Quantum error correction in multi-parameter quantum metrology

Wojciech Górecki,[1,*] Sisi Zhou,[2,3,*] Liang Jiang,[2,3] and Rafał Demkowicz-Dobrzański[1]

[1]*Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland*
[2]*Departments of Applied Physics and Physics, Yale University, New Haven, Connecticut 06511, USA*
[3]*Yale Quantum Institute, Yale University, New Haven, Connecticut 06511, USA*

We derive a necessary and sufficient condition for the possibility of preserving the Heisenberg scaling in general adaptive multi-parameter estimation schemes in presence of Markovian noise. In situations where the Heisenberg scaling can be preserved, we provide a quadratic semidefinite program to identify the optimal quantum error correcting (QEC) protocol that yields the best estimation precision, overcoming the technical challenges associated with potential incompatibility of the measurement optimally extracting information on different parameters. We provide examples of significant advantages offered by joint-parameter QEC protocols that sense all the parameters utilizing a single error-protected subspace over separate-parameter QEC protocols where each parameter is effectively sensed in a separate subspace.

Introduction. Quantum metrology aims at exploiting all possible features of quantum systems, such as coherence or entanglement, in order to boost the precision of measurements beyond that achievable by metrological schemes that operate within classical or semi-classical paradigms [1–9]. The most persuasive promise of quantum metrology is the possibility of obtaining the so-called Heisenberg scaling (HS), which manifests itself in the quadratically improved scaling of precision as a function of number of elementary probe systems involved in the experiment [10–18] or the total interrogation time of a probe system [19]. In either of these cases, the presence of decoherence typically restricts the quadratic improvement to small particle number or short-time regimes, whereas in the asymptotic regime the quantum-enhancement amount to constant factor improvements [20–25] even in case of the most general adaptive schemes [26]. Still, there are specific models where even in presence of decoherence the asymptotic HS may be preserved via application of appropriate quantum error correction (QEC) protocols [27–37].

Recently, a general theory providing a necessary and sufficient condition, the HNLS criterion (an acronym for "Hamiltonian-not-in-Lindblad span"), for achieving HS in a finite-dimensional system in the most general adaptive quantum metrological protocols under Markovian noise, has been developed [34, 35]. The theory allows for a quick identification of the most promising quantum metrological models and provides a clear recipe for designing the optimal adaptive schemes based on appropriately tailored QEC protocols. However, HNLS is restricted to the single parameter estimation case, while a lot of relevant metrological problems, like vector field sensing (e.g. magnetic field) [38], imaging [39] or multiple-arm interferometry [40, 41] or waveform estimation [42, 43] are inherently multi-parameter estimation problems. Multi-parameter estimation problems drew a lot of attention in recent years [44–49], yet no general theory that answers fundamental questions on possibility of preserving the HS in multiple-parameter estimation in presence of noise as well as the theory of designing optimal metrological protocols for this purpose has been developed so far. This aim of this paper is to fill this gap.

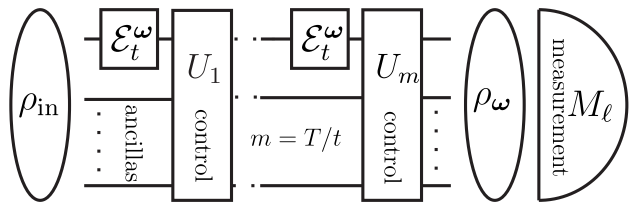

FIG. 1. General adaptive mutli-parameter quantum metrological scheme, where P parameters $\boldsymbol{\omega}=[\omega_i]_{i=1}^P$ are to be estimated. Total probe system evolution time T is divided into a number m of t-long steps of probe evolution $\mathcal{E}_t^{\boldsymbol{\omega}}$ interleaved with general unitary controls U_i. In the end a general collective measurement $\{M_\ell\}$ is performed yielding estimated value of all parameters $\tilde{\boldsymbol{\omega}}(\ell)$ with probability $p(\ell)=\mathrm{Tr}(\rho_{\boldsymbol{\omega}}M_\ell)$.

The main difficulty in dealing with fundamental metrological limits in multi-parameter scenarios is the fact that there are trade-off between the quality of estimating different parameters simultaneously. On one hand, different probe states may be optimal depending on which parameter we want to estimate and moreover measurements optimally extracting information about different parameters may be incompatible [44–46]. In particular, the widely used quantum Cramér-Rao (CR) bound is not in general saturable, due to the issues of measurement incompatibility, and as such the related quantum Fisher information (QFI) does not provide the full insight into the problem [45, 50–52]. On the other hand, stronger bounds, such as Holevo CR [52–55], are (except for specific cases [56]) more demanding computationally in their standard formulation and even if possible to compute for a given fixed protocol, they are notoriously hard to incorporate in the methods used to find the optimal metrological protocols designed with the QFI as the only figure of merit in mind.

In this paper, we generalize the HNLS condition to multi-parameter scenarios and also provide an algorithm to find the best possible protocol taking into account all the subtleties of incompatibility issues mentioned above, which is possible as we go beyond the typically used QFI-based formalism—the resulting optimal protocol yields estimation cost that is saturable with individual measurements on single probes.

---

* These two authors provided key and equal contributions to the project.

Formulation of the model. We assume the dynamics of a d-dimensional probe system is given by a general quantum master equation [57–59]:

$$\frac{d\rho}{dt} = -i[H,\rho] + \sum_{k=1}^{r}(L_k\rho L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k, \rho\}), \quad (1)$$

where the parameters to be estimated $\boldsymbol{\omega} = [\omega_1, \dots, \omega_P]$ enter linearly into the Hamiltonian of the evolution via Hermitian generators $\boldsymbol{G} = [G_1, \dots, G_P]^T$ (where ^{T} denotes transpose) so that $H = \boldsymbol{\omega} \cdot \boldsymbol{G} \equiv \sum_{k=1}^{P} \omega_k G_k$, and L_k are operators representing a general Markovian noise. Similarly, as in [34, 35] we consider the most general adaptive scheme (see Fig. 1) [26], where the probe system may be entangled with arbitrary number of ancillas, arbitrary number of intermediate unitary operators U_i may be used and a general collective measurement is performed on the final state $\rho_{\boldsymbol{\omega}}$. $\mathcal{E}_t^{\boldsymbol{\omega}}$ represents the probe system dynamics integrated over time t, whereas the total probe interrogation time is T. Such schemes are the most general schemes of probing quantum dynamics, assuming the total interrogation time is T, and encompass in particular all QEC procedures.

In single parameter estimation the optimal protocol is the one that yields the minimum estimation variance. In multiparameter case the estimator covariance matrix is the key object capturing estimation precision, defined as [51, 52]:

$$\Sigma_{ij} = \textstyle\sum_\ell \mathrm{Tr}(\rho_{\boldsymbol{\omega}} M_\ell)(\tilde{\omega}_i(\ell) - \omega_i)(\tilde{\omega}_j(\ell) - \omega_j) \quad (2)$$

for $i, j = 1, \dots, P$, where the estimator $\tilde{\boldsymbol{\omega}}(\ell)$ is a function mapping the measurement result ℓ to the parameter space, and measurement operator $M_\ell \geq 0$ and $\sum_\ell M_\ell = \mathbb{1}$ ("≥ 0" for matrices means positive semidefinite). Diagonal entries of Σ represent variances of estimators of respective parameters while off-diagonal terms represent correlations between the estimators. As a figure of merit one may simply choose $\mathrm{Tr}(\Sigma)$ which will be the sum of all individual parameter variance, or more generally $\mathrm{Tr}(W\Sigma)$, where W is a real positive cost matrix that determines the weight we associate with each parameter in the effective scalar cost function $\Delta_W^2 \tilde{\boldsymbol{\omega}} \equiv \mathrm{Tr}(W\Sigma)$. Note that we require strict positivity of W which is equivalent to saying that this is supposed to be an estimation problem of all P parameters, and not a problem where effectively only a smaller number of parameters are relevant. We assume that $\{M_\ell\}$ is locally unbiased, i.e. $\sum_\ell \tilde{\omega}_j(\ell)\mathrm{Tr}(\rho_{\boldsymbol{\omega}} M_\ell) = \omega_j$ and $\sum_\ell \tilde{\omega}_j(\ell)\mathrm{Tr}\big(\frac{\partial \rho_{\boldsymbol{\omega}}}{\partial \omega_i} M_\ell\big) = \delta_{ij}$, which is a standard assumption necessary to obtain meaningful precision bounds within the frequentist estimation framework [60, 61].

We will say that HS can be achieved in a multi-parameter estimation problem if and only if there exists an adaptive protocol that for every $W > 0$ yields $\Delta_W^2 \tilde{\boldsymbol{\omega}} \propto 1/T^2$ in the limit $T \to \infty$. This is equivalent to a requirement that all parameters (and any combination of parameters) are estimated with precision that scales like HS, generalizing HNLS to multiparameter scenarios.

Theorem 1 (Multi-parameter HNLS). *HS can be achieved in a multi-parameter estimation problem if and only if $\{(G_i)_\perp, i = 1, \dots, P\}$ are linearly independent operators. Here $(G_i)_\perp$ are orthogonal projections of G_i onto space $\mathcal{S}^\perp$ which is the orthogonal complement of the Lindblad span*

$$\mathcal{S} = \mathrm{span}_{\mathbb{R}}\{\mathbb{1}, L_k^H, iL_k^{AH}, (L_k^\dagger L_j)^H, i(L_k^\dagger L_j)^{AH}, \ \forall j,k\}, \quad (3)$$

in the Hilbert space of Hermitian matrices under the standard Hilbert-Schmidt scalar product, whereas the superscripts ^{H}, ^{AH} denote the Hermitian and anti-Hermitian part of an operator respectively.

Proof. Let us start with a brief reminder of the single parameter case solution, where $H = \omega G$ involves only a single generator G. As shown in [34, 35], the necessary and sufficient condition to achieve the HS is that $G \notin \mathcal{S}$, or in other words that $G_\perp \neq 0$. In [35] an explicit construction of the optimal QEC code was provided, where the code space $\mathcal{H}_\mathcal{C} \subseteq \mathcal{H}_S \otimes \mathcal{H}_A$ is defined on the Hilbert space of the probe system $\mathcal{H}_S$ extended by an ancillary space $\mathcal{H}_A$. The code space satisfies the QEC condition:

$$\Pi_{\mathcal{H}_\mathcal{C}} S \Pi_{\mathcal{H}_\mathcal{C}} \propto \Pi_{\mathcal{H}_\mathcal{C}}, \forall S \in \mathcal{S}, \quad (4)$$

where $\Pi_{\mathcal{H}_\mathcal{C}}$ denotes the projection onto $\mathcal{H}_\mathcal{C}$. Metrological sensitivity is guaranteed by the fact that G acts non-trivially on $\mathcal{H}_\mathcal{C}$:

$$G^{\mathcal{H}_\mathcal{C}} = \Pi_{\mathcal{H}_\mathcal{C}} G \Pi_{\mathcal{H}_\mathcal{C}} \not\propto \Pi_{\mathcal{H}_\mathcal{C}}. \quad (5)$$

As a result we obtain a noiseless unitary evolution generated by $G^{\mathcal{H}_\mathcal{C}}$ leading to HS of precision for estimation of ω. Note that, in the above formulas all operators (G, L_i) in $\mathcal{L}(\mathcal{H}_S)$ (where $\mathcal{L}(\circ)$ denotes the set of all linear operators acting on $\circ$) are trivially generalized to operators in $\mathcal{L}(\mathcal{H}_S \otimes \mathcal{H}_A)$ by tensoring with identity on $\mathcal{H}_A$.

