

Universal bounds for quantum metrology in the presence of correlated noise

Stanisław Kurdziałek,¹ Francesco Albarelli,² and Rafał Demkowicz-Dobrzański¹

¹*Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warszawa, Poland*

²*Scuola Normale Superiore, I-56126 Pisa, Italy*

We derive fundamental bounds for general quantum metrological models involving both temporal or spatial correlations (mathematically described by quantum combs), which may be effectively computed in the limit of a large number of probes or sensing channels involved. Although the bounds are not guaranteed to be tight in general, their tightness may be systematically increased by increasing numerical complexity of the procedure. Interestingly, this approach yields bounds tighter than the state of the art also for uncorrelated channels. We apply the bound to study the limits for the most general adaptive phase estimation models in the presence of temporally correlated dephasing. We consider dephasing both parallel (no Heisenberg scaling) and perpendicular (Heisenberg scaling possible) to the signal. In the former case our new bounds show that negative correlations are beneficial, for the latter we show evidence that the bounds are tight. We also apply the bounds to collisional thermometry, i.e. estimation of a parameter of the environment, showing evidence that entangled probes may provide only a limited advantage.

Introduction. Reducing noise effects is the foremost challenge in advancing quantum technologies [1]. Real-world quantum systems suffer from environmental interactions, leading to decoherence and particle loss, yet quantum error correction (QEC) codes often help to achieve quantum advantage [2].

In quantum metrology [3–13], the most spectacular quantum advantage is the Heisenberg scaling (HS), a precision that scales as $1/N$, where N is the number of resources (such as particles or channel uses), whereas for standard scaling (SS) the precision is proportional to $1/\sqrt{N}$. Strategies that attain the HS are known for noiseless [3, 14] and noisy [15–17] phase estimation—in the latter case, proper QEC is indispensable.

Recent developments in quantum channel estimation theory enable full characterization of precision limits under *uncorrelated* noise, allowing one to determine the optimal scaling (HS or SS) and the corresponding constant via a simple semidefinite program (SDP) [17–27]. However, noise and signal correlations are significant in many systems [28–37]. Despite various case studies including temporal [38–47] and spatial [48–53] correlations, or both [54, 55], metrological bounds for these cases are missing, especially in large- N limit. As such, one cannot assess the fundamental optimality of a specific protocol. In this work, we fill this gap and derive universal bounds for correlated noise models.

State-of-the-art bounds. Numerous metrological tasks, e.g. phase estimation or superresolution imaging [56], amount to the estimation of a single parameter θ —the goal is to find a (locally) unbiased estimator $\tilde{\theta}$ with minimal variance $\Delta^2\tilde{\theta}$. When the parameter is encoded in a quantum state ρ_θ , the *quantum Cramér-Rao Bound* (QCRB) says that $\Delta^2\tilde{\theta} \geq \frac{1}{F(\rho_\theta)}$, where $F(\rho_\theta)$ is the *quantum Fisher information* (QFI) [57, 58] (see Appendix A for the definition). The QFI is a local quantity that depends only on ρ_θ and its derivative $\dot{\rho}_\theta = \partial_\theta \rho_\theta$ evaluated at some specific value of θ .

Let $|\Psi_\theta\rangle$ be a purification of ρ_θ , then it holds that

$F(|\Psi_\theta\rangle) \geq F(\rho_\theta)$ because any measurement on ρ_θ can be implemented on $|\Psi_\theta\rangle$. Moreover, the QFI is equal to a minimization over all purifications of ρ_θ :

$$F(\rho_\theta) = 4 \min_{|\Psi_\theta\rangle} \langle \dot{\Psi}_\theta | \dot{\Psi}_\theta \rangle, \quad (1)$$

which follows from the existence of a *QFI non-increasing purification* (QFI NIP) satisfying [18, 59] $F(|\Psi_\theta\rangle) = F(\rho_\theta)$.

In quantum metrology, the parameter θ is encoded in a quantum channel $\Lambda_\theta(\bullet) = \sum_k K_{k,\theta} \bullet K_{k,\theta}^\dagger$, where $K_{k,\theta}$ are Kraus operators [20]. The ultimate precision of θ estimation is quantified by the *channel* QFI $\mathcal{F}(\Lambda_\theta) = \max_\rho F(\rho_\theta)$, where $\rho_\theta = (\Lambda_\theta \otimes \mathcal{I}_\mathcal{A})(\rho)$ and $\rho \in \mathcal{L}(\mathcal{H} \otimes \mathcal{A})$ is a possibly entangled input state of probe ($\mathcal{H}$) and noiseless ancillary ($\mathcal{A}$) systems; $\mathcal{L}(\bullet)$ is the set of linear operators acting on $\bullet$, and $\mathcal{I}$ the identity channel. Starting from (1) one finds [18]:

$$\mathcal{F}(\Lambda_\theta) = 4 \min_h \|\alpha\|, \quad \alpha = \sum_k \dot{K}_{k,\theta}^\dagger \dot{K}_{k,\theta}, \quad (2)$$

where $\|\bullet\|$ is the operator norm, $\dot{K}_{k,\theta} = \dot{K}_{k,\theta} - i h_{kk'} K_{k',\theta}$ and h is a hermitian matrix; the summation is performed over repeated indices. Importantly, (2) can be recast as a simple SDP [20].

The real potential of quantum metrology appears when $N > 1$ channel copies are probed collectively using *adaptive* or *active quantum feedback* (AD) strategies, where channels $\Lambda_\theta : \mathcal{L}(\mathcal{H}_{2i-1}) \rightarrow \mathcal{L}(\mathcal{H}_{2i})$ for $i \in \{1, 2, \dots, N\}$ are probed sequentially, and arbitrary quantum control channels $C_i : \mathcal{L}(\mathcal{H}_{2i} \otimes \mathcal{A}_i) \rightarrow \mathcal{L}(\mathcal{H}_{2i+1} \otimes \mathcal{A}_{i+1})$ can act on a probe and arbitrarily large ancilla $\mathcal{A}_i$ after each Λ_θ . The AD class covers all strategies besides those involving indefinite causal order [60]—in particular *parallel* strategies, i.e. all channels simultaneously probed by an entangled state, form a subset of AD [22].

A general AD strategy can be represented by a *quantum comb*, which models the sequence of channels sharing a common environment [61]. Formally, $E \in$

$\text{Comb}[(\mathcal{K}_1, \mathcal{K}_2), \dots, (\mathcal{K}_{2N-1}, \mathcal{K}_{2N})]$ when E is a quantum channel with inputs $\mathcal{K}_1, \mathcal{K}_3, \dots, \mathcal{K}_{2N-1}$ and outputs $\mathcal{K}_2, \mathcal{K}_4, \dots, \mathcal{K}_{2N}$, satisfying some extra linear conditions (see Appendix B) ensuring that output $\mathcal{K}_{2j}$ may only depend on input $\mathcal{K}_{2k-1}$ when $j \leq k$. The Hilbert spaces $\mathcal{K}_{2k-1}, \mathcal{K}_{2k}$ can be interpreted as input/output of k th comb tooth respectively.

Input state $\rho \in \mathcal{L}(\mathcal{H}_1)$ and controls C_i connected with ancillae $\mathcal{A}_i$ are represented using a comb $C^{(N)} \in \text{Comb}[(\emptyset, \mathcal{H}_1), (\mathcal{H}_2, \mathcal{H}_3), \dots, (\mathcal{H}_{2N-2}, \mathcal{H}_{2N-1} \otimes \mathcal{A}_N)]$, where $\emptyset$ is a trivial space (no input), see Fig. 1(a). The parameter encoding is described by a comb $\Lambda_\theta^{(N)} \in \text{Comb}[(\mathcal{H}_1, \mathcal{H}_2), \dots, (\mathcal{H}_{2N-1}, \mathcal{H}_{2N})]$ which can model any type of noise and signal correlations; for uncorrelated channels it reduces to $\Lambda_\theta^{(N)} = \Lambda_\theta^{\otimes N}$. The output state $\rho_\theta \in \mathcal{L}(\mathcal{H}_{2N} \otimes \mathcal{A}_N)$ is obtained by concatenating the corresponding inputs and outputs of $C^{(N)}$ and $\Lambda_\theta^{(N)}$ through the *link product* operation [61, 62], $\rho_\theta = C^{(N)} \star \Lambda_\theta^{(N)}$.

The *comb QFI* [46] $\mathcal{F}_{\text{AD}}^{(N)} = \max_{C^{(N)}} F(C^{(N)} \star \Lambda_\theta^{(N)})$ quantifies the ultimate estimation precision for a parameter encoded in a comb $\Lambda_\theta^{(N)}$. To evaluate $\mathcal{F}_{\text{AD}}^{(N)}$, we should decompose the Choi-Jamiołkowski (CJ) operator [63] of $\Lambda_\theta^{(N)}$ as $\Lambda_\theta^{(N)} = \sum_k |K_{k,\theta}^{(N)}\rangle \langle K_{k,\theta}^{(N)}|$, where $|K_{k,\theta}^{(N)}\rangle$ are vectorized Kraus operators; we use roman font for channels, and italics for the corresponding CJ operators. Analogously to α in (2), a *performance operator* is defined as $\alpha^{(N)} = \text{Tr}_{\text{out}} \left(\sum_k |\dot{K}_{k,\theta}^{(N)}\rangle \langle \dot{K}_{k,\theta}^{(N)}| \right)$ [64], where Tr_{out} is the partial trace over the last output space of $\Lambda_\theta^{(N)}$ ($\mathcal{H}_{2N}$), $|\dot{K}_{k,\theta}^{(N)}\rangle = |\dot{K}_{k,\theta}^{(N)}\rangle - ih_{kk'} |K_{k',\theta}^{(N)}\rangle$, h is a hermitian matrix. Thus, the comb QFI can be rewritten as [46, 60]

$$\mathcal{F}_{\text{AD}}^{(N)} = 4 \min_h \max_{\tilde{C}^{(N)}} \text{Tr} \left(\alpha^{(N)} \tilde{C}^{(N)} \right), \quad (3)$$

where $\tilde{C}^{(N)} = \text{Tr}_{\mathcal{A}_N} C^{(N)}$. This optimization can be formulated as an SDP [46, 60]. Unfortunately, its complexity grows exponentially with N , making it intractable for $N \gtrsim 5$.

For *uncorrelated* noise, one can circumvent this problem by computing bounds for $\mathcal{F}_{\text{AD}}$, using the following iteration [27]

$$\mathcal{F}_{\text{AD}}^{(l+1)} \leq \mathcal{F}_{\text{AD}}^{(l)} + 4 \min_h \left[\|\alpha\| + \sqrt{\mathcal{F}_{\text{AD}}^{(l)} \|\beta\|} \right], \quad (4)$$

where $\beta = \sum_k \dot{K}_k^\dagger K_k$ (see Appendix C1 for a new, simplified derivation). If there exists h for which $\beta = 0$, $\mathcal{F}_{\text{AD}}^{(N)}$ scales linearly with N for $N \gg 1$. Hence

$$\lim_{N \rightarrow \infty} \mathcal{F}_{\text{AD}}^{(N)} / N \leq 4 \min_h \|\alpha\| \quad \text{s.t.} \quad \beta = 0. \quad (5)$$

If instead $\beta \neq 0$ for all h , HS is in principle allowed, and the asymptotic form of (4) reads

$$\lim_{N \rightarrow \infty} \mathcal{F}_{\text{AD}}^{(N)} / N^2 \leq 4 \min_h \|\beta\|^2. \quad (6)$$

The iterative bound (4) and asymptotic bounds (5), (6) can be formulated as SDPs [21, 27]. The bound (4) is not tight in general, but its asymptotic limits (5) and (6) are always saturable using a parallel strategy [26].

Bound for correlated noise. In what follows, we present novel bounds analogous to (4), (5), (6) valid for all correlated noise models. Unlike the exact comb QFI formula (3), these bounds are efficiently calculable for arbitrarily large N .

The comb $\Lambda_\theta^{(N)}$ representing a correlated noise metrological model is a link product of its teeth $\overset{\leftrightarrow}{\Lambda}_\theta : \mathcal{L}(\mathcal{H}_{2i-1} \otimes \mathcal{R}_{i-1}) \rightarrow \mathcal{L}(\mathcal{H}_{2i} \otimes \mathcal{R}_i)$ for $i \in \{1, 2, \dots, N\}$, where $\mathcal{R}_i$ are inaccessible environmental spaces modeling correlations; the fixed state σ_{in} inputs the first register $\mathcal{R}_0$, the last register $\mathcal{R}_N$ is traced out—see Fig. 1(a). The $\leftrightarrow$ symbol in $\overset{\leftrightarrow}{\Lambda}_\theta$ indicates unconcatenated environmental spaces. Subsequent teeth are assumed identical, though this assumption can be easily dropped.

We divide $\Lambda_\theta^{(N)}$ into blocks $\overset{\leftrightarrow(m)}{\Lambda}_\theta$ of m teeth each (Fig. 1(b)), intertwined with control combs $C^{(l,m)}$ connected via ancillae $\mathcal{A}_l$. Any global strategy $C^{(N)}$ can be simulated using appropriate $C^{(l,m)}$ with sufficiently large $\mathcal{A}_l$.

Notably, combs $C^{(l,m)}$ control not only $\mathcal{H}$ and $\mathcal{A}$ at each protocol state, but also environmental subspaces $\mathcal{R}^{\leftrightarrow(m)}$ between different blocks $\overset{\leftrightarrow(m)}{\Lambda}_\theta$. In a normal AD scheme, the environment is always directly sent to a next tooth of $\Lambda_\theta^{(N)}$, but here we allow for arbitrary operations on environment every m teeth. These “environment leakages” often make the bound less tight, but are necessary when dividing N correlated channels into blocks of m while keeping the validity of the bound—neglecting the environmental information may lead to underestimated result.

