The theory of entanglement-assisted metrology for quantum channels

Sisi Zhou,1,2 and Liang Jiang2

1Department of Physics, Yale University, New Haven, Connecticut 06511, USA
2Pritzker School of Molecular Engineering, The University of Chicago, Illinois 60637, USA
(Dated: March 25, 2020)

The quantum Fisher information (QFI) measures the amount of information that a quantum state carries about an unknown parameter. The (entanglement-assisted) QFI of a quantum channel is defined to be the maximum QFI of the output state assuming an entangled input state over a single probe and an ancilla. Both the channel QFI and the optimal input state could be solved via a semidefinite program (SDP). In quantum metrology, people are interested in calculating the QFI of N identical copies of a quantum channel when N → ∞, which we call the asymptotic QFI. It was known that the asymptotic QFI grows either linearly or quadratically with N. Here we obtain a simple criterion that determines whether the scaling is linear or quadratic. In both cases, we found a quantum error correction protocol achieving the asymptotic QFI and an SDP to solve the optimal code. When the asymptotic QFI is quadratic, the Heisenberg limit, a feature once thought unique to unitary quantum channels, is recovered. When the asymptotic QFI is linear, we show it is still in general larger than N times the channel QFI, showing the non-additivity of the channel QFI of general quantum channels.

I. INTRODUCTION

Quantum metrology studies parameter estimation in a quantum system [1–5]. Usually, a quantum probe interacts with a physical system and the experimentalist performs measurements on the final probe state and infers the value of the unknown parameter(s) in the system from the measurement outcomes. It has wide applications in frequency spectroscopy [6–9], gravitational-wave detectors [10–13] and other high-precision measurements [14–17].

The quantum Fisher information (QFI), whose operational meaning is given by the quantum Cramér-Rao bound, characterizes the amount of information a quantum state carries about an unknown parameter [18–21]. To explore the fundamental limit on parameter estimation, we usually consider the situation where the number of quantum channels N (or the probing time t) is large. The Heisenberg limit (HL), an O(N2) (or O(t2)) scaling of the QFI, is the ultimate estimation limit allowed by quantum mechanics. It could be obtained, for example, using GHZ states in noiseless systems [9, 22]. On the other hand, the standard quantum limit (SQL), an O(N) (or O(t)) scaling of the QFI, usually appears in noisy systems and could be achieved using product states. Much work has been done towards determining whether or not the HL is achievable for a given quantum channel [23–35]. Some necessary conditions were derived—for example, it was shown that the HL cannot be achieved for programmable channels [24]. Using the channel extension method, another necessary condition, which we will call the HNKS condition, was derived [23, 27]. In particular, for a special type of quantum channel where we estimate the Hamiltonian parameter under Markovian noise, HNKS becomes sufficient [30, 31].

In general, the optimal QFI achievable in a quantum system always has a leading term equal to \( \tilde{\gamma}_{\text{HL}} \cdot N^2 \) or \( \tilde{\gamma}_{\text{SQL}} \cdot N \), corresponding to either the HL or SQL scaling. We call the leading term the asymptotic QFI. As pointed out earlier, for general quantum channels, there was not a unified approach to determine whether the scaling is HL or SQL. For quantum channels where the scalings are known, it is also crucial to understand how to achieve the asymptotic QFI. For example, for unitary channels, the HL is achievable and a GHZ state in the multipartite two-level systems consisting of the lowest and highest energy states achieves the asymptotic QFI [22]. Under the effect of noise, a variety of quantum strategies were also proposed to enhance the QFI [8, 10, 36–48], but no conclusions for general quantum channels were drew. One natural question to ask is whether entangled probes can improve the QFI, compared to product states. When estimating the noise parameter in teleportation-covariant channels (e.g. Pauli or erasure channels) [49, 50], it was shown that entanglement between probes are unnecessary and product states are sufficient to achieve the asymptotic QFI. However, when estimating the phase parameter in dephasing or erasure channels, the channel QFI is no longer additive. The asymptotic QFI is achievable using spin-squeezed states [8, 28, 37]. When viewing the QFI as a function of the probing time, it was also shown that the asymptotic QFI with respect to (wrt) the Hamiltonian parameter is achieved using the approximate quantum error correction technique under general Markovian dynamics [51].

Given a quantum channel, we aim to answer the following two important questions: does it follow the HL or the SQL, and how to achieve the asymptotic QFI? In this paper, we answer these two open problems using a quantum error correction (QEC) protocol. QEC has been a powerful tool widely used in quantum computing and quantum communication to protect quantum information from noise [52–55]. In quantum metrology, we also need to protect the quantum signal from noise [28–34, 56–64]. Here is a typical example where QEC was proven useful: when a qubit is subject to a \( \sigma_z \) signal and a \( \sigma_x \) noise, the QFI follows the SQL if no quantum control is added. However, the HL could be recovered using fast and frequent QEC [56–61]. The result could be generalized to any system with a signal Hamiltonian and Markovian noise [30, 31]. These QEC protocols, however, all rely on fast and frequent quantum operations and have limited practical applications.

In this paper, we consider general quantum channels that are arbitrary functions of the unknown parameter, which could be (but not necessarily is) either the Hamiltonian parameter or
the symmetric logarithmic derivative (SLD) satisfying
\[ \rho \] where \( N \) is the number of repeated experiments and \( \rho \) is the QFI of the state \( \rho \). The leading term \( \delta_{\text{opt}}(\omega) \) is equal to either \( \delta_{\text{opt}}(\omega)N^2 \) or \( \delta_{\text{opt}}(\omega) \) when \( N \to \infty \). The ancillary system is assumed to be arbitrarily large in both cases.

**II. MAIN RESULTS**

The quantum Cramér-Rao bound is a lower bound of the estimation precision [18–21],

\[ \delta \omega \geq \frac{1}{\sqrt{N_{\text{expr}}F(\rho)}} \] where \( \delta \omega \) is the standard deviation of any unbiased estimator of \( \omega \), \( N_{\text{expr}} \) is the number of repeated experiments and \( F(\rho) \) is the QFI of the state \( \rho \). The quantum Cramér-Rao bound is saturable asymptotically (\( N_{\text{expr}} \gg 1 \)) using maximum likelihood estimators [65, 66]. Therefore, the QFI is a good measure of the amount of information a quantum state \( \rho \) carries about an unknown parameter. It is defined by \( F(\rho) = \text{Tr}(L^2 \rho) \), where \( L \) is a Hermitian operator called the symmetric logarithmic derivative (SLD) satisfying

\[ \rho_\omega = \frac{1}{2}(\rho_\omega L + L \rho_\omega), \] where \( \cdot \) denotes \( \frac{\partial}{\partial \omega} \). We will use \( L_A[B] \) to represent Hermitian operators satisfying \( B = \frac{1}{2}(LA + AL) \). Here \( L = L_{\rho_\omega}[\rho_\omega] \). The QFI could also be equivalently defined through purification [23]:

\[ F(\rho_\omega) = 4 \min_{|\psi_\omega\rangle, \text{Tr}(|\psi_\omega\rangle\langle\psi_\omega|) = \rho_\omega} \langle \psi_\omega | \hat{J}_\omega | \psi_\omega \rangle, \] where \( \rho_\omega \in \mathcal{S}(\mathcal{H}_P), |\psi_\omega\rangle \in \mathcal{S}(\mathcal{H}_P \otimes \mathcal{H}_E) \), \( \mathcal{H}_P \) is the probe system which we assume to be finite-dimensional, \( \mathcal{H}_E \) is an arbitrarily large environment and \( \mathcal{S}(\bullet) \) denotes the set of density operators in \( \bullet \).

We consider a quantum channel \( \mathcal{E}_\omega(\rho) = \sum_{i=1}^{r} K_i \rho K_i^T \), where \( r \) is the rank of the channel. The entanglement-assisted QFI of \( \mathcal{E}_\omega \) (see Fig. 1a) is defined by,

\[ \tilde{\mathcal{F}}_1(\mathcal{E}_\omega) := \max_{\rho \in \mathcal{S}(\mathcal{H}_P \otimes \mathcal{H}_A)} F((\mathcal{E}_\omega \otimes I)(\rho)). \] Here we utilize the entanglement between the probe and an arbitrarily large ancillary system \( \mathcal{H}_A \). We will omit the word “entanglement-assisted” in the definitions below for simplicity. Practically, the ancilla is a quantum system with a long coherence time, e.g. nuclear spins [60] or any QEC-protected system [31]. It also helps simplify the complicated calculation of the QFI. The convexity of QFI implies the optimal input state is always pure. Using the purification-based definition of the QFI (Eq. (3)), we have [23]

\[ \tilde{\mathcal{F}}_1(\mathcal{E}_\omega) = 4 \max_{\rho \in \mathcal{S}(\mathcal{H}_P)} \min_{K' = u K} \text{Tr}(\rho K' K'') \] \[ = 4 \min_{K' = u K} \| K' K'' \| = 4 \min_{h \in \mathbb{C}} \| \alpha \|, \] where \( \| \cdot \| \) is the operator norm, \( \mathbb{H}_r \) is the space of \( r \times r \) Hermitian matrices and \( K = (K_1, \ldots, K_r)^T, K' = (K'_1, \ldots, K'_r)^T = u K \) represents all possible Kraus representations of \( \mathcal{E}_\omega \) via isometric transformations \( u \) [23]. Let \( h = i u u' \) and \( \alpha = K' K'' = (K - i h K)(K - i h K) \). The minimization could be performed over arbitrary Hermitian operator \( h \) in \( \mathbb{C}^{r \times r} \) [27]. Any purification of the optimal \( \rho \) in Eq. (5) is an optimal input state in \( \mathcal{H}_P \otimes \mathcal{H}_A \). The problem could be solve via a (quadratic) SDP [27, 35] (see also Appx. F). Note that the optimal input state would in general depend on the true value of \( \omega \) and in practice should be chosen adaptively throughout the experiment [67, 68].

Consider \( N \) identical copies of the quantum channel \( \mathcal{E}_\omega \) [23, 27] (see Fig. 1b), let

\[ \tilde{\mathcal{F}}_N(\mathcal{E}_\omega) := \tilde{\mathcal{F}}_1(\mathcal{E}_\omega^N) = \max_{\rho} F((\mathcal{E}_\omega^N \otimes I)(\rho)). \] Clearly \( \tilde{\mathcal{F}}_N \geq N \tilde{\mathcal{F}}_1 \) using the additivity of the QFI. An upper bound of \( \tilde{\mathcal{F}}_N(\mathcal{E}_\omega) \) could be derived from Eq. (6) (see Appx. A),

\[ \tilde{\mathcal{F}}_N(\mathcal{E}_\omega) \leq 4 \min_{h} \left(N \| \alpha \| + N(N-1) \| \beta \|^2 \right), \] where \( \beta = i K' K'' \). If there is an \( h \) such that \( \beta = 0 \),

\[ \tilde{\mathcal{F}}_N(\mathcal{E}_\omega) \leq 4 \min_{h, \beta=0} N \| \alpha \|, \]
and $\mathfrak{S}_N(\mathcal{E}_\omega)$ follows the SQL asymptotically. Therefore, it is only possible to achievable the HL if $H \notin S$, where

$$H = i K^\dagger K, \quad S = \text{span}_\mathbb{C}\{K_i^j, \forall i, j\}. \quad (10)$$

Here $\text{span}_\mathbb{C}\{\cdot\}$ represents all Hermitian operators which are linear combinations of operators in $\{\cdot\}$. We call it the HNKS condition, an acronym for “Hamiltonian-not-in-Kraus-span”. One can check that $H$ and $\beta$ are always Hermitian by taking the derivative of $K^\dagger K = I$. Note that different Kraus representations may lead to different $H$, but it does not affect the validity of $H \notin S$. For a unitary channel $r = 1$ and $K_1 = U_\omega = e^{-iH\omega}$, $H = iU_\omega^\dagger U_\omega$ exactly the Hamiltonian for $\omega$, explaining its name. The HL is achievable for unitary channels because $S = \text{span}_\mathbb{C}\{I\}$ and we always have $H \notin S$ for nontrivial $H$.

We will show in Sec. V that HNKS is also a sufficient condition to achieve the HL, giving the following theorem:

**Theorem 1.** $\mathfrak{S}_N(\mathcal{E}_\omega) = \Theta(N^2)$ if and only if $H \notin S$. Otherwise, $\mathfrak{S}_N(\mathcal{E}_\omega) = \Theta(N)$.