(Necessity) Suppose $(G_i)_\perp$'s are linearly dependent. Then there exists a linear (invertible) transformation on the parameter space $A \in \mathbb{R}^{P\times P}$: $\boldsymbol{\omega}' = \boldsymbol{\omega} A^{-1}$, (where we also modify accordingly the generators $\boldsymbol{G}' = A\boldsymbol{G}$ and cost matrix $W' = AWA^T$, so that H and $\Delta_W^2 \tilde{\boldsymbol{\omega}}$ remain unchanged), such that $(G_i')_\perp = 0$ for some i. Then, from single-parameter theorem, ω_i' cannot be estimated with precision better than $\Delta^2 \tilde{\omega}_i' \sim 1/T$ which contradicts the HS requirements.

(Sufficiency) Suppose $(G_i)_\perp$'s are linearly independent. We assume the ancilla space to be a direct sum of P subspaces $\mathcal{H}_{A_i}$ so that the whole Hilbert space is $\mathcal{H}_S \otimes (\mathcal{H}_{A_1} \oplus \cdots \oplus \mathcal{H}_{A_P})$. Using the single parameter construction from [35], we may construct separate code spaces for each parameter using orthogonal ancillas $\mathcal{H}_{\mathcal{C}_i} \subseteq \mathcal{H}_S \otimes \mathcal{H}_{A_i}$ so that the QEC condition Eq. (4) are satisfied within each code space $\mathcal{H}_{\mathcal{C}_i}$ separately. While constructing the code space for the i-th parameter, we include all the remaining generators G_j ($j \neq i$) in the Lindblad span, so effectively treating them as noise. As a result thanks to the QEC condition it follows that $\forall_{i\neq j} \Pi_{\mathcal{H}_{\mathcal{C}_i}} (G_j)_\perp \Pi_{\mathcal{H}_{\mathcal{C}_i}} \propto \Pi_{\mathcal{H}_{\mathcal{C}_i}}$ and hence within a given subspace only one parameter is being sensed via the effective generator $G_i^{\mathcal{H}_{\mathcal{C}_i}} = \Pi_{\mathcal{H}_{\mathcal{C}_i}} G_i \Pi_{\mathcal{H}_{\mathcal{C}_i}}$, while all other generators act trivially. If $|\psi_i\rangle \in \mathcal{H}_{\mathcal{C}_i}$ is the optimal state for measuring ω_i, the state to be used in order to obtain HS for all parameters which is not affected by noise reads

$\rho_{\rm in} = \frac{1}{P}\sum_{i=1}^{P}\left|\psi_i\right\rangle\left\langle\psi_i\right| \in \mathcal{H}_S\otimes\left(\oplus_{i=1}^{P}\mathcal{H}_{A_i}\right)$—there is no measurement incompatibility issue because different parameters are encoded on orthogonal subspaces. □

Note that the protocol used in the sufficiency part of the proof above may be optimized by applying at first transformation $\boldsymbol{\omega}' = \boldsymbol{\omega}A^{-1}$ (and choosing proper codespaces $\mathcal{H}_{\mathcal{C}_i'}$) which leads to minimal total cost $\Delta^2_W\tilde{\boldsymbol{\omega}}$. We will refer to such a scheme as the optimal separate-parameter QEC scheme (SEP-QEC). In contrast to this construction, we will consider QEC strategies which allow for simultaneous estimation of all the parameters in a single coherent protocol by utilizing states within a single protected code space, which we will call the joint-parameter QEC scheme (JNT-QEC).

Since in the SEP-QEC protocol, we effectively measure each parameter only once in every P repetitions of an experiment (reflected by the $1/P$ factor in the $\rho_{\rm in}$), for a fixed total number of measurements, the uncertainty of estimating a given parameter will grow proportionally to P. Moreover, since the cost $\Delta^2_W\tilde{\boldsymbol{\omega}}$ is effectively a sum of P variances of different parameters the final cost $\Delta^2_W\tilde{\boldsymbol{\omega}}$ will scale as P^2 with the number of parameters (assuming the precision of the optimal estimation of a single parameter does not scale with P). The largest gains we may expect from the JNT-QEC protocol is a reduction of the cost by a factor of P, as in principle we might be able to estimate many parameters simultaneously without scarifying the precision. As we will see below, this will be an unlikely situation and typically we will face some unavoidable trade-offs—we will attempt to push to their limits. From now on, we assume the multi-parameter HNLS condition is satisfied. Without loss of generality, we assume the generators $\{G_i\}_{i=1}^{P} \subset \mathcal{S}^\perp$ are orthonormal, since the component in $\mathcal{S}$ does not contribute and there is an transformation A on parameters providing orthonormality. The following theorem provides a recipe to find the optimal JNT-QEC protocol, which can be implemented as a quadratic semidefinite program [62] (see Appendix A of [63]).

Theorem 2 (Optimal JNT-QEC). *Given a cost matrix W, the minimum cost $\Delta^2_W\tilde{\boldsymbol{\omega}}$ that can be achieved in a JNT-QEC scheme reads*

$$\Delta^2_W\tilde{\boldsymbol{\omega}} = \frac{1}{4T^2}\min_{G_i^{\mathcal{C}},B_i,\nu_i} \mathrm{Tr}\left[W(\Gamma^T\Gamma)^{-1}\right], \textit{ subject to}$$
$$\Gamma = \left(\mathrm{Im}[G_1^{\mathcal{C}}\left|0\right\rangle],\dots,\mathrm{Im}[G_P^{\mathcal{C}}\left|0\right\rangle]\right) \textit{ being full-rank,}$$
$$\frac{\mathbb{1}}{d}+\sum_{i=1}^{P}(G_i^{\mathcal{C}})^T\otimes G_i+\sum_{i=P+1}^{P'}\nu_i\mathbb{1}\otimes S_i+\sum_{i=P'+1}^{d^2-1}B_i\otimes R_i\geq 0, \tag{6}$$

where $\mathbb{1}/\sqrt{d}$, $\{G_i\}_{i=1}^{P}$, $\{S_i\}_{i=P+1}^{P'}$, $\{R_i\}_{i=P'}^{d^2-1}$ form an orthonormal basis of Hermitian operators in $\mathcal{L}(\mathcal{H}_S)$ such that $\mathcal{S} = \mathrm{span}_{\mathbb{R}}\{\mathbb{1},(S_i)_{i=P+1}^{P'}\}$. Moreover, $G_i^{\mathcal{C}}$, B_i are Hermitian matrices in $\mathcal{L}(\mathcal{C})$ where $\mathcal{C}$ is an abstract $P+1$ dimensional code space $\mathcal{C} = \mathrm{span}\{\left|0\right\rangle,\dots,\left|P\right\rangle\}$ and $\nu_i\in\mathbb{R}$.

Proof. The proof contains two major steps. First we formulate the QEC conditions as a semidefinite constraint, allowing us to determine the class of achievable effective generators $G_i^{\mathcal{C}}$. It will then become clear that for any valid $\{G_i^{\mathcal{C}}\}_{i=1}^{P}$, $\{UG_i^{\mathcal{C}}U^\dagger\}_{i=1}^{P}$ is also valid for arbitrary unitary U. Therefore we can set the initial state to be the logical zero state in $\mathcal{H}_{\mathcal{C}}$. Then, fixing the initial state and the system dynamics, we search for the optimal measurement using improved multi-parameter CR-like bounds for pure states, which we discuss in detail in Appendix B of [63].

We focus on estimation around point $\boldsymbol{\omega} = [0,\dots,0]$, which could always be achieved by applying inverse Hamiltonian dynamics [46]. Note that only $\mathrm{span}\{\left|\psi_{\boldsymbol{\omega}}\right\rangle, G_1^{\mathcal{H}_{\mathcal{C}}}\left|\psi_{\boldsymbol{\omega}}\right\rangle,\dots,G_P^{\mathcal{H}_{\mathcal{C}}}\left|\psi_{\boldsymbol{\omega}}\right\rangle\}$ is relevant to the estimation of $\boldsymbol{\omega}$, therefore we set $\dim\mathcal{H}_{\mathcal{C}} = P+1$.

Let $\{\left|c_k\right\rangle\}_{k=0}^{P}$ be an orthornormal basis in $\mathcal{H}_{\mathcal{C}}$. The effective generators are clearly defined (up to a term proportional to identity) as $[G_i^{\mathcal{H}_{\mathcal{C}}}]_l^k = \left\langle c_k\right|G_i\otimes\mathbb{1}\left|c_l\right\rangle$. From now on we will regard these generators as acting in the abstract $P+1$ dimensional code space $\mathcal{C} = \mathrm{span}\{\left|0\right\rangle,\dots,\left|P\right\rangle\}$, and we will write them as $G_i^{\mathcal{C}}$, with the same matrix element as the original physical generators $[G_i^{\mathcal{C}}]_l^k = [G_i^{\mathcal{H}_{\mathcal{C}}}]_l^k$.

Let us introduce the matrix $C\in\mathcal{L}(\mathcal{C}\otimes\mathcal{H}_S)$ characterizing a code $C = \mathrm{Tr}_{\mathcal{H}_A}(\sum_{k,l=0}^{P}\left|k\right\rangle\left\langle l\right|_{\mathcal{C}}\otimes\left|c_k\right\rangle\left\langle c_l\right|_{\mathcal{H}_S\otimes\mathcal{H}_A})$, which is proportional to the reduced density matrix of the maximal entangled state between $\mathcal{C}$ and $\mathcal{H}_{\mathcal{C}}$. By its construction $C\geq 0$. The effective evolution generators are given as:

$$(G_i^{\mathcal{C}})^T = \mathrm{Tr}_{\mathcal{H}_S}\left[C(\mathbb{1}\otimes G_i)\right] \quad i=1,\dots,P. \tag{7}$$

Taking into account orthonormality of $\left|c_k\right\rangle$ and the QEC condition Eq. (4), we obtain the following constraints:

$$\mathrm{Tr}_{\mathcal{H}_S}(C) = \mathbb{1},\ \forall_{S_i\in\mathcal{S}}\ \mathrm{Tr}_{\mathcal{H}_S}\left[C(\mathbb{1}\otimes S_i)\right]\propto\mathbb{1}. \tag{8}$$

Let $\mathbb{1}/\sqrt{d}$, $\{G_i\}_{i=1}^{P}$, $\{S_i\}_{i=P+1}^{P'}$, $\{R_i\}_{i=P'}^{d^2-1}$ form an orthonormal basis of Hermitian operators in $\mathcal{L}(\mathcal{H}_S)$ such that $\mathcal{S} = \mathrm{span}_{\mathbb{R}}\{\mathbb{1},(S_i)_{i=P+1}^{P'}\}$. Any $C\geq 0$ satisfying Eqs. (7)-(8) has the following form:

$$C = \frac{\mathbb{1}}{d}+\textstyle\sum_{i=1}^{P}(G_i^{\mathcal{C}})^T\otimes G_i + \sum_{i=P+1}^{P'}\nu_i\mathbb{1}\otimes S_i+\sum_{i=P'+1}^{d^2-1}B_i\otimes R_i, \tag{9}$$

where $\nu_i\in\mathbb{R}$ and B_i are Hermitian. Conversely, for any nonnegative defined $C\geq 0$, we can consider its purification $\left|C\right\rangle\in\mathcal{C}\otimes\mathcal{H}_S\otimes\mathcal{H}_A$, which when written as $\left|C\right\rangle = \sum_{k=0}^{P}\left|k\right\rangle_{\mathcal{C}}\otimes\left|c_k\right\rangle_{\mathcal{H}_S\otimes\mathcal{H}_A}$ yields the code states $\left|c_k\right\rangle$. Note that it means an ancillary space with $\dim\mathcal{H}_A = (P+1)d$ is sufficiently large. Therefore $\{G_i^{\mathcal{C}}\}$ is an achievable set of effective generators in $\mathcal{L}(\mathcal{C})$ (satisfying the QEC condition) if and only if there exist such $\nu_i\in\mathbb{R}$ and B_i, for which $C\geq 0$. Having formulated the QEC conditions as a semidefinite constraint we now move on to discuss measurement optimization.