Additionally, let us replace the joint mixed state $\rho_\theta^{(l)}$ of $\mathcal{H}, \mathcal{R}, \mathcal{A}$ after each block $\overset{\leftrightarrow(m)}{\Lambda}_\theta$ with its QFI NIP $|\Psi_\theta^{(l)}\rangle$. This can only increase the final QFI, since any operation possible with a state is also possible with its purification.

Through purification, the space carrying the information about $\rho_\theta^{(l)}$ and $\dot{\rho}_\theta^{(l)}$ reduces to the qubit subspace $\mathcal{V}_l = \text{span}(|\Psi_\theta^{(l)}\rangle, |\dot{\Psi}_\theta^{(l)}\rangle)$. After fixing a basis of $\mathcal{V}_l$ and constructing a comb $\overset{\leftrightarrow(m)}{\Lambda}_\theta^{(l,m)} = |\Psi_\theta^{(l)}\rangle \langle \Psi_\theta^{(l)}| \otimes \overset{\leftrightarrow(m)}{\Lambda}_\theta$ containing the purified and “qubitized” $\rho_\theta^{(l)}$ together with next m teeth, we get (see Appendix D1 for a detailed derivation) $\mathcal{F}_{\text{AD}}^{(i)} \leq F^{(i)}$, where $F^{(0)} = 0$,

$$F^{(l+m)} = \max_{C^{(l,m)}} F(C^{(l,m)} \star \overset{\leftrightarrow(m)}{\Lambda}_\theta^{(l,m)}), \quad \overset{\leftrightarrow(m)}{\Lambda}_\theta^{(l,m)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \overset{\leftrightarrow(m)}{\Lambda}_\theta, \\ \dot{\overset{\leftrightarrow(m)}{\Lambda}}_\theta^{(l,m)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \dot{\overset{\leftrightarrow(m)}{\Lambda}}_\theta + \frac{\sqrt{F^{(l)}}}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \overset{\leftrightarrow(m)}{\Lambda}_\theta, \quad (7)$$

maximizing over $C^{(l,m)}$ amounts to computing the comb QFI of $\overset{\leftrightarrow(m)}{\Lambda}_\theta^{(l,m)}$.

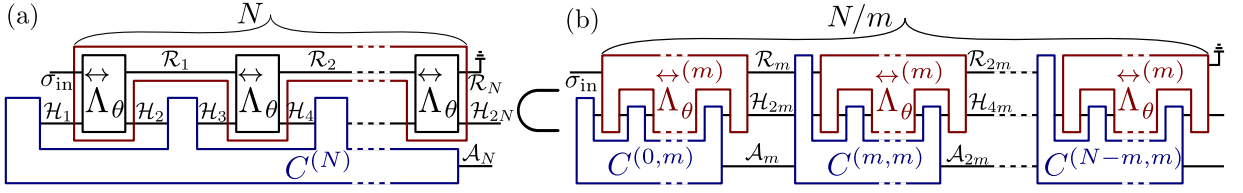


FIG. 1. The comb $\Lambda_\theta^{(N)}$ consists of teeth $\Lambda_\theta^{\leftrightarrow}$ connected with environmental spaces $\mathcal{R}$; physical estimation strategies $C^{(N)}$ act only on probe subspaces $\mathcal{H}$ (a). To derive an upper bound, we divide $\Lambda_\theta^{(N)}$ into blocks of m teeth $\Lambda_\theta^{\leftrightarrow(m)}$ (b), and allow for more powerful strategies, in which a purified state of $\mathcal{R}, \mathcal{H}, \mathcal{A}$ after each block is an input of the next control $C^{(l,m)}$.

The bound generated by (7) can be applied to uncorrelated noise—then additional QFI NIPs are the only reason for the bound untightness, making it at least as tight as (4). The new bound can be then tighter than the old one even for $m = 1$, increasing m further tightens (7), but only for finite N —see Appendix C 2.

For correlated noise, for which (4) does not apply, the role of m is more significant, since information leaks from the environment every m steps. Consequently, increasing m tightens the bound also asymptotically.

Using (7), one can prove that (see Appendix D 2) :

$$\begin{aligned} F^{(l+m)} &\leq F^{(l)} + 4 \min_h \left[a^{(m)} + \sqrt{F^{(l)}} b^{(m)} \right], a^{(m)} = \\ &= \max_{\tilde{C}^{(m)}} \text{Tr} \left(\alpha^{(m)} \tilde{C}_{00}^{(m)} \right), b^{(m)} = \max_{\tilde{C}^{(m)}} \text{Re Tr} \left(\beta^{(m)} \tilde{C}_{10}^{(m)} \right), \end{aligned} \quad (8)$$

where $\tilde{C}^{(m)}$ is a comb of the same type as any $\tilde{C}^{(l,m)} = \text{Tr}_{\mathcal{A}_{l+m}} C^{(l,m)}$, $\alpha^{(m)}$ is a performance operator of $\Lambda_\theta^{\leftrightarrow(m)}$, $\beta^{(m)}$ is a comb-adapted analogue of β (defined as $\alpha^{(m)}$, but without derivative acting on $\langle \bullet |$ part), $\tilde{C}_{ij}^{(m)} = \nu_i \langle i | \tilde{C}^{(m)} | j \rangle \nu_i$. Equation (8) is similar to (4), but $\|\alpha\|$ and $\|\beta\|$ are replaced with their comb-adapted analogues $a^{(m)}$, $b^{(m)}$. Consequently, we get asymptotic bounds analogous to (5) and (6) valid for the correlated case; when $b^{(m)} = 0$ for some h , then

$$\lim_{N \rightarrow \infty} \frac{\mathcal{F}_{\text{AD}}^{(N)}}{N} \leq \frac{4}{m} \min_h a^{(m)} \text{ s.t. } b^{(m)} = 0, \quad (9)$$

so the QFI scales at most linearly. If $b^{(m)} \neq 0$ for all h , then HS is allowed:

$$\lim_{N \rightarrow \infty} \frac{\mathcal{F}_{\text{AD}}^{(N)}}{N^2} \leq \frac{4}{m^2} \min_h b^{(m)^2}. \quad (10)$$

In Appendix D 5 we show how to translate (7), (9) and (10) into single SDPs using the technique from Refs. [46, 60]. Even though bounds (7), (9), (10) are not generally tight, they can be tightened by increasing m (which increases the complexity of the resulting SDPs).

Example: correlated dephasing. To illustrate the practical relevance of the new bounds, we consider

phase estimation in the presence of correlated dephasing. Single-qubit dephasing shrinks the Bloch vector by a factor $\eta = \cos(\epsilon)$ towards the axis defined by a unit vector $\vec{n}$. It corresponds to a random rotation by angle $+\epsilon$ or $-\epsilon$ around $\vec{n}$, each with probability $1/2$.

To introduce basic correlations, we assume the consecutive rotational signs $r_i \in \{+, -\}$ follow a binary Markov chain given by conditional probabilities [65]

$$S_{i|i-1}(r_i|r_{i-1}) = (1 + r_i r_{i-1} C)/2, \quad (11)$$

where $C \in [-1, 1]$ controls correlation: $C = 0$ means no correlation, $C = \pm 1$ gives maximal (anti-)correlation. The corresponding quantum model $\Lambda_\theta^{(N)}$ is constructed by alternating unitaries $V_\theta = R_{\hat{z}}(\theta) R_{\vec{n}}(+\epsilon) \otimes |+\rangle \langle +| + R_{\hat{z}}(\theta) R_{\vec{n}}(-\epsilon) \otimes |-\rangle \langle -|$ acting on $\mathcal{H}$ and $\mathcal{R}$ with mixing operations S performing stochastic map (11) on basis $|\pm\rangle$ of $\mathcal{R}$ [65], $R_{\vec{n}}(\varphi) = e^{-\frac{i}{2} \varphi \vec{n} \cdot \vec{\sigma}}$, $\vec{\sigma} = [\sigma_x, \sigma_y, \sigma_z]$, notice that noise (random rotation by angle $\pm\epsilon$) precedes signal (rotation by θ).

The comb $\Lambda_\theta^{(N)}$ can be split into pieces $\Lambda_\theta^{\leftrightarrow(m)}$ in different ways, and this choice affects the tightness of the resulting bound. First, we consider *parallel dephasing* $\vec{n} = \hat{z}$, when noise commutes with signal. If we cut the chain after S and before V_θ , we get information on the sign of ϵ for the first dephasing channel in each block, while cutting the chain after V_θ and before S , we get this information for the *last* channel in each block. Consequently, the first/last channel in each block is effectively noiseless, so the bound manifests an unrealistic HS.

To resolve this issue, we write S as a concatenation of two identical stochastic maps, $S = \sqrt{S} \circ \sqrt{S}$, see Appendix E 1 for exact construction of $\sqrt{S}$. The chain cuts are then made in the middle of every m -th map S , the resulting $\Lambda_\theta^{\leftrightarrow(m)}$ consists of V_θ with S between them and $\sqrt{S}$ at both ends, see Fig. 2(a). We insert $\Lambda_\theta^{\leftrightarrow(m)}$ into the recursive procedure (7) to derive bounds, sending the maximally mixed state as input for $\mathcal{R}$ for the first iteration (probabilities of $\pm$ are equal in the first channel).

We observe that the HS is not possible for correlated parallel dephasing for any $-1 < C < 1$ and $0 \leq \eta < 1$. In Fig. 2(a) we compare upper bounds, calculated for positive ($C = 0.4$) and negative ($C = -0.4$) correlations for $\eta = 0.75$, with lower bounds—QFIs achievable using protocols with small ancilla dimension d_A , calculated using

To assess tightness, we compare our upper bounds with lower bounds obtained by optimizing the output QFI over input product states. This optimization, performed efficiently via tensor networks [66, 67], shows near saturation of the upper bound for $\tau = 10\tau_{\text{opt}}$, confirming that product states are optimal in this regime. For $\tau = \tau_{\text{opt}}$, a gap remains, but the bounds still closely follow the QFI trend, indicating limited potential gain from entanglement or adaptivity.

Conclusions. These examples represent only a fraction of the possible applications of the new bounds. The formalism allows one to handle both classical and inherently quantum correlations, i.e. generic non-Markovian open quantum systems [29, 30, 32, 34]. Although the sequential scheme suggests a focus on temporal correlations, the bound is also valid for analogous spatial correlations, where channels are sampled in parallel, since this is a subset of adaptive strategies. Furthermore, this approach provides a tighter bound for uncorrelated noise than the current state of the art [27].

This work is complementary to Refs. [65, 76], where a tensor network approach was proposed to find AD esti-

mation strategies with a limited ancilla dimension, providing a *lower* bound on $\mathcal{F}_{\text{AD}}$. In combination, these works establish a comprehensive framework to study metrological protocols in the presence of correlated noise, for a large number of probed channels. The package containing the functions to compute the bounds as well as perform tensor-network optimization to obtain the lower bounds as in Fig. 2 is now publicly available [67].

A challenging extension will be to proceed towards the continuous time limit, where quantum combs (known as process tensors [34]) are routinely used in numerical studies of the dynamics and control of non-Markovian open systems [32, 77–80]. Finally, it will be interesting to see if similar ideas apply to the related problem of discriminating quantum combs [81–88].

Acknowledgements. We thank Wojciech Górecki and Andrea Smirne for helpful discussions. This work was supported by the National Science Center (Poland) grant No.2020/37/B/ST2/02134. FA acknowledges support from Marie Skłodowska-Curie Action EUHORIZON-MSCA-2021PF-01 (project QECANM, grant No. 101068347).