Furthermore, in Sec. V and Sec. VI, we will provide a QEC protocol which achieves the QFI upper bound (Eq. (8)) asymptotically both when $H \in S$ or $H \notin S$:

**Theorem 2.** When $H \notin S$,

$$\mathfrak{S}_\text{HL}(\mathcal{E}_\omega) := \lim_{N \to \infty} \mathfrak{S}_N(\mathcal{E}_\omega)/N^2 = 4 \min_h \|e^{\beta_i}\|^2. \quad (11)$$

There exists an input state $|\psi_N\rangle$ solvable via an SDP such that

$$\lim_{N \to \infty} F((\mathcal{E}_\omega^\otimes N \otimes \mathcal{I})(|\psi_N\rangle))/N^2 = \mathfrak{S}_\text{HL}(\mathcal{E}_\omega).$$

**Theorem 3.** When $H \in S$,

$$\mathfrak{S}_\text{SQL}(\mathcal{E}_\omega) := \lim_{N \to \infty} \mathfrak{S}_N(\mathcal{E}_\omega)/N = 4 \min_{h: \beta = 0} \|\alpha\|. \quad (12)$$

For any $\eta > 0$, there exists an input state $|\psi_{N, \eta}\rangle$ solvable via an SDP such that

$$\lim_{N \to \infty} F((\mathcal{E}_\omega^\otimes N \otimes \mathcal{I})(|\psi_{N, \eta}\rangle))/N > \mathfrak{S}_\text{SQL}(\mathcal{E}_\omega) - \eta.$$

In the following, we will first prove Theorem 2 and Theorem 3 for a single qubit dephasing channel where both the phase and the noise parameter vary wrt $\omega$. Then we will generalize the results to arbitrary quantum channels $\mathcal{E}_\omega$ using a QEC protocol. Theorem 1 will be a corollary of Theorem 2. The roadmap to achieve the asymptotic QFI is illustrated in Fig. 2.

**III. DEPHASING CHANNELS**

According to Eq. (8), $\mathfrak{S}_\text{HL} \leq \mathfrak{S}^{(u)}_{\text{HL}}$ and $\mathfrak{S}_\text{SQL} \leq \mathfrak{S}^{(u)}_{\text{SQL}}$, where $\mathfrak{S}^{(u)}_{\text{HL}} := 4 \min_h \|\beta\|^2$ and $\mathfrak{S}^{(u)}_{\text{SQL}} := 4 \min_{h: \beta = 0} \|\alpha\|$. $(u)$ refers to the upper bounds here. In this section, we will show the above equalities hold for any single dephasing channel

$$D_\omega(\rho) = (1-p)e^{-i\frac{\omega}{2}}\sigma_zpe^{i\frac{\omega}{2}}\sigma_z + p\sigma_ze^{-i\frac{\omega}{2}}\sigma_zpe^{i\frac{\omega}{2}}\sigma_z, \quad (13)$$

which is the composition of the conventional dephasing channel $\rho \to (1-p)\rho + p\sigma_z\rho_0\sigma_z$ ($0 \leq p < 1$) and the rotation in the $z$-direction $\rho \to e^{-i\frac{\omega}{2}}\sigma_zpe^{i\frac{\omega}{2}}\sigma_z$. Both $p$ and $\phi$ are functions of an unknown parameter $\omega$. As shown in Appx. B, the HNKS condition is equivalent to $p = 0$ and the QFI upper bounds of $D_\omega$ are

$$\mathfrak{S}^{(u)}_{\text{HL}}(D_\omega) = |\xi|^2, \quad \mathfrak{S}^{(u)}_{\text{SQL}}(D_\omega) = \frac{1}{2|\xi|^2}, \quad (14)$$

where $\xi = (1 - 2p)e^{-i\phi}$.

Now we show that $\mathfrak{S}^{(u)}_{\text{HL}}(D_\omega) = \mathfrak{S}^{(u)}_{\text{SQL}}(D_\omega)$ and provide the optimal input states in both cases. When HNKS is satisfied ($p = 0$), $D_\omega$ is unitary. Using the GHZ state $|\psi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle^\otimes N + |1\rangle^\otimes N)$ as the input state, we could achieve

$$F(D_\omega^\otimes N (|\psi_0\rangle \langle \psi_0|)) = |\xi|^2 N^2, \quad (15)$$

which implies $\mathfrak{S}_{\text{HL}}(D_\omega) = \mathfrak{S}_{\text{SQL}}(D_\omega)$.

To calculate the optimal QFI when HNKS is violated ($p > 0$), we will use the following two useful formulæ. For any pure state input $|\psi_0\rangle$ and output $\rho_\omega = D_\omega^\otimes N (|\psi_0\rangle \langle \psi_0|)$, we have, for all $N$,

$$F(\rho_\omega) = F_p(\rho_\omega) + F_\phi(\rho_\omega), \quad (16)$$

where $F_p(\rho_\omega) = \text{Tr}(L_p^2 \rho_\omega)$ is the QFI wrt $\omega$ when only the noise parameter $p$ varies wrt $\omega$, where the SLD $L_p$ satisfies

$$\frac{1}{2} \frac{\partial^2}{\partial p^2} \tilde{p} = L_p \rho_\omega + \rho_\omega L_p.$$
\( \omega \) when only the phase parameter \( \varphi \) varies wrt \( \omega \). The proof of Eq. (16) is provided in Appx. C. Another useful formula is [70],

\[
F(\rho) \geq \frac{1}{(\Delta J^2)_{\rho}} \left( \frac{\partial \langle J \rangle_{\rho}}{\partial \omega} \right)^2,
\]

for arbitrary \( \rho \) as a function of \( \omega \) and arbitrary Hermitian operator \( J \) where \( \langle J \rangle_{\rho} = Tr(J \rho) \) and \( (\Delta J^2)_{\rho} = \langle J^2 \rangle_{\rho} - \langle J \rangle_{\rho}^2 \).

Consider an \( N \)-qubit spin-squeezed state [37, 69]:

\[
|\psi_{\mu,\nu}\rangle = e^{-ir_{J_y} e^{-\frac{\omega}{2} J_x e^{-i J_y} J_y}} |0\rangle^N, \tag{18}
\]

where \( J_{x,y,z} \) are the Pauli operators and \( \omega \) is the phase parameter. Let \( |\psi_0\rangle = e^{i\phi J_y} |\psi_{\mu,\nu}\rangle \). Using Eq. (16) and Eq. (17), we have for \( \rho_{\omega} = \mathcal{D}^N_\omega (|\psi_0\rangle \langle \psi_0|) \),

\[
F(\rho_{\omega}) \geq \frac{1}{(\Delta J_x^2)_{\rho_{\omega}}} \left( \frac{\partial \langle J_x \rangle_{\rho_{\omega}}}{\partial \omega} \right)^2 \\
+ \frac{1}{(\Delta J_y^2)_{\rho_{\omega}}} \left( \frac{\partial \langle J_y \rangle_{\rho_{\omega}}}{\partial \phi} \right)^2. \tag{19}
\]

As shown in Appx. D, as \( N \to \infty \), with suitable choices of \( (\mu, \nu) \), we have (up to the lowest order of \( N \)), \( (\Delta J_x^2)_{\rho_{\omega}} \approx (\Delta J_y^2)_{\rho_{\omega}} \approx p(1-p)N \), \( \frac{\partial \langle J_x \rangle_{\rho_{\omega}}}{\partial \omega} \approx -\hat{p}N \) and \( \frac{\partial \langle J_y \rangle_{\rho_{\omega}}}{\partial \phi} \approx (1-2p)\hat{\phi}N/2 \). For example, we can choose \( \mu = 4(\frac{N}{4})^{5/6} \) and \( \nu = \frac{1}{2} \arctan \frac{4N}{1-4N^{1/2}} \). The corresponding \( |\psi_{\mu,\nu}\rangle \) is illustrated in Fig. 2e using the quasi-probability distribution \( Q(\theta, \phi) = |(\theta, \varphi)|^2 \) on a sphere [69]. Therefore,

\[
F(\rho_{\omega}) \geq \frac{1}{1 - |\xi|^2} N + o(N), \tag{20}
\]

which implies \( \mathfrak{f}_{\mathcal{D}_\omega} (\mathcal{D}_\omega) = \mathfrak{f}_{\mathcal{D}_\omega} (\mathcal{D}_\omega) \) compared with \( \mathfrak{f}_{\mathcal{S}_\omega} (\mathcal{D}_\omega) \) (see Appx. B), \( \mathfrak{f}_{\mathcal{S}_\omega} (\mathcal{D}_\omega) \) has a factor of \( 1/(4p(1-p)) \) enhancement when we estimate the phase parameter (\( \hat{p} = 0 \)). When we estimate the noise parameter (\( \hat{\phi} = 0 \)), however, \( \mathfrak{f}_{\mathcal{S}_\omega} (\mathcal{D}_\omega) = \mathfrak{f}_{\mathcal{S}_\omega} (\mathcal{D}_\omega) \). In general, \( \mathfrak{f}_{\mathcal{S}_\omega}/\mathfrak{f}_{\mathcal{S}_\omega} \) is between 1 and \( 1/(4p(1-p)) \).

To sum up, we proved Theorem 2 and Theorem 3 are true for dephasing channels. The ancilla is not required here. When the noise is non-zero, the QFI must follow the SQL for dephasing channels. The ancilla is not required here.

\[ \mathcal{R} \circ \mathcal{E}_\omega \circ \mathcal{E}_{\text{enc}} = \mathcal{D}_{t,\omega}. \tag{21} \]

The contraction fully utilizes the advantage of the ancilla. Let \( \dim \mathcal{H}_\mathcal{P} = d \) and \( \dim \mathcal{H}_\mathcal{A} = 2d \). We pick a QEC code

\[
|0_i\rangle = \sum_{i,j=1}^d A_{0,i,j} |i\rangle_p |j\rangle_\mathcal{A}, \quad |1_i\rangle = \sum_{i,j=1}^d A_{1,i,j} |i\rangle_p |j\rangle_\mathcal{A}, \tag{22}
\]

with the encoding channel \( \mathcal{E}_{\text{enc}} (\cdot) = V (\cdot) V^\dagger \) where \( V = |0_i\rangle \langle 0_i| + |1_i\rangle \langle 1_i| \), and a recovery channel

\[
\mathcal{R}(\cdot) = \sum_{m=1}^M (|0_m\rangle \langle R_m, 0| + |1_m\rangle \langle Q_m, 1|) \cdot \tag{23}
\]

Here \( A_{0,1} \) is matrices in \( \mathbb{C}^{d \times d} \) satisfying \( \text{Tr}(A_{0,1} A_{0,1}^\dagger) = 1, R = (|R_m\rangle \langle R_m|) \) and \( Q = (|Q_m\rangle \langle Q_m|) \) are matrices satisfying \( R R^\dagger = Q Q^\dagger = I \). The last ancillary qubit in \( \mathcal{H}_\mathcal{A} \) guarantees the logical channel to be dephasing, which satisfies

\[
\xi = \sum_{i,m} \langle R_m, 0| K_i |0_i\rangle \langle 1_i| K_i^\dagger |Q_m, 1 \rangle, \tag{24}
\]

and \( \mathfrak{f}_{\mathcal{S}_\mathcal{H}_{\text{SQL}}}(\mathcal{D}_{t,\omega}) \) could then be directly calculated using Eq. (14). Below, we will show that by optimizing \( \mathfrak{f}_{\mathcal{S}_\mathcal{H}_{\text{SQL}}}(\mathcal{D}_{t,\omega}) \) over both the recovery channel \( (R,Q) \) and the QEC code \( (A_{0,1}) \), the QFI upper bounds \( \mathfrak{f}_{\mathcal{S}_\mathcal{H}_{\text{SQL}}}(\mathcal{E}_\omega) \) are achievable.

V. ACHIEVING THE HL UPPER BOUND

When \( H \notin \mathcal{S} \), we construct a QEC code such that the HL upper bound \( \mathfrak{f}_{\mathcal{S}_\mathcal{H}}(\mathcal{E}_\omega) \) is achieved. For dephasing channels, the HL is achievable only if \( \xi = 1 \). Since any transformation \( R \leftarrow e^{i\varphi} R \) does not affect the QFI, without loss of generality (WLOG), we assume \( \xi = 1 \). It means that the QEC has to be perfect, i.e. satisfies the Knill-Laflamme condition [53]

\[
P K_j^i P \propto P, \quad \forall i,j, \tag{25}
\]

where \( P = |0_i\rangle \langle 0_i| + |1_i\rangle \langle 1_i| \). Moreover, there exists a Kraus representation \( \{ K_j^i \}_{j=1}^r \) such that \( P K_j^i K_j^i P = \mu_i \delta_{ij} P \) and \( K_j^i P = U_i \sqrt{\mu_i} P \). The unitary \( U_i \) has the form

\[
U_i = U_{0,i} \otimes |0\rangle \langle 0| + U_{1,i} \otimes |1\rangle \langle 1|, \tag{26}
\]

where \( U_{0,i} \) and \( U_{1,i} \) are also unitary. Let

\[
|R_i\rangle = \langle 0| U_{0,i} |0_i\rangle, \quad |Q_i\rangle = \langle 0| U_{1,i} |0_i\rangle, \tag{27}
\]

for \( 1 \leq i \leq r' \). We could also add some additional \( |R_i\rangle \) and \( |Q_i\rangle \) to them to make sure they are two complete and orthonormal bases. Then one could verify that \( \xi = 1 \)

\[
\tilde{\xi} = -i \text{Tr}((H \otimes I) \sigma_{z,i}), \tag{28}
\]

where \( \sigma_{z,i} = |0_i\rangle \langle 0_i| - |1_i\rangle \langle 1_i| \). Let \( \tilde{C} = A_0 A_0^\dagger - A_1 A_1^\dagger \), \( \tilde{\xi} = -i \text{Tr}(H \tilde{C}) \) and the Knill-Laflamme condition is equivalent \( \text{Tr}(C S) = 0 \), \( \forall S \in \mathcal{S} \). The optimization of the QFI over
the QEC code becomes

\[
\text{maximize } |\xi| = |\text{Tr}(H\tilde{C})|,
\]

subject to \(\|\tilde{C}\|_1 \leq 2, \text{Tr}(\tilde{C}S) = 0, \forall \tilde{C} \in \mathbb{H}_d, S \in \mathcal{S},\) \(\text{(30)}\)

where \(\|\cdot\|_1\) is the trace norm. A similar SDP problem was considered in Ref. [31]. The optimal \(|\xi|\) is equal to

\[
2 \min_{S \in \mathcal{S}} \|H - S\|^2
\]

where we used the fact that for any \(S \in \mathcal{S}\) there is an \(h \in \mathbb{H}_d\) such that \(S = KHK^\dagger\) and vice versa. \textbf{Theorem 2} is then proven. Note that, given the optimal \(C\), we can always choose \(A_0, A_1\) with orthogonal supports and the last ancillary qubit in \(H_A\) could be removed because \(|0\rangle\) and \(|1\rangle\) in this case could be distinguished by projections onto the orthogonal supports in \(H_A\) [31]. Therefore a \(d\)-dimensional ancillary system is sufficient.