Looking for the optimal JNT-QEC protocol, we may restrict ourselves to pure states—for any initial state ρ we can consider its purification (on properly redefined code space using additional ancillas) which will not increase the final cost function. Combining the Holevo CR bound for general mixed states [52] and the Matsumoto CR bound for pure states [44], we show in Appendix B of [63] that the minimum cost $\Delta^2_W\tilde{\boldsymbol{\omega}}$ (for a concrete code space) is given by:

$$\begin{aligned} &\min_{|x_i\rangle\in\mathcal{C}\oplus\mathbb{C}^P} \mathrm{Tr}(W\mathcal{X}^T\mathcal{X}), \text{ where } \mathcal{X}=(|x_1\rangle,\dots,|x_P\rangle), \\ &\text{subject to } \ 2T\mathrm{Im}[\langle x_i|G_j^{\mathcal{C}}|0\rangle]=\delta_{ij}, \\ &\qquad\quad \text{and } \langle 0|x_i\rangle=0,\ \mathrm{Im}[\langle x_i|x_j\rangle]=0. \end{aligned} \tag{10}$$

where $|0\rangle \in \mathcal{C}$ is the initial state, corresponding to $|c_0\rangle \in \mathcal{H}_{\mathcal{C}}$. Due to freedom in choosing the code space $\mathcal{H}_{\mathcal{C}} \subset \mathcal{H}_S \otimes \mathcal{H}_A$, we may also restrict $|x_i\rangle \in \mathcal{C}$ with no loss of generality. Since $\forall_i\, \langle 0|x_i\rangle = 0$ and $\forall_{i,j}\mathrm{Im}[\langle x_i|x_j\rangle] = 0$, we may assume without loss of generality that $\forall_{i,j}\mathrm{Im}[\langle i|x_j\rangle] = 0$ (such basis $\{|i\rangle\}_{i=0}^P$ may always be constructed by the Gram-Schmidt orthonormalization process). Then $\mathrm{Im}[\langle x_i|G_j^{\mathcal{H}_\mathcal{C}}|0\rangle] = \langle x_i|\mathrm{Im}[G_j^{\mathcal{H}_\mathcal{C}}|0\rangle]$. Therefore we define $\Gamma = \left(\mathrm{Im}[G_1^{\mathcal{C}}|0\rangle],\dots,\mathrm{Im}[G_P^{\mathcal{C}}|0\rangle]\right)$ which satisfy $\mathcal{X}\Gamma = \frac{1}{2T}\mathbb{1}$. From that we have $\mathrm{Tr}(W\mathcal{X}^T\mathcal{X}) = \frac{1}{4T^2}\mathrm{Tr}(W(\Gamma^T\Gamma)^{-1})$. Note also, that all vectors $\mathrm{Im}[G_i^{\mathcal{C}}|0\rangle] \in \mathrm{span}\{|1\rangle,\dots,|P\rangle\}$, therefore after removing the rows which by constructions are equal to 0, we may see Γ as $P\times P$ real matrix. $\square$

It is worth noting that the above algorithm is applicable in decoherence-free cases, to identify QEC codes to solve the measurement incompatibility problem. In addition, in Appendix C of [63] we discuss generalizations of this algorithm where the QEC condition is satisfied separately in different code subspaces $\mathcal{H}_{\mathcal{C}_i}$ instead of the whole code space $\mathcal{H}_{\mathcal{C}}$, similar to the case in the SEP-QEC construction.

Examples. Consider first the simplest single-qubit case with $d=2$. HS is achievable via QEC only in case of single-rank Pauli noise (specified by a single Hermitian Lindbladian L) [33]. Without loss of generality we set $L=\sigma_z$ (the Pauli-Z matrix). Since $\mathcal{S}=\mathrm{span}\{\mathbb{1},\sigma_z\}$, at most two parameters may be estimated in a qubit system with HS (as $\dim(\mathcal{S}^\perp)=2$). However, as proven in Appendix D of [63] when the multi-parameter HNLS criterion is met, SEP-QEC provides the optimal cost $\Delta^2_W\tilde{\boldsymbol{\omega}} = 1/T^2$ and there is no benefit in performing the more sophisticated JNT-QEC.

In order to appreciate the superiority of JNT-QEC over SEP-QEC, let us consider a two-qubit model which is a multi-parameter generalization of the one from [36]. Consider two localized qubits, coupled to a magnetic field, which is constant in both time and space, apart from some small fluctuations in the z direction. These fluctuations are assumed to be uncorrelated in time, but maximally anticorrelated in space (for two qubits they have always opposite signs). Such a system may be effectively described by Eq. (1) with $H=\frac{1}{2}\sum_{i=1}^2 \boldsymbol{\omega}\cdot\boldsymbol{\sigma}^{(i)}$ (where $\boldsymbol{\sigma}^{(i)}=[\sigma_x^{(i)},\sigma_y^{(i)},\sigma_z^{(i)}]$ acts on the i^{th} atom) and single Lindblad operator $L=\sqrt{2\gamma}(\sigma_z^{(1)}-\sigma_z^{(2)})$. It can be shown that the optimal cost with respect to individual parameters $\Delta^2\omega_{x,y,z}=\frac{1}{4T^2}$. Assuming the standard cost matrix $W=\mathbb{1}$, we immediately conclude that the precision achievable using SEP-QEC is $\Delta^2_W\tilde{\boldsymbol{\omega}}_{\text{SEP}}=\frac{P^2}{4T^2}=\frac{9}{4T^2}$. On the other hand, our algorithm for optimal JNT-QEC yields a significantly smaller

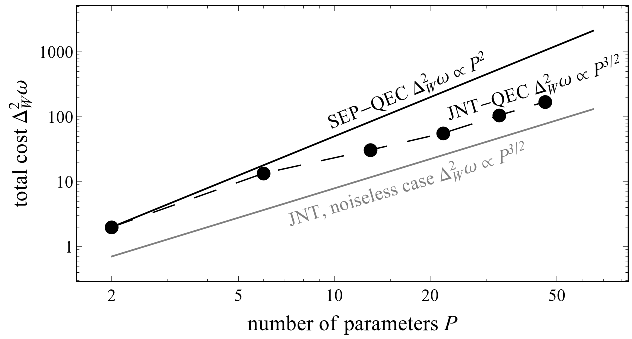

FIG. 2. Numerical results for the optimal JNT-QEC strategy for estimating the $SU(d)$ generators under the noise J_z (black points joined by dashed line), in contrast with the optimal precision asymptotically achievable by SEP-QEC $\Delta^2_W\tilde{\boldsymbol{\omega}}=\frac{P^2}{2T^2}$ (black solid line) and the lower bound $\Delta^2_W\tilde{\boldsymbol{\omega}}=\frac{P^{3/2}}{4T^2}$ asymptotically achievable only in the noiseless case (solid gray line). Here we set $T=1$.

cost $\Delta^2_W\tilde{\boldsymbol{\omega}}_{\text{JNT}} \approx \frac{5.3}{4T^2}$. A closed-form QEC code utilizing a 4-dimensional ancilla is provided in Appendix E of [63].

Finally let us consider an example which shows an asymptotic advantage (with the number of parameters) of JNT-QEC over SEP-QEC. Given a d-dimensional Hilbert space. The Hamiltonian $H=\sum_{i=1}^P\omega_i G_i$ is composed of all $SU(d)$ generators (i.e. orthonormal d-dimensional traceless Hermitian matrices, $P=d^2-1$). In the noiseless case, it can be shown that the smallest achievable cost is equal to $\sum_{i=1}^P\Delta^2\tilde{\omega}_i = \frac{d(d^2-1)}{4T^2}=\Theta(P^{\frac{3}{2}})$ [46, 64] (see also Appendix F of [63] for the proof). Compared to SEP-QEC, where the cost will necessarily scales like P^2, we observe an improvement by a factor of $\sqrt{P}$.

Let us now turn to a noisy version. To use a clear and intuitive notation, we may see the d-dimensional Hilbert space as the one associated with a a spin-j particle (where $d=2j+1$). Consider an example with a single Lindblad operator $J_z=\sum_{k=-j}^{j}k\,|k\rangle\langle k|$ (which is an interesting case as there is no two-dimensional subspace of $\mathcal{H}_S$ for which J_z acts trivial). From Theorem 1 we know that only parameters associated with generators $G_i \notin \mathrm{span}_{\mathbb{R}}\{\mathbb{1},J_z,J_z^2\}$ may be measured with HS, therefore we restrict ourselves to measuring only $P=d^2-3$ corresponding parameters. In Fig. 2, we present numerical results for such a problem, and we observe a significant advantage over the SEP-QEC protocol as well as strong indication of the asymptotic $P^{3/2}$ scaling identical to the noiseless case. Even though the optimal JNT-QEC code cannot be written down analytically in a concise way, in Appendix G of [63] we provide an analytical suboptimal construction achieving the $P^{3/2}$ scaling, supporting the numerical results.

Summary. This paper provides a complete framework to identify the possibility of achieving the HS scaling in multi-parameter estimation in presence of Markovian noise as well as determining the optimal QEC protocols. The results are obtained within the frequentist estimation approach, and we may expect that if the problem was approached using a Bayesian approach we might observe a typical π factor discrepancy between the two approaches characteristic for the HS regime [65, 66]. Rigorous Bayesian treatment of the problem is, however, beyond the scope of this work.

ACKNOWLEDGMENTS

We thank Mengzhen Zhang and John Preskill for inspiring discussions. WG and RDD acknowledge support from the National Science Center (Poland) grant No. 2016/22/E/ST2/00559. SZ and LJ acknowledge support from the ARL-CDQI (W911NF-15-2-0067, W911NF-18-2-0237), ARO (W911NF-18-1-0020, W911NF-18-1-0212), ARO MURI (W911NF-16-1-0349), AFOSR MURI (FA9550-14-1-0052, FA9550-15-1-0015), DOE (DE-SC0019406), NSF (EFMA-1640959), and the Packard Foundation (2013-39273).