-
- [1] D. Suter and G. A. Álvarez, Colloquium : Protecting quantum information against environmental noise, *Rev. Mod. Phys.* **88**, 041001 (2016).
 - [2] D. A. Lidar and T. A. Brun, eds., *Quantum Error Correction*, 1st ed. (Cambridge University Press, 2013).
 - [3] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum-Enhanced Measurements: Beating the Standard Quantum Limit, *Science* **306**, 1330 (2004).
 - [4] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, *Nat. Photonics* **5**, 222 (2011).
 - [5] R. Demkowicz-Dobrzański, M. Jarzyna, and J. Kołodyński, Quantum Limits in Optical Interferometry, in *Progress in Optics, Volume 60*, edited by E. Wolf (Elsevier, Amsterdam, 2015) pp. 345–435, [arXiv:1405.7703](https://arxiv.org/abs/1405.7703).
 - [6] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, *Rev. Mod. Phys.* **89**, 035002 (2017).
 - [7] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, Quantum metrology with nonclassical states of atomic ensembles, *Rev. Mod. Phys.* **90**, 035005 (2018).
 - [8] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Advances in photonic quantum sensing, *Nat. Photonics* **12**, 724 (2018).
 - [9] E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino, Photonic quantum metrology, *AVS Quantum Sci.* **2**, 024703 (2020).
 - [10] J. Liu, M. Zhang, H. Chen, L. Wang, and H. Yuan, Optimal Scheme for Quantum Metrology, *Adv Quantum Tech* **5**, 2100080 (2022).
 - [11] L. Jiao, W. Wu, S.-Y. Bai, and J.-H. An, Quantum Metrology in the Noisy Intermediate-Scale Quantum Era, *Adv Quantum Tech* , 2300218 (2023).
 - [12] Q. Liu, Z. Hu, H. Yuan, and Y. Yang, Fully-Optimized Quantum Metrology: Framework, Tools, and Applications, *Adv Quantum Tech* , 2400094 (2024).
 - [13] V. Montenegro, C. Mukhopadhyay, R. Yousefjani, S. Sarkar, U. Mishra, M. G. Paris, and A. Bayat, Review: Quantum metrology and sensing with many-body systems, *Physics Reports* **1134**, 1 (2025), review: Quantum metrology and sensing with many-body systems.
 - [14] Z. Y. Ou, Fundamental quantum limit in precision phase measurement, *Phys. Rev. A* **55**, 2598 (1997).
 - [15] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin, Quantum error correction for metrology, *Phys. Rev. Lett.* **112**, 150802 (2014).
 - [16] W. Dür, M. Skotiniotis, F. Fröwis, and B. Kraus, Improved quantum metrology using quantum error correction, *Phys. Rev. Lett.* **112**, 080801 (2014).
 - [17] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Achieving the Heisenberg limit in quantum metrology using quantum error correction, *Nat. Commun.* **9**, 78 (2018).
 - [18] A. Fujiwara and H. Imai, A fibre bundle over manifolds of quantum channels and its application to quantum statistics, *J. Phys. A* **41**, 255304 (2008).
 - [19] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology, *Nat. Phys.* **7**, 406 (2011).
 - [20] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, The elusive Heisenberg limit in quantum-enhanced metrology, *Nat. Commun.* **3**, 1063 (2012).
 - [21] J. Kołodyński and R. Demkowicz-Dobrzański, Efficient tools for quantum metrology with uncorrelated noise, *New J. Phys.* **15**, 073043 (2013).
 - [22] R. Demkowicz-Dobrzański and L. Maccone, Using Entanglement Against Noise in Quantum Metrology, *Phys. Rev. Lett.* **113**, 250801 (2014).
 - [23] P. Sekatski, M. Skotiniotis, J. Kołodyński, and W. Dür, Quantum metrology with full and fast quantum control, *Quantum* **1**, 27 (2017).

- [24] R. Demkowicz-Dobrzański, J. Czakowski, and P. Sekatski, Adaptive Quantum Metrology under General Markovian Noise, *Phys. Rev. X* **7**, 041009 (2017).
- [25] S. Zhou and L. Jiang, Optimal approximate quantum error correction for quantum metrology, *Phys. Rev. Res.* **2**, 013235 (2020).
- [26] S. Zhou and L. Jiang, Asymptotic Theory of Quantum Channel Estimation, *PRX Quantum* **2**, 010343 (2021).
- [27] S. Kurdzialek, W. Górecki, F. Albarelli, and R. Demkowicz-Dobrzański, Using adaptiveness and causal superpositions against noise in quantum metrology, *Phys. Rev. Lett.* **131**, 090801 (2023).
- [28] K. Banaszek, A. Dragan, W. Wasilewski, and C. Radzewicz, Experimental Demonstration of Entanglement-Enhanced Classical Communication over a Quantum Channel with Correlated Noise, *Phys. Rev. Lett.* **92**, 257901 (2004).
- [29] F. Caruso, V. Giovannetti, C. Lupo, and S. Mancini, Quantum channels and memory effects, *Rev. Mod. Phys.* **86**, 1203 (2014).
- [30] I. de Vega and D. Alonso, Dynamics of non-Markovian open quantum systems, *Rev. Mod. Phys.* **89**, 015001 (2017).
- [31] P. Szańkowski, G. Ramon, J. Krzywda, D. Kwiatkowski, and L. Cywiński, Environmental noise spectroscopy with qubits subjected to dynamical decoupling, *J. Phys. Condens. Matter* **29**, 333001 (2017).
- [32] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, Non-Markovian quantum processes: Complete framework and efficient characterization, *Phys. Rev. A* **97**, 012127 (2018).
- [33] U. von Lüpke, F. Beaudoin, L. M. Norris, Y. Sung, R. Winik, J. Y. Qiu, M. Kjaergaard, D. Kim, J. Yoder, S. Gustavsson, L. Viola, and W. D. Oliver, Two-qubit spectroscopy of spatiotemporally correlated quantum noise in superconducting qubits, *Phys. Rev. X Quantum* **1**, 010305 (2020).
- [34] S. Milz and K. Modi, Quantum Stochastic Processes and Quantum non-Markovian Phenomena, *PRX Quantum* **2**, 030201 (2021).
- [35] R. Harper and S. T. Flammia, Learning Correlated Noise in a 39-Qubit Quantum Processor, *PRX Quantum* **4**, 040311 (2023).
- [36] F. A. Mele, G. D. Palma, M. Fanizza, V. Giovannetti, and L. Lami, Optical Fibers With Memory Effects and Their Quantum Communication Capacities, *IEEE Trans. Inf. Theory* **70**, 8844 (2024).
- [37] J. Preskill, Sufficient condition on noise correlations for scalable quantum computing, *QIC* **13**, 181 (2013).
- [38] Y. Matsuzaki, S. C. Benjamin, and J. F. Fitzsimons, Magnetic field sensing beyond the standard quantum limit under the effect of decoherence, *Phys. Rev. A* **84**, 012103 (2011).
- [39] A. W. Chin, S. F. Huelga, and M. B. Plenio, Quantum Metrology in Non-Markovian Environments, *Phys. Rev. Lett.* **109**, 233601 (2012).
- [40] K. Macieszczak, Zeno limit in frequency estimation with non-Markovian environments, *Phys. Rev. A* **92**, 010102 (2015).
- [41] A. Smirne, J. Kołodyński, S. F. Huelga, and R. Demkowicz-Dobrzański, Ultimate Precision Limits for Noisy Frequency Estimation, *Phys. Rev. Lett.* **116**, 120801 (2016).
- [42] J. F. Haase, A. Smirne, J. Kołodyński, R. Demkowicz-Dobrzański, and S. F. Huelga, Fundamental limits to frequency estimation: A comprehensive microscopic perspective, *New J. Phys.* **20**, 053009 (2018).
- [43] P. Szańkowski, M. Trippenbach, and J. Chwedeńczuk, Parameter estimation in memory-assisted noisy quantum interferometry, *Phys. Rev. A* **90**, 063619 (2014).
- [44] A. Smirne, A. Lemmer, M. B. Plenio, and S. F. Huelga, Improving the precision of frequency estimation via long-time coherences, *Quantum Sci. Technol.* **4**, 025004 (2019).
- [45] D. Tamascelli, C. Benedetti, H.-P. Breuer, and M. G. A. Paris, Quantum probing beyond pure dephasing, *New J. Phys.* **22**, 083027 (2020).
- [46] A. Altherr and Y. Yang, Quantum Metrology for Non-Markovian Processes, *Phys. Rev. Lett.* **127**, 060501 (2021).
- [47] X. Yang, X. Long, R. Liu, K. Tang, Y. Zhai, X. Nie, T. Xin, J. Li, and D. Lu, Control-enhanced non-Markovian quantum metrology, *Commun. Phys.* **7**, 282 (2024).
- [48] J. Jeske, J. H. Cole, and S. F. Huelga, Quantum metrology subject to spatially correlated Markovian noise: Restoring the Heisenberg limit, *New J. Phys.* **16**, 073039 (2014).
- [49] D. Layden and P. Cappelaro, Spatial noise filtering through error correction for quantum sensing, *Npj Quantum Inf.* **4**, 30 (2018).
- [50] J. Czakowski, K. Pawłowski, and R. Demkowicz-Dobrzański, Many-body effects in quantum metrology, *New J. Phys.* **21**, 053031 (2019).
- [51] D. Layden, S. Zhou, P. Cappelaro, and L. Jiang, Ancilla-Free Quantum Error Correction Codes for Quantum Metrology, *Phys. Rev. Lett.* **122**, 040502 (2019).
- [52] D. Layden, M. Chen, and P. Cappelaro, Efficient Quantum Error Correction of Dephasing Induced by a Common Fluctuator, *Phys. Rev. Lett.* **124**, 020504 (2020).
- [53] G. Planella, M. F. B. Cenni, A. Acín, and M. Mehboudi, Bath-Induced Correlations Enhance Thermometry Precision at Low Temperatures, *Phys. Rev. Lett.* **128**, 040502 (2022).
- [54] F. Beaudoin, L. M. Norris, and L. Viola, Ramsey interferometry in correlated quantum noise environments, *Phys. Rev. A* **98**, 020102 (2018).
- [55] F. Riberi, L. M. Norris, F. Beaudoin, and L. Viola, Frequency estimation under non-Markovian spatially correlated quantum noise: Restoring superclassical precision scaling, *New J. Phys.* **24**, 103011 (2022).
- [56] M. Tsang, R. Nair, and X.-M. Lu, Quantum theory of superresolution for two incoherent optical point sources, *Phys. Rev. X* **6**, 031033 (2016).
- [57] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
- [58] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, 2nd ed. (Edizioni della Normale, Pisa, 2011).
- [59] J. Kołodyński, *Precision Bounds in Noisy Quantum Metrology*, Ph.D. thesis, University of Warsaw (2014), [arXiv:1409.0535](https://arxiv.org/abs/1409.0535).
- [60] Q. Liu, Z. Hu, H. Yuan, and Y. Yang, Optimal Strategies of Quantum Metrology with a Strict Hierarchy, *Phys. Rev. Lett.* **130**, 070803 (2023).
- [61] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum Circuit Architecture, *Phys. Rev. Lett.* **101**, 060401 (2008).

- (2008).
- [62] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, *Phys. Rev. A* **80**, 022339 (2009).
 - [63] The CJ operator of a channel $\Lambda : \mathcal{H}_{\text{in}} \rightarrow \mathcal{H}_{\text{out}}$ is defined as $\Lambda = \Lambda \otimes \mathcal{I}(|\Psi_+\rangle\langle\Psi_+|)$, where $|\Psi_+\rangle = \sum_i |i\rangle |i\rangle$, $|i\rangle$ is orthonormal basis of $\mathcal{H}_{\text{in}}$, $\mathcal{I}$ is identity channel.
 - [64] The performance operator is often defined as transposition of this expression [46]. It does not affect the further results—for example, (3) remains valid irrespectively of the convention because $(C^{(N)})^T$ is a comb iff $C^{(N)}$ is a comb. The typical notation for performance operator is Ω , here we use α because of analogy with a single-channel object.
 - [65] S. Kurdzialek, P. Dulian, J. Majsak, S. Chakraborty, and R. Demkowicz-Dobrzański, Quantum metrology using quantum combs and tensor network formalism, *New Journal of Physics* **10.1088/1367-2630/ada8d1** (2025).
 - [66] K. Chabuda, J. Dziarmaga, T. J. Osborne, and R. Demkowicz-Dobrzański, Tensor-network approach for quantum metrology in many-body quantum systems, *Nat. Commun.* **11**, 250 (2020).
 - [67] P. Dulian, S. Kurdzialek, and R. Demkowicz-Dobrzański, Qmetro++, <https://github.com/pdulian/qmetro> (2025).
 - [68] F. Ciccarello, S. Lorenzo, V. Giovannetti, and G. M. Palma, Quantum collision models: Open system dynamics from repeated interactions, *Phys. Rep.* **954**, 1 (2022).
 - [69] M. Guță, Fisher information and asymptotic normality in system identification for quantum Markov chains, *Phys. Rev. A* **83**, 062324 (2011).
 - [70] A. Godley and M. Guta, Adaptive measurement filter: Efficient strategy for optimal estimation of quantum Markov chains, *Quantum* **7**, 973 (2023).
 - [71] S. Seah, S. Nimmrichter, D. Grimmer, J. P. Santos, V. Scarani, and G. T. Landi, Collisional Quantum Thermometry, *Phys. Rev. Lett.* **123**, 180602 (2019).
 - [72] A. Shu, S. Seah, and V. Scarani, Surpassing the thermal Cramér-Rao bound with collisional thermometry, *Phys. Rev. A* **102**, 042417 (2020).
 - [73] E. O'Connor, B. Vacchini, and S. Campbell, Stochastic Collisional Quantum Thermometry, *Entropy* **23**, 1634 (2021).
 - [74] T. M. Mendonça, D. O. Soares-Pinto, and M. Paterostro, Information flow-enhanced precision in collisional quantum thermometry (2024), [arXiv:2407.21618](https://arxiv.org/abs/2407.21618).
 - [75] This assumption differs from Ref. [71, 72], in which the initial state was assumed to be stationary state of both S_T^x and V . However, this assumption does not affect the asymptotic results.
 - [76] Q. Liu and Y. Yang, Efficient tensor networks for control-enhanced quantum metrology, *Quantum* **8**, 1571 (2024).
 - [77] M. R. Jørgensen and F. A. Pollock, Exploiting the Causal Tensor Network Structure of Quantum Processes to Efficiently Simulate Non-Markovian Path Integrals, *Phys. Rev. Lett.* **123**, 240602 (2019).
 - [78] G. E. Fux, E. P. Butler, P. R. Eastham, B. W. Lovett, and J. Keeling, Efficient Exploration of Hamiltonian Parameter Space for Optimal Control of Non-Markovian Open Quantum Systems, *Phys. Rev. Lett.* **126**, 200401 (2021).
 - [79] E. P. Butler, G. E. Fux, C. Ortega-Taberner, B. W. Lovett, J. Keeling, and P. R. Eastham, Optimizing Performance of Quantum Operations with Non-Markovian Decoherence: The Tortoise or the Hare?, *Phys. Rev. Lett.* **132**, 060401 (2024).
 - [80] C. Ortega-Taberner, E. O'Neill, E. Butler, G. E. Fux, and P. R. Eastham, Unifying methods for optimal control in non-Markovian quantum systems via process tensors, *J. Chem. Phys.* **161**, 124119 (2024).
 - [81] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Memory Effects in Quantum Channel Discrimination, *Phys. Rev. Lett.* **101**, 180501 (2008).
 - [82] G. Gutoski, On a measure of distance for quantum strategies, *J. Math. Phys.* **53**, 032202 (2012).
 - [83] X. Wang and M. M. Wilde, Resource theory of asymmetric distinguishability for quantum channels, *Phys. Rev. Research* **1**, 033169 (2019).
 - [84] K. Nakahira and K. Kato, Generalized quantum process discrimination problems, *Phys. Rev. A* **103**, 062606 (2021).
 - [85] K. Nakahira and K. Kato, Simple Upper and Lower Bounds on the Ultimate Success Probability for Discriminating Arbitrary Finite-Dimensional Quantum Processes, *Phys. Rev. Lett.* **126**, 200502 (2021).
 - [86] C. Hirche, Quantum Network Discrimination, *Quantum* **7**, 1064 (2023).
 - [87] G. Zambon, Process tensor distinguishability measures, *Phys. Rev. A* **110**, 042210 (2024).
 - [88] J. Bavaresco, M. Murao, and M. T. Quintino, Strict Hierarchy between Parallel, Sequential, and Indefinite-Causal-Order Strategies for Channel Discrimination, *Phys. Rev. Lett.* **127**, 200504 (2021).

