Optimal function estimation with photonic quantum sensor networks

Jacob Bringewatt,[1,2,*] Adam Ehrenberg [1,2,*] Tarushii Goel,[1,2,3,*] and Alexey V. Gorshkov [1,2]

[1]*Joint Center for Quantum Information and Computer Science, NIST/University of Maryland College Park, Maryland 20742, USA*
[2]*Joint Quantum Institute, NIST/University of Maryland College Park, Maryland 20742, USA*
[3]*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

The problem of optimally measuring an analytic function of unknown local parameters each linearly coupled to a qubit sensor is well understood, with applications ranging from field interpolation to noise characterization. Here we resolve a number of open questions that arise when extending this framework to Mach-Zehnder interferometers and quadrature displacement sensing. In particular, we derive lower bounds on the achievable mean square error in estimating a linear function of either local phase shifts or quadrature displacements. In the case of local phase shifts, these results prove, and somewhat generalize, a conjecture by Proctor *et al.* [arXiv:1702.04271]. For quadrature displacements, we extend proofs of lower bounds to the case of arbitrary linear functions. We provide optimal protocols achieving these bounds up to small (multiplicative) constants and describe an algebraic approach to deriving new optimal protocols, possibly subject to additional constraints. Using this approach, we prove necessary conditions for the amount of entanglement needed for any optimal protocol for both local phase and displacement sensing.

I. INTRODUCTION

In quantum metrology, entangled states of quantum sensors are used to try to obtain a performance advantage in estimating an unknown parameter or parameters (e.g., field amplitudes) coupled to the sensors. In addition to this practical advantage of quantum sensing, the theory of the ultimate performance limits for parameter estimation tasks is deeply related to a number of topics of theoretical interest in quantum information science, such as resource theories [1], the geometry of quantum state space [2], quantum speed limits [3–5], and quantum control theory [4].

Initial experimental and theoretical work on quantum sensing focused on optimizing the estimation of a single unknown parameter (see, e.g., Ref. [6] for a review). More recently, the problem of distributed quantum sensing has become an area of particular interest [7]. Here one considers a network of quantum sensors, each coupled to a local unknown parameter. The prototypical task in this setting is to measure some function or functions of these parameters. In this context, the task of optimally measuring a single linear function $q(\boldsymbol{\theta})$ of d independent local parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)^T$ is particularly well studied both theoretically [8–21] and experimentally [22–25]. In addition to its independent utility (i.e., for measuring an average of local fields in some region), linear

function estimation serves as a key subtask of more general metrological tasks, such as measuring an analytic function of the unknown parameters [26], measuring an analytic function of dependent parameters [27,28], or measuring multiple functions [29,30].

For qubit sensors, the asymptotic limits on performance for these function estimation tasks are rigorously understood, and techniques for generating optimal protocols subject to various constraints, such as limited entanglement between sensors, are known [18]. However, despite extensive theoretical and experimental research on distributed quantum sensing for photonic quantum sensors (see, e.g., [7,31] for reviews), the asymptotic performance limits for function estimation are not yet rigorously established. Here we close this gap, proving an ultimate bound on asymptotic performance, as measured by the mean square error of the estimator, for measuring a linear function of unknown parameters each coupled to a different photonic mode via either (1) the number operator $\hat{n}$ or (2) a field-quadrature operator, chosen without loss of generality to be the momentum quadrature $\hat{p} := i(\hat{a}^\dagger - \hat{a})/2$. That is, we are interested in determining a function of either unknown local phase shifts or unknown quadrature displacements. For case (1), our primary focus, we derive this bound subject to a strict constraint on photon number, proving a long-standing conjecture appearing in Ref. [8]. In case (2), we derive our bound subject to a constraint on the average photon number, which is more natural in this setting as quadrature displacements are not photon-number conserving. Here our results are consistent with existing bounds in the literature [13], but, for completeness, we include derivations in this setting using an equivalent mathematical framework to the number operator case and the qubit sensor case [18]. This allows for a natural comparison of the various performance limits and resource requirements

of function estimation in quantum sensor networks and opens the door to designing new, information-theoretically optimal protocols in the asymptotic limit of sufficient data.