[1] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).
[2] M. G. A. Paris, Int. J. Quantum Inf. **07**, 125 (2009).
[3] V. Giovannetti, S. Lloyd, and L. Maccone, Nat. Photonics **5**, 222 (2011).
[4] G. Toth and I. Apellaniz, J. Phys. A: Math. Theor. **47**, 424006 (2014).
[5] R. Demkowicz-Dobrzanski, M. Jarzyna, and J. Kołodyński, in *Prog. Optics*, Vol. 60, edited by E. Wolf (Elsevier, 2015) pp. 345–435.
[6] R. Schnabel, Phys. Rep. **684**, 1 (2017).
[7] C. L. Degen, F. Reinhard, and P. Cappellaro, Rev. Mod. Phys. **89**, 035002 (2017).
[8] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Rev. Mod. Phys. **90**, 035005 (2018).
[9] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Nat. Photonics **12**, 724 (2018).
[10] C. M. Caves, Phys. Rev. D **23**, 1693 (1981).
[11] M. Holland and K. Burnett, Phys. Rev. Lett. **71**, 1355 (1993).
[12] H. Lee, P. Kok, and J. P. Dowling, J. Mod. Optic. **49**, 2325 (2002).
[13] D. Wineland, J. Bollinger, W. Itano, F. Moore, and D. Heinzen, Phys. Rev. A **46**, R6797 (1992).
[14] K. McKenzie, D. A. Shaddock, D. E. McClelland, B. C. Buchler, and P. K. Lam, Phys. Rev. Lett. **88**, 231102 (2002).
[15] J. Bollinger, W. M. Itano, D. Wineland, and D. Heinzen, Phys. Rev. A **54**, R4649 (1996).
[16] D. Leibfried, M. Barrett, T. Schaetz, J. Britton, J. Chiaverini, W. Itano, J. Jost, C. Langer, and D. Wineland, Science **304**, 1476 (2004).
[17] V. Giovannetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004).
[18] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Phys. Rev. Lett. **79**, 3865 (1997).
[19] M. de Burgh and S. D. Bartlett, Phys. Rev. A **72**, 042301 (2005).
[20] A. Fujiwara and H. Imai, J. Phys. A: Math. Theor. **41**, 255304 (2008).
[21] R. Demkowicz-Dobrzański, U. Dorner, B. Smith, J. Lundeen, W. Wasilewski, K. Banaszek, and I. Walmsley, Phys. Rev. A **80**, 013825 (2009).
[22] B. Escher, R. de Matos Filho, and L. Davidovich, Nat. Phys. **7**, 406 (2011).
[23] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, Nat. Commun. **3**, 1063 (2012).
[24] J. Kołodyński and R. Demkowicz-Dobrzański, New J. Phys. **15**, 073043 (2013).
[25] S. I. Knysh, E. H. Chen, and G. A. Durkin, arXiv:1402.0495 (2014).
[26] R. Demkowicz-Dobrzański and L. Maccone, Phys. Rev. Lett. **113**, 250801 (2014).
[27] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin, Phys. Rev. Lett. **112**, 150802 (2014).
[28] W. Dür, M. Skotiniotis, F. Froewis, and B. Kraus, Phys. Rev. Lett. **112**, 080801 (2014).
[29] R. Ozeri, arXiv:1310.3432 (2013).
[30] G. Arrad, Y. Vinkler, D. Aharonov, and A. Retzker, Phys. Rev. Lett. **112**, 150801 (2014).
[31] T. Unden, P. Balasubramanian, D. Louzon, Y. Vinkler, M. B. Plenio, M. Markham, D. Twitchen, A. Stacey, I. Lovchinsky, A. O. Sushkov, *et al.*, Phys. Rev. Lett. **116**, 230502 (2016).
[32] F. Reiter, A. S. Sørensen, P. Zoller, and C. A. Muschik, Nat. Commun. **8**, 1822 (2017).
[33] P. Sekatski, M. Skotiniotis, J. Kołodyński, and W. Dür, Quantum **1**, 27 (2017).
[34] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Phys. Rev. X **7**, 041009 (2017).
[35] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Nat. Commun. **9**, 78 (2018).
[36] D. Layden and P. Cappellaro, npj Quantum Inf. **4**, 30 (2018).
[37] D. Layden, S. Zhou, P. Cappellaro, and L. Jiang, Phys. Rev. Lett. **122**, 040502 (2019).
[38] T. Baumgratz and A. Datta, Phys. Rev. Lett. **116**, 030801 (2016).
[39] M. Tsang, R. Nair, and X.-M. Lu, Phys. Rev. X **6**, 031033 (2016).
[40] P. C. Humphreys, M. Barbieri, A. Datta, and I. A. Walmsley, Phys. Rev. Lett. **111**, 070403 (2013).
[41] M. Gessner, L. Pezzè, and A. Smerzi, Phys. Rev. Lett. **121**, 130503 (2018).
[42] M. Tsang, H. M. Wiseman, and C. M. Caves, Phys. Rev. Lett. **106**, 090401 (2011).
[43] D. W. Berry, M. J. W. Hall, and H. M. Wiseman, Phys. Rev. Lett. **111**, 113601 (2013).
[44] K. Matsumoto, J. Phys. A.: Math. Theor. **35**, 3111 (2002).
[45] S. Ragy, M. Jarzyna, and R. Demkowicz-Dobrzański, Phys. Rev. A **94**, 052108 (2016).
[46] H. Yuan, Phys. Rev. Lett. **117**, 160801 (2016).
[47] J. Liu and H. Yuan, Phys. Rev. A **96**, 042114 (2017).
[48] R. Nichols, P. Liuzzo-Scorpo, P. A. Knott, and G. Adesso, Phys. Rev. A **98**, 012114 (2018).
[49] W. Ge, K. Jacobs, Z. Eldredge, A. V. Gorshkov, and M. Foss-Feig, Phys. Rev. Lett. **121**, 043604 (2018).
[50] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
[51] C. W. Helstrom, *Quantum detection and estimation theory* (Academic press, 1976).
[52] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam, 1982).
[53] H. Nagaoka and M. Hayashi (World Scientific Singapore, 2005) Chap. 8.

[54] J. Suzuki, J. Math. Phys. **57**, 042201 (2016).
[55] M. Guţă and A. Jenčová, Comm. Math. Phys. **276**, 341 (2007).
[56] A. Fujiwara, Math. Eng. Tech. Rep **94**, 94 (1994).
[57] G. Lindblad, Comm. Math. Phys. **48**, 119 (1976).
[58] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, J. Math. Phys. **17**, 821 (1976).
[59] H.-P. Breuer, F. Petruccione, *et al.*, *The theory of open quantum systems* (Oxford University Press on Demand, 2002).
[60] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory* (Prentice Hall, 1993).
[61] R. Gill and S. Massar, Phys. Rev. A **61**, 042312 (2000).
[62] M. Grant and S. Boyd, "Cvx: Matlab software for disciplined convex programming," .
[63] See Supplemental Material for detailed proofs and generalizations.
[64] N. Kura and M. Ueda, Phys. Rev. A **97**, 012101 (2018).
[65] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. **85**, 5098 (2000).
[66] M. Jarzyna and R. Demkowicz-Dobrzański, New J. Phys. **17**, 013010 (2015).
[67] D. A. Lidar, I. L. Chuang, and K. B. Whaley, Phys. Rev. Lett. **81**, 2594 (1998).

Appendix A: Eq. (6) as a quadratic semidefinite program

For simplicity, let us denote by $|g_i\rangle$ the P-dimensional vector $\mathrm{Im}[G_i^{\mathcal{C}}|0\rangle] \in \mathrm{span}\{|1\rangle,\ldots,|P\rangle\}$ (in such notation $\Gamma = (|g_1\rangle,\ldots,|g_P\rangle)$ is a real $P\times P$ matrix). Then

$$G_i^{\mathcal{C}} = i\mathrm{Im}[G_i^{\mathcal{C}}]+\mathrm{Re}[G_i^{\mathcal{C}}] = i\begin{pmatrix} 0 & \langle g_i| \\ -|g_i\rangle & \widetilde{G}_i^{\mathcal{C}}\end{pmatrix}+\mathrm{Re}[G_i^{\mathcal{C}}] \quad \text{(A1)}$$

where $\widetilde{G}_i^{\mathcal{C}}$ is a real antisymmetric matrix clearly defined by above. The optimization problem given in Eq. (6) reads:

$$\frac{1}{4T^2}\min_{\Gamma,\widetilde{G}_i^{\mathcal{C}},\mathrm{Re}[G_i^{\mathcal{C}}],B_i,\nu_i} \mathrm{Tr}\left(W(\Gamma^T\Gamma)^{-1}\right),$$
subject to
$$\Gamma = \left(\mathrm{Im}[G_1^{\mathcal{C}}|0\rangle],\ldots,\mathrm{Im}[G_P^{\mathcal{C}}|0\rangle]\right) \text{ being full-rank},$$
$$\frac{\mathbb{1}}{d}+\sum_{i=1}^{P}(G_i^{\mathcal{C}})^T\otimes G_i+\sum_{i=P+1}^{P'}\nu_i\mathbb{1}\otimes S_i+\sum_{i=P'+1}^{d^2-1}B_i\otimes R_i\geq 0. \quad \text{(A2)}$$

Now we show that it is sufficient to restrict Γ to matrices such that $\Gamma\sqrt{W^{-1}} > 0$. Suppose the optimal solution of above is Γ_*. Using the polar decomposition of $\Gamma_*\sqrt{W^{-1}}$,

$$\Gamma_*\sqrt{W^{-1}} = Q\Gamma_*'\sqrt{W^{-1}}, \quad \text{(A3)}$$

where Q is orthogonal and $\Gamma_*'\sqrt{W^{-1}} > 0$. Let $\Gamma_*' = (|g'^1\rangle\cdots|g'^P\rangle)$. Then we have

$$\sum_{i=1}^{P}(G_i^{\mathcal{C}})^T\otimes G_i = \sum_{i=1}^{P}Q_1(G_i^{\mathcal{C}'})^TQ_1^T\otimes G_i, \quad \text{(A4)}$$

where

$$(G_i^{\mathcal{C}'})^T = -i\begin{pmatrix} 0 & \langle g'^i| \\ -|g'^i\rangle & (\widetilde{G}_i^{\mathcal{C}})'\end{pmatrix}+\mathrm{Re}[G_i^{\mathcal{C}}]', \quad \text{(A5)}$$

$(\widetilde{G}_i^{\mathcal{C}})' = Q^T\widetilde{G}_i^{\mathcal{C}}Q$, $\mathrm{Re}[G_i^{\mathcal{C}}]' = Q_1^T\mathrm{Re}[G_i^{\mathcal{C}}]Q_1$, and

$$Q_1 = \begin{pmatrix} 1 & 0 \\ 0 & Q\end{pmatrix}. \quad \text{(A6)}$$

We can also perform a similar transformations on B_i. It is then clear that Γ_*' is an optimal solution too, because both the semi-definite constraint as well as the figure of merit to be optimized are invariant under Q_1 transformation. It means restricting $\Gamma\sqrt{W^{-1}}$ to positive matrices does not change the solution of Eq. (A2). Note that with this constraint, the objective function could be written as $\mathrm{Tr}\left(\left(\Gamma\sqrt{W^{-1}}\right)^{-2}\right)$.

Adding the constraint $\Gamma\sqrt{W^{-1}} > 0$, we rewrite Eq. (6) as the following quadratic semidefinite program

$$\frac{1}{4T^2}\min_{\Gamma,\widetilde{G}_i^{\mathcal{C}},\mathrm{Re}[G_i^{\mathcal{C}}],B_i,\nu_i,D} \mathrm{Tr}\left(D^2\right),$$
subject to
$$\frac{\mathbb{1}}{d}+\sum_{i=1}^{P}(G_i^{\mathcal{C}})^T\otimes G_i+\sum_{i=P+1}^{P'}\nu_i\mathbb{1}\otimes S_i+\sum_{i=P'+1}^{d^2-1}B_i\otimes R_i\geq 0,$$
$$\begin{pmatrix} D & \mathbb{1} \\ \mathbb{1} & \Gamma\sqrt{W^{-1}}\end{pmatrix}\geq 0, \quad \text{(A7)}$$

where D is a $P\times P$ real symmetric matrix and the last matrix inequality imposes effectively the constraint that $D \geq \left(\Gamma\sqrt{W^{-1}}\right)^{-1}$, and since we minimize $\mathrm{Tr}(D^2)$ it is equivalent to setting $D = \left(\Gamma\sqrt{W^{-1}}\right)^{-1}$. At the same time this inequality guarantees that $\Gamma\sqrt{W^{-1}} \geq 0$. For completeness, let us reiterate meaning of all the objects in the above formulation. $\mathbb{1}/\sqrt{d}$, $\{G_i\}_{i=1}^P$, $\{S_i\}_{i=P+1}^{P'}$, $\{R_i\}_{i=P'}^{d^2-1}$ form an orthonormal basis of Hermitian operators in $\mathcal{L}(\mathcal{H}_S)$ such that $\mathcal{S} = \mathrm{span}_{\mathbb{R}}\{\mathbb{1},(S_i)_{i=P+1}^{P'}\}$. Γ is a $P\times P$ real matrix, $\widetilde{G}_i^{\mathcal{C}}$ a is $P\times P$ real antisymmetric matrix, and $\mathrm{Re}[G_i^{\mathcal{C}}]$ is a $(P+1)\times(P+1)$ real symmetric matrix, where $G_i^{\mathcal{C}}$ is defined by Eq. (A1). Moreover, $G_i^{\mathcal{C}}$, B_i are Hermitian matrices in $\mathcal{L}(\mathcal{C})$ where $\mathcal{C}$ is an abstract $P+1$ dimensional code space $\mathcal{C} = \mathrm{span}\{|0\rangle,\ldots,|P\rangle\}$ and $\nu_i\in\mathbb{R}$. In such a formulation, the optimization problem could be easily solve numerically, for example, using the Matlab-based package CVX [62].