Appendix A: Basics of Quantum Fisher Information

The quantum Fisher information (QFI) of an arbitrary mixed state ρ_θ can be calculated using the formula

$$F(\rho_\theta) = \text{Tr}(\rho_\theta L_\theta^2), \quad \dot{\rho}_\theta = \frac{1}{2}(\rho_\theta L_\theta + L_\theta \rho_\theta), \quad (\text{A1})$$

where dot denotes the derivative over θ , and the symmetric logarithmic derivative (SLD) matrix L_θ can be calculated by solving the rightmost equation from (A1). For pure state models $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$, the closed analytical formula for the SLD and the QFI can be easily derived:

$$L_\theta = 2 \left(|\dot{\psi}_\theta\rangle\langle\psi_\theta| + |\psi_\theta\rangle\langle\dot{\psi}_\theta| \right), \quad F(|\psi_\theta\rangle) = 4(|\dot{\psi}_\theta\rangle\langle\dot{\psi}_\theta| - |\langle\psi_\theta|\dot{\psi}_\theta\rangle|^2). \quad (\text{A2})$$

When $|\Psi_\theta\rangle$ is a purification of ρ_θ , then $F(|\Psi_\theta\rangle) \geq F(\rho_\theta)$. Additionally, for each ρ_θ we can construct its QFI non-increasing purification (QFI NIP) satisfying

$$F(|\Psi_\theta\rangle) = F(\rho_\theta), \quad \langle \Psi_\theta | \dot{\Psi}_\theta \rangle = 0 \quad (\text{A3})$$

The second condition can be always satisfied by multiplying $|\Psi_\theta\rangle$ with a global, θ -dependent phase.

Appendix B: Basics of quantum combs

Let us introduce the formal definition of quantum combs [61]. We use roman font for channels and italics for corresponding Choi-Jamiołkowski (CJ) operators: the CJ operator of a channel $\Lambda : \mathcal{H}_{\text{in}} \rightarrow \mathcal{H}_{\text{out}}$ is defined as $\Lambda = \Lambda \otimes \mathcal{I}(|\Psi_+\rangle\langle\Psi_+|)$, where $|\Psi_+\rangle = \sum_i |i\rangle |i\rangle$, $|i\rangle$ is an orthonormal basis of $\mathcal{H}_{\text{in}}$, $\mathcal{I}$ is the identity channel. Let E be a quantum channel with inputs $\mathcal{K}_1, \mathcal{K}_3, \dots, \mathcal{K}_{2N-1}$ and outputs $\mathcal{K}_2, \mathcal{K}_4, \dots, \mathcal{K}_{2N}$. Then, $E \in \text{Comb}[(\mathcal{K}_1, \mathcal{K}_2), (\mathcal{K}_3, \mathcal{K}_4), \dots, (\mathcal{K}_{2N-1}, \mathcal{K}_{2N})]$ if and only if (iff) there exists a sequence of operators $E^{(k)}$ for $k = 1, 2, \dots, N$ such that $E = E^{(N)}$ and:

1. $\text{Tr}_{2k} E^{(k)} = E^{(k-1)} \otimes \mathbb{1}_{2k-1}$ for $k = 2, 3, \dots, N$
2. $\text{Tr}_2 E^{(1)} = \mathbb{1}_1$
3. $E \succeq 0$

Intuitively, $\mathcal{K}_{2k-1}, \mathcal{K}_{2k}$ are input/output spaces of the k th tooth of a comb. The above conditions imply the causal order of the teeth—the output $\mathcal{K}_{2l}$ can depend on the input $\mathcal{K}_{2j-1}$ only if $l \geq j$. In other words, quantum combs can be used to simulate channels with memory, when each output may depend on all preceding inputs. Therefore, they are the ideal tool to represent sequences of correlated channels or adaptive quantum strategies. Notably, a sequence $E^{(k)}$ consists of quantum combs with increasing number of teeth: $E^{(k)} \in \text{Comb}[(\mathcal{K}_1, \mathcal{K}_2), \dots, (\mathcal{K}_{2k-1}, \mathcal{K}_{2k})]$. Two combs can be connected with each other using a *link product* operation $\star$ [61]. When E, F are two quantum combs, then $G = E \star F$ is a comb created by connecting the input spaces of E with the corresponding output spaces of F and the output spaces of E with the corresponding input spaces of F (only common spaces of E and F get connected). Let E, F act on sets of spaces M, N respectively. The link product is then formally defined using CJ operators :

$$G = E \star F \iff G = \text{Tr}_{M \cap N} [(E^{\text{T}_{M \cap N}} \otimes \mathbb{1}_{N \setminus M}) (\mathbb{1}_{M \setminus N} \otimes F)],$$

where $M \cap N$ is the set of common (linked) spaces, $\text{T}_{M \cap N}$ is a partial transposition (with respect to common spaces only). When $M = N$, then $E \star F$ is a scalar, using the comb conditions and the fact that each output of E is an input of F and *vice versa*, it can be then shown that $E \star F = 1$. For more information about quantum combs and link product see Refs. [61, 62, 81].

Appendix C: Uncorrelated noise bound

The bound (4) has been derived in Ref. [27], and the asymptotic bounds (5) and (6) are its direct consequences, see equations (11), (12) from Ref. [27]. In what follows, we provide a simpler derivation of (4). From the new derivation it is clear, that the bound generated by (7) for uncorrelated noise is at least as tight as the one given by (4). The advantage of the old derivation is that it can be easily generalized to strategies involving causal superpositions.

1. Simplified derivation

We consider an adaptive (AD) scheme in which N independent channels Λ_θ are probed (see Fig. 3) and keep the notation from the main text. Let $\rho_\theta^{(l)} \in \mathcal{L}(\mathcal{H}_{2l} \otimes \mathcal{A}_l)$ be the probe and ancilla state after l th use of Λ_θ , and let $F(\rho_\theta^{(l)}) = F^{(l)}$. The state after the next control is $C_l(\rho_\theta^{(l)})$, we construct its QFI NIP $|\Psi_\theta^{(l)}\rangle \in \mathcal{H}_{2l+1} \otimes \mathcal{A}_{l+1} \otimes \mathcal{E}_l$, where $\mathcal{E}_l$ is an artificial space added for purification purposes. Using (A3) and (A2) and the fact that a θ -independent channel cannot increase the QFI [59, p. 57], we obtain

$$4 \langle \dot{\Psi}_\theta^{(l)} | \dot{\Psi}_\theta^{(l)} \rangle = F(C_l(\rho_\theta^{(l)})) \leq F^{(l)}. \quad (\text{C1})$$

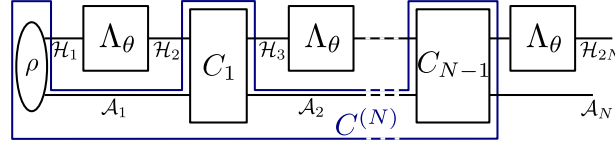


FIG. 3. Adaptive scheme for uncorrelated channels

Let $\bar{K}_{k,\theta} = K_{k,\theta} \otimes \mathbb{1}_{\mathcal{A}_{l+1} \otimes \mathcal{E}_l}$, the state of probe and ancilla after the action of the $(l+1)$ th channel can be written as

$$\rho_\theta^{(l+1)} = \text{Tr}_{\mathcal{E}_l} \left(\sum_k \bar{K}_{k,\theta} |\Psi_\theta^{(l)}\rangle \langle \Psi_\theta^{(l)}| \bar{K}_{k,\theta}^\dagger \right).$$

Since the QFI of a subsystem is smaller or equal than the QFI of the whole system (i) and the QFI of a purification is larger or equal than the QFI of a purified state (ii), we obtain

$$F^{(l+1)} \stackrel{(i)}{\leq} F \left(\sum_k \bar{K}_{k,\theta} |\Psi_\theta^{(l)}\rangle \langle \Psi_\theta^{(l)}| \bar{K}_{k,\theta}^\dagger \right) \stackrel{(ii)}{\leq} F \left(\sum_k \bar{K}_{k,\theta} |\Psi_\theta^{(l)}\rangle \otimes |k\rangle_{\mathcal{E}'_l} \langle k|_{\mathcal{E}'_l} \right), \quad (\text{C2})$$

where $\mathcal{E}'_l$ is an additional space added for purification, and the vectors $|k\rangle_{\mathcal{E}'_l}$ form its o.-n. basis. After expanding the rightmost part of (C2) using (A2), we obtain

$$\begin{aligned} F^{(l+1)}/4 &\leq \langle \dot{\Psi}_\theta^{(l)} | \sum_k \bar{K}_{\theta,k}^\dagger \bar{K}_{\theta,k} | \dot{\Psi}_\theta^{(l)} \rangle + \langle \dot{\Psi}_\theta^{(l)} | \sum_k \bar{K}_{\theta,k}^\dagger \dot{\bar{K}}_{\theta,k} | \Psi_\theta^{(l)} \rangle + \langle \Psi_\theta^{(l)} | \sum_k \dot{\bar{K}}_{\theta,k}^\dagger \bar{K}_{\theta,k} | \dot{\Psi}_\theta^{(l)} \rangle + \\ &+ \langle \Psi_\theta^{(l)} | \sum_k \dot{\bar{K}}_{\theta,k}^\dagger \dot{\bar{K}}_{\theta,k} | \Psi_\theta^{(l)} \rangle \end{aligned} \quad (\text{C3})$$

From (C1) and the identity $\sum_k \bar{K}_{\theta,k}^\dagger \bar{K}_{\theta,k} = \mathbb{1}$ we deduce that the first term is upper bounded by $F^{(l)}/4$; the last term is upper-bounded by $\|\alpha\|$ because $\langle \Psi_\theta^{(l)} | \Psi_\theta^{(l)} \rangle = 1$ and $\|A \otimes \mathbb{1}\| = \|A\|$. Both the second and the third term are upper-bounded by $\sqrt{F^{(l)}/4} \|\beta\|$, this follows from the inequality

$$\langle x | A | y \rangle \leq \sqrt{\langle x | x \rangle} \|A\| \sqrt{\langle y | y \rangle},$$

which is a less general version of (7) from [27]. After taking this all together, we obtain

$$\mathcal{F}_{\text{AD}}^{(l+1)} \leq \mathcal{F}_{\text{AD}}^{(l)} + 4 \min_h \left[\|\alpha\| + \sqrt{\mathcal{F}_{\text{AD}}^{(l)}} \|\beta\| \right], \quad (\text{C4})$$

which is (4) from the main text.

To derive this bound, we assumed to have access to the QFI NIP of probe and ancilla after each control. Moreover, we maximized each term of (C3) independently—this is another factor making the bound not tight, since usually all the terms cannot be maximal at the same time. Interestingly, to derive (7), we only assumed to have access to the QFI NIP every m steps, no additional assumptions were required (see the exact derivation in the next Section). Therefore, the bound generated by (7) is guaranteed to be at least as tight as the one generated by (4) already for $m = 1$. Increasing m may tighten the bound even more.

2. Example

To illustrate the effectiveness of the new bound for uncorrelated noise, let us calculate it for $m \in \{1, 2, 3\}$ for phase estimation in the presence of amplitude damping noise, for which $K_{k,\theta} = K_k e^{-\frac{i\theta\sigma_z}{2}}$, with

$$K_1 = |0\rangle \langle 0| + \sqrt{p} |1\rangle \langle 1|, \quad K_2 = \sqrt{1-p} |0\rangle \langle 1|.$$

In Fig. 4, we demonstrate the results for $p = 0.5$. The newly introduced bound is tighter than the old one already for $m = 1$. For larger m , the bound becomes even tighter, and very close to the optimal QFI calculated exactly for $m \leq 4$ using the algorithm from Ref. [46].