The rest of the paper proceeds as follows. In Sec. II we formally set up the problem of interest and provide useful notation. In Sec. III we prove lower bounds on the mean squared error of an estimator for arbitrary linear functions for both number operator and displacement operator generators. We then study protocols that saturate these bounds in Sec. IV. Finally, we discuss other entanglement-restricted optimal protocols in Sec. V.

II. PROBLEM SETUP

Consider a sensor network of d optical modes each coupled to an unknown parameter θ_j for $j \in \{1, \ldots, d\}$ via

$$\hat{H}(s) = \sum_{j=1}^{d} \theta_j \hat{g}_j + \hat{H}_c(s) =: \boldsymbol{\theta} \cdot \hat{\boldsymbol{g}} + \hat{H}_c(s), \qquad (1)$$

where $\hat{g}_j$ is the local coupling Hamiltonian and boldface denotes vectors. Here we consider the following two cases:

$$\hat{g}_j := \hat{n}_j = \hat{a}_j^\dagger \hat{a}_j, \qquad (2a)$$

$$\hat{g}_j := \hat{p}_j = \frac{i}{2}(\hat{a}_j^\dagger - \hat{a}_j), \qquad (2b)$$

where $\hat{a}_j^\dagger, \hat{a}_j$ are the bosonic creation and annihilation operators acting on mode j, $\hat{n}_j$ is the number operator acting on mode j, and $\hat{p}_j$ is the momentum- ($\hat{p}$-) quadrature on mode j. The choice of $\hat{p}$ quadrature is, of course, arbitrary. All results apply equally well for coupling to any quadrature. The $\boldsymbol{\theta}$-independent, time-dependent Hamiltonian $\hat{H}_c(s)$ is a control Hamiltonian, possibly including coupling to an arbitrary number of ancilla modes. Here $s \in [0, t]$, where t is the total sensing time.

In either case, our task is to measure a linear function $q(\boldsymbol{\theta}) = \boldsymbol{\alpha} \cdot \boldsymbol{\theta}$ of the local field amplitudes $\boldsymbol{\theta}$ where $\boldsymbol{\alpha} \in \mathbb{Q}^d$ is a vector of rational coefficients. (The restriction to rational coefficients is due to the discreteness of the resources—the number of photons—available in this problem; in the case we are interested in—large photon numbers—this is only a technical point.) To accomplish this task, we consider probe states with either fixed photon number N or fixed average photon number $\overline{N}$. Given such probe states, we consider encoding the unknown parameters into the state via the unitary evolution generated by the Hamiltonian in Eq. (1).

We will consider both an unrestricted control Hamiltonian and a control Hamiltonian fixed to have the form

$$\hat{H}_c(s) = \hat{h}_c(s)\delta(s - j\Delta t), \qquad (3)$$

where $\hat{h}_c(s)$ is a (unitless) Hermitian operator, $\delta(s)$ is the Dirac delta function, $\Delta t := t/M$ is the time for a single application of the encoding unitary $\exp(-iH\Delta t)$. The index $j \in \{1, \ldots, M\}$ indexes these applications, where M is the total number of applications. This construction is motivated by the fact that typical physical implementations of a number operator coupling, e.g., in a Mach-Zehnder interferometer, and displacement operator coupling, e.g., via an electro-optical modulator (EOM), often do not allow for intermediate

controls at arbitrary times. Therefore, when we fix our control Hamiltonians to be described by Eq. (3), we have limited any controls to be applied between each pass through these optical elements; for simplicity, we have assumed that these control operations can be implemented on a timescale much shorter than the timescale of phase accumulation. Without loss of generality, we will let $\Delta t = 1$ for the rest of this paper, implying that (in this setting) $t = M$. Therefore, the parameter encoding procedure for the photon number coupling is done via the unitary

$$U = U^{(M)}V U^{(M-1)}V \cdots U^{(1)}V = \prod_{m=1}^{M}(U^{(m)}V), \qquad (4)$$

where $V := \exp(-i\hat{\boldsymbol{g}} \cdot \boldsymbol{\theta})$ and $U^{(m)}$ for $m \in \{1, \ldots, M\}$ denote the unitaries applied between passes. Here, by pass, we mean a single application of the unitary V. We use the convention that the product operation left multiplies.