Appendix B: Derivation of the saturable bounds in multi-parameter estimation — Eq. (10)

Typically quantum multi-parameter estimation problems are analyzed utilizing the CR bound [50–52]:

$$\mathrm{Tr}(W\cdot\Sigma)\geq\mathrm{Tr}(W\mathbf{F}^{-1}),\quad \mathbf{F}_{ij} = \mathrm{Re}(\mathrm{Tr}(\rho_{\boldsymbol{\omega}}\Lambda_i\Lambda_j)), \quad \text{(B1)}$$

where $\mathbf{F}$ is the $P\times P$ QFI matrix and Λ_i (symmetric logarithmic derivatives) satisfy $\frac{\partial\rho_{\boldsymbol{\omega}}}{\partial\omega_i} = \frac{1}{2}(\Lambda_i\rho_{\boldsymbol{\omega}}+\rho_{\boldsymbol{\omega}}\Lambda_i)$. This bound is not saturable in general, due to potential non-compatibility of the optimal measurements, unless $\mathrm{Im}(\mathrm{Tr}(\rho_{\boldsymbol{\omega}}\Lambda_i\Lambda_j)) = 0$ [45]. It should be emphasized, that direct minimization of the CR bound with the saturability constraint does not guarantee identification of the optimal protocol—the optimal protocol might correspond to the situation when the CR bound is not saturable.

Below we present a derivation of a stronger Holevo CR bound [52–54] in its standard form (Theorem B1) and in a slightly changed version (Theorem B2) (which is a generalization of Matsumoto CR bound [44] for mixed states). In contrast to the standard CR bound, this bound is saturable in general – for pure states using individual measurements [44] (Theorem B3), while for mixed states using collective measurements on many copies [55]. The second version, when applied to pure states, gives exactly Eq. (10) which we make use of in the proof of Theorem 2..

In whole this section we restrict to locally unbiased measurements.

Theorem B1 (Holevo CR bound [52]). *Given a family of states $\rho_{\boldsymbol{\omega}} \in \mathcal{H}$ where $\boldsymbol{\omega} = [\omega_i]_{i=1}^P$, we have that for any cost matrix W:*

$$\mathrm{Tr}(W \cdot \Sigma) \geq \min_{X_i \in \mathcal{L}(\mathcal{H})} \mathrm{Tr}(W \cdot \mathrm{Re}V) + \mathrm{Tr}[\mathrm{abs}(W \cdot \mathrm{Im}V)], \tag{B2}$$

where Σ is the estimation covariance matrix, $\mathrm{Tr}(\mathrm{abs}(\cdot))$ is the sum of the absolute values of the eigenvalues of a matrix, $V_{ij} := \mathrm{Tr}(\rho_{\boldsymbol{\omega}} X_i X_j)$ and the minimization is performed over Hermitian matrices X_j satisfying $\mathrm{Tr}(\frac{\partial \rho_{\boldsymbol{\omega}}}{\partial \omega_i} X_j) = \delta_{ij}$ for all i, j. The last term may be equivalently written as

$$\mathrm{Tr}(\mathrm{abs}(W \cdot \mathrm{Im}V)) = \|\sqrt{W} \cdot \mathrm{Im}V \cdot \sqrt{W}\|_1, \tag{B3}$$

where $\|\cdot\|_1$ denotes the trace norm.

Proof. We start with clarifying equality $\mathrm{Tr}[\mathrm{abs}(W \cdot \mathrm{Im}V)] = \|\sqrt{W} \cdot \mathrm{Im}V \cdot \sqrt{W}\|_1$. For any diagonalizable matrix

$$B = T \begin{bmatrix} \beta_1 & 0 & \\ & \ddots & \\ 0 & & \beta_P \end{bmatrix} T^{-1}, \tag{B4}$$

$\mathrm{abs}(B)$ is defined as[1]:

$$\mathrm{abs}(B) := T \begin{bmatrix} |\beta_1| & 0 & \\ & \ddots & \\ 0 & & |\beta_P| \end{bmatrix} T^{-1}. \tag{B6}$$

As $\mathrm{Im}V$ is anti-hermitian and W is positive-defined, we have $\mathrm{abs}(W \cdot \mathrm{Im}V) = \sqrt{W}|\sqrt{W} \cdot \mathrm{Im}V \cdot \sqrt{W}|\sqrt{W}^{-1}$, which gives Eq. (B3). Now we may focus on the essential part of the proof.

First, for any measurement M_ℓ and estimator $\tilde{\boldsymbol{\omega}}(\ell)$ we may define

$$X_i := \sum_\ell (\tilde{\omega}_i(\ell) - \omega_i) M_\ell. \tag{B7}$$

Therefore $\min_{X_i}$ in Eq. (B2) corresponds to optimization over measurements. Condition $\mathrm{Tr}(\frac{\partial \rho_{\boldsymbol{\omega}}}{\partial \omega_i} X_j) = \delta_{ij}$ holds true according to the local unbiasedness assumption.

For simplicity of notation we assume $\boldsymbol{\omega} = [0, \ldots, 0]$. First we prove the matrix inequality $\Sigma \geq V$ (for any set of X_i). To do this we consider the following matrix in $\mathcal{L}(\mathbb{C}^P \otimes \mathcal{H})$,

$$\sum_\ell \begin{bmatrix} \tilde{\omega}_1(\ell) - X_1 \\ \vdots \\ \tilde{\omega}_P(\ell) - X_P \end{bmatrix} M_\ell \begin{bmatrix} \cdots & \tilde{\omega}_i(\ell) - X_i & \cdots \end{bmatrix} \geq 0. \tag{B8}$$

Any ℓ-th element of this sum is positive semidefinite, and hence is the whole matrix. After applying $\mathrm{Tr}_{\mathcal{H}}(\rho_{\boldsymbol{\omega}} \cdot\ )$, we arrive at $\Sigma \geq V$. Note that this inequality holds on the complex space $\mathbb{C}^P$ and, as it may happen that $\mathrm{Im}V \neq 0$, it is stronger than simply $\mathrm{Tr}(W \cdot \Sigma) \geq \mathrm{Tr}(W \cdot V)$ (which is equivalent to the standard CR bound). Therefore, in principle a stronger bound may be derived. Note that from $\Sigma \geq V$ we have $\sqrt{W}\Sigma\sqrt{W} \geq \sqrt{W}V\sqrt{W}$. Inequality is still valid after transposition $\sqrt{W}\Sigma\sqrt{W} \geq (\sqrt{W}V\sqrt{W})^T$ and from that

$$\sqrt{W}(\Sigma - \mathrm{Re}V)\sqrt{W} \geq \pm i\sqrt{W}\mathrm{Im}V\sqrt{W}. \tag{B9}$$

That means for any column vector $\boldsymbol{v}$,

$$\boldsymbol{v}^\dagger(\sqrt{W}(\Sigma - \mathrm{Re}V)\sqrt{W})\boldsymbol{v} \geq |\boldsymbol{v}^\dagger(\sqrt{W}(\mathrm{Im}V)\sqrt{W})\boldsymbol{v}|. \tag{B10}$$

Applying the above for all eigenvectors of $\sqrt{W}\mathrm{Im}V\sqrt{W}$ and adding together we end up with

$$\mathrm{Tr}(\sqrt{W}(\Sigma - \mathrm{Re}V)\sqrt{W}) \geq \mathrm{Tr}(|\sqrt{W}(\mathrm{Im}V)\sqrt{W}|), \tag{B11}$$

which together with Eq. (B3) gives

$$\mathrm{Tr}(W(\Sigma - \mathrm{Re}V)) \geq \mathrm{Tr}(\mathrm{abs}(W \cdot \mathrm{Im}V)). \tag{B12}$$

The above inequality holds for any locally unbiased measurements and the corresponding X_i, leading to Eq. (B2). ☐

It is worth to mention, that when the second term in Eq. (B2) is dropped, the Holevo CR bound reduces to the standard CR bound [45, 52].

Note also that for projective measurements (i.e. $M_\ell^2 = M_\ell$) inequality $\Sigma \geq V$ becomes equality $\Sigma = V$ and that in particular implies that $\mathrm{Im}V = 0$. Since any POVM on $\mathcal{H}$ may be modeled as projective measurement on $\mathcal{H} \oplus \mathcal{H}_M$, another version of the above theorem may be formulated:

Theorem B2 (Matsumoto CR bound [44] generalized for mixed states). *Given a family of states $\rho_{\boldsymbol{\omega}} \in \mathcal{H}$ with $\boldsymbol{\omega} = [\omega_i]_{i=1}^P$, we have that for any cost matrix W:*

$$\mathrm{Tr}(W \cdot \Sigma) \geq \min_{X_i \in \mathcal{L}(\mathcal{H} \oplus \mathbb{C}^P):\, \mathrm{Im}V = 0} \mathrm{Tr}(W \cdot V), \tag{B13}$$

where $V_{ij} := \mathrm{Tr}(\rho_{\boldsymbol{\omega}} X_i X_j)$ and $\{X_j\}$ satisfy $\mathrm{Tr}(\frac{\partial \rho_{\boldsymbol{\omega}}}{\partial \omega_i} X_j) = \delta_{ij}$. Here $\rho_{\boldsymbol{\omega}} \in \mathcal{L}(\mathcal{H})$ is trivially extended to $\rho_{\boldsymbol{\omega}} \in \mathcal{L}(\mathcal{H} \oplus \mathbb{C}^P)$ by letting it be zero outside $\mathcal{H}$. Moreover, this bound is equivalent to the Holevo CR bound.