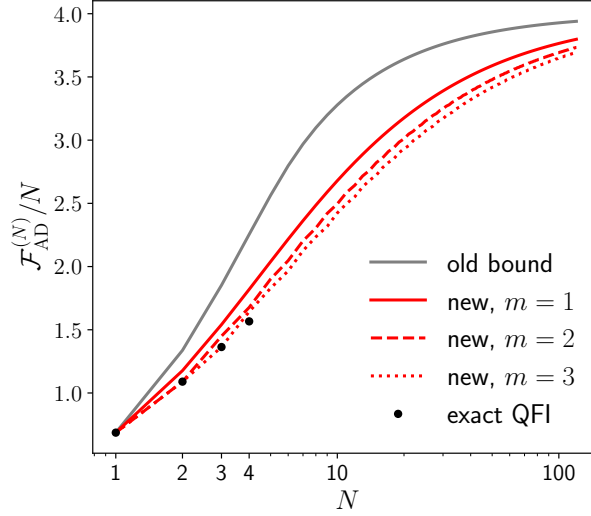


FIG. 4. Precision bounds for phase estimation in the presence of amplitude damping noise.

Appendix D: Correlated noise bounds

In this Section, we provide a detailed derivation of the recursive bounds (7) and (8) valid for correlated noise. Then, we derive the asymptotic bounds (9) and (10) using (8). Finally, we demonstrate how to formulate these new recursive and asymptotic bounds as SDPs.

1. Derivation of recursive bound (7)

Consider the joint state of the system, ancilla, and environment after l teeth of the comb $\Lambda_\theta^{(l)}$: $\rho_\theta^{(l)} \in \mathcal{H}_{2l} \otimes \mathcal{R}_l \otimes \mathcal{A}_l$, see Fig. 1 in the main text. To upper bound $F(\rho_\theta^{(l+m)})$ in terms of $F(\rho_\theta^{(l)})$, we replace $\rho_\theta^{(l)}$ with its QFI NIP $|\Psi_\theta^{(l)}\rangle$. This substitution can only increase the resulting QFI, since any operation possible on the original state is also possible on its purification. Next, we derive an upper bound on the QFI of the state obtained by evolving $|\Psi_\theta^{(l)}\rangle$ through the next m teeth, represented by $\Lambda_\theta^{\leftrightarrow(m)}$. Arbitrary adaptive control is allowed during this evolution. To calculate this upper bound, let us construct a comb $\Lambda_\theta^{(l,m)} = |\Psi_\theta^{(l)}\rangle \langle \Psi_\theta^{(l)}| \otimes \Lambda_\theta^{\leftrightarrow(m)}$, such that $|\Psi_\theta^{(l)}\rangle \langle \Psi_\theta^{(l)}|$ is its first tooth (with trivial input), see Fig. 5.

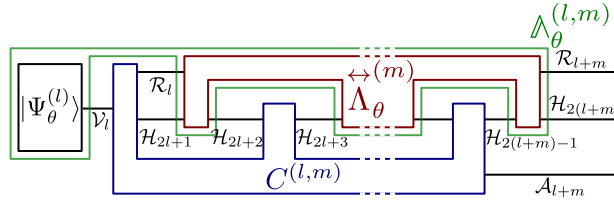


FIG. 5. The scheme of comb $\Lambda_\theta^{(l,m)}$ and construction of the output state $\rho_\theta^{(l+m)}$

Then, the total output state of $\mathcal{H}_{2(l+m)} \otimes \mathcal{R}_{l+m} \otimes \mathcal{A}_{l+m}$ can be written as $C^{(l,m)} \star \Lambda_\theta^{(l,m)}$, where $C^{(l,m)}$ is an arbitrary comb representing adaptive control, $C^{(l,m)}$ acts on the spaces depicted in Fig. 5. Notice, that $C^{(l,m)}$ acts also on environmental subspace $\mathcal{R}_l$, which is impossible in a normal AD scheme. Therefore, the QFI of $C^{(l,m)} \star \Lambda_\theta^{(l,m)}$ might be larger than $F(\rho_\theta^{(l+m)})$, moreover $F(\rho_\theta^{(l+m)}) \leq \max_{C^{(l,m)}} F(C^{(l,m)} \star \Lambda_\theta^{(l,m)})$ because all types of controls allowed in a normal AD scheme can be simulated by a proper choice of $C^{(l,m)}$. These “environmental leakages” that happen every m teeth are one of the reasons for the potential looseness of the bound.

We are now nearly ready to compute the upper bound for $F(\rho_\theta^{(l+m)})$, expressed as the comb QFI of $\mathbb{A}_\theta^{(l,m)}$, which can be evaluated for any comb using the results of Refs. [46, 60]. To fully specify $\mathbb{A}_\theta^{(l,m)}$, we define the basis of $\mathcal{V}_l = \text{span}(|\Psi_\theta^{(l)}\rangle, |\dot{\Psi}_\theta^{(l)}\rangle)$:

$$|0\rangle = |\Psi_\theta^{(l)}\rangle, \quad |1\rangle = 2|\dot{\Psi}_\theta^{(l)}\rangle / \sqrt{F(\rho_\theta^{(l)})}. \quad (\text{D1})$$

This basis is orthonormal, which follows directly from quantum state normalization and from properties of the QFI NIP (A3). The density matrices of the state and its derivative can be expressed in this basis as

$$|\Psi_\theta^{(l)}\rangle \langle \Psi_\theta^{(l)}| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \frac{d}{d\theta} (|\Psi_\theta^{(l)}\rangle \langle \Psi_\theta^{(l)}|) = \frac{\sqrt{F(\rho_\theta^{(l)})}}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Finally, we can explicitly write the bound for $F(\rho_\theta^{(l+m)})$ as $F(\rho_\theta^{(l+m)}) \leq F^{(l+m)}$, where $F^{(km)}$ is defined by the following iteration

$$F^{(0)} = 0, \quad F^{(l+m)} = \max_{C^{(l,m)}} F(C^{(l,m)} \star \mathbb{A}_\theta^{(l,m)}), \quad \mathbb{A}_\theta^{(l,m)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \overset{\leftrightarrow}{\Lambda}_\theta^{(m)}, \quad \dot{\mathbb{A}}_\theta^{(l,m)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \dot{\overset{\leftrightarrow}{\Lambda}}_\theta^{(m)} + \frac{\sqrt{F^{(l)}}}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \overset{\leftrightarrow}{\Lambda}_\theta^{(m)}, \quad (\text{D2})$$

where $F(\rho_\theta^{(l)})$ was changed to its upper bound $F^{(l)}$ in the expression for $\frac{d}{d\theta} (|\Psi_\theta^{(l)}\rangle \langle \Psi_\theta^{(l)}|)$.

The maximization over $C^{(l,m)}$ can be performed using the SDP described in Refs. [46, 60]. The iteration D2 gives an upper bound for $F(\rho_\theta^{(l)})$ valid for all control combs C , so it is also an upper bound for $\mathcal{F}_{\text{AD}}^{(l)}$. Interestingly, for uncorrelated cases there are no “environmental leakages”, so the only reason for the lack of tightness of the bound is that the state $\rho_\theta^{(l)}$ is replaced with its QFI NIP every m steps. This makes the bound (D2) at least as tight as (C4) for uncorrelated cases.

2. Derivation of recursive bound (8)

Let $\overset{\leftrightarrow}{\Lambda}_\theta^{(m)} = \sum_k |K_{k,\theta}^{(m)}\rangle \langle K_{k,\theta}^{(m)}|$, since $\mathbb{A}_\theta^{(l,m)} = |\Psi_\theta^{(l)}\rangle \langle \Psi_\theta^{(l)}| \otimes \overset{\leftrightarrow}{\Lambda}_\theta^{(m)}$, we can decompose

$$\mathbb{A}_\theta^{(l,m)} = \sum_k |L_{k,\theta}^{(l,m)}\rangle \langle L_{k,\theta}^{(l,m)}|, \quad (\text{D3})$$

where $|L_{k,\theta}^{(l,m)}\rangle = |\Psi_\theta^{(l)}\rangle \otimes |K_{k,\theta}^{(m)}\rangle$. Using Leibniz rule and (D1) we obtain

$$|L_{k,\theta}^{(l,m)}\rangle = |0\rangle \otimes |K_{k,\theta}^{(m)}\rangle, \quad |\dot{L}_{k,\theta}^{(l,m)}\rangle = |0\rangle \otimes |\dot{K}_{k,\theta}^{(m)}\rangle + \sqrt{F^{(l)}}/2 |1\rangle \otimes |K_{k,\theta}^{(m)}\rangle. \quad (\text{D4})$$

Let $|\dot{\tilde{K}}_{k,\theta}^{(m)}\rangle = |\dot{K}_{k,\theta}^{(m)}\rangle - ih_{kk'} |K_{k',\theta}^{(m)}\rangle$, where h is a hermitian matrix generating equivalent Kraus representations of a channel; the summation is performed over repeated indices. Then, $|\dot{\tilde{L}}_{k,\theta}^{(l,m)}\rangle = |0\rangle \otimes |\dot{\tilde{K}}_{k,\theta}^{(m)}\rangle + \sqrt{F^{(l)}}/2 |1\rangle \otimes |K_{k,\theta}^{(m)}\rangle$, where $|\dot{\tilde{L}}_{k,\theta}^{(l,m)}\rangle = |\dot{L}_{k,\theta}^{(l,m)}\rangle - ih_{kk'} |L_{k',\theta}^{(l,m)}\rangle$. The performance operator of $\overset{\leftrightarrow}{\Lambda}_\theta^{(m)}$ is $\alpha^{(m)} = \text{Tr}_{\text{out}} \left(\sum_k |\dot{\tilde{K}}_{k,\theta}^{(m)}\rangle \langle \dot{\tilde{K}}_{k,\theta}^{(m)}| \right)$, the performance operator of $\mathbb{A}_\theta^{(l,m)}$ is $\alpha_{\mathbb{A}}^{(l,m)} = \text{Tr}_{\text{out}} \left(\sum_k |\dot{\tilde{L}}_{k,\theta}^{(l,m)}\rangle \langle \dot{\tilde{L}}_{k,\theta}^{(l,m)}| \right)$, here Tr_{out} is the partial trace over the subspace $\mathcal{H}_{2(l+m)} \otimes \mathcal{R}_{l+m}$, which is the last output of $\overset{\leftrightarrow}{\Lambda}_\theta^{(m)}$ and of $\mathbb{A}_\theta^{(l,m)}$. Using (D4), we get :

$$\alpha_{\mathbb{A}}^{(l,m)} = |0\rangle \langle 0| \otimes \alpha^{(m)} + \frac{1}{2} \sqrt{F^{(l)}} |0\rangle \langle 1| \otimes \beta^{(m)} + \frac{1}{2} \sqrt{F^{(l)}} |1\rangle \langle 0| \otimes \beta^{(m)\dagger} + \frac{1}{4} F^{(l)} |1\rangle \langle 1| \otimes \tilde{\Lambda}^{(m)}, \quad (\text{D5})$$

where $\tilde{\Lambda}^{(m)} = \text{Tr}_{\text{out}} \overset{\leftrightarrow}{\Lambda}_\theta^{(m)}$, $\beta^{(m)} = \text{Tr}_{\text{out}} \left(\sum_k |\dot{\tilde{K}}_{k,\theta}^{(m)}\rangle \langle K_{k,\theta}^{(m)}| \right)$. After writing the maximization in (D2) using the expression for the comb QFI (3) in terms of the performance operator, we get

$$F^{(l+m)} = \min_h \max_{\tilde{C}^{(m,l)}} 4\text{Tr} \left(\alpha_{\mathbb{A}}^{(l,m)} \tilde{C}^{(l,m)} \right),$$

where $\tilde{C}^{(l,m)} = \text{Tr}_{\mathcal{A}_{l+m}} C^{(l,m)}$. After decomposing $\tilde{C}^{(l,m)} = \sum_{i,j=0}^1 |i\rangle_{\mathcal{V}_i} \langle j| \otimes \tilde{C}_{ij}^{(l,m)}$, using (D5) and the normalization condition $\text{Tr}(\tilde{C}_{11}^{(l,m)} \tilde{\Lambda}^{(m)}) = \tilde{C}_{11}^{(l,m)} \star \Lambda^{(m)} = 1$, we obtain

$$4\text{Tr}\left(\alpha_{\mathcal{A}}^{(l,m)} \tilde{C}^{(l,m)}\right) = F^{(l)} + 4\text{Tr}\left(\alpha^{(m)} \tilde{C}_{00}^{(l,m)}\right) + 4\sqrt{F^{(l)}} \text{Re}\left(\text{Tr}\left(\beta^{(m)} \tilde{C}_{10}^{(l,m)}\right)\right),$$

and consequently

$$F^{(l+m)} = F^{(l)} + 4 \min_h \max_{\tilde{C}^{(l,m)}} \left[\text{Tr}\left(\alpha^{(m)} \tilde{C}_{00}^{(l,m)}\right) + \sqrt{F^{(l)}} \text{Re}\left(\text{Tr}\left(\beta^{(m)} \tilde{C}_{10}^{(l,m)}\right)\right) \right] \quad (\text{D6})$$

The normalization condition was derived using the fact that $\tilde{C}_{11}^{(l,m)} \in \text{Comb}[(\emptyset, \mathcal{H}_{2l+1} \otimes \mathcal{R}_l), (\mathcal{H}_{2l+2}, \mathcal{H}_{2l+3}), \dots, (\mathcal{H}_{2(l+m)-2}, \mathcal{H}_{2(l+m)-1})]$ and $\tilde{\Lambda}^{(m)} \in \text{Comb}[(\mathcal{H}_{2l+1} \otimes \mathcal{R}_l, \mathcal{H}_{2l+2}), (\mathcal{H}_{2l+3}, \mathcal{H}_{2l+4}), \dots, (\mathcal{H}_{2(l+m)-1}, \emptyset)]$, so all outputs of $\tilde{C}_{11}^{(l,m)}$ are inputs of $\tilde{\Lambda}^{(m)}$ and *vice versa*.