We have demonstrated the QEC code achieving the optimal HL for arbitrary quantum channels. The code is designed to satisfy the Knill-Laflamme condition and optimize the QFI. The logical dephasing channel is exactly the identity channel at the true value of \(\omega\) and any change in \(\omega\) results in a detectable phase, allowing it to be estimated at the HL.

\section{Achieving the SQL Upper Bound}

When \(H \in \mathcal{S}\), the situation is much more complicated because when \(|\xi| = 1\) we must also have \(|\xi| = 0\) and no signal could be detected. Therefore we must consider the trade-off between maximizing the signal and minimizing the noise. To be exact, we want to maximize

\[
\tilde{\mathfrak{g}}_{\text{SQL}}(D_{L,\omega}) = \frac{|\xi|^2}{1 - |\xi|^2}.
\]

We will show for any \(\eta > 0\), there exists a near-optimal code and recovery such that \(\tilde{\mathfrak{g}}_{\text{SQL}}(D_{L,\omega}) > \tilde{\mathfrak{g}}^{(\eta)}_{\text{SQL}}(\varepsilon_{\omega}) - \eta\), proving \textbf{Theorem 3}. We only consider the case where \(\tilde{\mathfrak{g}}_{\text{SQL}}(\varepsilon_{\omega}) > \tilde{\mathfrak{g}}_1(\varepsilon_{\omega}) > 0\) because otherwise \(\tilde{\mathfrak{g}}_1(\varepsilon_{\omega}) = \tilde{\mathfrak{g}}_{\text{SQL}}(\varepsilon_{\omega})\) and product states are sufficient to achieve \(\tilde{\mathfrak{g}}_{\text{SQL}}(\varepsilon_{\omega})\). Detailed derivations could be found in \textbf{Appx. E} and we sketch the proof here. To simplify the calculation, we consider a special type of code, first introduced in Ref. [51], where

\[
A_0 = \sqrt{1 - \varepsilon^2} C + \varepsilon D, \quad A_1 = \sqrt{1 - \varepsilon^2} C - \varepsilon D, \quad \text{(33)}
\]

satisfying \(\text{Tr}(C^\dagger D) = 0\) and \(\text{Tr}(C^\dagger C) = \text{Tr}(D^\dagger D) = 1\). In this section, we define \(\bar{C} = C^\dagger D + D^\dagger C\) (differed by a factor of \(\varepsilon\sqrt{1 - \varepsilon^2}\) from the \(\tilde{C}\) defined in \textbf{Sec. V}) and also assume \(C\) is full rank so that \(\bar{C}\) could be an arbitrary Hermitian matrix. \(\varepsilon\) is a small parameter and we will calculate \(\mathfrak{g}_{\text{SQL}}(D_{L,\omega})\) up to the lowest order of \(\varepsilon\).

To proceed, we first introduce the vectorization of matrices \(|\psi\rangle = \sum_{ij} *_{ij} |i \rangle |j\rangle\) for all \(* \in C^{d \times d}\) to simplify the notation. We define \(E_{0,1} = \sqrt{1 - \varepsilon^2} E \pm \varepsilon F \in C^{d^2 \times d^2}\) where

\[
E = (|K_1C\rangle \cdots |K_rC\rangle), \quad F = (|K_1D\rangle \cdots |K_rD\rangle), \quad \text{(34)}
\]

satisfying \(\text{Tr}(E^\dagger F) = 0\) and \(\text{Tr}(E^\dagger E) = \text{Tr}(F^\dagger F) = 1\). Let the recovery matrix \(T = QR^\dagger \in C^{d^2 \times d^2}\), then

\[
\xi = \text{Tr}(TE_0E_1^\dagger), \quad \dot{\xi} = \text{Tr}(TE_0E_1^\dagger + TE_1E_0^\dagger). \quad \text{(35)}
\]

We consider the regime where both the signal and the noise are sufficiently small—both the denominator and the numerator in \textbf{Eq. (32)} will be \(O(\varepsilon^2)\). The recovery matrix \(T\) should also be close to the identity operator. We assume \(T = e^{\varepsilon G}\) where \(G\) is Hermitian and let \(\sigma = EE^\dagger\), \(\tilde{\sigma} = i(FE^\dagger - EF^\dagger)\). Expanding \(T, E_0, E_1\) around \(\varepsilon = 0\), we first optimize \(\mathfrak{g}_{\text{SQL}}(D_{L,\omega})\) over all possible \(G\), which gives (up to the lowest order of \(\varepsilon\)),

\[
\tilde{\mathfrak{g}}_{\text{SQL}}(D_{L,\omega}) \approx \max_G |\text{Tr}(G\tilde{\sigma})|^2 \sqrt{\frac{|\text{Tr}(G\tilde{\sigma})|^2}{4 - |\text{Tr}(G\tilde{\sigma})|^2}} + (\Delta G^2)_{\sigma}. \quad \text{(36)}
\]

The maximization could be calculated by taking the derivative with respect to \(G\). We can show that the optimal \(G\) is

\[
G_{\text{opt}} = (4 - \text{Tr}(L_\sigma[\tilde{\sigma}])L_\sigma[\tilde{\sigma}])L_\sigma[\tilde{\sigma}] + \text{Tr}(L_\sigma[\tilde{\sigma}])L_\sigma[\tilde{\sigma}], \quad \text{(37)}
\]

and the corresponding optimal QFI is

\[
\tilde{\mathfrak{g}}_{\text{SQL}}(D_{L,\omega}) \approx \text{Tr}(L_\sigma[\tilde{\sigma}]) + \frac{\text{Tr}(L_\sigma[\tilde{\sigma}])^2}{4 - \text{Tr}(L_\sigma[\tilde{\sigma}])}. \quad \text{(38)}
\]

Now \(\tilde{\mathfrak{g}}_{\text{SQL}}(D_{L,\omega})\) is a function of the code \((C\) and \(D\)) only. We will further simplify it such that it is a function of only \(C\) and \(\tilde{C}\). Let \(\tau = E^\dagger E, \tilde{\tau} = E^\dagger F + F^\dagger E\), \(\tau' = iE^\dagger E - iE^\dagger E\) such that

\[
\tau_{ij} = \text{Tr}(C^\dagger K_i K_j C), \quad \tilde{\tau}_{ij} = \text{Tr}(\tilde{C} K_i^\dagger K_j), \quad \tau'_{ij} = i\text{Tr}(C^\dagger K_i^\dagger \tilde{K}_j C) - i\text{Tr}(C^\dagger \tilde{K}_i K_j C). \quad \text{(39)}
\]

Then we can verify that

\[
\text{Tr}(L_\sigma[\tilde{\sigma}]) = 4\text{Tr}(C^\dagger \tilde{K}_i^\dagger K_j C) - \text{Tr}(L_\tau[\tau']) , \quad \text{(41)}
\]

\[
\text{Tr}(L_\sigma[\tilde{\sigma}]) = -2\text{Tr}(\tilde{C} H) + \text{Tr}(L_\tau[\tilde{\tau}]) , \quad \text{(42)}
\]

\[
\text{Tr}(L_\sigma[\tilde{\sigma}]) = 4 - \text{Tr}(L_\tau[\tau']) , \quad \text{(43)}
\]

and

\[
\tilde{\mathfrak{g}}_{\text{SQL}}(D_{L,\omega}) \approx f(C, \tilde{C}) = 4\text{Tr}(C^\dagger \tilde{K}_i^\dagger K_j C) - \text{Tr}(L_\tau[\tau']) + \frac{(-2\text{Tr}(\tilde{C} H) + \text{Tr}(L_\tau[\tilde{\tau}]))^2}{\text{Tr}(L_\tau[\tau'])}. \quad \text{(44)}
\]

At this stage, it is not obvious why the maximization of \(\tilde{\mathfrak{g}}_{\text{SQL}}(D_{L,\omega})\) over \(C\) and \(\tilde{C}\) is equal to \(\tilde{\mathfrak{g}}_{\text{SQL}}(\varepsilon_{\omega})\). To see that,
we need to reformulate the SQL upper bound using its dual program. First we note that
\[ \tilde{\Phi}_{\text{sql}}^{(u)}(\mathcal{E}_\omega) = \max_{C: \text{Tr}(C^\dagger C) = 1} \min_{\alpha, \beta = 0} 4\text{Tr}(C^\dagger \alpha C), \]
where we are allowed to exchange the order of maximization and minimization thanks to Sion’s minimax theorem [71, 72]. Fixing \( C \), we consider the optimization problem \( \min_{\alpha, \beta = 0} 4\text{Tr}(C^\dagger \alpha C) \). When \( C \) is full rank, we can show that it is equivalent to \( \max_{\alpha, \beta = 0} f(C, \tilde{C}) \), where \( \tilde{C} \) is introduced as the Lagrange multiplier associated with the constraint \( \beta = 0 \) [73].

The procedure to find a near-optimal code such that \( \tilde{\Phi}_{\text{sql}}^{(u)}(\mathcal{D}_{1, \omega}) > \tilde{\Phi}_{\text{sql}}^{(u)}(\mathcal{E}_\omega) \) for any \( \eta > 0 \) goes as follows:

1. Find a full rank \( C^0 \) such that \( \text{Tr}(C^0 C^0) = 1 \) and \( \min_{\alpha, \beta = 0} 4\text{Tr}(C^0 \alpha C^0) < \tilde{\Phi}_{\text{sql}}^{(u)}(\mathcal{E}_\omega) - \eta/2 \).

2. Find a Hermitian \( C^0 \) such that \( f(C^0, \tilde{C}) \) is maximized and let \( D^0 = \frac{1}{2} C^0 - \tilde{C}^0 \). Rescale \( D^0 \) such that \( \text{Tr}(D^0 D^0) = 1 \).

3. Calculate \( \tilde{\Phi}_{\text{sql}}^{(u)}(\mathcal{D}_{1, \omega}) \) using Eqs. (33)-(35) and Eq. (37). Find a small \( \epsilon^o > 0 \) such that \( \tilde{\Phi}_{\text{sql}}^{(u)}(\mathcal{D}_{1, \omega}) > f(C^0, \tilde{C}^0) + \eta/2 \).

The numerical algorithms for step (1) and (2) are provided in Appx. F, where the most computationally intensive part is a SDP.

To conclude, we proposed a perturbation code which could achieve the SQL upper bound with an arbitrarily small error. We take the limit where the parameter \( \epsilon \) which distinguishes the logical zero and one states is sufficiently small. Note that if we take \( \epsilon = 0 \), the probe state will be a product state and we can only achieve \( \tilde{\Phi}_{\text{sql}}^{(u)}(\mathcal{D}_{1, \omega}) \). This discontinuity appears because we must first take the limit \( N \rightarrow \infty \) before taking the limit \( \epsilon \rightarrow 0 \) and the impact of a small \( \epsilon \) becomes significant in the asymptotic limit.