Proof. It is enough to prove the equivalence with the Holevo CR bound. Let us decompose $X_i \in \mathcal{L}(\mathcal{H} \oplus \mathbb{C}^P)$ into $X_i = Y_i + Z_i$, where $\Pi_{\mathcal{H}} Y \Pi_{\mathcal{H}} = Y$ and $\Pi_{\mathcal{H}} Z \Pi_{\mathcal{H}} = 0$ ($\Pi_{\mathcal{H}}$ is the projection onto $\mathcal{H}$). We define matrices V_X, V_Y, V_Z as $(V_X)_{ij} = \mathrm{Tr}(\rho_{\boldsymbol{\omega}} X_i X_j)$ (similarly for Y and Z). Note, that

[1] Note that $\mathrm{abs}(B) \neq |B| := \sqrt{B^\dagger B}$, unless B is a normal matrix. This happens when W commutes with $\mathrm{Im}V$. This is, in particulary, true whenever the cost matrix W is proportional to the identity. In such cases we may rewrite:

$$\mathrm{Tr}\left(\mathrm{abs}(W \cdot \mathrm{Im}V)\right) = \|W \cdot \mathrm{Im}V\|_1, \tag{B5}$$

all $V_X, V_Y, V_Z \geq 0$. To prove the equivalence between the Matsumoto and the Holevo CR bound, we will show that for any fixed V_Y the following equality holds:

$$\min_{Z_i:\mathrm{Im}V_X=0} \mathrm{Tr}(WV_X) = \mathrm{Tr}(W\cdot\mathrm{Re}V_Y)+\mathrm{Tr}(\mathrm{abs}(W\cdot\mathrm{Im}V_Y)). \tag{B14}$$

Directly from the definitions of Y_i and Z_i, we have $V_X = V_Y + V_Z$. Using that and applying Eq. (B3) we see that Eq. (B14) is equivalent to

$$\min_{Z_i:\mathrm{Im}V_Z=-\mathrm{Im}V_Y} \mathrm{Tr}(W\cdot\mathrm{Re}V_Z) = \|\sqrt{W}\cdot\mathrm{Im}V_Y\cdot\sqrt{W}\|_1. \tag{B15}$$

To prove Eq. (B15), first note that from $\mathrm{Im}V_X = 0$, we have $\mathrm{Im}V_Z = -\mathrm{Im}V_Y$. Next we use a similar reasoning like in the proof of (Theorem B1). From $V_Z \geq 0$ we have $\sqrt{W}V_Z\sqrt{W} \geq 0$ and $(\sqrt{W}V_Z\sqrt{W})^T \geq 0$, which leads to $\sqrt{W}\mathrm{Re}V_Z\sqrt{W} \geq \pm i\sqrt{W}\mathrm{Im}V_Z\sqrt{W}$. From that $\mathrm{Tr}(W\mathrm{Re}V_Z) \geq \|\sqrt{W}\mathrm{Im}V_Z\sqrt{W}\|_1$. Using $\mathrm{Im}V_Z = -\mathrm{Im}V_Y$, we have $\mathrm{Tr}(W\mathrm{Re}V_Z) \geq \|\sqrt{W}\mathrm{Im}V_Y\sqrt{W}\|_1$. To attain the lower bound, we may take $V_Z = |\mathrm{Im}V_Y| - i\mathrm{Im}V_Y$. Moreover, since $|\mathrm{Im}V_Y| - i\mathrm{Im}V_Y \geq 0$ there exists a corresponding set of $Z_i \in \mathcal{L}(\mathbb{C}^P)$ such that $V_Z = |\mathrm{Im}V_Y| - i\mathrm{Im}V_Y$ is attained. □

While for any projective measurement we can define X_i, there is no guarantee that for any set of X_i there exists a proper projective measurement (end estimator) satisfying Eq. (B7). However, the existence may be proved when we deal with pure states $\rho_{\boldsymbol{\omega}} = |\psi_{\boldsymbol{\omega}}\rangle\langle\psi_{\boldsymbol{\omega}}|$:

Theorem B3 (Saturability of the Holevo and the Matsumoto CR bounds for pure states [44])**.** *Given a family of pure states $\rho_{\boldsymbol{\omega}} = |\psi_{\boldsymbol{\omega}}\rangle\langle\psi_{\boldsymbol{\omega}}| \in \mathcal{H}$ with $\boldsymbol{\omega} = [\omega_i]_{i=1}^P$, we have that for any cost matrix W:*

$$\begin{aligned}
&\min_{\text{locally unbiased } \{M_\ell\}} \mathrm{Tr}(W\cdot\Sigma) \\
&= \min_{X_i\in\mathcal{L}(\mathcal{H}\oplus\mathbb{C}^P):\,\mathrm{Im}V=0} \mathrm{Tr}(W\cdot V) \\
&= \min_{X_i\in\mathcal{L}(\mathcal{H})} \mathrm{Tr}(W\cdot\mathrm{Re}V) + \mathrm{Tr}(\mathrm{abs}(W\cdot\mathrm{Im}V)),
\end{aligned} \tag{B16}$$

where $V_{ij} := \mathrm{Tr}(\rho_{\boldsymbol{\omega}}X_iX_j)$ and X_j satisfy $\mathrm{Tr}(\frac{\partial\rho_{\boldsymbol{\omega}}}{\partial\omega_i}X_j) = \delta_{ij}$. (Again, when $X_j \in \mathcal{L}(\mathcal{H}\oplus\mathbb{C}^P)$, $|\psi_{\boldsymbol{\omega}}\rangle \in \mathcal{H}$ is trivially extended to $|\psi_{\boldsymbol{\omega}}\rangle \in \mathcal{H}\oplus\mathbb{C}^P$ by letting it be zero outside $\mathcal{H}$.)

Proof. We will prove the first equality (the second one comes directly from the Theorem B2).

As we only deal with pure states, let us introduce a simplified notation $|x_i\rangle := X_i|\psi_{\boldsymbol{\omega}}\rangle$. Then $V_{ij} = \langle x_i|x_j\rangle$ (which may be also written with $\mathcal{X} = (|x_1\rangle,\ldots,|x_P\rangle)$ as matrix equality $V = \mathcal{X}^T\mathcal{X}$). Note that $\forall_i\,\langle\psi_{\boldsymbol{\omega}}|x_i\rangle = 0$ is satisfied automatically for any $|x_i\rangle$ which minimizes the above formula—taking $|x_i'\rangle = |x_i\rangle + \alpha_i|\psi_{\boldsymbol{\omega}}\rangle$ may only increase $\mathrm{Tr}(WV)$ and it does not change the constraint $\mathrm{Tr}(\frac{\partial\rho_{\boldsymbol{\omega}}}{\partial\omega_i}X_j) = \delta_{ij}$. Therefore, we get exactly the form from Eq. (10).

As $\forall_i\,\langle\psi_{\boldsymbol{\omega}}|x_i\rangle = 0$ and $\forall_{i,j}\,\langle x_i|x_j\rangle \in \mathbb{R}$ one may choose a basis $\{|b_i\rangle\}$ of $\mathrm{span}\{|\psi_{\boldsymbol{\omega}}\rangle, |x_1\rangle,\ldots,|x_P\rangle\}$ satisfying: $\forall_i\,\langle\psi_{\boldsymbol{\omega}}|b_i\rangle \in \mathbb{R}\backslash\{0\}$ and $\forall_{i,j}\,\langle x_i|b_j\rangle \in \mathbb{R}$.

Then one can define a projective measurement:

$$M_\ell = |b_\ell\rangle\langle b_\ell| \;\; (\ell = 1,\ldots,P+1), \tag{B17}$$

$$M_0 = \mathbb{1} - \textstyle\sum_{\ell=1}^{P+1}|b_\ell\rangle\langle b_\ell|, \tag{B18}$$

with the corresponding estimator:

$$\tilde{\omega}_i(\ell) = \frac{\langle b_\ell|x_i\rangle}{\langle b_\ell|\psi_{\boldsymbol{\omega}}\rangle} + \omega_i,\; \ell \geq 1, \quad \tilde{\omega}_i(0) = 0, \tag{B19}$$

which is locally unbiased and satisfies

$$|x_i\rangle = \sum_{\ell=0}^{P+1}(\tilde{\omega}_i(\ell) - \omega_i)M_\ell|\psi_{\boldsymbol{\omega}}\rangle. \tag{B20}$$

Therefore, by virtue of Theorem B1 and Theorem B2, this concludes the proof. □

Appendix C: Generalization of the algorithm, where QEC is performed within subspaces sensitive to subsets of parameters

Let $\mathcal{C}$ be $\mathcal{C} = \bigoplus_{k=1}^K \mathcal{C}_k$ ($\dim(\mathcal{C}_k) = d_k$). Again, we use $\mathcal{C}$ to represent the $\sum_{k=1}^K d_k$ dimensional abstract code space and $\mathcal{H}_\mathcal{C}$ to represent the code space as a subspace of $\mathcal{H}_S\otimes\mathcal{H}_A$. Assuming each $\mathcal{H}_{\mathcal{C}_i}$ satisfies the QEC condition

$$\Pi_{\mathcal{H}_{\mathcal{C}_i}} S \Pi_{\mathcal{H}_{\mathcal{C}_i}} \propto \Pi_{\mathcal{H}_{\mathcal{C}_i}},\; \forall S \in \mathcal{S} \tag{C1}$$

and any operator acting within the system does not flip states between different code spaces, i.e.

$$\Pi_{\mathcal{H}_{\mathcal{C}_i}} O \Pi_{\mathcal{H}_{\mathcal{C}_j}} = 0,\; \forall i \neq j, \tag{C2}$$

for any Hermitian operator $O \in \mathcal{L}(\mathcal{H}_S)$, then for initial states of the form

$$\rho = \sum_{k=1}^K p_k|\psi_k\rangle\langle\psi_k|, \quad |\psi_k\rangle \in \mathcal{C}_k, \tag{C3}$$

the evolution is unitary, generated by $\Pi_{\mathcal{H}_\mathcal{C}} H \Pi_{\mathcal{H}_\mathcal{C}}$.

The proper C matrix for such situations has the form

$$\begin{aligned}
C_K \equiv {} & \textstyle\sum_{i=1}^P (G_i^\mathcal{C})^T\otimes G_i + \sum_{i=P+1}^{P'}\big(\sum_{k=1}^K \nu_{ik}\Pi_{\mathcal{C}_k}\big)\otimes S_i \\
& + \textstyle\sum_{i=P'+1}^{d^2-1} B_i\otimes R_i + \frac{1}{d}\mathbb{1}
\end{aligned} \tag{C4}$$

where $G_i^\mathcal{C}$ satisfies $G_i^\mathcal{C} = \sum_k \Pi_{\mathcal{C}_k} G_i^\mathcal{C} \Pi_{\mathcal{C}_k}$ and B_i satisfies $B_i = \sum_k \Pi_{\mathcal{C}_k} B_i \Pi_{\mathcal{C}_k}$.

Note that replacing the final state $\rho_{\boldsymbol{\omega}} = \sum_{k=1}^K p_k|\psi_{i\boldsymbol{\omega}}\rangle\langle\psi_{i\boldsymbol{\omega}}|$ by the pure one $|\psi_{\boldsymbol{\omega}}\rangle = \sum_{k=1}^K\sqrt{p_k}|\psi_{k\boldsymbol{\omega}}\rangle$ will not change the cost (though it will not be the actual state we obtain). Then we can use the algorithm Eq. (6) in Theorem 2 in an almost unchanged form.

Let us introduce a basis

$$\mathcal{C} = \mathrm{span}\{|k, i_k\rangle\}_{k=1,\ldots,K,\, i_k=0,\ldots,d_k-1}. \tag{C5}$$

so that $G_i^{\mathcal{C}}$, B_i are all block diagonal. Then Eq. (6) should be transformed into

$$\begin{aligned}&\text{minimize }\ \mathrm{Tr}(W(\Gamma^T\Gamma)^{-1}), \quad \text{subject to }\ C_K \geq 0,\\ &\Gamma = \big(\mathrm{Im}[G_1^{\mathcal{C}}\left|\psi\right>], \ldots, \mathrm{Im}[G_P^{\mathcal{C}}\left|\psi\right>]\big) \text{ being full-rank,}\end{aligned} \tag{C6}$$

where the variables are $\left|\psi\right> = \sum_{k=1}^{K} \sqrt{p_k}\left|k,0\right>$ with $p_k > 0$, $\sum_{k=1}^{K} p_k = 1$, block diagonal Hermitian matrices $G_i^{\mathcal{C}}$, B_i and $v_{ik} \in \mathbb{R}$.