To further simplify the bound, we maximize each term of the RHS of (D6) independently and use the inequality between the maximum of a sum and the sum of maxima, which leads to (8) from the main text:

$$F^{(l+m)} \leq F^{(l)} + 4 \min_h \left[a^{(m)} + \sqrt{F^{(l)}} b^{(m)} \right], \quad a^{(m)} = \max_{\tilde{C}^{(m)}} \text{Tr}\left(\alpha^{(m)} \tilde{C}_{00}^{(m)}\right), \quad b^{(m)} = \max_{\tilde{C}^{(m)}} \text{Re} \text{Tr}\left(\beta^{(m)} \tilde{C}_{10}^{(m)}\right). \quad (\text{D7})$$

The bound (D7) is less tight than (D2), but, as we will prove in the next section, there is no difference between bounds generated by (D7) and (D2) asymptotically (for large N).

For simplicity, starting from (D7), we replace $\tilde{C}^{(l,m)}$ with $\tilde{C}^{(m)}$ under maximization, since $\tilde{C}^{(l,m)}$ has the same structure regardless of l . This is possible, when subsequent teeth of a comb $\overset{\leftrightarrow}{\Lambda}_\theta$ are the same, as we assumed from the very beginning in the main text. Notice that this assumption is not necessary if we just want to derive recursive bounds as in (D2), but it is important to derive an asymptotic bound. For simplicity and for the sake of deriving an asymptotic bound, let us also adopt convention $l = 0$ for naming of spaces $\mathcal{H}$, $\mathcal{R}$, $\mathcal{A}$. Then we have $\tilde{C}^{(m)} \in \text{Comb}[(\mathcal{V}_0, \mathcal{R}_0 \otimes \mathcal{H}_1), (\mathcal{H}_2, \mathcal{H}_3), \dots, (\mathcal{H}_{2m-2}, \mathcal{H}_{2m-1} \otimes \mathcal{A}_m)]$, see Fig. 5 for a comparison.

3. Derivation of asymptotic bounds (9) and (10)

To derive the asymptotic bounds we will use the following lemma.

Lemma 1 Let $x^{(n)}$ be a sequence of real numbers satisfying $x^{(n+1)} = x^{(n)} + A + 2B\sqrt{x^{(n)}}$ for any integer n , $x^{(0)} = 0$, $A \geq B^2$. Then

$$\lim_{n \rightarrow \infty} x^{(n)}/n = A \quad \text{for } B = 0; \quad \lim_{n \rightarrow \infty} x^{(n)}/n^2 = B^2 \quad \text{for } B \geq 0.$$

Proof When $B = 0$, then $x^{(n)} = An$, so the first part of lemma is obvious. When $B > 0$, then for any $n \geq 1$

$$x^{(n)} \leq f(n) = An + B^2 n(n-1) + (A - B^2)n \log n, \quad (\text{D8})$$

which is proven in the Supplementary Material of Ref. [27], Appendix E, in which $\|\alpha\|$ and $\|\beta\|$ play the role of A and B . Let us now prove that for any $n \geq 1$

$$x^{(n)} \geq B^2 n^2. \quad (\text{D9})$$

This can be shown by induction, we have $x^{(1)} = A \geq B^2 1^2$, and when (D9) holds for some n , then for $n+1$ we have

$$x^{(n+1)} \geq B^2 n^2 + 2B^2 n + A \geq B^2 n^2 + 2B^2 n + B^2 = B^2 (n+1)^2.$$

From (D8), (D9), and the fact that $\lim_{n \rightarrow \infty} f(n)/n^2 = B^2$ follows the second part of the lemma. ■

The asymptotic bounds for uncorrelated noise (5) and (6) are direct consequences of the iterative bound (4) and Lemma 1—notice that asymptotically (for large $\mathcal{F}_{\text{AD}}^{(i)}$) the minimum over h is always achieved when $\|\beta\|$ is minimal, since it is multiplied by $\sqrt{\mathcal{F}_{\text{AD}}^{(i)}}$ and consequently eventually dominates over the term with $\|\alpha\|$, unless $\|\beta\| = 0$.

The assumptions of Lemma 1 are true for the uncorrelated case because of the inequality $\|\alpha\| \geq \|\beta\|^2$. Let us show that the analogous inequality holds for the correlated case as well.

Lemma 2 $a^{(m)} \geq b^{(m)2}$

Proof Let us use the notation $\tilde{C}^A, \tilde{C}^B$ for combs $\tilde{C}^{(m)}$ maximizing $a^{(m)}$ and $b^{(m)}$ respectively. Because $a^{(m)} = \text{Tr}(\alpha^{(m)} \tilde{C}_{00}^A) \geq \text{Tr}(\alpha^{(m)} \tilde{C}_{00}^B)$ and $b^{(m)2} \leq |\text{Tr}(\beta^{(m)} \tilde{C}_{10}^B)|^2$, the inequality

$$\text{Tr}(\alpha^{(m)} \tilde{C}_{00}^B) \geq |\text{Tr}(\beta^{(m)} \tilde{C}_{10}^B)|^2$$

implies the thesis of our lemma. To prove this inequality, notice that

$$\begin{aligned} |\text{Tr}(\beta^{(m)} \tilde{C}_{10}^B)| &= \left| \text{Tr} \left(\text{Tr}_{\text{out}} \left(\sum_k |\dot{K}_{k,\theta}^{(m)}\rangle \langle K_{k,\theta}^{(m)}| \right) \tilde{C}_{10}^B \right) \right| = \left| \sum_k \langle K_{k,\theta}^{(m)} | \tilde{C}_{10}^B \otimes \mathbb{1}_{\text{out}} | \dot{K}_{k,\theta}^{(m)} \rangle \right| = \\ &= \left| \sum_k \langle K_{k,\theta}^{(m)} | \left(\tilde{C}_{11}^B \right)^{\frac{1}{2}} \left(\tilde{C}_{11}^B \right)^{-\frac{1}{2}} \tilde{C}_{10}^B \otimes \mathbb{1}_{\text{out}} | \dot{K}_{k,\theta}^{(m)} \rangle \right| \leq \\ &\stackrel{(i)}{\leq} \sqrt{\sum_k \langle K_{k,\theta}^{(m)} | \tilde{C}_{11}^B \otimes \mathbb{1}_{\text{out}} | K_{k,\theta}^{(m)} \rangle} \cdot \sqrt{\sum_k \langle \dot{K}_{k,\theta}^{(m)} | \tilde{C}_{10}^{B\dagger} \left(\tilde{C}_{11}^B \right)^{-1} \tilde{C}_{10}^B \otimes \mathbb{1}_{\text{out}} | \dot{K}_{k,\theta}^{(m)} \rangle} \stackrel{(ii)}{\leq} \sqrt{\sum_k \langle \dot{K}_{k,\theta}^{(m)} | \tilde{C}_{00}^B \otimes \mathbb{1}_{\text{out}} | \dot{K}_{k,\theta}^{(m)} \rangle} = \\ &= \sqrt{\text{Tr}(\alpha^{(m)} \tilde{C}_{00}^B)} \end{aligned}$$

In (i) we used Cauchy-Schwarz inequality $|\sum_k \langle x_k | y_k \rangle| \leq \sqrt{\sum_k \langle x_k | x_k \rangle} \sqrt{\sum_k \langle y_k | y_k \rangle}$ with $|x_k\rangle = \left(\tilde{C}_{11}^B \right)^{\frac{1}{2}} \otimes \mathbb{1}_{\text{out}} |K_{k,\theta}^{(m)}\rangle$, $|y_k\rangle = \left(\tilde{C}_{11}^B \right)^{-\frac{1}{2}} \tilde{C}_{10}^B \otimes \mathbb{1}_{\text{out}} |\dot{K}_{k,\theta}^{(m)}\rangle$. In (ii) we used the fact that the expression under the first square root is $\text{Tr} \left(\tilde{A}_\theta^{(m)} \tilde{C}_{11}^B \right) = 1$, since $\tilde{A}_\theta^{(m)}$ and $\tilde{C}_{11}^B$ are two combs whose inputs and outputs are compatible. We also used the fact that the matrix $\tilde{C}^B = \begin{bmatrix} \tilde{C}_{00}^B & \tilde{C}_{01}^B \\ \tilde{C}_{10}^B & \tilde{C}_{11}^B \end{bmatrix}$ is hermitian and positive-semidefinite, which implies that $\tilde{C}_{01}^B = \tilde{C}_{10}^{B\dagger}$, and due to Schur's complement condition $\tilde{C}_{00}^B \succeq \tilde{C}_{10}^{B\dagger} \left(\tilde{C}_{11}^B \right)^{-1} \tilde{C}_{10}^B$, which means that $\langle u | \tilde{C}_{00}^B \otimes \mathbb{1} | u \rangle \geq \langle u | \tilde{C}_{10}^{B\dagger} \left(\tilde{C}_{11}^B \right)^{-1} \tilde{C}_{10}^B \otimes \mathbb{1} | u \rangle$ for any $|u\rangle$. ■

If we perform the minimization over h in (D7) for $F^{(l)} \rightarrow \infty$, then the term with $b^{(m)}$ dominates the term with $a^{(m)}$. When there is an h for which $b^{(m)} = 0$, then for large enough l the minimum over h in recursive steps is achieved when $b^{(m)} = 0$, and, according to Lemma 1 applied for a sequence $F^{(0)}, F^{(m)}, F^{(2m)}, \dots$, we get

$$\lim_{N \rightarrow \infty} \frac{\mathcal{F}_{\text{AD}}^{(N)}}{N} \leq \frac{4}{m} \min_h a^{(m)} \text{ s.t. } b^{(m)} = 0, \quad (\text{D10})$$

which is (9) from the main text. When there is no such h , in the large l limit the minimum over h in (D7) is achieved when $b^{(m)}$ is minimal. Therefore, using Lemmas 1 and 2 for a sequence $F^{(0)}, F^{(m)}, F^{(2m)}, \dots$ we get

$$\lim_{N \rightarrow \infty} \frac{\mathcal{F}_{\text{AD}}^{(N)}}{N^2} \leq \frac{4}{m^2} \min_h b^{(m)2}. \quad (\text{D11})$$

Notably, according to Lemma 1, we did not loose any tightness while going from the recursion (D7) to the asymptotic bounds (D10) and (D11). However, the tightness could be potentially lost when we replaced maximum of sum in (D6) with sum of maxima in (D7). Let us now demonstrate that this has no effect asymptotically, so after iterating (D2) many times we get an asymptotic behavior predicted by (D10) or (D11). To show this, let us consider two cases:

1. Heisenberg scaling is possible. Then, for $F^{(l)} \rightarrow \infty$ the term $\sqrt{F^{(l)}} \text{Re} \left(\text{Tr} \left(\beta^{(m)} \tilde{C}_{10} \right) \right)$ dominates over the term $\text{Tr}(\alpha^{(m)} \tilde{C}_{00})$ in the RHS of (D6). When we pick $\tilde{C}^{(l+m)}$ such that $\tilde{C}_{10}$ maximizes the dominating term, then the result of a maximization over $\tilde{C}^{(l+m)}$ is arbitrarily close to the result of independent maximization over $\tilde{C}^A, \tilde{C}^B$ for $F^{(l)} \rightarrow \infty$ (the ratio between the two results $\rightarrow 1$).

2. Heisenberg scaling is not possible. Then, to perform minimization over h in (D6) for $F^{(l)} \rightarrow \infty$, we should choose h for which $\text{Re} \left(\text{Tr} \left(\beta^{(m)} \tilde{C}_{10} \right) \right) = 0$ for any $\tilde{C}_{10}$. Then, we can choose $\tilde{C}_{10} = 0$ without affecting the result. This makes the condition $\tilde{C}^{(m)} \succeq 0$ equivalent to $\tilde{C}_{00}, \tilde{C}_{11} \succeq 0$, and the maximization over $\tilde{C}^{(m)}$ of the sum of terms is equivalent to maximization of the first term over $\tilde{C}_{00}$.

4. Simpler form of the asymptotic bound for standard scaling

The asymptotic bound (D10) is expressed as a minimization over all hermitian matrices h for which the constraint $b^{(m)} = 0$ is satisfied. This constraint is hard to deal with because $b^{(m)}$ is defined as a result of a non-trivial maximization, see (D7)—let us therefore try to formulate it in a more manageable way.

Let $\mathcal{X}_m$ be a linear subspace of space $\mathcal{X}_m^{\text{tot}} = \mathcal{L}(\mathcal{R}_0 \otimes \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_{2m-1})$ such that $X \in \mathcal{X}_m$ iff

1. $\text{Tr}(X) = 0$,
2. There exists a sequence of operators $X^{(1)}, X^{(2)}, \dots, X^{(m-1)}, X^{(m)}$ for which $X = X^{(m)}$, $\forall_{2 \leq k \leq m} \text{Tr}_{\mathcal{H}_{2k-1}} X^{(k)} = X^{(k-1)} \otimes \mathbb{1}_{\mathcal{H}_{2k-2}}$, $X^{(1)} \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{R}_0)$.

Notice that condition 2. is similar to the linear comb conditions, a space $\mathcal{X}_m$ has a similar structure to the set of CJ operators of combs, but without positivity constraint and with zero trace. Let us also define $\mathcal{X}_m^\perp$ as an orthogonal complement of $\mathcal{X}_m$ in the space $\mathcal{X}_m^{\text{tot}}$ with respect to Hilbert-Schmidt scalar product $(A|B) = \text{Tr}(AB^\dagger)$. Since $\beta^{(m)} \in \mathcal{X}_m^{\text{tot}}$, we can uniquely decompose it as $\beta^{(m)} = \beta_1^{(m)} + \beta_2^{(m)}$, where $\beta_1^{(m)} \in \mathcal{X}_m$, $\beta_2^{(m)} \in \mathcal{X}_m^\perp$.