**VII. EXAMPLES**

**A. Depolarizing channels**

In this section, we calculate \( \tilde{\Phi}_{1}(N_{\omega}) \), \( \tilde{\Phi}_{\text{sql}}(N_{\omega}) \), and \( \tilde{\Phi}_{\text{hl}}(N_{\omega}) \) for depolarizing channels \( N_{\omega}(\rho) = N(U_{\omega}(\rho)) \) where
\[ N(\rho) = (1 - p)\rho + p_x \sigma_x \rho \sigma_x + p_y \sigma_y \rho \sigma_y + p_z \sigma_z \rho \sigma_z, \]
\[ p_{x,y,z} \geq 0, p = p_x + p_y + p_z < 1 \text{ and } U_{\omega}(\cdot) = e^{-i\omega \sigma_z (\cdot)} e^{i\omega \sigma_z}. \]

First, we notice that HNKS is satisfied if and only if \( p_x = p_z = 0 \) or \( p_x = p_y = 0 \). When HNKS is satisfied, \( \tilde{\Phi}_{\text{sql}}(N_{\omega}) = 1 \). It is the same as the \( \tilde{\Phi}_{\text{hl}} \) when there is no noise \( p = 0 \) because the Kraus operator \( \{\sigma_x, \sigma_y\} \) is perpendicular to the Hamiltonian \( \{\sigma_z\} \) and could be fully corrected. It is consistent with previous results for single qubit Hamiltonian estimation that the HL is achievable if and only if the Markovian noise is rank-one and not parallel to the Hamiltonian [29, 56–61]. As calculated in Appx. G,
\[ \tilde{\Phi}_{1}(N_{\omega}) = 1 - w, \]
where \( w = 4 \left( \frac{p_x p_y}{p_x + p_y} + \frac{1 - p_x p_y}{1 - p_x p_y} \right) \leq 1 \). When HNKS is violated,
\[ \tilde{\Phi}_{\text{sql}}(N_{\omega}) = (1 - w)/w. \]

In the equations above, when \( p_x = p_y = 0 \), we take \( \frac{p_x p_y}{p_x + p_y} = 0 \), in which case \( N_{\omega} \) becomes the dephasing channel introduced in Sec. III where \( \phi = \omega \) and \( p \) is independent of \( \omega \).

We observe that
\[ \tilde{\Phi}_{\text{sql}}(N_{\omega}) = \tilde{\Phi}_{1}(N_{\omega})/\omega \geq \tilde{\Phi}_{1}(N_{\omega}), \]
and the equality \( (w = 1) \) holds if and only if \( p_x = p_y = 1/2 \), in which case \( \tilde{\Phi}_{\text{sql}}(N_{\omega}) = \tilde{\Phi}_{1}(N_{\omega}) = 0 \) as \( \omega = \frac{N}{\omega} \) becomes a mixture of a completely dephasing channel and a completely depolarizing channel [74] where \( \omega \) cannot be detected.

The asymptotic QFI is in general non-additive. In particular, when \( p \ll 1 \), we have \( w \ll 1 \) and \( \tilde{\Phi}_{\text{sql}}(N_{\omega}) \gg \tilde{\Phi}_{1}(N_{\omega}) \). We also illustrate the difference between \( \tilde{\Phi}_{\text{sql}}(N_{\omega}) \) and \( \tilde{\Phi}_{1}(N_{\omega}) \) as a function of \( p_x \) and \( p_y \) when \( p_x = 0.1 \). \( \tilde{\Phi}_{\text{sql}}(N_{\omega}) = \tilde{\Phi}_{1}(N_{\omega}) = 0 \) at \( (p_x, p_y, p_z) = (0.4, 0.4, 0.1) \). The ratio \( \tilde{\Phi}_{\text{sql}}(N_{\omega})/\tilde{\Phi}_{1}(N_{\omega}) \) increases near the boundary of \( p_x + p_y < 0.9 \).

**B. U-covariant channels**

Let \( U = \{U_i\}_{i=1}^n \subset \mathbb{C}^{d \times d} \) be a set of unitary operators such that for some probability distribution \( \{p_i\}_{i=1}^n \), \( \{p_i, U_i\}_{i=1}^n \) is a unitary 1-design [75], satisfying
\[ \sum_{i=1}^n p_i U_i A U_i^\dagger = \text{Tr}(A) \frac{I}{d}, \forall A \in \mathbb{C}^{d \times d}. \]
be a solution of
\[ \langle \alpha | U \rho U^\dagger | \alpha \rangle = \rho_{\alpha} \]
for all \( \alpha \) and \( h \), where \( \alpha^* = \alpha |_{h=k=\alpha} \). Then \( \{|C_i\} \in \mathcal{H}_P \otimes \mathcal{H}_A \) is an optimal input state of a single quantum channel \( \mathcal{T}_\omega \), and if only if \( \rho^* = C^* C^{-1} \) satisfies Eq. (52). According to Eq. (51), if \(|C_i\) is an optimal input, \(|U C_i\rangle = (U \otimes I) |C_i\rangle\) is also an optimal input for all \( U \in \mathbb{U} \) and satisfies Eq. (52). Then \( \sum_{i=1}^n p_i U_i \rho U_i^\dagger = \frac{1}{n} \) also satisfies Eq. (52), implying the maximally entangled state \(|\frac{1}{\sqrt{n}}\rangle\) is an optimal input for \( \mathcal{T}_\omega \). The discussion above also works for \( \mathcal{T}_\omega^{\otimes N} \) because \( \mathcal{T}_\omega^{\otimes N} \) is \( \mathbb{U}^{\otimes N} \)-covariant and \( \{|\frac{1}{\sqrt{n}}\rangle\} \) is a unitary 1-design on \( C^{N_d \times N_d} \). Therefore \(|\frac{1}{\sqrt{n}}\rangle\) is an optimal input for \( \mathcal{T}_\omega^{\otimes N} \), which implies \( \hat{\mathcal{S}}_N^{\otimes N} \) is an optimal input for \( \mathcal{T}_\omega^{\otimes N} \).

VIII. CONCLUSIONS AND OUTLOOK

In this paper, we focus on the asymptotic behaviour of the QFI of a quantum channel when the number of identical channels \( N \) is infinitely large. We consolidate the HNKS condition by showing it unambiguously determines whether or not the scaling of the asymptotic QFI is quadratic or linear. In both cases, we show that the optimal input state achieving the asymptotic QFI could be solved via a SDP. To find the optimal input state, we reduce every quantum channel to a single qubit dephasing channel where both the phase and the noise parameter vary wrt the unknown parameter and then optimize the asymptotic QFI of the logical dephasing channel over the encoding and the recovery channel. The optimal input state is either the logical GHZ state (when HNKS is satisfied) or the logical spin-squeezed state (when HNKS is violated). This provides a unified framework for channel estimation while previous known results are centered on either Hamiltonian or noise estimation separately.

The metrological protocol we considered in Fig. 1b is usually called parallel strategies where \( N \) identical quantum channels act in parallel on a quantum state. Researchers also consider a more powerful protocol called sequential strategies where we allow arbitrary quantum controls between each quantum channel [28]. The QFI optimized over all possible inputs and quantum controls has a similar (but different) upper bound [78] to Eq. (8), from which we can see that Theorem 1 and Theorem 3 still hold true for sequential strategies. The conclusions in Ref. [30, 31] and Ref. [51] could be viewed as an instance of Theorem 1 and Theorem 3 for sequential strategies where we estimate the Hamiltonian parameter under Markovian noise in an infinitely small time interval. Theorem 3, which holds for both parallel and sequential strategies also implies when HNKS is violated (for example when we estimate the proportion of two quantum channels in a mixture of them), there is no advantage of sequential strategies over parallel strategies asymptotically. Related problems were also considered in the asymmetric channel discrimination setting where it was also shown that there is no advantage of sequential strategies over parallel strategies asymptotically [79–81]. Finally, it would also be interesting to see whether our results are generalizable to scenarios where memory effect is considered [82, 83].

ACKNOWLEDGMENTS

We thank Chang-Ling Zou, Haining Pan, Zi-Wen Liu and Yuxiang Yang for helpful discussions. We acknowledge support from the ARL-CDQI (W911NF-15-2-0067), ARO (W911NF-18-1-0020, W911NF-18-1-0212), ARO MURI (W911NF-16-1-0349), AFOSR MURI (FA9550-15-1-0015, FA9550-19-1-0399), DOE (DE-SC0019406), NSF (EFMA-1640959, OMA-1936118), and the Packard Foundation (2013-2027).

[1] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nature Photonics 5, 222 (2011).
[2] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, Reviews of Modern Physics 89, 035002 (2017).
[3] D. Braun, G. Adesso, F. Benatti, R. Floreanini, U. Marzolino, M. W. Mitchell, and S. Pirandola, Quantum-enhanced measurements without entanglement, Reviews of Modern Physics 90, 035006 (2018).
[4] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Quantum metrology with nonclassical states of atomic ensembles, Reviews of Modern Physics 90, 035005 (2018).
[5] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Advances in photonic quantum sensing, Nature Photonics 12, 724 (2018).
[6] B. Sanders and G. Milburn, Optimal quantum measurements for phase estimation, Physical Review Letters 75, 2944 (1995).
[7] J. Bollinger, W. M. Itano, D. Wineland, and D. Heinzen, Optimal frequency measurements with maximally correlated states, Physical Review A 54, R4649 (1996).
[8] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Improvement of frequency standards with quantum entanglement, Physical Review Letters 79, 3865 (1997).

[9] D. Leibfried, M. D. Barrett, T. Schaetz, J. Britton, J. Chiaverini, W. M. Itano, J. D. Jost, C. Langer, and D. J. Wineland, Toward heisenberg-limited spectroscopy with multiparticle entangled states, Science 304, 1476 (2004).

[10] C. M. Caves, Quantum-mechanical noise in an interferometer, Physical Review D 23, 1693 (1981).

[11] B. Yurke, S. L. McCall, and J. R. Klauder, Su (2) and su (1, 1) interferometers, Physical Review A 33, 4033 (1986).

[12] D. Berry and H. Wiseman, Optimal states and almost optimal adaptive measurements for quantum interferometry, Physical Review Letters 85, 5098 (2000).

[13] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Entanglement-free heisenberg-limited phase estimation, Nature 450, 393 (2007).

[14] V. Buzek, R. Derka, and S. Massar, Optimal quantum clocks, Physical Review Letters 82, 2207 (1999).

[15] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum-enhanced measurements: beating the standard quantum limit, Science 306, 1330 (2004).

[16] A. Valencia, G. Scarcelli, and Y. Shih, Distant clock synchronization using entangled photon pairs, Applied Physics Letters 85, 2655 (2004).

[17] M. de Burgh and S. D. Bartlett, Quantum methods for clock synchronization: Beating the standard quantum limit without entanglement, Physical Review A 72, 042301 (2005).

[18] C. W. Helstrom, Quantum detection and estimation theory (Academic press 1976).

[19] A. S. Holevo, Probabilistic and statistical aspects of quantum theory, volume 1 (Springer Science & Business Media 2011).

[20] M. G. Paris, Quantum estimation for quantum technology, International Journal of Quantum Information 7, 125 (2009).

[21] S. L. Braunstein and C. M. Caves, Statistical distance and the geometry of quantum states, Physical Review Letters 72, 3439 (1994).

[22] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum metrology, Physical Review Letters 96, 010401 (2006).

[23] A. Fujiwara and H. Imai, A fibre bundle over manifolds of quantum channels and its application to quantum statistics, Journal of Physics A: Mathematical and Theoretical 41, 255304 (2008).

[24] Z. Ji, G. Wang, R. Duan, Y. Feng, and M. Ying, Parameter estimation of quantum channels, IEEE Transactions on Information Theory 54, 5172 (2008).

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

[26] M. Hayashi, Comparison between the cramer-rao and the minimax approaches in quantum channel estimation, Communications in Mathematical Physics 304, 689 (2011).

[27] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, The elusive heisenberg limit in quantum-enhanced metrology, Nature Communications 3, 1063 (2012).

[28] R. Demkowicz-Dobrzański and L. Maccone, Using entanglement against noise in quantum metrology, Physical Review Letters 113, 250801 (2014).

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

[30] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Adaptive quantum metrology under general markovian noise, Physical Review X 7, 041009 (2017).

[31] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Achieving the heisenberg limit in quantum metrology using quantum error correction, Nature Communications 9, 78 (2018).

[32] D. Layden and P. Cappellaro, Spatial noise filtering through error correction for quantum sensing, npj Quantum Information 4, 30 (2018).

[33] D. Layden, S. Zhou, P. Cappellaro, and L. Jiang, Ancilla-free quantum error correction codes for quantum metrology, Physical Review Letters 122, 040502 (2019).

[34] W. Górecki, S. Zhou, L. Jiang, and R. Demkowicz-Dobrzański, Optimal probes and error-correction schemes in multi-parameter quantum metrology, arXiv preprint arXiv:1901.00896 (2019).

[35] H. Yuan and C.-H. F. Fung, Quantum parameter estimation with general dynamics, npj Quantum Information 3, 1 (2017).

[36] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. Moore, and D. Heinzen, Spin squeezing and reduced quantum noise in spectroscopy, Physical Review A 46, R6797 (1992).

[37] D. Ulam-Orgikh and M. Kitagawa, Spin squeezing and decoherence limit in Ramsey spectroscopy, Physical Review A 64, 052106 (2001).

[38] R. Demkowicz-Dobrzański, K. Banaszek, and R. Schnabel, Fundamental quantum interferometry bound for the squeezed-light-enhanced gravitational wave detector geo 600, Physical Review A 88, 041802 (2013).

[39] R. Chaves, J. Brask, M. Markiewicz, J. Kołodyński, and A. Acín, Noisy metrology beyond the standard quantum limit, Physical Review Letters 111, 120401 (2013).

[40] M. B. Plenio and S. F. Huelga, Sensing in the presence of an observed environment, Physical Review A 93, 032123 (2016).