In both settings, it is worth emphasizing that, while our information-theoretic results lower bounding the asymptotically achievable mean square error of an estimate $\tilde{q}$ of q will apply to any protocol within the framework(s) described above, the explicit protocols we will develop will use finite ancillary modes and finite controls.

III. LOWER BOUNDS

Following the approach of Refs. [10,18], we compute lower bounds on the mean square error $\mathcal{M}$ of an estimator $\tilde{q}$ of q by rewriting the Hamiltonian in Eq. (1) as

$$H(s) = \sum_{j=1}^{d} (\boldsymbol{\alpha}^{(j)} \cdot \boldsymbol{\theta})(\boldsymbol{\beta}^{(j)} \cdot \hat{\boldsymbol{g}}) + \hat{H}_c(s), \qquad (5)$$

for some (time-independent) choice of basis vectors $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{d}$, where $\boldsymbol{\alpha}^{(1)} := \boldsymbol{\alpha}$ and $\{\boldsymbol{\beta}^{(j)}\}_{j=1}^{d}$ is a dual basis such that $\boldsymbol{\alpha}^{(i)} \cdot \boldsymbol{\beta}^{(j)} = \delta_{ij}$. The vectors $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{d}$ are associated with a change of basis $\boldsymbol{\theta} \to \boldsymbol{q}$ where $q_j := \boldsymbol{\alpha}^{(j)} \cdot \boldsymbol{\theta}$ such that $q_1 = q$; that is, $\boldsymbol{\alpha}^{(1)} =: \boldsymbol{\alpha}$ with corresponding dual vector $\boldsymbol{\beta}^{(1)} =: \boldsymbol{\beta}$. Then we can define a $\boldsymbol{\beta}$-parameterized generator of translations with respect to the quantity q as

$$\hat{g}_{q,\boldsymbol{\beta}} := \min_{q^{(2)}, \cdots, q^{(d)}} \frac{\partial \hat{H}}{\partial q} \bigg|_{q^{(2)}, \cdots, q^{(d)}} = \boldsymbol{\beta} \cdot \hat{\boldsymbol{g}}. \qquad (6)$$

Armed with Eq. (6), we can write a bound on $\mathcal{M}$ in terms of a single-parameter quantum Cramér-Rao bound [31–33]

$$\mathcal{M} \geqslant \frac{1}{\mu \mathcal{F}(q|\boldsymbol{\beta})}, \qquad (7)$$

where $\mathcal{F}(q|\boldsymbol{\beta})$ is the quantum Fisher information with respect to q, given some choice of fixing the extra $d - 1$ degrees of freedom in our problem, as specified by the vector $\boldsymbol{\beta} \in \mathbb{R}^d$ such that $\boldsymbol{\alpha} \cdot \boldsymbol{\beta} = 1$. Any such single-parameter bound is a valid lower bound as fixing extra degrees of freedom can only give us more information about the parameter q (see below for mathematical details). μ is the number of experimental repetitions. This bound holds for an unbiased estimator $\tilde{q}$. When deriving our bounds, we will restrict ourselves to single-shot Fisher information and set $\mu = 1$ [34]. Quantum

Fisher information is maximized for pure states, so restricting ourselves to pure states and unitary encoding of the unknown parameters into the state we can write

$$\mathcal{F}(q|\boldsymbol{\beta}) \leqslant 4t^2 \max_{\rho}[(\Delta \hat{g}_{q,\boldsymbol{\beta}})_{\rho}]^2, \tag{8}$$

where $\hat{g}_{q,\boldsymbol{\beta}}$ is the $\boldsymbol{\beta}$-parameterized generator of translations with respect to the unknown function q. The variance $[\Delta(\hat{g}_{q,\boldsymbol{\beta}})_{\rho}]^2$ is taken with respect to a pure probe state $\rho = |\psi\rangle\langle\psi|$.