We are now ready to prove the following statement.

Lemma 3 $b^{(m)} \geq 0$, moreover $b^{(m)} = 0$ iff $\beta_1^{(m)} = 0$.

Proof The comb conditions for $\tilde{C}^{(m)}$ are equivalent to the following set of conditions for its blocks $\tilde{C}_{ij}^{(m)}$: (i) $\tilde{C}_{00}^{(m)}, \tilde{C}_{11}^{(m)} \in \text{Comb}[(\emptyset, \mathcal{H}_1 \otimes \mathcal{R}_0), (\mathcal{H}_2, \mathcal{H}_3), \dots, (\mathcal{H}_{2m-2}, \mathcal{H}_{2m-1})]$; (ii) $\tilde{C}_{01}^{(m)}, \tilde{C}_{10}^{(m)} \in \mathcal{X}_m$; (iii) $\tilde{C}_{01}^{(m)} = \tilde{C}_{10}^{(m)\dagger}$ and $\tilde{C}_{00}^{(m)} \succeq \tilde{C}_{10}^{(m)\dagger} \left(\tilde{C}_{11}^{(m)} \right)^{-1} \tilde{C}_{10}^{(m)}$. When $\tilde{C}_{11}^{(m)}$ is singular, then its inverse should be replaced with pseudo inverse. Conditions (i) and (ii) are a consequence of the linear constraints for the comb $\tilde{C}^{(m)}$, and condition (iii) is a consequence of the positivity constraint and Schur's complement condition. If we choose any $\tilde{C}_{00}^{(m)}, \tilde{C}_{11}^{(m)}$ satisfying (i) and set $\tilde{C}_{01}^{(m)} = \tilde{C}_{10}^{(m)} = 0$, then conditions (ii)-(iii) are also satisfied and $\text{Tr} \left(\beta^{(m)} \tilde{C}_{10}^{(m)} \right) = 0$ which proves that $b^{(m)} \geq 0$. Since for all cases $\tilde{C}_{10}^{(m)\dagger} \in \mathcal{X}_m$ and $\beta_2^{(m)} \in \mathcal{X}_m^\perp$, we have $\text{Tr} \left(\tilde{C}_{10}^{(m)} \beta_2^{(m)} \right) = 0$, so $b^{(m)} = 0$ when $\beta_1^{(m)} = 0$. Let us now assume that $\beta_1^{(m)} \neq 0$, and fix strictly positive definite $\tilde{C}_{00}^{(m)}$ and $\tilde{C}_{11}^{(m)}$ satisfying condition (i), and $\tilde{C}_{01}^{(m)} = \epsilon \beta_1^{(m)}$, $\tilde{C}_{10}^{(m)} = \epsilon \beta_1^{(m)\dagger}$. Then, condition (ii) is satisfied since $\beta_1^{(m)} \in \mathcal{X}_m$, and for small enough $\epsilon > 0$ condition (iii) is also satisfied since $\tilde{C}_{00}^{(m)}$ is strictly positive. Therefore, there exists a comb $\tilde{C}^{(m)}$ for which $\text{Re} \left(\text{Tr} \left(\beta^{(m)} \tilde{C}_{10}^{(m)} \right) \right) = \epsilon \text{Re} \left(\text{Tr} \left(\beta_1^{(m)} \beta_1^{(m)\dagger} \right) \right) > 0$, so $b^{(m)} > 0$, which finishes the proof of the second part of the lemma. ■

According to Lemma 3 the condition $b^{(m)} = 0$ can be replaced with condition $\beta_1^{(m)} = 0$, or, equivalently, with the condition $\beta^{(m)} \in \mathcal{X}_m^\perp$.

5. Formulations as SDPs

To formulate the bounds as SDPs, we will use the following Lemma.

Lemma 4 Let us consider the following primal SDP maximization problem:

$$\begin{aligned} & \max_C \text{Tr}(AC) \\ & \text{s.t.} \quad C \in \text{Comb}[(\mathcal{K}_1, \mathcal{K}_2), \dots, (\mathcal{K}_{2N-1}, \mathcal{K}_{2N})] \end{aligned}$$

A is a hermitian matrix, here $\text{Comb}[\dots]$ is the set of CJ operators of combs. Then the dual problem is

$$\begin{aligned} & \min_{Q^{(1)}, Q^{(2)}, \dots, Q^{(N)}} \text{Tr} \left(Q^{(1)} \right) \\ & \text{s.t.} \quad Q^{(N)} \otimes \mathbb{1}_{2N} \succeq A, \\ & \quad \text{Tr}_{2k-1} Q^{(k)} = Q^{(k-1)} \otimes \mathbb{1}_{2k-2} \quad \text{for } k \in \{2, 3, \dots, N\} \end{aligned} \quad (\text{D12})$$

where $Q^{(k)}$ are hermitian matrices, $Q^{(k)} \in \mathcal{L}(\mathcal{K}_1 \otimes \mathcal{K}_2 \otimes \dots \otimes \mathcal{K}_{2k-1})$. The optimal value of the dual problem is equal to optimal value of the primal problem (strong duality).

Proof A very similar statement was proven in the Supplementary Material of Ref. [46], see Section I.B, Lemmas 5 and 6. There, A was assumed to be equal to performance operator, but this assumption was not used, and the proof remains correct for any hermitian A . Moreover, the proof provided in Ref. [46] can be only directly applied to cases when $\mathcal{K}_1$ is a trivial one-dimensional space $\emptyset$. To generalize the proof for a general case, it is enough to apply it for a comb with an additional artificial empty teeth, so we consider $C \in \text{Comb}[(\emptyset, \emptyset), (\mathcal{K}_1, \mathcal{K}_2), \dots, (\mathcal{K}_{2N-1}, \mathcal{K}_{2N})]$ instead of $C \in \text{Comb}[(\mathcal{K}_1, \mathcal{K}_2), \dots, (\mathcal{K}_{2N-1}, \mathcal{K}_{2N})]$. Then, we end up with a dual problem in the form (D12). ■

The iterative bound (D2) can be formulated as an SDP using Algorithm 1 from Ref. [46] directly applied for parameter-dependent comb $\mathbb{A}_\theta^{(l,m)}$. Let us use the decomposition (D3), (D4), and define $|\dot{c}_{k,j}^{(l,m)}(h)\rangle = \langle j | \dot{\tilde{L}}_{k,\theta}^{(l,m)} \rangle$, where $|j\rangle$ is some orthonormal basis of $\mathcal{H}_{2(l+m)} \otimes \mathcal{R}_{l+m}$ (notice that $|\dot{c}_{k,j}^{(l,m)}(h)\rangle$ depend linearly on the mixing hermitian matrix h). The recursive step (D2) can be calculated using the following SDP:

$$\begin{aligned} F^{(l+m)} &= 4 \min_{h, \{Q^{(k)}\}_{k \in \{1, \dots, m\}}} \text{Tr} \left(Q^{(1)} \right), \\ &\text{subject to} \\ &A \succeq 0, \\ &\forall_{2 \leq k \leq m-1} \text{Tr}_{\mathcal{H}_{2(l+k)}} Q^{(k+1)} = Q^{(k)} \otimes \mathbb{1}_{\mathcal{H}_{2(l+k)-1}}, \\ &\text{Tr}_{\mathcal{H}_{2l+2}} Q^{(2)} = Q^{(1)} \otimes \mathbb{1}_{\mathcal{H}_{2l+1} \otimes \mathcal{R}_l}, \end{aligned}$$

where

$$A = \left(\begin{array}{c|ccc} Q^{(m)} \otimes \mathbb{1}_{\mathcal{H}_{2(l+m)-1}} & |\dot{c}_{1,1}^{(l,m)}(h)\rangle & \dots & |\dot{c}_{r,d}^{(l,m)}(h)\rangle \\ \hline \langle \dot{c}_{1,1}^{(l,m)}(h) | & & & \\ \vdots & & & \\ \langle \dot{c}_{r,d}^{(l,m)}(h) | & & \mathbb{1}_{dr} & \end{array} \right),$$

d is the dimension of $\mathcal{H}_{2(l+m)} \otimes \mathcal{R}_{l+m}$, r is the rank of $\mathbb{A}_\theta^{(l,m)}$, $Q^{(1)} \in \mathcal{L}(\mathcal{V}_l)$, notice that $|\dot{c}_{k,j}^{(l,m)}(h)\rangle$ depend on $F^{(l)}$.

The asymptotic bound (D10) can be written as a similar SDP with additional constraints for h coming from condition $b^{(m)} = 0$. First, let us notice that the maximization over $\tilde{C}^{(m)}$ in $a^{(m)}$ in (D10) boils down to a maximization over $\tilde{C}_{00}^{(m)} \in \text{Comb}[(\emptyset, \mathcal{H}_1 \otimes \mathcal{R}_0), (\mathcal{H}_2, \mathcal{H}_3), \dots, (\mathcal{H}_{2m-2}, \mathcal{H}_{2m-1})]$. Secondly, the condition $b^{(m)} = 0$ can be written as $\beta^{(m)} \in \mathcal{X}_m^\perp$ (see Lemma 3 and the remark below). Since the dual affine space to comb space is another comb space (with outputs and inputs interchanged) [60, 88], it can be shown that $\beta^{(m)} \in \mathcal{X}_m^\perp$ iff there exists a sequence of operators $Y^{(1)}, Y^{(2)}, \dots, Y^{(m-1)}$ for which $\beta^{(m)} = Y^{(m-1)} \otimes \mathbb{1}_{\mathcal{H}_{2m-1}}$, $\forall_{2 \leq k \leq m-1} \text{Tr}_{\mathcal{H}_{2k}} Y^{(k)} = Y^{(k-1)} \otimes \mathbb{1}_{\mathcal{H}_{2k-1}}$, $\text{Tr}_{\mathcal{H}_2} Y^{(1)} = Y^{(0)} \mathbb{1}_{\mathcal{H}_1 \otimes \mathcal{R}_0}$, $Y^{(0)} \in \mathbb{C}$. By introducing the notation $|c_{k,j}^{(m)}(h)\rangle = \langle j | K_{k,\theta}^{(m)} \rangle$, $|\dot{c}_{k,j}^{(m)}(h)\rangle = \langle j | \dot{\tilde{K}}_{k,\theta}^{(m)} \rangle$, we get

$$\beta^{(m)}(h) = \sum_{k,j} |\dot{c}_{k,j}^{(m)}(h)\rangle \langle c_{k,j}^{(m)}|, \quad (\text{D13})$$

notice that $\beta^{(m)}(h)$ depends linearly on h . Finally, after supplementing the SDP from Ref. [46] with the condition

for $\beta^{(m)}$, we get the following SDP for the asymptotic bound:

$$\begin{aligned}
\lim_{N \rightarrow \infty} \mathcal{F}_{\text{AD}}^{(N)} / N &\leq 4/m \min_{h, \{Q^{(k)}, Y^{(k)}\}_{k \in \{0, 1, \dots, m-1\}}} Q^{(0)}, \\
&\text{subject to} \\
A &\succeq 0, \\
\forall_{2 \leq k \leq m-1} \text{Tr}_{\mathcal{H}_{2k}} Q^{(k)} &= Q^{(k-1)} \otimes \mathbb{1}_{\mathcal{H}_{2k-1}}, \\
\text{Tr}_{\mathcal{H}_2} Q^{(1)} &= Q^{(0)} \mathbb{1}_{\mathcal{H}_1 \otimes \mathcal{R}_0}, \quad Q^{(0)} \in \mathbb{R}, \\
\beta^{(m)}(h) &= Y^{(m-1)} \otimes \mathbb{1}_{\mathcal{H}_{2m-1}}, \\
\forall_{2 \leq k \leq m-1} \text{Tr}_{\mathcal{H}_{2k}} Y^{(k)} &= Y^{(k-1)} \otimes \mathbb{1}_{\mathcal{H}_{2k-1}}, \\
\text{Tr}_{\mathcal{H}_2} Y^{(1)} &= Y^{(0)} \mathbb{1}_{\mathcal{H}_1 \otimes \mathcal{R}_0}, \quad Y^{(0)} \in \mathbb{C},
\end{aligned}$$

where

$$A = \left(\begin{array}{c|ccc} Q^{(m-1)} \otimes \mathbb{1}_{\mathcal{H}_{2m-1}} & |\dot{c}_{1,1}^{(m)}(h)\rangle & \dots & |\dot{c}_{r,d}^{(m)}(h)\rangle \\ \hline \langle \dot{c}_{1,1}^{(m)}(h) | & & & \\ \vdots & & & \\ \langle \dot{c}_{r,d}^{(m)}(h) | & & \mathbb{1}_{dr} & \end{array} \right),$$

d is dimension of $\mathcal{H}_{2m} \otimes \mathcal{R}_m$, r is the rank of $\Lambda_{\theta}^{\leftrightarrow(m)}$.