[41] F. Albarelli, M. A. Rossi, M. G. Paris, and M. G. Genoni, Ultimate limits for quantum magnetometry via time-continuous measurements, New Journal of Physics 19, 123011 (2017).

[42] F. Albarelli, M. A. Rossi, D. Tamascelli, and M. G. Genoni, Restoring heisenberg scaling in noisy quantum metrology by monitoring the environment, Quantum 2, 110 (2018).

[43] Y. Matsuzaki, S. C. Benjamin, and J. Fitzsimons, Magnetic field sensing beyond the standard quantum limit under the effect of decoherence, Physical Review A 84, 012103 (2011).

[44] A. W. Chin, S. F. Huelga, and M. B. Plenio, Quantum metrology in non-markovian environments, Physical Review Letters 109, 233601 (2012).

[45] A. Smirne, J. Kołodyński, S. F. Huelga, and R. Demkowicz-Dobrzański, Ultimate precision limits for noisy frequency estimation, Physical Review Letters 116, 120801 (2016).

[46] J. Liu and H. Yuan, Quantum parameter estimation with optimal control, Physical Review A 96, 012117 (2017).

[47] H. Xu, J. Li, L. Liu, Y. Wang, H. Yuan, and X. Wang, Generalizable control for quantum parameter estimation through reinforcement learning, npj Quantum Information 5, 1 (2019).

[48] K. Chabuda, J. Dziarmaga, T. J. Osborne, and R. Demkowicz-Dobrzański, Tensor-network approach for quantum metrology in many-body quantum systems, Nature Communications 11, 1 (2020).

[49] S. Pirandola and C. Lupo, Ultimate precision of adaptive noise estimation, Physical Review Letters 118, 100502 (2017).

[50] R. Laurenza, C. Lupo, G. Spedalieri, S. L. Braunstein, and S. Pirandola, Channel simulation in quantum metrology, Quantum Measurements and Quantum Technology 5, 1 (2018).

[51] S. Zhou and L. Jiang, Optimal approximate quantum error correction for quantum metrology, Phys. Rev. Research 2, 013235 (2020).
Appendix A: Deriving the upper bound of \( \tilde{S}_N(\mathcal{E}_\omega) \)

For completeness, we provide a proof \([23]\) of Eq. (8) in the main text. Let \( K_i^{(1)} = K_i \) for \( i \in [r] \), where \( [r] = \{1, 2, \ldots, r\} \). Inductively, let

\[
K_i^{(n+1)} = K_i^{(n)} \otimes K_i^{(1)}, \quad \forall t = (t_1, t_2) \in [r]^n \times [r].
\]

\{ \(K_i^{(n)}\)\}_{i \in [r]^n} \text{ is a Kraus representation of } \mathcal{E}_\omega \text{ for all } n. \text{ Then let } \alpha^{(n)} = \sum_{t_1} K_i^{(n)} K_i^{(1)}, \beta^{(n)} = i \sum_{t_1} K_i^{(n)} \, \tilde{K}_i^{(n)} \text{, we have}

\[
\alpha^{(n+1)} = \sum_{t_1, t_2} \left( \frac{\partial (K_i^{(n)} \otimes K_i^{(1)})}{\partial \omega} \right)^\dagger \left( \frac{\partial (K_i^{(n)} \otimes K_i^{(1)})}{\partial \omega} \right) = \alpha^{(n)} \otimes I + 2 \beta^{(n)} \otimes \beta^{(1)} + I \otimes \alpha^{(1)}, \quad \beta^{(n+1)} = i \sum_{t_1, t_2} \left( \frac{\partial (K_i^{(n)} \otimes K_i^{(1)})}{\partial \omega} \right)^\dagger (K_i^{(n)} \otimes K_i^{(1)}) = \beta^{(n)} \otimes I + I \otimes \beta^{(1)}. \tag{A2}
\]

The solution is \( \beta^{(N)} = \sum_{k=0}^{N-1} I \otimes \beta^{(1)} \otimes I \otimes \otimes \otimes \quad \text{and}

\[
\alpha^{(N)} = \sum_{k=0}^{N-1} I \otimes \alpha^{(1)} \otimes I \otimes \otimes \otimes + 2 \sum_{k_1=0}^{N-2} \sum_{k_2=0}^{N-k_1} I \otimes k_1 \otimes \beta^{(1)} \otimes I \otimes k_2 \otimes \beta^{(1)} \otimes I \otimes \otimes \otimes . \tag{A4}
\]

Therefore, \( \mathcal{S}_N(\mathcal{E}_\omega) \leq 4 \| \alpha^{(N)} \| \leq 4N \| \alpha^{(1)} \| + 4N(N-1) \| \beta^{(1)} \|^2 \) and the inequality holds for any Kraus representation of \( \mathcal{E}_\omega \). We can choose \( \mathbf{K}' = \mathbf{nK} \), then

\[
\mathcal{S}_N(\mathcal{E}_\omega) \leq 4 \min_h (N \| \alpha \| + N(N-1) \| \beta \|^2), \tag{A5}
\]

where \( h = iu^\dagger \hat{u} \) is an arbitrary Hermitian matrix, \( \alpha = \mathbf{K}^\dagger \mathbf{K}' = (\mathbf{K} - ih\mathbf{K})(\mathbf{K} - ih\mathbf{K})^\dagger \) and \( \beta = i \mathbf{K}^\dagger \mathbf{K}' = i \mathbf{K}^\dagger (\mathbf{K} - ih\mathbf{K}) \).

Appendix B: Calculating the QFI upper bounds for dephasing channels

Here we calculate \( \mathcal{S}_{\text{inh}}^{(a)} = 4 \min_h \| \beta \|^2 \) and \( \mathcal{S}_{\text{sqz}}^{(a)} = 4 \min_h, \beta = 0 \| \alpha \| \) for dephasing channels

\[
\mathcal{D}_\omega(\rho) = (1 - p)e^{-\frac{i\sigma_z}{2}}\rho e^{\frac{i\sigma_z}{2}} + p\sigma_z e^{-\frac{i\sigma_z}{2}}\rho e^{\frac{i\sigma_z}{2}}\sigma_z = \sum_{i=1}^{2} K_i \rho K_i^\dagger, \tag{B1}
\]

where \( K_1 = \sqrt{1 - p} e^{-\frac{i\sigma_z}{2}}, K_2 = \sqrt{p} \sigma_z e^{-\frac{i\sigma_z}{2}}. \) Assume \( p > 0, \) then

\[
\mathbf{K} = \left( \sqrt{1 - p} e^{-\frac{i\sigma_z}{2}}, \sqrt{p} \sigma_z e^{-\frac{i\sigma_z}{2}} \right), \quad \mathbf{K} = \left( \begin{array}{c} \sqrt{\frac{p}{2}} \sqrt{1 - p} - \sqrt{\frac{1}{2}} \sigma_z \\ -\sqrt{\frac{1}{2}} \sigma_z \end{array} \right) , \quad \mathbf{K} = \frac{e^{-\frac{i\sigma_z}{2}}}{\sqrt{1 - p}} \left( \begin{array}{c} \sqrt{\frac{p}{2}} \sqrt{1 - p} - \sqrt{\frac{1}{2}} \sigma_z \\ -\sqrt{\frac{1}{2}} \sigma_z \end{array} \right) \tag{B2}
\]

\[
\mathbf{K} - ih\mathbf{K} = \frac{1}{\sqrt{1 - p}} \left( \sqrt{\frac{p}{2}} \sqrt{1 - p} - i h_1 \sqrt{1 - p} - \sqrt{\frac{1}{2}} \sigma_z \right) + \frac{\sqrt{1 - p}}{2} \sigma_z - i h_2 \sqrt{\frac{1}{2}} \sigma_z - i h_2 \sqrt{1 - p} e^{-\frac{i\sigma_z}{2}} \sigma_z, \tag{B3}
\]

\[
\beta = i \mathbf{K}^\dagger (\mathbf{K} - ih\mathbf{K}) = \frac{e^{-\frac{i\sigma_z}{2}}}{h_1} \left( \begin{array}{c} \sqrt{\frac{p}{2}} \sqrt{1 - p} - \sqrt{\frac{1}{2}} \sigma_z \\ -\sqrt{\frac{1}{2}} \sigma_z \end{array} \right) \left( \begin{array}{c} \sqrt{\frac{p}{2}} \sqrt{1 - p} - \sqrt{\frac{1}{2}} \sigma_z \\ -\sqrt{\frac{1}{2}} \sigma_z \end{array} \right) \sigma_z + (1 - p) h_{11} + ph_{22} + \sqrt{p(1 - p)}(h_{12} + h_{21}) \sigma_z. \tag{B4}
\]

\[
\alpha = (\mathbf{K} - ih\mathbf{K})^\dagger (\mathbf{K} - ih\mathbf{K}) = \frac{\hat{p}^2}{4(1 - p)} + \frac{|h_{12}|^2}{4p(1 - p)} + \frac{\phi^2}{4} + |h_{12}|^2 + 2 \sqrt{p(1 - p)} \phi \text{Re}[h_{12}] \tag{B5}
\]

\[
+ 2 \text{Re} \left[ - \frac{\sqrt{p}}{\sqrt{1 - p}} ih_{12} + ((1 - p) h_{11} + h_{22}) \frac{\phi}{2} + (h_{11} + h_{22}) \sqrt{p(1 - p)} - ih_{12} + h_{22} \sqrt{p(1 - p)} - i \hat{p} \sqrt{1 - p} \right] \sigma_z.
\]

\( \beta = 0 \) is equivalent to \( (1 - p) h_{11} + ph_{22} = 0 \) and \( \frac{\phi}{2} + \sqrt{p(1 - p)}(h_{12} + h_{21}) = 0 \), which is achievable for any \( p > 0 \). When \( h_{11} = h_{22} = 0 \) and \( h_{12} = h_{21} = -\frac{\phi}{2 \sqrt{p(1 - p)}} \), \( \| \alpha \| = \| \mathcal{S}_{\text{inh}}^{(a)}(\mathcal{D}_\omega) \| = \frac{(1 - 2p)^2 \phi^2}{4p(1 - p)} + \frac{\hat{p}^2}{1 - |\xi|^2}, \tag{B6}
\]
where $\xi = (1 - 2p)e^{-i\phi} = \langle 0 | D_\omega(|0\rangle (1) | 1 \rangle$ is a complex number completely determining the channel.

When $p = 0$, we must also have $\bar{p} = 0$. Then $\beta = \frac{\phi}{2} \sigma_z + h_{11}$ and

$$\mathcal{S}_0^{(\mu)}(D_\omega) = 4 \min_h \|\beta\|^2 = |\phi|^2 = |\xi|^2. \quad (B7)$$

We can also calculate the channel QFI

$$\mathcal{S}_1^{(\mu)}(D_\omega) = 4 \min_h \|\alpha\| = \begin{cases} (1 - 2p)^2 \phi^2 + \frac{\bar{p}^2}{(1 - p)^2}, & p > 0, \\ (1 - 2p)^2 \bar{p}^2, & p = 0. \end{cases} \quad (B8)$$

It could be achieved using $|\psi_0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$.

\textbf{Appendix C: A useful formula for calculating the QFI of dephasing channels}

In this appendix, we prove Eq. (16) in the main text. Let $|\psi\rangle = e^{-i\phi J_z} |\psi_0\rangle$ and a subspace

$$\mathcal{Z} = \text{span}\{ \prod_{j=1}^N (\sigma_z^{(k)})^{j_k} |\psi\rangle, (j_1, \ldots, j_N) \in \{0, 1\}^N \}. \quad (C1)$$

Assume $\dim \mathcal{Z} = n$. $\mathcal{Z}$ must have an orthonormal basis $\{ |e_{\ell}\rangle \}_{\ell=1}^n$ where $|e_{\ell}\rangle = \sum_{j_1, \ldots, j_N = 0}^1 r_{\ell, (j_1, \ldots, j_N)} \prod_{k=1}^N (\sigma_z^{(k)})^{j_k} |\psi\rangle$ with real $r_{\ell, (j_1, \ldots, j_N)}$. For example, one can use the Gram-schmidt procedure to find $\{ |e_{\ell}\rangle \}_{\ell=1}^n$ because $|\psi\rangle \prod_{k=1}^N (\sigma_z^{(k)})^{j_k} |\psi\rangle \in \mathbb{R}$ for all $(j_1, \ldots, j_N) \in \{0, 1\}^\otimes N$. Then

$$\rho_\omega = D_\omega^\otimes N(|\psi\rangle \langle \psi|) = (D_\omega |\phi=0\rangle^\otimes N(|\psi\rangle \langle \psi|)$$

$$= \sum_{j_1, \ldots, j_N = 0}^1 (1 - p)^{(N - \sum_{k=1}^N j_k)} p^\sum_{k=1}^N j_k \prod_{k=1}^N (\sigma_z^{(k)})^{j_k} |\psi\rangle \langle \psi| \prod_{k=1}^N (\sigma_z^{(k)})^{j_k} = \sum_{\ell, \ell' = 1}^n \chi_{\ell\ell'} |e_{\ell}\rangle \langle e_{\ell'}| \quad (C2)$$

where $\chi \in \mathbb{R}^{n \times n}$ is a symmetric matrix. $\chi = \sum_{i=1}^n \mu_i v_i v_i^T$ where $v_i$ are real orthonormal eigenvectors of $\chi$. Then we can write $\rho_\omega = \sum_{\ell=1}^n \mu_{\ell} |\psi_\ell\rangle \langle \psi_\ell|$ where $|\psi_\ell\rangle = \sum_{\ell'=1}^n \chi_{\ell\ell'} |e_{\ell'}\rangle$. Then according to the definition of QFI,

$$F(\rho_\omega) = 2 \sum_{\ell, \ell' : \mu_\ell + \mu_{\ell'} \neq 0} \frac{|\langle \psi_\ell | \hat{\rho}_\omega | \psi_{\ell'}\rangle|^2}{\mu_\ell + \mu_{\ell'}}. \quad (C3)$$