Ultimately, we seek a choice of new basis that yields the tightest possible bound on the quantum Fisher information $\mathcal{F}(q)$. This choice is determined by the solution to [35]

$$\min_{\boldsymbol{\beta}} \max_{\rho}[\Delta(\boldsymbol{\beta} \cdot \hat{\boldsymbol{g}})_{\rho}]^2, \quad \text{subject to } \boldsymbol{\alpha} \cdot \boldsymbol{\beta} = 1. \tag{9}$$

Let $(\boldsymbol{\beta}^*, \rho^*)$ be a solution for this optimization problem. Then we can rewrite the single-shot version of Eq. (7) as

$$\mathcal{M} \geqslant \frac{1}{4t^2[\Delta(\boldsymbol{\beta}^* \cdot \hat{\boldsymbol{g}})_{\rho^*}]^2}. \tag{10}$$

This bound can be understood as corresponding to the optimal choice of an imaginary single parameter scenario, where we have fixed $d-1$ of the d parameters controlling the evolution of the state, leaving only the parameter of interest q free to vary. While this requires giving ourselves information that we do not have, additional information can only reduce $\mathcal{M}$, and, therefore, any such choice provides a lower bound on $\mathcal{M}$ (via single-parameter bounds) when we do not have such information. While not guaranteed by this method of derivation, we shall see that such bounds are saturable, up to small multiplicative constants.

Constraints can be placed on the probe state ρ depending on the physical generators coupled to the parameters of interest: as previously discussed, in this work we consider the constraints of fixed photon number N for the generator $\hat{n}_j$ and fixed average photon number $\overline{N}$ for the generator $\hat{p}_j$. The rationale behind these constraints is as follows. $\hat{p}$ does not conserve photon number, hence it does not make sense to restrict to a fixed photon number sector when coupling to quadrature operators, and, thus, average photon number is the natural constraint. For $\hat{n}$, on the other hand, we must work in the fixed photon sector, as using fixed average photon number allows for the construction of pathological probe states enabling arbitrarily precise sensing. In particular, consider the state

$$|\psi_a\rangle = \sqrt{\frac{a-1}{a}}|0\rangle + \sqrt{\frac{1}{a}}|a\overline{N}\rangle. \tag{11}$$

It is easy to see that $|\psi_a\rangle$ has mean photon number $\overline{N}$ and variance $(a-1)\overline{N}^2$. Hence, even for fixed $\overline{N}$, letting a get arbitrarily large allows for an arbitrarily large variance, and hence arbitrarily precise sensing.

Leaving the details of the calculation to Appendix A, solving the above optimization problem for $\hat{g}_j = \hat{n}_j$ restricted to probe states with exactly N photons yields

$$\mathcal{M} \geqslant \frac{\max\left\{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}^2, \|\boldsymbol{\alpha}\|_{1,\mathcal{N}}^2\right\}}{N^2 t^2}, \tag{12}$$

where $\mathcal{P} := \{j \,|\, \alpha_j \geqslant 0\}$ and $\mathcal{N} := \{j \,|\, \alpha_j < 0\}$. In the second line, we use the notation

$$\|\boldsymbol{\alpha}\|_{1,\mathcal{S}} := \sum_{i \in \mathcal{S}} |\alpha_i|, \tag{13}$$

where $\mathcal{S} \in \{\mathcal{P}, \mathcal{N}\}$. For the rest of the paper, we assume without loss of generality that we are in the case that $\|\boldsymbol{\alpha}\|_{1,\mathcal{P}} \geqslant \|\boldsymbol{\alpha}\|_{1,\mathcal{N}}$ to simplify our expressions. In the special case where $\boldsymbol{\alpha}$ possesses only positive coefficients (i.e., $\mathcal{N} = \emptyset$),

$$\mathcal{M} \geqslant \frac{\|\boldsymbol{\alpha}\|_1^2}{N^2 t^2}, \tag{14}$$

proving a long-standing conjecture from Ref. [8] that this is the minimum attainable variance for $\boldsymbol{\alpha} \in \mathbb{Q}^d$ with $\boldsymbol{\alpha} \geqslant 0$ and $N\boldsymbol{\alpha} \in \mathbb{N}^d$. This is our primary result.

