left[\binom{2J}{J+M}\right]^{-1} \sum_{m_1=-j_1}^{j_1} \binom{2j_1}{j_1+m_1}\left[\binom{2J-2j_1}{J-j_1+M-m_1}\binom{2J-2j_1}{J-j_1-M-m_1}\right]^{\frac{1}{2}} \ , \tag{S52}$$

and the asymptotic expansion for A can be carried out like in App. S5. We expect $A = 1 - O(j_1 M^2/J^2)$, which gives us $\Delta\theta \sim M^{-1} = N^{-b}$. This scaling matches with the QFI scaling in the red-shaded parameter regime of Fig. 1(b) in the main text.

A global measurement is generally hard to implement so it is usually more desirable to use a local measurement. Below, we show that the local measurement $\mathbf{D}_{\bar{d}} := \bigotimes_{j=d+1}^{N} D_j$ can be used to obtain θ with a $\Delta\theta$ scaling beyond the standard quantum limit, although it does not saturate the scaling given by the QFI. Note that, for $\rho_\theta = \sum_{m=-j_1}^{j_1} \lambda_m |m\rangle\langle m|$, if j_1 is a half-integer, then we can throw out one more qudit so that j_1 is an integer. We thus see that $\mathbf{D}_{\bar{d}}$ will pick up the signal from the coherence between states $|j_2, M-m\rangle$ and $|j_2, -M-m\rangle$ only when $m = 0$. More explicitly, we have $\langle \mathbf{D}_{\bar{d}} \rangle = \text{Tr}[\rho_\theta \mathbf{D}_{\bar{d}}] = a_0 b_0 \cos(2M\theta)$. Comparing this result with Eq. (S51) and noting that $a_0 b_0 \sim (j_1)^{-\frac{1}{2}}$, we obtain $\Delta\theta \sim \sqrt{j_1}/M \sim N^{-(b-c/2)}$ if $j_1 \sim N^c$ and $M \sim N^b$, exhibiting a scaling beyond the standard quantum limit if $2b - c > 1$, together with the conditions $b > c$ and $c < 2(1 - b)$. For $c = 0$, the local measurement would give a variance scaling matching the QFI scaling.

S8. SPIN-1 XY-DZYALOSHINSKII-MORIYA SCAR STATES AS AQECC

In this appendix, we show that the quantum many-body scar states in the Spin-1 XY-Dzyaloshinskii-Moriya model [65–67] can form an AQECC. While the original model considers a system with N spin-1 degrees of freedom, here we generalize it to N spin-s degrees of freedom. Specifically, we consider the spin-s local Hilbert space with the basis $\{|m\rangle, m = -s, \cdots, s\}$. We define the operators

$$J^z = \frac{1}{2}\sum_{j=1}^{N}(|s\rangle\langle s| - |-s\rangle\langle -s|)_j \ , \quad J^+ = \sum_{j=1}^{N} e^{i\phi_j}|s\rangle\langle -s|_j \ , \quad J^- = \sum_{j=1}^{N} e^{-i\phi_j}|-s\rangle\langle s|_j \ , \tag{S53}$$

where ϕ_j is some site-dependent phase. It is easy to verify that these operators generate the $SU(2)$ algebra $[J^+, J^-] = 2J^z$ and $[J^z, J^\pm] = \pm J^\pm$. The quantum many-body scar states are $|S_M\rangle := Z_M(J^+)^M|S_0\rangle$, where $|S_0\rangle = \bigotimes_{j=1}^{N} |-s\rangle_j$ and Z_M is the normalization constant. The scar states in the Spin-1 XY-Dzyaloshinskii-Moriya model correspond to $s = 1$ with some specific choice of ϕ_j.

In fact, these scar states are Dicke states in disguise. This can be seen by redefining $|s\rangle_j \to e^{-\phi_j/2}|s\rangle_j$ and $|-s\rangle_j \to e^{\phi_j/2}|s\rangle_j$, which can be achieved by applying a tensor product of single-site unitaries $U = \bigotimes_{j=1}^{N} u_j$, where $u_j = |s\rangle\langle s|_j e^{-\phi_j/2} + |-s\rangle\langle -s|_j e^{\phi_j/2} + \sum_{m=-s+1\cdots s-1} |m\rangle\langle m|_j$. We therefore see that $|S_M\rangle$ are Dicke states $|J,M\rangle$ $(s = \frac{1}{2}, J = \frac{N}{2})$ embedded in a larger local Hilbert space, up to a tensor product of single-site unitary. Therefore,

we conclude that the quantum many-body scar states $\mathfrak{C} = \text{span}\{|S_M\rangle; M = M_{\min}, M_{\min} + \Delta, \cdots, M_{\max}\}$ forms an AQECC against d-local noise or erasures at known locations if we take $\Delta \geq 2sd + 1$. The code parameters and the inaccuracy bound against different errors can be readily obtained from Theorems 1 and 2 with $s = \frac{1}{2}$, $J = \frac{N}{2}$.