In particular, by setting $K = P$ and $\forall_k, d_k = 2$, we reconstruct the SEP-QEC protocol; and by setting $K = 1$ and $d_1 = P + 1$, we reconstruct the JNT-QEC protocol. The most general protocol beyond SEP-QEC and JNT-QEC in this framework would be $K = P$ and $\forall_k, d_k = P + 1$. However, the problem is no longer quadratic as optimization over p_k is required.

Appendix D: Optimality of SEP-QEC in the single-qubit model

Below we show that when the Hamiltonian is $H = \omega_x\sigma_x + \omega_y\sigma_y$ and the Lindblad operator is $L = \sigma_z$, there is no advantage of using the JNT-QEC protocol in comparison to the SEP-QEC protocol for ω_x, ω_y estimation with the standard cost $W = \mathbb{1}$.

As each parameter may be estimated separately with precision $\Delta^2\omega_{x/y} = \frac{1}{4T^2}$ [33], the optimal precision of SEP-QEC is $\Delta^2_W\tilde{\boldsymbol{\omega}} = \frac{1}{T^2}$. We will show that it cannot be outperformed by JNT-QEC.

First we note that the diagonal elements of the QFI matrix for the state $\left|\psi_{\boldsymbol{\omega}}\right> = \left|\psi\right>$ are

$$\begin{aligned}\mathbf{F}_{ii} &= 4T^2(\left<\psi|\sigma_i\Pi_{\mathcal{H}_{\mathcal{C}}}\sigma_i|\psi\right> - |\left<\psi|\Pi_{\mathcal{H}_{\mathcal{C}}}\sigma_i|\psi\right>|^2)\\ &\leq 4T^2\left<\psi|\sigma_i\Pi_{\mathcal{H}_{\mathcal{C}}}\sigma_i|\psi\right> \quad (i = x, y).\end{aligned} \tag{D1}$$

As $\Delta^2_W\tilde{\boldsymbol{\omega}} = \Delta^2\tilde{\omega}_x + \Delta^2\tilde{\omega}_y \geq \frac{1}{\mathbf{F}_{xx}} + \frac{1}{\mathbf{F}_{yy}} \geq \frac{4}{\mathbf{F}_{xx}+\mathbf{F}_{yy}}$, we will focus on finding an upper bound of $\sum_{i=x,y}\left<\psi|\sigma_i\Pi_{\mathcal{H}_{\mathcal{C}}}\sigma_i|\psi\right>$.

Let $\{\left|c_0\right>, \left|c_1\right>, \left|c_2\right>\}$ be an orthonormal basis of $\mathcal{H}_{\mathcal{C}} \subseteq \mathcal{H}_S \otimes \mathcal{H}_A$. They could be written down as

$$\left|c_i\right> = \cos(\varphi^i)\left|0\right>\left|A_0^i\right> + \sin(\varphi^i)\left|1\right>\left|A_1^i\right>, \tag{D2}$$

where $\left|A_{0/1}^i\right>$ are normalized states in $\mathcal{H}_A$ and $\varphi^i \in [0, \frac{\pi}{2}]$ (the phase is hidden in choosing $\left|A_{0/1}^i\right>$). The QEC condition demands $\forall_{i,j}\left<c_i|\sigma_z|c_j\right> = \lambda\delta_{ij}$, which leads to the following two constraints. First, $\forall_i \cos^2(\varphi^i) - \sin^2(\varphi^i) = \lambda$ means all φ^i are equal (therefore superscript i will be omitted). Secondly, $\forall_{i\neq j}\cos^2(\varphi)\left<A_0^i|A_0^j\right> - \sin^2(\varphi)\left<A_1^i|A_1^j\right> = 0$. Together with the orthonormality condition $\forall_{i\neq j}\cos^2(\varphi)\left<A_0^i|A_0^j\right> + \sin^2(\varphi)\left<A_1^i|A_1^j\right> = 0$, we have

$$\begin{aligned}&\left|c_i\right> = \cos(\varphi)\left|0\right>\left|A_0^i\right> + \sin(\varphi)\left|1\right>\left|A_1^i\right>,\\ &\forall_{i,j}\left<A_0^i|A_0^j\right> = \delta_{ij}, \quad \left<A_1^i|A_1^j\right> = \delta_{ij}.\end{aligned} \tag{D3}$$

Note that there is no fixed relationship between sets $\{\left|A_0^i\right>\}_{i=0}^{i=2}$ and $\{\left|A_1^i\right>\}_{i=0}^{i=2}$—in particular it may happen that $\mathrm{span}\{\left|A_0^i\right>\} \neq \mathrm{span}\{\left|A_1^i\right>\}$. Effective generators in the chosen basis are given as:

$$(G_x^{\mathcal{C}})_{ji} = \left<c_j|\sigma_x|c_i\right> = \frac{\sin(2\varphi)}{2}(\left<A_0^j|A_1^i\right> + \left<A_1^j|A_0^i\right>), \tag{D4}$$

$$(G_y^{\mathcal{C}})_{ji} = \left<c_j|\sigma_y|c_i\right> = i\frac{\sin(2\varphi)}{2}(\left<A_1^j|A_0^i\right> - \left<A_0^j|A_1^i\right>). \tag{D5}$$

Without loss of generality, we assume that $\left|\psi\right> = \left|c_0\right>$. Then

$$\begin{aligned}&\left<\psi|\sigma_{x/y}\Pi_{\mathcal{H}_{\mathcal{C}}}\sigma_{x/y}|\psi\right> = \sum_{i=0}^{2}\left<c_0|\sigma_{x/y}|c_i\right>\left<c_i|\sigma_{x/y}|c_0\right>\\ &= \frac{\sin^2(2\varphi)}{4}\Big(\sum_{i=0}^{2}\left|\left<A_0^0|A_1^i\right>\right|^2 + \left|\left<A_1^0|A_0^i\right>\right|^2\\ &\qquad \pm 2\mathrm{Re}(\left<A_1^0|A_0^i\right>\left<A_1^i|A_0^0\right>)\Big).\end{aligned} \tag{D6}$$

Since for each $k = 0/1$, states $\{\left|A_k^i\right>\}_{i=0}^{2}$ are mutually orthonormal,

$$\begin{aligned}&\sum_{i=x,y}\left<\psi|\sigma_i\Pi_{\mathcal{H}_{\mathcal{C}}}\sigma_i|\psi\right> =\\ &\frac{\sin^2(2\varphi)}{2}\sum_{i=0}^{2}(\left|\left<A_0^0|A_1^i\right>\right|^2 + \left|\left<A_1^0|A_1^i\right>\right|^2) \leq \sin^2(2\varphi) \leq 1,\end{aligned} \tag{D7}$$

where the first inequality is saturated if and only if both $\left|A_0^0\right> \in \mathrm{span}\{\left|A_1^i\right>\}_{i=0}^{2}$ and $\left|A_1^0\right> \in \mathrm{span}\{\left|A_0^i\right>\}_{i=0}^{2}$. Using Eq. (D1) we get $\Delta^2_W\tilde{\boldsymbol{\omega}} \geq 1/T^2$.

Appendix E: The optimal protocol for sensing all magnetic field components in presence of correlated dephasing noise in the two-qubit model

Here we discuss the second example in details. First we note that the optimal precisions for estimating each $\omega_{x,y,z}$ separately are $\Delta^2\omega_{x,y,z} = \frac{1}{4T^2}$ by calculating the optimal QFIs [35]. Next we provide a SEP-QEC code achieving the precision $\Delta^2_W\tilde{\boldsymbol{\omega}} = \frac{9}{4T^2}$ which is the best achievable by SEP-QEC. ω_z can be estimated using a decoherence-free subspace [67] by letting $\left|\psi_z\right> = \frac{1}{\sqrt{2}}(\left|\uparrow\uparrow\right> + \left|\downarrow\downarrow\right>)$. To measure ω_x one may take $\left|\psi_x\right> = \frac{1}{2}\big((\left|\uparrow\uparrow\right> + \left|\downarrow\downarrow\right>)\left|1\right>_A + (\left|\uparrow\downarrow\right> + \left|\downarrow\uparrow\right>)\left|2\right>_A\big)$ and define code space as $\mathcal{C}_x = \mathrm{span}\{\left|\psi_x\right>, G_2\left|\psi_x\right>\}$. It satisfies Eq. (4), therefore noise may be corrected. The same situation holds for ω_y. Let $\rho_{\mathrm{in}} = \frac{1}{P}\sum_{i=x,y,z}\left|\psi_i\right>\left<\psi_i\right|$, and we have $\Delta^2_W\tilde{\omega} = \frac{9}{4T^2}$ in line with general considerations on the performance of the SEP-QEC codes as given in the discussion after Theorem 1.

In contrast, below we present the optimal JNT-QEC protocol, which was numerically found by the algorithm presented in Theorem 2 and reconstructed to its analytical form. We will use the standard Bell states notation:

$$\left|\Phi_+\right> = \frac{1}{\sqrt{2}}(\left|\uparrow\uparrow\right> + \left|\downarrow\downarrow\right>), \quad \left|\Phi_-\right> = \frac{1}{\sqrt{2}}(\left|\uparrow\uparrow\right> - \left|\downarrow\downarrow\right>),$$
$$\left|\Psi_+\right> = \frac{1}{\sqrt{2}}(\left|\uparrow\downarrow\right> + \left|\uparrow\downarrow\right>), \quad \left|\Psi_-\right> = \frac{1}{\sqrt{2}}(\left|\uparrow\downarrow\right> - \left|\uparrow\downarrow\right>).$$

Entanglement with ancilla will be abbreviated in the subscript $|\psi\rangle \otimes |i\rangle_A \equiv |\psi\rangle_i$. Using the numerical algorithm we have found out, that optimal code space is one of the form:

$$\begin{aligned}
|c_0\rangle &= -\cos(\varphi)\,|\Phi_+\rangle_1 + \frac{i}{\sqrt{2}}\sin(\varphi)(|\Phi_+\rangle_2 + |\Phi_-\rangle_3),\\
|c_1\rangle &= -i\sin(\varphi)\,|\Phi_+\rangle_1 - \cos(\varphi)\,|\Psi_+\rangle_2\,,\\
|c_2\rangle &= -\sin(\varphi)\,|\Phi_-\rangle_1 - i\cos(\varphi)\,|\Psi_+\rangle_3\,,\\
|c_3\rangle &= -\frac{1}{\sqrt{2}}\sin(\varphi)(|\Phi_-\rangle_2 + |\Phi_+\rangle_3) + \cos(\varphi)\,|\Psi_+\rangle_4\,,
\end{aligned} \tag{E1}$$

where the initial state is $|\psi_{\boldsymbol{\omega}=\mathbf{0}}\rangle = |c_0\rangle$. Note that the presence of the last term in $|c_3\rangle$ (entangled with $|4\rangle_A$) is necessary to satisfy QEC conditions. The value of φ could be easily solved analytically and the minimum total cost of estimation $\Delta^2_W\tilde{\boldsymbol{\omega}}$ is achieved when:

$$\cos(\varphi) = \sqrt{\frac{\sqrt{7+4\sqrt{2}}-3}{4\sqrt{2}-2}} \approx 0.39, \tag{E2}$$

and the corresponding optimal cost is:

$$\Delta^2_W\tilde{\boldsymbol{\omega}} \approx \frac{5.31}{4T^2}. \tag{E3}$$

Appendix F: Saturable bound for precision in the noiseless model involving all $SU(d)$ generators