Note that in principle Eq. (C3) only holds true when $\{ |\psi_\ell\rangle \}$ is a complete basis of $\mathcal{H}_\ell^\otimes N$, that is, $\text{span} \{ |\psi_\ell\rangle \} = \mathcal{H}_\ell^\otimes N$. However, here we only consider all states in the subspace $\mathcal{Z}$ because $\Pi_{\mathcal{Z}} \rho_\omega \Pi_{\mathcal{Z}} = \hat{\rho}_\omega$. The derivative of $\rho_\omega$ wrt $\omega$ is

$$\dot{\rho}_\omega = \frac{\partial \rho_\omega}{\partial \phi} \dot{\phi} + \frac{\partial \rho_\omega}{\partial \phi} \dot{\phi} + \sum_{j_1, \ldots, j_N = 0}^1 (1 - p)^{(N - \sum_{k=1}^N j_k)} p^\sum_{k=1}^N j_k \prod_{k=1}^N (\sigma_z^{(k)})^{j_k} |\psi\rangle \langle \psi| \prod_{k=1}^N (\sigma_z^{(k)})^{j_k}$$

$$+ \sum_{j_1, \ldots, j_N = 0}^1 (1 - p)^{(N - \sum_{k=1}^N j_k)} p^\sum_{k=1}^N j_k \prod_{k=1}^N (\sigma_z^{(k)})^{j_k} |\psi\rangle \langle \psi| \prod_{k=1}^N (\sigma_z^{(k)})^{j_k}.$$  

Then we have

$$\langle \psi_\ell | \dot{\rho}_\omega | \psi_{\ell'}\rangle = a_{\ell\ell'} + ib_{\ell\ell'}, \quad (C5)$$

where $a_{\ell\ell'} = \langle \psi_\ell | \frac{\partial \rho_\omega}{\partial \phi} | \psi_{\ell'}\rangle \in \mathbb{R}$, $b_{\ell\ell'} = \langle \psi_\ell | \frac{\partial \rho_\omega}{\partial \phi} | \phi_{\ell'}\rangle \in \mathbb{R}$. Therefore,

$$F(\rho_\omega) = 2 \sum_{\ell, \ell' : \mu_\ell + \mu_{\ell'} \neq 0} \frac{|\langle \psi_\ell | \dot{\rho}_\omega | \psi_{\ell'}\rangle|^2}{\mu_\ell + \mu_{\ell'}} = \sum_{\ell, \ell' : \mu_\ell + \mu_{\ell'} \neq 0} \frac{|a_{\ell\ell'}|^2 + |b_{\ell\ell'}|^2}{\mu_\ell + \mu_{\ell'}} = F_p(\rho_\omega) + F_\phi(\rho_\omega), \quad (C6)$$

which is the same as Eq. (16) in the main text.
Appendix D: The optimal squeezed state for dephasing channels

Let the input state $|\psi_0\rangle = e^{i\varphi_1} J_x |\psi_{\mu,\nu}\rangle$, where $|\psi_{\mu,\nu}\rangle$ is an $N$-qubit spin-squeezed state

$$|\psi_{\mu,\nu}\rangle = e^{-i\nu J_x} e^{-i\varphi_2 J_y} |0\rangle^\otimes N.$$  \hspace{1cm} (D1)

The output state is $\rho_\omega = D_{\omega}^\otimes N (|\psi_0\rangle \langle \psi_0|) = (D_{\omega, |\varphi_0=0})^\otimes N (|\psi\rangle \langle \psi|)$. Then

$$\langle J_{x,y}\rangle_{\rho_\omega} = (1 - 2p) \langle J_{x,y}\rangle_{|\psi_{\mu,\nu}\rangle},$$  \hspace{1cm} (D2)

$$\langle J_{2,x,y}\rangle_{\rho_\omega} = \frac{N}{4} + (1 - 2p)^2 \left( \langle J_{2,y}^\rho_{\psi_{\mu,\nu}} \rangle - \frac{N}{4} \right),$$  \hspace{1cm} (D3)

$$\frac{\partial \langle J_x \rangle_{\rho_\omega}}{\partial p} \dot{p} = -2p \langle J_x \rangle_{\rho_{\psi_{\mu,\nu}}}, \hspace{1cm} \frac{\partial \langle J_y \rangle_{\rho_\omega}}{\partial \phi} \dot{\phi} = (1 - 2p) \dot{\phi} \langle J_x \rangle_{\rho_{\psi_{\mu,\nu}}}.$$  \hspace{1cm} (D4)

It was shown in Ref. [69] that choosing $\nu = \frac{\pi}{2} - \frac{1}{2} \arctan \frac{b}{a}$,

$$\langle J_x \rangle_{|\psi_{\mu,\nu}\rangle} = \frac{N}{2} \cos(\mu/2) N^{-1}, \hspace{1cm} \langle J_y \rangle_{|\psi_{\mu,\nu}\rangle} = 0,$$  \hspace{1cm} (D5)

$$\langle \Delta J_{x,y}^2 \rangle_{|\psi_{\mu,\nu}\rangle} = \frac{N}{4} \left( N \left( 1 - \cos^2(N - 1) \right) - \frac{N - 1}{2} a \right),$$  \hspace{1cm} (D6)

$$\langle \Delta J_{x,y}^2 \rangle_{|\psi_{\mu,\nu}\rangle} = \frac{N}{4} \left( 1 + \frac{N - 1}{4} \left( a - \sqrt{a^2 + b^2} \right) \right),$$  \hspace{1cm} (D7)

where $a = 1 - \cos^{N-2} \mu$, $b = 4 \sin \frac{\mu}{2} \cos^{N-2} \frac{\mu}{2}$. Let $N \gg 1$, $\mu = \Theta(N^{-5/6})$, then

$$\langle J_x \rangle_{|\psi_{\mu,\nu}\rangle} \approx \frac{N}{2}, \hspace{1cm} \langle \Delta J_{x,y}^2 \rangle_{|\psi_{\mu,\nu}\rangle} \approx O(N^{2/3}), \hspace{1cm} \langle \Delta J_{y}^2 \rangle_{|\psi_{\mu,\nu}\rangle} \approx O(N^{2/3}),$$  \hspace{1cm} (D8)

and $\langle \Delta J_{x}^2 \rangle_{\rho_{\psi_{\mu,\nu}}} \approx \langle \Delta J_{y}^2 \rangle_{\rho_{\psi_{\mu,\nu}}} \approx (1 - p) N$, $\frac{\partial \langle J_x \rangle_{\rho_{\psi_{\mu,\nu}}}}{\partial p} \dot{p} \approx -\dot{p} N$ and $\frac{\partial \langle J_y \rangle_{\rho_{\psi_{\mu,\nu}}}}{\partial \phi} \dot{\phi} \approx (1 - 2p) \dot{\phi} N/2$.

Appendix E: Optimizing the QFI when HNKS is violated

In this appendix, we optimize the QFI

$$\mathcal{F}_{\text{SQL}}(D_{\omega,\varphi}) = \frac{|\dot{\xi}|^2}{1 - |\xi|^2}$$  \hspace{1cm} (E1)

using Eqs. (34)-(35). We expand $T$ and $E_0 E_1$ around $\varepsilon = 0$

$$T = e^{i\varepsilon G} = 1 + i\varepsilon G - \frac{\varepsilon^2}{2} G^2 + O(\varepsilon^3),$$  \hspace{1cm} (E2)

$$E_0 E_1^\dagger = (1 - \varepsilon^2) E E^\dagger + \varepsilon \sqrt{1 - \varepsilon^2 (E F^\dagger - F E^\dagger)} - \varepsilon^2 F F^\dagger = \sigma + i\varepsilon \tilde{\sigma} - \varepsilon^2 (F F^\dagger + E E^\dagger) + O(\varepsilon^3),$$  \hspace{1cm} (E3)

where $\sigma = EE^\dagger$ and $\tilde{\sigma} = i(F F^\dagger - E E^\dagger)$. Then

$$\text{Tr}(T E_0 E_1^\dagger) = 1 - 2\varepsilon^2 - \frac{\varepsilon^2}{2} \text{Tr}(G^2 \sigma) + i \varepsilon \text{Tr}(G \sigma) - \varepsilon^2 \text{Tr}(G \tilde{\sigma}) + O(\varepsilon^3),$$  \hspace{1cm} (E4)

$$\text{Tr}(T (\tilde{E_0} E_1^\dagger + E_0 \tilde{E_1}^\dagger)) = i \varepsilon \text{Tr}(G \tilde{\sigma}) + O(\varepsilon^2),$$  \hspace{1cm} (E5)

where we used $\text{Tr}(F^\dagger F) = 1$ and $\text{Tr}(\tilde{\sigma}) = 0$ because $\text{Tr}(E^\dagger F) = 0$. Then

$$\mathcal{F}_{\text{SQL}}(D_{\omega,\varphi}) = \max_{G} \frac{|\text{Tr}(G \tilde{\sigma})|^2}{4 + 2 \text{Tr}(G \tilde{\sigma}) + \text{Tr}(G^2 \sigma) - |\text{Tr}(G \sigma)|^2} + O(\varepsilon)$$  \hspace{1cm} (E6)

$$= \max_{G,x} \frac{|\text{Tr}(G \tilde{\sigma})|^2}{4 x^2 + 2 x \text{Tr}(G \tilde{\sigma}) + \text{Tr}(G^2 \sigma) - |\text{Tr}(G \sigma)|^2} + O(\varepsilon)$$  \hspace{1cm} (E7)

$$= \max_{G} \frac{|\text{Tr}(G \tilde{\sigma})|^2}{4 - |\text{Tr}(G \sigma)|^2 + (\text{Tr}(G^2 \sigma) - |\text{Tr}(G \sigma)|^2)} + O(\varepsilon),$$  \hspace{1cm} (E8)
shown as Eq. (36) in the main text, where in the second step we used the fact that any rescaling of \( G (G \leftarrow G / x) \) should not change the optimal QFI.

To find the optimal \( G \), we first observe that \( \text{Tr}(\hat{\sigma}) = \text{Tr}(\hat{\sigma}) = 0 \). Therefore, WLOG, we assume \( \text{Tr}(G \hat{\sigma}) = 0 \) because \( G \leftarrow G - \text{Tr}(G) \frac{1}{r} \) does not change the target function. Let the derivative of Eq. (E8) be zero, we have

\[
2\hat{\sigma} \left( \text{Tr}(G^2 \sigma) - \frac{[\text{Tr}(G \hat{\sigma})]^2}{4} \right) - \text{Tr}(G \hat{\sigma}) \left( (\sigma G + G \sigma) - \frac{2\text{Tr}(G \hat{\sigma}) \hat{\sigma}}{4} \right) = 0, 
\]

\[
\Rightarrow \frac{\hat{\sigma}}{\text{Tr}(G \hat{\sigma})} \left( \text{Tr}(G^2 \sigma) - \frac{[\text{Tr}(G \hat{\sigma})]^2}{4} \right) + \frac{\text{Tr}(G \hat{\sigma}) \hat{\sigma}}{4} = \frac{1}{2} (\sigma G + G \sigma), 
\]

\[
\Leftrightarrow G = L_{\sigma} [x \hat{\sigma} + y \hat{\sigma}], \quad 4y = \text{Tr}(G \hat{\sigma}) = \text{Tr}(L_{\sigma} [x \hat{\sigma} + y \hat{\sigma}]), \quad \Leftrightarrow \quad x = 4 - \text{Tr}(L_{\sigma} [\hat{\sigma} \hat{\sigma}]), \quad y = \text{Tr}(L_{\sigma} [\hat{\sigma} \hat{\sigma}]). 
\]

Note that in Eq. (E11) we used \( x \hat{\sigma} + y \hat{\sigma} = \frac{1}{2} (G \sigma + \sigma G) \) and \( \text{Tr}(G^2 \sigma) = \text{Tr}(G (x \hat{\sigma} + y \hat{\sigma})) \). Plug the optimal \( G = L_{\sigma} [x \hat{\sigma} + y \hat{\sigma}] \) into Eq. (E8) where \( x, y \) satisfies Eq. (E12), we get

\[
\mathcal{F}_{\text{SQL}} (D_{t, \omega}) = \text{Tr}(L_{\sigma} [\hat{\sigma} \hat{\sigma}]) + \frac{\text{Tr}(L_{\sigma} [\hat{\sigma} \hat{\sigma}]^2)}{4 - \text{Tr}(L_{\sigma} [\hat{\sigma} \hat{\sigma}])} + O(\varepsilon), 
\]

shown as Eq. (38) in the main text.