The asymptotic bound in the presence of HS (D11) can be written as

$$\lim_{N \rightarrow \infty} \mathcal{F}_{\text{AD}}^{(N)} / N^2 \leq \left[2/m \min_h \max_{\tilde{C}^{(m)} \in \text{Comb}[(\mathcal{V}_0, \mathcal{H}_1 \otimes \mathcal{R}_0), (\mathcal{H}_2, \mathcal{H}_3), \dots, (\mathcal{H}_{2m-2}, \mathcal{H}_{2m-1})]} \text{Tr} \left(\tilde{C}^{(m)} \begin{bmatrix} 0 & \frac{1}{2} \beta^{(m)} \\ \frac{1}{2} \beta^{(m)\dagger} & 0 \end{bmatrix} \right) \right]^2,$$

where we used the block decomposition $\tilde{C}^{(m)} = \begin{bmatrix} \tilde{C}_{00}^{(m)} & \tilde{C}_{01}^{(m)} \\ \tilde{C}_{10}^{(m)} & \tilde{C}_{11}^{(m)} \end{bmatrix}$ where $\tilde{C}_{01}^{(m)\dagger} = \tilde{C}_{10}^{(m)}$. After dualizing the maximization problem using Lemma 4, we can write the whole min max problem as a single SDP minimization:

$$\begin{aligned}
\sqrt{\lim_{N \rightarrow \infty} \mathcal{F}_{\text{AD}}^{(N)} / N^2} &\leq 2/m \min_{h, \{Q^{(k)}\}_{k \in \{1, \dots, m\}}} \text{Tr}(Q^{(1)}), \\
&\text{subject to} \\
Q^{(m)} \otimes \mathbb{1}_{\mathcal{H}_{2m-1}} &\succeq \begin{bmatrix} 0 & \frac{1}{2} \beta^{(m)} \\ \frac{1}{2} \beta^{(m)\dagger} & 0 \end{bmatrix}, \\
\forall_{2 \leq k \leq m-1} \text{Tr}_{\mathcal{H}_{2k}} Q^{(k+1)} &= Q^{(k)} \otimes \mathbb{1}_{\mathcal{H}_{2k-1}}, \\
\text{Tr}_{\mathcal{H}_2} Q^{(2)} &= Q^{(1)} \otimes \mathbb{1}_{\mathcal{R}_0 \otimes \mathcal{H}_1},
\end{aligned}$$

where $Q^{(1)} \in \mathcal{L}(\mathcal{V}_0)$, $\beta^{(m)}$ is given by (D13).

Appendix E: Examples

The code we used to generate upper bounds introduced in this work and tensor-network based lower bounds introduced in [65, 66] is available in a public repository [67]. To generate upper bounds, we used the function `ad_asym_bound_correlated` from file `qmetro/bounds.py`. To generate lower bounds for correlated dephasing examples, we used tensor-network based optimization over adaptive strategies with limited ancilla dimension d_A , this was done using the function `iss_tnet_adaptive_qfi` from the file `qmetro/protocols/iss.py`. We calculated lower bounds for the collisional thermometry example by considering a parallel scheme, when input state and output measurement are represented using tensor networks—the optimization was performed using the function `iss_tnet_parallel_qfi` from the file `qmetro/protocols/iss.py`.

1. Correlated dephasing

All the details related to the correlated dephasing model are described in Ref. [65] (note that here we use notation S instead of T for a mixing map). The only additional difficulty is that the tightness of an upper bound depends on the way in which we cut the chain of correlated channels into pieces. As we checked numerically, for parallel dephasing the tightest bound is obtained when we decompose mixing map as $S = \sqrt{S} \circ \sqrt{S}$, and make cuts between two maps $\sqrt{S}$, see Fig. 2(a) in the main text. The Kraus operators of channel S implementing a stochastic map $S_{i|i-1}$ on the basis $|\pm\rangle$ are $S_{sr} = \sqrt{\frac{1+srC}{2}} |s\rangle\langle r|$ for $s, r \in \{+, -\}$. It can be shown by direct calculations, that the square root map for $C > 0$ is

$$\sqrt{S}_{i|i-1}(r_i|r_{i-1}) = (1 + r_i r_{i-1} \sqrt{|C|})/2,$$

the corresponding Kraus operators of a quantum map are $\sqrt{S}_{sr} = \sqrt{\frac{1+sr\sqrt{|C|}}{2}} |s\rangle\langle r|$ for $s, r \in \{+, -\}$. When $C < 0$, we should construct a map like for positive C of the same absolute value, but then additionally add a flip between elements of basis $|+\rangle, |-\rangle$ between two maps $\sqrt{S}$.

2. Collisional thermometry

The collisional quantum thermometry model was introduced in Ref. [71]. The thermal bath of temperature T is probed by a qubit thermometer whose hamiltonian is $H = \hbar\Omega\sigma_z/2$. The thermalization of a thermometer is described by a quantum master equation in Lindblad form

$$\frac{d\rho}{dt} = \mathcal{L}_T(\rho) = \gamma(\bar{n} + 1)\mathcal{D}(\sigma_-) + \gamma\bar{n}\mathcal{D}(\sigma_+),$$

where $\mathcal{D}(X) = X\rho X^\dagger - \frac{1}{2}\{X^\dagger X, \rho\}$, $\sigma_\pm = (\sigma_x \pm i\sigma_y)$, $\bar{n} = (e^{\hbar\Omega/k_B T} - 1)^{-1}$, γ is a coupling constant between the thermometer and the thermal bath. The channel describing the thermalization for a time τ is given by integrating the above equation

$$S_T^\tau(\rho) = e^{\mathcal{L}_T \tau}(\rho),$$

which is a generalized amplitude damping channel. When $\tau \rightarrow \infty$, then the output of a channel S_T^τ is a Gibbs state $\rho_{\text{th},T} = e^{-H_\varepsilon/k_B T}/Z$, the QFI of this state is a thermal Fisher information $F_{\text{th},T} = F(\rho_{\text{th},T})$. Interestingly, the channel QFI of S_T^τ can be larger than $F_{\text{th},T}$. We numerically study a case when $k_B T/\hbar\Omega = 2$, which means that $\bar{n} = 1.514$ (the same value of $\bar{n}$ was chosen in Ref. [71]). We observed that the maximal channel QFI is achieved for $\tau_{\text{opt}} \approx 0.417\gamma^{-1}$, then $\mathcal{F}(S_T^\tau) \approx 5.66F_{\text{th},T}$, see Fig. 7(a). Notably, one needs to entangle the input state with ancilla to achieve this optimal channel QFI.

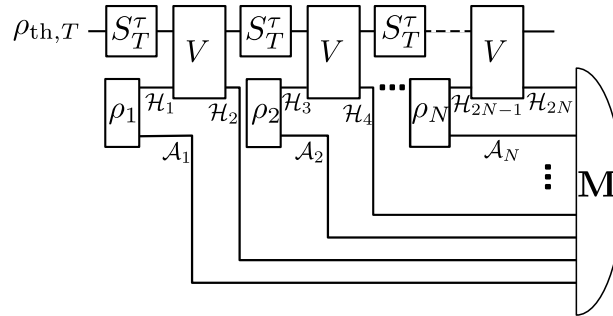


FIG. 6. Collisional quantum thermometry: thermalizing channels S_T^τ acting on thermometer are intertwined with unitary channels V describing energy exchange between mediators ($\mathcal{H}$) and thermometer. At the end, the joint measurement $\mathbf{M}$ is performed on all mediators and ancillary qubits (initially entangled with mediators). Our upper bounds cover strategies with all possible entangled states of mediators and ancillas. We also calculate lower bounds for the strategy in which initially mediators are not entangled with each other, but each mediator ($\mathcal{H}_{2i-1}$) can be entangled with the corresponding ancilla ($\mathcal{A}_i$).

In a collisional scheme, there is no direct access to a thermometer, one only has access to mediator qubits that exchange energy with thermometer after each thermalization process, the interaction between thermometer and mediator is governed by the unitary $V = e^{-igt(\sigma_+ \otimes \sigma_- + \sigma_- \otimes \sigma_+)}$, g is a coupling constant, t is an interaction time; a full

energy swap occurs for $gt = 0.5\pi$, in this work we numerically analyze the case $gt = 0.35\pi$. In Fig. 6, we present a collisional thermometry scheme, in which all mediators are prepared in a product state, but each mediator is assisted with a qubit ancilla, which can be entangled with it. Mediators interact with thermometer, and at the end, all mediators and ancillary qubits are measured. Note, that the thermometer is initialized in a thermal state, whereas in Ref. [71] the initial state of a thermometer is a stationary state of concatenation of maps V and S_T^τ . This stationary state depends on the input mediator of V , so we cannot fix a stationary state when computing universal bounds. However, the choice of initial state of thermometer does not affect the asymptotic results.

Importantly, our upper bounds cover a much broader class of strategies, in which mediators can be prepared in an arbitrary entangled state, and adaptive control between subsequent probeings can be applied.

To complement these general upper bounds, we calculate lower bounds using tensor network techniques [66, 67], assuming that the input state of mediators is a product state, each mediator is entangled with an ancillary qubit, and that the final state of mediators and ancillas is measured using a collective measurement $\mathbf{M}$, which is described by a matrix product operator with bound dimension 4 (which means that some non-locality in a final measurement is allowed). As demonstrated in Fig. 2(c) in the main text, this strategy is optimal (up to numerical precision) for $\tau = 10\tau_{\text{opt}}$, when correlations between subsequent channels are negligible. For $\tau = \tau_{\text{opt}}$, there is some potential advantage from using entangled states—however, this potential advantage is relatively small.

To make the bounds as tight as possible, we need to judiciously cut the whole chain of channels V and S_T^τ ($\Lambda_\theta^{(N)}$) into smaller pieces $\Lambda_\theta^{\leftrightarrow(m)}$. More precisely, we need to find combs $\Lambda_\theta^{\leftrightarrow(m)}$ such that their link product is $\Lambda_\theta^{(N)}$ (environmental spaces of neighbouring combs are linked). Therefore, we can concatenate $\Lambda_\theta^{\leftrightarrow(m)}$ with some map X acting on its output environment and the inverse of this map X^{-1} acting on an input environment—then, after linking $\Lambda_\theta^{\leftrightarrow(m)}$ with each other, each X will be concatenated with X^{-1} , so we will obtain the same resulting comb $\Lambda_\theta^{(N)}$. However, the bound tightness might be affected by the choice of X . Importantly, X must be chosen such that the resulting $\Lambda_\theta^{\leftrightarrow(m)}$ are still valid combs (in particular, they must be completely positive).

In our example we choose $X = S_{T_0}^{\tau_0}$, which is a map that does not differ from $S_T^{\tau_0}$ at the value of temperature around which the estimation is performed, but $\frac{d}{dT}S_{T_0}^{\tau_0} = 0$, which means that $S_{T_0}^{\tau_0}$ is insensitive to the changes of T , so its channel QFI is 0. The intuition behind this choice is that such T -insensitive thermalization may hide some information that leaks from environment when we cut the whole chain of channels into pieces.

Obviously, the inverse map $S_{T_0}^{-\tau_0}$ is not completely positive. However, the total map acting on an input environment is $S_T^\tau \circ S_{T_0}^{-\tau_0}$ is completely positive assuming that $\tau_0 < \tau$ —the resulting channel at $T = T_0$ is a thermalization with time $\tau - \tau_0$, the channel is full-rank, so it will remain completely positive in the first order expansion irrespectively of its derivative.

The large values of τ_0 allow to reduce the role of environmental leakages at the output of $\Lambda_\theta^{\leftrightarrow(m)}$, but at the same time, extra anti-thermalization map $S_{T_0}^{-\tau_0}$ at the beginning may increase the QFI. Therefore, one needs to optimize over τ_0 (assuming $\tau_0 < \tau$) to get the tightest possible bound. Interestingly, for $\tau = \tau_{\text{opt}}$ the optimal value is $\tau_0 = 0$, so it is not advantageous to use the described trick. The situation is completely different for $\tau = 10\tau_{\text{opt}}$, see Fig. 7(b), then choosing $\tau_0 > 0$ can tighten the bound significantly. With the optimal value $\tau_0 \approx 0.62\gamma^{-1}$, we get an upper bound that coincides with lower bounds already for $m = 1$.

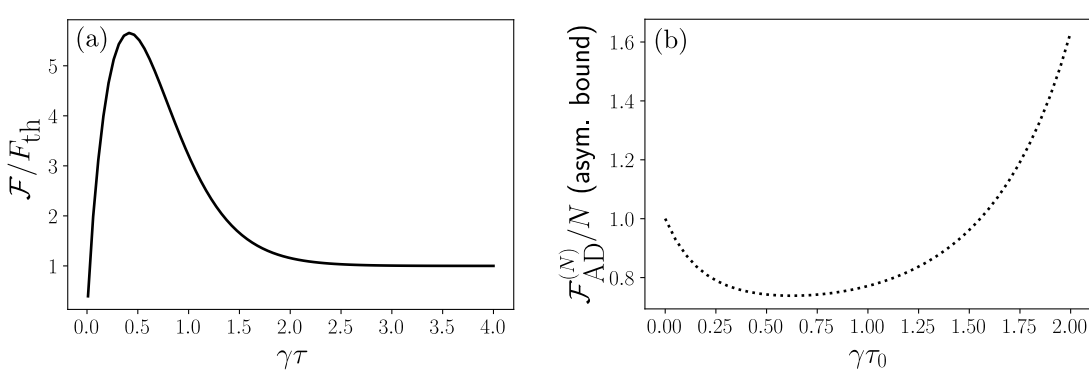


FIG. 7. (a): The ratio between channel QFI of a thermal channel S_T^τ and the QFI of a thermal state as a function of thermalization time multiplied by a coupling constant $\gamma\tau$. (b) The tightness of the bound is affected by the choice of time τ_0 in an inverse map $S_T^{-\tau_0}$. Here we show how an asymptotic bound depends on $\gamma\tau_0$ for $\tau = 10\tau_{\text{opt}}$ $m = 1$.