To derive the bound we use the QFI and the following chain of inequalities:

$$\sum_{i=1}^{d^2-1}\Delta^2\tilde{\omega}_i \geq \sum_{i=1}^{d^2-1}(\mathbf{F}^{-1})_{ii} \geq \sum_{i=1}^{d^2-1}\frac{1}{\mathbf{F}_{ii}} \geq \frac{(d^2-1)^2}{\sum_{i=1}^{d^2-1}\mathbf{F}_{ii}}, \tag{F1}$$

where the first one is the CR inequality and the rest are general algebraic properties of positive semidefinite matrices. What remains to do is to derive a proper bound for the trace of the QFI matrix. We focus on the estimation around point $\boldsymbol{\omega} = [0,\dots,0]$. For any initial state $|\psi\rangle \in \mathcal{H}_S\otimes\mathcal{H}_A$ we have:

$$\begin{aligned}
\mathbf{F}_{ii} = 4T^2\left(\langle\psi|G_i^2\otimes\mathbb{1}|\psi\rangle - \langle\psi|G_i\otimes\mathbb{1}|\psi\rangle^2\right)\\
\leq 4T^2\,\langle\psi|G_i^2\otimes\mathbb{1}|\psi\rangle\,,
\end{aligned} \tag{F2}$$

where G_i are d^2-1 generators of the SU(d) group, which we normalize so that $\{\frac{1}{\sqrt{d}}\mathbb{1}, G_1,\dots,G_{d^2-1}\}$ is an orthonormal basis of Hermitian operators on $\mathcal{H}_S$. Taking into account the normalization and realizing that $\sum_{i=1}^{d^2-1}G_i^2$ is the Casimir operator of the su(d) algebra, and hence is proportional to the identity, we get that $\sum_{i=1}^{d^2-1}G_i^2 = \frac{d^2-1}{d}\mathbb{1}$. Therefore:

$$\sum_{i=1}^{d^2-1}\mathbf{F}_{ii} \leq 4T^2\,\langle\psi|\sum_{i=1}^{d^2-1}G_i^2\otimes\mathbb{1}|\psi\rangle = 4T^2\frac{d^2-1}{d}. \tag{F3}$$

After substituting the above to Eq. (F1) we get

$$\sum_{i=1}^{d^2-1}\Delta^2\tilde{\omega}_i \geq \frac{d(d^2-1)}{4T^2}, \tag{F4}$$

which proves the bound.

The example of a state which saturates the above bound is $|\psi\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^d |i\rangle\otimes|i\rangle_A$. For such a state thet QFI matrix is given by $\mathbf{F}_{ij} = \delta_{ij}\frac{4T^2}{d}$, so the second and third inequalities in Eq. (F1) become equalities. As $\mathrm{Im}(\langle\psi|\Lambda_i\Lambda_j|\psi\rangle \propto \langle\psi|[G_i,G_j]|\psi\rangle = 0$, the first one (the CR bound) is saturable as well.

Appendix G: An analytical construction of a JNT-QEC protocol yielding the cost scaling $\Theta(P^{\frac{3}{2}})$ in the model involving $SU(d)$ generators and J_z noise

Below we present an example of JNT-QEC protocol allowing one to achieve a total cost $\Delta^2_W\tilde{\boldsymbol{\omega}} = \Theta(P^{\frac{3}{2}})$ for the last example in the main text. For clarification, we treat d-dimensional Hilbert space as a single spin-j particle ($d = 2j+1$) and we use the notation where $\{|k\rangle\}_{k=-j}^{j}$ is the eigenbasis of the J_z operator.

We consider a problem where the noise generator J_z and the unitary evolution H read:

$$J_z = \sum_{k=-j}^{j} k\,|k\rangle\langle k|\,, \qquad H = \sum_{i=1}^{P}\omega_i G_i, \tag{G1}$$

where G_i is an orthonormal basis of $\mathcal{S}^\perp$—the orthogonal complement of $\mathcal{S} = \mathrm{span}\{\mathbb{1}, J_z, J_z^2\}$ (therefore $P = d^2-3$). For technical reasons we distinguish three group of operators that form the basis $\{G_i\}$:

- Real off-diagonal: $G^R_{kl} = \frac{1}{\sqrt{2}}(|k\rangle\langle l| + |l\rangle\langle k|)$
- Imaginary off-diagonal: $G^I_{kl} = \frac{i}{\sqrt{2}}(|k\rangle\langle l| - |l\rangle\langle k|)$
- Diagonal: $G^D_i = \sum_{k=-j}^{j} g_i^k\,|k\rangle\langle k|$

and in what follows we prove the scaling for each group. For simplicity, we assume, that j is an integer (for half-integer j the proof remains almost the same) and in this section we focus on the estimation around point $\boldsymbol{\omega} = [0,\dots,0]$ and set $T = 1$.

Real off-diagonal generators. We take $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_S)$ and the state $|\psi_{\boldsymbol{\omega}}\rangle = |\psi^R\rangle = \frac{1}{\sqrt{2j+1}}\sum_{k=-j}^{j}|k\rangle|k\rangle_A \in \mathcal{H}_S\otimes\mathcal{H}_{A,R}$, we have

$$\langle\psi^R|J_z|\psi^R\rangle = 0, \quad \langle\psi^R|J_z^2|\psi^R\rangle = \frac{j(j+1)}{3}. \tag{G2}$$

We construct the code space in the following way. First, we act on $|\psi\rangle$ with generators: $G^R_{kl}|\psi\rangle = \frac{1}{\sqrt{2}}(|k\rangle|l\rangle_A + |l\rangle|k\rangle_A)$

and then we "fix it" to satisfy the QEC condition by extending ancilla $\mathcal{H}_A \to \mathcal{H}_A \oplus \mathcal{H}_B$ and adding more terms:

$$\begin{aligned}|c^R_{kl}\rangle = \frac{p}{\sqrt{2}}(|k\rangle|l\rangle_A + |l\rangle|k\rangle_A)+\\ + q|j\rangle|klj\rangle_B + r|{-j}\rangle|kl(-j)\rangle_B + s|0\rangle|kl0\rangle_B\,,\end{aligned} \tag{G3}$$

where $\langle klm|k'l'm'\rangle_B = \delta_{(klm)(k'l'm')}$. Then the QEC condition is equivalent to:

$$\begin{aligned}\langle c^R_{kl}|c^R_{kl}\rangle &= p^2+q^2+r^2+s^2=1,\\ \langle c^R_{kl}|J_z|c^R_{kl}\rangle &= \frac{p^2}{2}(k+l)+(q^2-r^2)j=0,\\ \langle c^R_{kl}|J_z^2|c^R_{kl}\rangle &= \frac{p^2}{2}(k^2+l^2)+(q^2+r^2)j^2=\frac{j(j+1)}{3}.\end{aligned} \tag{G4}$$

The off-diagonal terms are automatically zero, no matter what p, q, r, s are. We can write down q^2, r^2 and s^2 as linear functions of p^2:

$$\begin{aligned}q^2 &= \frac{1}{2j^2}\Big(\frac{j(j+1)}{3}-\frac{p^2}{2}(k^2+l^2+j(k+l))\Big),\\ r^2 &= \frac{1}{2j^2}\Big(\frac{j(j+1)}{3}-\frac{p^2}{2}(k^2+l^2-j(k+l))\Big),\\ s^2 &= 1-p^2-\frac{1}{j^2}\Big(\frac{j(j+1)}{3}-\frac{p^2}{2}(k^2+l^2)\Big).\end{aligned} \tag{G5}$$

Note that p is a valid coefficient if the above set of equations has a solution (i.e. if the right-hand sides are positive). As $-2j \le k+l \le 2j$, $k^2+l^2 \le 2j^2$, this always holds provided $p^2 = \frac{1}{6}$. For the code space spanned by vectors constructed in such a way, we have

$$G^{\mathcal{C}_R}_{kl}|\psi^R\rangle = \frac{p}{\sqrt{2j+1}}|c^R_{kl}\rangle\,. \tag{G6}$$

The QFIs are $\mathbf{F}^R_{(kl)(k'l')} = 4\mathrm{Re}(\langle\psi^R|G^{\mathcal{C}_R}_{kl}G^{\mathcal{C}_R}_{k'l'}|\psi^R\rangle - \langle\psi^R|G^{\mathcal{C}_R}_{kl}|\psi^R\rangle\langle\psi^R|G^{\mathcal{C}_R}_{k'l'}|\psi^R\rangle)$ which in our case simplifies to:

$$\mathbf{F}^R_{(kl)(k'l')} = \delta_{(kl)(k'l')}4\langle\psi^R|(G^{\mathcal{C}_R}_{kl})^2|\psi^R\rangle = \frac{4p^2}{2j+1}. \tag{G7}$$

As $\langle\psi^R|[G^{\mathcal{C}_R}_{kl}, G^{\mathcal{C}_R}_{k'l'}]|\psi^R\rangle = 0$, the CR bound is saturable and the total cost is

$$\sum_{k>l}\Delta^2\omega^R_{kl} = \frac{2j+1}{4}\sum_{k>l}\frac{1}{p^2} = \frac{3j(2j+1)^2}{2} = \Theta(P^{\frac{3}{2}}). \tag{G8}$$

Imaginary off-diagonal generators. The reasoning is analogous to the previous case. Note that using different ancillary spaces for real and imaginary generators is needed. Even though $\langle\psi^R|G^R_{kl}J_zG^R_{k'l'}|\psi^R\rangle \propto \delta_{kl,k'l'}$ and $\langle\psi^R|G^I_{kl}J_zG^I_{k'l'}|\psi^R\rangle \propto \delta_{kl,k'l'}$ are satisfied automatically, $\langle\psi^R|G^R_{kl}J_zG^I_{k'l'}|\psi^R\rangle \propto \delta_{kl,k'l'}$ may not be true.

Diagonal generators. The number of diagonal generators scales like $\Theta(j)$ (when for off-diagonal the scaling is $\Theta(j^2)$), implying that the estimation with respect to diagonal generators does not contribute significantly to the overall scaling. Therefore we could simply use the SEP-QEC approach. Following [35], any traceless generator may by written down as:

$$G^D_i = \frac{1}{2}\mathrm{Tr}(|G^D_i|)(\rho_{i+}-\rho_{i-}). \tag{G9}$$

We define states $|c_{i+}\rangle, |c_{i-}\rangle$ as purification of these density matrices by using mutually orthogonal ancillas $\mathcal{H}_{Ai+}, \mathcal{H}_{Ai-}$:

$$\rho_{i+/-} = \mathrm{Tr}_{Ai+/-}(|c_{i+/-}\rangle\langle c_{i+/-}|), \tag{G10}$$

Therefore

$$\begin{aligned}\langle c_{i+}|\,G^D_i\,|c_{i+}\rangle &= \frac{1}{2}\mathrm{Tr}(|G^D_i|) \ge \frac{1}{2\sqrt{2j+1}},\\ \langle c_{i-}|\,G^D_i\,|c_{i-}\rangle &= -\frac{1}{2}\mathrm{Tr}(|G^D_i|) \le -\frac{1}{2\sqrt{2j+1}},\end{aligned} \tag{G11}$$

and from that $F_{\omega_i} \ge \frac{1}{2j+1}$. For a single parameter problem CR bound is always saturable, so we have

$$\begin{aligned}\sum_{i=1}^{2j-2}\Delta^2\omega^D_i &= (2j-2)\sum_{i=1}^{2j-2}\frac{1}{F_{\omega_i}}\\ &\le (2j+1)(2j-2)^2 = \Theta(P^{\frac{3}{2}}).\end{aligned} \tag{G12}$$