Next we express \( \text{Tr}(L_{\sigma} [\hat{\sigma} \hat{\sigma}]) \), \( \text{Tr}(L_{\sigma} [\hat{\sigma} \hat{\sigma}]) \) and \( \text{Tr}(L_{\sigma} [\hat{\sigma} \hat{\sigma}]) \) in terms of \( C \) and \( \tilde{C} \). Let \( \tau = E^\dagger E \), \( \tilde{\tau} = E^\dagger F + F^\dagger E \), \( \tau' = i E^\dagger \tilde{E} - i E^\dagger E \) such that

\[
\tau_{ij} = \text{Tr}(C^\dagger K^\dagger_i K_j C), \quad \tilde{\tau}_{ij} = \text{Tr}(\tilde{C} K^\dagger_i K_j), \quad \tau'_{ij} = i \text{Tr}(C^\dagger K^\dagger_i K_j C) - i \text{Tr}(C^\dagger K^\dagger_i K_j C). 
\]

WLOG, assume \( \tau_{ij} = \text{Tr}(C^\dagger K^\dagger_i K_j C) = \lambda_i \delta_{ij} \), which could always be achieved by a unitary transformation on \( K \). We also have \( \lambda_i > 0 \) for all \( i \) because \( C \) is full rank and \( \{|i\rangle\}_{i=1}^d \) are linearly independent. Using an orthonormal basis \( \{|i\rangle\}_{i=1}^d \), where \( |i\rangle = \frac{1}{\sqrt{\lambda_i}} |K_i C\rangle \) for \( 1 \leq i \leq r \). We have

\[
\sigma = \left( \begin{array}{c} \left( \lambda_i \delta_{ij} \right) \\ 0 \end{array} \right), \quad \hat{\sigma} = \left( \begin{array}{c} \left( \langle \tilde{K}_i C | \tilde{K}_j C \rangle \sqrt{\frac{\lambda_i}{\lambda_j}} \right)^2 + \left( \langle \tilde{K}_i C | \tilde{K}_j C \rangle \sqrt{\frac{\lambda_j}{\lambda_i}} \right)^2 \\ 0 \end{array} \right), \quad \hat{\sigma}' = \left( \begin{array}{c} \left( i \langle \tilde{K}_i C | \tilde{K}_j D \rangle \sqrt{\frac{\lambda_i}{\lambda_j}} \right)^2 - i \left( i \langle \tilde{K}_i C | \tilde{K}_j D \rangle \sqrt{\frac{\lambda_j}{\lambda_i}} \right)^2 \\ 0 \end{array} \right), 
\]

where \( 1 \leq i, j \leq r \) and \( r + 1 \leq i', j \leq d^2 \). Then we can show Eqs. (41)-(43) in the main text.

\[
\text{Tr}(L_{\sigma} [\hat{\sigma} \hat{\sigma}]) = 2 \sum_{i,j; \lambda_i, \lambda_j > 0} \frac{|\langle \hat{\sigma} \rangle_{ij}|^2}{\lambda_i + \lambda_j} 
= 2 \sum_{i,j=1}^r \frac{|\langle \tilde{K}_i C | \tilde{K}_j C \rangle \sqrt{\frac{\lambda_i}{\lambda_j}} + \left( \langle \tilde{K}_i C | \tilde{K}_j C \rangle \sqrt{\frac{\lambda_j}{\lambda_i}} \right)^2}{\lambda_i + \lambda_j} + 4 \sum_{i'=r+1}^d \sum_{j=1}^r \frac{|\langle \tilde{K}_i C | \tilde{K}_j C \rangle \sqrt{\frac{\lambda_j}{\lambda_i}}|^2}{\lambda_j} 
= 4 \text{Tr}(C^\dagger \tilde{K}^\dagger \tilde{K} C) + 2 \sum_{i,j=1}^r \frac{|\langle \tilde{K}_i C | \tilde{K}_j C \rangle \sqrt{\frac{\lambda_i}{\lambda_j}} + \left( \langle \tilde{K}_i C | \tilde{K}_j C \rangle \sqrt{\frac{\lambda_j}{\lambda_i}} \right)^2}{\lambda_i + \lambda_j} - 2 \frac{|\langle \tilde{K}_i C | \tilde{K}_j C \rangle|^2}{\lambda_i} 
= 4 \text{Tr}(C^\dagger \tilde{K}^\dagger \tilde{K} C) - 2 \sum_{i,j=1}^r \frac{\tau_{ij}'}{\lambda_i + \lambda_j} = 4 \text{Tr}(C^\dagger \tilde{K}^\dagger \tilde{K} C) - \text{Tr}(L_{\tau'} [\tau']'), 
\]
\[
\text{Tr}(L_\sigma[\delta]\tilde{\sigma}) = 2 \sum_{i,j: \lambda_i + \lambda_j > 0} \frac{|(\delta)_{ij}|^2}{\lambda_i + \lambda_j}
\]
\[
= 2 \sum_{i,j=1}^{r} \frac{|i\langle K_i C|K_j D\rangle\sqrt{\frac{\lambda_j}{\lambda_i}} - i\sqrt{\frac{\lambda_i}{\lambda_j}}\langle K_i D|K_j C\rangle|^2}{\lambda_i + \lambda_j} + 4 \sum_{i'= r+1}^{d} \sum_{j=1}^{r} \frac{|i\langle i'| K_j D\rangle\sqrt{\lambda_j}}{\lambda_j}^2
\]
\[
= 4 + 2 \sum_{i,j=1}^{r} \frac{|i\langle K_i C|K_j D\rangle\sqrt{\frac{\lambda_j}{\lambda_i}} - i\sqrt{\frac{\lambda_i}{\lambda_j}}\langle K_i D|K_j C\rangle|^2}{\lambda_i + \lambda_j} - 2\frac{|\langle K_i C|K_j D\rangle|^2}{\lambda_i}
\]
\[
= 4 - 2 \sum_{i,j} |\tilde{\tau}_{ij}|^2 \frac{\lambda_i + \lambda_j}{\lambda_i + \lambda_j} = 4 - \text{Tr}(L_\tau[\tilde{\tau}]\tilde{\tau}),
\]
and
\[
\text{Tr}(L_\sigma[\delta]\tilde{\delta}) = 2 \sum_{i,j: \lambda_i + \lambda_j > 0} \frac{\delta_{ij}\tilde{\delta}_{ij}}{\lambda_i + \lambda_j}
\]
\[
= 2 \sum_{i,j=1}^{r} \frac{\delta_{ij}\tilde{\delta}_{ij}}{\lambda_i + \lambda_j} + 2 \sum_{i'= r+1}^{d} \sum_{j=1}^{r} \frac{\delta_{ij}\tilde{\delta}_{ij}}{\lambda_j} + 2 \sum_{i'= r+1}^{d} \sum_{j=1}^{r} \frac{\delta_{ij}\tilde{\delta}_{ij}}{\lambda_j}
\]
\[
= -2\text{Tr}(\tilde{C}H) + 2 \sum_{i,j=1}^{r} \frac{\delta_{ij}\tilde{\delta}_{ij}}{\lambda_i + \lambda_j} + 2i \sum_{i,j=1}^{r} \frac{\langle K_j D|K_i C\rangle\langle K_i C|K_j D\rangle - \langle \tilde{K}_j C|K_i C\rangle\langle K_i C|\tilde{K}_j D\rangle}{\lambda_i}
\]
\[
= -2\text{Tr}(\tilde{C}H) + 2 \sum_{i,j=1}^{r} \frac{\tau'_{ij}\tilde{\tau}_{ij}}{\lambda_i + \lambda_j} = -2\text{Tr}(\tilde{C}H) + \text{Tr}(L_\tau[\tau']\tilde{\tau}).
\]
Therefore, we conclude that
\[
\mathfrak{g}_{\text{QSU}}(D_{\omega,\omega}) \approx f(C, \tilde{C}) = 4\text{Tr}(C^\dagger \tilde{K}^\dagger KC) - \text{Tr}(L_\tau[\tau']\tilde{\tau}) + \frac{(-2\text{Tr}(\tilde{C}H) + \text{Tr}(L_\tau[\tau']\tilde{\tau}))^2}{\text{Tr}(L_\tau[\tau']\tilde{\tau})}.
\]
Next, we want to show
\[
\max_{C \in \mathbb{C}^d} f(C, \tilde{C}) = \min_{h:\beta=0} 4\text{Tr}(C^\dagger \alpha C)
\]
when $C$ is full rank. To calculate the dual program of the RHS, we introduce a Hermitian matrix $\tilde{C}$ as a Lagrange multiplier of $\beta = 0$ [73]. The Lagrange function is
\[
L(\tilde{C}, h) = 4\text{Tr}(C^\dagger (\tilde{K} - ih \tilde{K})^\dagger (\tilde{K} - ih \tilde{K})C) + \text{Tr}(\tilde{C}(H + K^\dagger hK))
\]
then
\[
\min_h L(\tilde{C}, h) = \min_h 4\text{Tr}(C^\dagger (\tilde{K} - ih \tilde{K})^\dagger (\tilde{K} - ih \tilde{K})C) + \text{Tr}(\tilde{C}(H + K^\dagger hK))
\]
\[
= \min_h 4\text{Tr}(C^\dagger \tilde{K}^\dagger KC) + 4\text{Tr}(\tau h^2) + 4\text{Tr}(iC^\dagger h\tilde{K}C - iC^\dagger \tilde{K}hKC) + \text{Tr}(\tilde{C}(H + K^\dagger hK))
\]
\[
= \min_h 4\text{Tr}(C^\dagger \tilde{K}^\dagger KC) + 4\text{Tr}(\tau h^2) + 4\text{Tr}(h^T \tau') + \text{Tr}(\tilde{C}H) + \text{Tr}(h^T \tilde{\tau})
\]
\[
= 4\text{Tr}(C^\dagger \tilde{K}^\dagger KC) + \text{Tr}(\tilde{C}H) - \frac{1}{8} \sum_{i,j=1}^{r} \frac{4\tau'_{ij} + \tilde{\tau}_{ij}}{\lambda_i + \lambda_j}. 
\]
The dual program is
\[
\max_{C} \min_h L(\tilde{C}, h) = \max_{C} 4\text{Tr}(C^\dagger \tilde{K}^\dagger KC) + \text{Tr}(\tilde{C}H) - \frac{1}{8} \sum_{i,j=1}^{r} \frac{16\tau'_{ij}^2 + \tilde{\tau}_{ij}^2 + 4\tau'_{ij}\tilde{\tau}_{ij} + \tilde{\tau}_{ij}\tau'_{ij}}{\lambda_i + \lambda_j}
\]
\[
= \max_{C} 4\text{Tr}(C^\dagger \tilde{K}^\dagger KC) + x\text{Tr}(\tilde{C}H) - \frac{1}{8} \sum_{i,j=1}^{r} \frac{16\tau'_{ij}^2 + x^2\tilde{\tau}_{ij}^2 + 8x\tau'_{ij}\tilde{\tau}_{ij}}{\lambda_i + \lambda_j}
\]
\[
= \max_{C} 4\text{Tr}(C^\dagger \tilde{K}^\dagger KC) - 2 \sum_{i,j=1}^{r} \frac{|\tau'_{ij}|^2}{\lambda_i + \lambda_j} + \left( -\text{Tr}(\tilde{C}H) + \sum_{i,j=1}^{r} \frac{\tilde{\tau}_{ij}}{\lambda_i + \lambda_j} \right)^2 = \max_{C} f(C, \tilde{C}),
\]
Appendix F: The numerical algorithm to find the optimal code when HNKS is violated

1. Find the optimal C

We first describe a numerical algorithm finding a full rank $C^\circ$ such that $\text{Tr}(C^{\circ\dagger}C^{\circ}) = 1$ and

$$\min_{h,\beta=0} 4\text{Tr}(C^{\circ\dagger}\alpha C^{\circ}) > \tilde{\mathcal{F}}_{\text{sqk}}(\mathcal{E}_\omega) - \eta/2. \quad (F1)$$

for any $\eta > 0$. We first note that $\tilde{\mathcal{F}}_{\text{sqk}}(\mathcal{E}_\omega) = \min_{h,\beta=0} 4\|\alpha\|$ could be solved via the following (quadratic) SDP [27],

$$\min_{x} x^2, \text{ subject to } \begin{pmatrix} x I_d & K_1^\dagger \cdots K_r^\dagger \\ K_1 & x I_{d'} & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ K_r & 0 & \cdots & x I_{d'} \end{pmatrix} \succeq 0, \quad \beta = 0. \quad (F2)$$

where $d$ and $d'$ are the input and output dimension of $\mathcal{E}_\omega$, $I_n$ is a $n \times n$ identity matrix and $\tilde{K} = \tilde{K} - i\hbar K$.