Similarly, for the case of local quadrature displacements restricted to probe states with average photon number $\overline{N}$, we obtain the following bound:

$$\mathcal{M} \geqslant \frac{\|\boldsymbol{\alpha}\|_2^2}{4\overline{N}t^2} - O\left(\frac{d\|\boldsymbol{\alpha}\|_2^2}{\overline{N}^2 t^2}\right). \tag{15}$$

Equation (15) is a minor generalization of the results in Refs. [7,13], extended to allow for negative coefficients and for arbitrary non-Gaussian probe states. Therefore, for completeness, we include a reminder of the arguments from Refs. [7,13] along with our more general derivation in Appendix B.

We can compare the bounds in Eqs. (12) and (15) to the corresponding bounds on the mean square error obtainable by separable protocols—that is, those using separable probe states such that each parameter θ_i is measured individually using an optimized partition of the available photons, and then these estimates are used to compute q. In particular, for number operator coupling and fixed photon number states, using $\eta_j = \frac{|\alpha_j'|}{\|\boldsymbol{\alpha}'\|_1} N$ photons ($\alpha_j' := \alpha_j^{2/3}$) in mode j, it holds that [8]

$$\mathcal{M}_{\text{sep}} \geqslant \frac{\|\boldsymbol{\alpha}'\|_{2/3}^2}{N^2 t^2}, \tag{16}$$

where $\|\cdot\|_{2/3}$ denotes the Schatten p function

$$\|\boldsymbol{v}\|_p = \left(\sum_i v_i^p\right)^{1/p} \tag{17}$$

with $p = 2/3$. When $p \in [1, \infty]$, this function is a norm, but for $p \in (0, 1)$ it is not, as it does not satisfy the property of absolute homogeneity, but it still provides a convenient notational shorthand.

Performing a similar optimization for the case of displacement coupling and fixed average photon number, one obtains

$$\mathcal{M}_{\text{sep}} \geqslant \frac{\|\boldsymbol{\alpha}\|_1^2}{4\overline{N}t^2} + O\left(\frac{1}{\overline{N}^2 t^2}\right), \tag{18}$$

where the optimum division of photons is given by using $\eta_j = \frac{|\alpha_j|}{\|\boldsymbol{\alpha}\|_1} N$ photons in mode j. A non-closed-form version of this bound can be found in Ref. [12] in the case where $\overline{N}$ is finite. One recovers our result in the asymptotic in $\overline{N}$ limit.

Consequently, in both the phase and displacement sensing settings, the achievable advantage due to entanglement between modes is fully characterized by the difference between the vector p norm of $\boldsymbol{\alpha}$ with $p = \frac{2}{3}, 1$ or $p = 1, 2$, respectively. By generalized Hölder's inequality, $\|\boldsymbol{\alpha}\|_{2/3}^2 \leqslant d\|\boldsymbol{\alpha}\|_1^2$ and $\|\boldsymbol{\alpha}\|_1^2 \leqslant d\|\boldsymbol{\alpha}\|_2^2$. Both inequalities are saturated for any "average-like" function with $|\boldsymbol{\alpha}| \propto (1, 1, \ldots, 1)^T$. In both cases, we obtain a $O(1/d)$ improvement in precision due to entanglement, consistent with the so-called Heisenberg scaling in the number of sensors d. This is consistent with results for qubits in Ref. [10], where the best improvement between the separable and entangled bounds occurs when measuring an average-like function. For the case of phase sensing, the optimal performance, including constants, is obtained when $\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}^2 = \|\boldsymbol{\alpha}\|_{1,\mathcal{N}}^2 = \|\boldsymbol{\alpha}\|_1/2$ (which occurs when the vector $\boldsymbol{\alpha}$ is half positive ones and half negative ones).