To find the full rank $C^\circ$, we first find a density matrix $\rho^\circ$ satisfying

$$\min_{h,\beta=0} 4\text{Tr}(\rho^\circ\alpha) = \min_{h,\beta=0} 4\|\alpha\|. \quad (F3)$$

It could be done via the following two-step algorithm [51]:

1) Find an $h^\circ$ using the SDP (Eq. (F2)), such that $\alpha^\circ = \alpha|_{h=h^\circ}$ satisfies $\|\alpha^\circ\| = \min_{h,\beta=0} \|\alpha\|$.

2) Let $\Pi^\circ$ be the projection onto the subspace spanned by all eigenstates corresponding to the largest eigenvalue of $\alpha^\circ$, we find an optimal density matrix $\rho^\circ$ satisfying $\Pi^\circ \rho^\circ \Pi^\circ = \rho^\circ$ and

$$\text{Re}[\text{Tr}(\rho^\circ(iK^\dagger \delta h)(\tilde{K} - i\hbar \alpha^\circ K))] = 0, \quad \forall \delta h \in \mathbb{H}_r, \text{ s.t. } K^\dagger \delta h K = 0. \quad (F4)$$

Then $C^\circ = ((1 - \eta')\rho^\circ + \eta' I_d)^{1/2}$ where $\eta' = \eta/(2\tilde{\mathcal{F}}_{\text{sqk}}(\mathcal{E}_\omega))$ is a full-rank matrix satisfying

$$\min_{h,\beta=0} 4\text{Tr}(C^{\circ\dagger}\alpha C^{\circ}) \geq (1 - \eta') \tilde{\mathcal{F}}_{\text{sqk}}(\mathcal{E}_\omega) = \tilde{\mathcal{F}}_{\text{sqk}}(\mathcal{E}_\omega) - \eta/2. \quad (F5)$$

This two-step algorithm could also be used to find $\rho^\bullet$ whose purification is the optimal input state of a single quantum channel $\mathcal{E}_\omega$ achieving $\tilde{\mathcal{F}}_1(\mathcal{E}_\omega)$:

1) Find an $h^\bullet$ using the SDP in Eq. (F2) without the requirement $\beta = 0$, such that $\alpha^\bullet = \alpha|_{h=h^\bullet}$ satisfies $\|\alpha^\bullet\| = \min_h \|\alpha\|$.

2) Let $\Pi^\bullet$ be the projection onto the subspace spanned by all eigenstates corresponding to the largest eigenvalue of $\alpha^\bullet$, we find an optimal density matrix $\rho^\bullet$ satisfying $\Pi^\bullet \rho^\bullet \Pi^\bullet = \rho^\bullet$ and

$$\text{Re}[\text{Tr}(\rho^\bullet(iK^\dagger \delta h)(\tilde{K} - i\hbar \alpha^\bullet K))] = 0, \quad \forall \delta h \in \mathbb{H}_r. \quad (F6)$$

Note that Ref. [35] provides another SDP algorithm which could be used to solve $\rho^\circ$ and $\alpha^\circ$.

2. The validity of the algorithm to find the optimal $C$

For completeness, we prove the validity of the above two-step algorithm. According to Sion’s minimax theorem [71, 72], for convex compact sets $\mathcal{P} \subset \mathbb{R}^m$ and $\mathcal{Q} \subset \mathbb{R}^n$ and $g : P \times Q \to \mathbb{R}$ such that $g(x, y)$ is a continuous convex (concave) function in $x$ ($y$) for every fixed $y$ ($x$), then

$$\max_{y \in \mathcal{Q}} \min_{x \in \mathcal{P}} g(x, y) = \min_{x \in \mathcal{P}} \max_{y \in \mathcal{Q}} g(x, y). \quad (F7)$$
In particular, if \((x^*, y^*)\) is a solution of \(\max_{\rho \in \Omega} \min_{x, y} g(x, y)\), then there must exists an \(x^*\) such that \((x^*, y^*)\) is a saddle point. Let \((x^*, y^*)\) be a solution of \(\min_{x, y} \max_{\rho \in \Omega} g(x, y)\). Then we must have

\[
g(x^*, y^*) \leq g(x^*, y^*),
\]

\(\text{Eq. (F8)}\)

According to Eq. (F7), \(g(x^*, y^*) = g(x^*, y^*)\) and all equalities must hold for the above equation. Moreover,

\[
g(x, y) \leq g(x^*, y^*), \quad \forall (x, y) \in \mathcal{Q} \times \Omega,
\]

\(\text{Eq. (F9)}\)

which means \((x^*, y^*)\) is a saddle point. For example, we can take \(x = h \in \mathbb{H}_r\), \(y = C C^\dagger = \rho \in \mathcal{S}(\mathcal{H})\) and \(g(x, y) = 4\text{Tr}(\rho\alpha)\). (We can also add the constraint \(\beta = 0\) on \(h\) which does not affect our discussion below). Then the solution of the above optimization problem is \(\bar{g}_1(\mathcal{E}_x)\) (or \(\bar{g}_{SQL}(\mathcal{E}_x)\) with the constraint \(\beta = 0\)). Note that we can always confine \(h\) in a compact set such that the solutions are not altered and the minimax theorem is applicable [51]. Let \((h^*, \rho^*)\) be any solution of the optimization problem \(\max_{\rho \in \Omega} \min_{h} 4\text{Tr}(\rho\alpha)\). Then there exists an \(h^*\) such that \((h^*, \rho^*)\) is a saddle point. Similarly, if \(g(x^*, y^*)\) is a solution of \(\min_{x, y} \max_{\rho \in \Omega} g(x, y)\), which in our case is an SDP (Eq. (F2)). There must exists a \(h\) such that \((h^*, \rho^*)\) is a saddle point. For example, we can take \((h^*, \rho^*)\) to be any solution of the optimization problem \(\min_{h} \max_{\rho} 4\text{Tr}(\rho\alpha)\). Then there exists an \(\rho^*\) such that \((h^*, \rho^*)\) is a saddle point. Moreover, \((h^*, \rho^*)\) is a saddle point if and only if

(i) \(\text{Tr}(\rho^* \alpha^*) = ||\alpha^*|| \iff \text{Tr}(\rho^* \alpha^*) \geq \text{Tr}(\rho \alpha^*), \forall \rho\).

(ii) \(\text{Re}[\text{Tr}(\rho^* (iK^\dagger \delta h)(K - i \hbar^\dagger K))] = 0, \forall \delta h \in \mathbb{H}_r, \iff \text{Tr}(\rho^* \alpha^*) \leq \text{Tr}(\rho^* \alpha^*), \forall h\).

It justifies the validity of the two-step algorithm we described above.

3. Find the optimal \(\hat{C}\)

Next, we describe how to find \(\hat{C}^o\) such that \(f(C^o, \hat{C}^o) = \max_{\hat{C}} f(C^o, \hat{C}) = \min_{h, \beta} 4\text{Tr}(C^o \alpha C^o)\). According to Appx. E,

\[
f(C, \hat{C}) = 4\text{Tr}(C^\dagger \hat{C}^\dagger \hat{K} K \hat{C}) - 2 \sum_{i,j=1}^{r} \frac{\tau_{ij}}{\lambda_i + \lambda_j} + \frac{\left(- \text{Tr}(\hat{C}^\dagger \hat{C} H) + \sum_{i,j=1}^{r} \frac{\tau_{ij}}{\lambda_i + \lambda_j}\right)^2}{2 \sum_{i,j=1}^{r} \frac{\tau_{ij}^2}{\lambda_i + \lambda_j}},
\]

\(\text{Eq. (F10)}\)

where we have assumed \(\tau_{ij} = \text{Tr}(C^\dagger K_i^\dagger K_j C) = \lambda_i \delta_{ij}\). For a fixed \(C\), \(\hat{C}\) is a linear function in \(\hat{C}\). We could always write

\[
f(C, \hat{C}) = f_1(C) + \frac{||\langle \hat{C} | f_2(C) \rangle ||^2}{\langle C | f_2(C) | C \rangle},
\]

\(\text{Eq. (F11)}\)

where \(f_1(C) \in \mathbb{R}, f_2(C) \in \mathbb{C}^{d \times d}\) is Hermitian and \(f_3(C) \in \mathbb{C}^{d \times d}\) is positive semidefinite. Moreover, \(||f_2(C)||\) is in the support of \(f_3(C)\). \(f_{1,2,3}(C)\) are functions of \(C\) only. According to Cauchy-Schwarz inequality,

\[
\max_{\hat{C}} f(C, \hat{C}) = f_1(C) + \langle f_2(C) | f_3(C)^{-1} | f_2(C) \rangle,
\]

\(\text{Eq. (F12)}\)

where the maximum is attained when \(\langle \hat{C} \rangle = f_3(C)^{-1} | f_2(C) \rangle\) and \(-1\) here means the Moore-Penrose pseudoinverse. Therefore, we take

\[
|\langle \hat{C} \rangle|^2 = f_3(C)^{-1} | f_2(C) \rangle.
\]

\(\text{Eq. (F13)}\)

Appendix G: The asymptotic QFI for the depolarizing channels

Here we calculate \(\bar{g}_1, \bar{g}_{SQL}, \bar{g}_{HL}\) for depolarizing channels

\[
\mathcal{N}_{\omega}(\rho) = (1-p)e^{-i\omega \sigma_z \rho e^{i\omega \sigma_z}} + p_x \sigma_x e^{-i\omega \sigma_x \rho e^{i\omega \sigma_x}} + p_y \sigma_y e^{-i\omega \sigma_y \rho e^{i\omega \sigma_y}} + p_z \sigma_z e^{-i\omega \sigma_z \rho e^{i\omega \sigma_z}} = \sum_{i=1}^{4} K_i \rho K_i^\dagger,
\]

\(\text{Eq. (G1)}\)
where $K_1 = \sqrt{1 - p} e^{-\frac{i\pi}{2} \sigma_z}$, $K_2 = \sqrt{p_x} \sigma_x e^{-\frac{i\pi}{2} \sigma_z}$, $K_3 = \sqrt{p_y} \sigma_y e^{-\frac{i\pi}{2} \sigma_z}$, $K_4 = \sqrt{p_z} \sigma_z e^{-\frac{i\pi}{2} \sigma_z}$.

\[ K = \begin{pmatrix} \sqrt{1 - p} \\ \sqrt{p_x} \sigma_x \\ \sqrt{p_y} \sigma_y \\ \sqrt{p_z} \sigma_z \end{pmatrix} e^{-\frac{i\pi}{2} \sigma_z}, \quad \dot{K} = \begin{pmatrix} -\frac{1}{2} \sqrt{1 - p} \sigma_z \\ -\frac{1}{2} \sqrt{p_x} \sigma_y \\ \frac{1}{2} \sqrt{p_y} \sigma_x \\ -\frac{1}{2} \sqrt{p_z} \sigma_z \end{pmatrix} e^{-\frac{i\pi}{2} \sigma_z}, \quad (G2) \]

\[ \beta = iK^\dagger (\dot{K} - ih) = \frac{1}{2} \sigma_z + K^\dagger h K. \quad (G3) \]

\[ \beta = 0 \Rightarrow \begin{cases} (1 - p) h_{11} + p_x h_{22} + p_y h_{33} + p_z h_{44} = 0, \\ \sqrt{(1 - p)} p_z (h_{12} + h_{21}) + i \sqrt{p_y} p_z h_{34} - i \sqrt{p_y} p_z h_{43} = 0, \\ \sqrt{(1 - p)} p_y (h_{13} + h_{31}) - i \sqrt{p_x} p_z h_{24} + i \sqrt{p_x} p_z h_{42} = 0, \\ \frac{1}{2} + \sqrt{(1 - p)} p_z (h_{14} + h_{41}) + i \sqrt{p_x} p_y h_{23} - i \sqrt{p_x} p_y h_{32} = 0. \end{cases} \quad (G4) \]

Clearly, HNKS is satisfied if and only if $p_x = p_z = 0$ or $p_y = p_z = 0$. It is easy to see that when $h_{ij} = 0$ for all $i, j$ except $h_{32}$, $h_{41}$ and $h_{41}, \alpha = ||\alpha|| I$, $||\alpha||$ takes its minimum and

\[ ||\alpha|| = \frac{1}{4} + \sqrt{(1 - p)} p_z (h_{14} + h_{41}) + i \sqrt{p_x} p_y (h_{23} - h_{32}) + (1 - p + p_z) |h_{14}|^2 + (p_x + p_y) |h_{23}|^2. \quad (G5) \]

Then

\[ \tilde{F}_1(N_\omega) = 4 \min_h ||\alpha|| = 1 - 4 \left( \frac{p_x p_y}{p_x + p_y} + \frac{(1 - p)p_z}{1 - p + p_z} \right). \quad (G6) \]

When HNKS is satisfied,

\[ \tilde{F}_{\text{HN}}(N_\omega) = 4 \min_h ||\beta||^2 = 1, \quad (G7) \]

and when HNKS is violated,

\[ \tilde{F}_{\text{SQL}}(N_\omega) = 4 \min_{h, \beta = 0} ||\alpha|| = -1 + \frac{1}{4} \left( \frac{p_x p_y}{p_x + p_y} + \frac{(1 - p)p_z}{1 - p + p_z} \right)^{-1}. \quad (G8) \]