IV. PROTOCOLS

A. Existing protocols

The bounds established in the previous section are all saturable, up to small multiplicative constants, using protocols that exist in the literature, or slight variations thereof. In particular, Refs. [8,11] present a protocol for estimating a linear function of local phase shifts with positive coefficients (i.e., $\boldsymbol{\alpha} \geqslant 0$) which achieves the bound in Eq. (12) up to a small multiplicative constant. This protocol makes use of a so-called proportionally weighted N00N state over $d + 1$ modes,

$$|\psi\rangle \propto \left| N \frac{\alpha_1}{\|\boldsymbol{\alpha}\|_1}, \ldots, N \frac{\alpha_d}{\|\boldsymbol{\alpha}\|_1}, 0 \right\rangle + |0, \ldots, 0, N\rangle, \quad (19)$$

where we have expressed the state in an occupation number basis over $d + 1$ modes and have dropped the normalization for concision. The last mode serves as a reference mode. Observe that, for this state to be well defined, it is essential that $\boldsymbol{\alpha}/\|\boldsymbol{\alpha}\|_1 \in \mathbb{Q}^d$ and that N is sufficiently large that the resulting occupation numbers are integers. Details of how protocols using this probe state work and how they generalize to the case of negative coefficients are provided in Appendix D. A description of how to achieve the separable bound in Eq. (16) is provided in Appendix B.

Similarly, in the case of measuring a linear function of displacements using states with fixed average photon number, Ref. [12] provides a protocol that, up to small multiplicative constants, saturates the bound in Eq. (15), and a separable protocol that, again up to small constants, achieves the bound in Eq. (18). Interestingly, these protocols require only Gaussian probe states, indicating that these states are optimal. In particular, these protocols make use of an initial single-mode squeezed state, followed by a properly constructed beam-splitter array to prepare a multimode entangled probe state with the appropriate sensitivity to quadrature displacements in each mode. Homodyne measurements on each mode can then be used to extract the function of interest. Consistent with this fact, our separable lower bound matches the Gaussian state-restricted bound obtained in Ref. [12] and the bound for arbitrary states derived in Ref. [13] for the particular case of measuring an average.

B. Algebraic conditions for new protocols

Other protocols are possible and can be derived via a simple set of algebraic conditions. In particular, for a probe state to exist saturating the bound in Eq. (10), or its specific versions in Eqs. (12) and (15), we require the existence of an optimal choice of basis transformation $\boldsymbol{\theta} \to \boldsymbol{q}$ such that knowing q_j for $j > 1$ yields no information about $q = q_1$. Mathematically, this means that the quantum Fisher information matrix [36] with respect to the parameters $\boldsymbol{q}$ must have the following properties:

$$\mathcal{F}(\boldsymbol{q})_{11} = 4t^2[\Delta(\boldsymbol{\beta}^* \cdot \hat{\boldsymbol{g}})_{\rho^*}]^2, \quad (20a)$$

$$\mathcal{F}(\boldsymbol{q})_{1i} = \mathcal{F}(\boldsymbol{q})_{i1} = 0 \quad (\forall i \neq 1), \quad (20b)$$

Recall that $(\boldsymbol{\beta}^*, \rho^*)$ are the solution to the minimax problem in Eq. (9). We can reexpress these conditions in terms of the quantum Fisher information matrix with respect to $\boldsymbol{\theta}$ as

$$(\boldsymbol{\beta}^*)^T \mathcal{F}(\boldsymbol{\theta}) \boldsymbol{\beta}^* = 4t^2[\Delta(\boldsymbol{\beta}^* \cdot \hat{\boldsymbol{g}})_{\rho^*}]^2, \quad (21a)$$

$$(\boldsymbol{\beta}^*)^T \mathcal{F}(\boldsymbol{\theta}) \boldsymbol{\beta}^{(i)} = (\boldsymbol{\beta}^{(i)})^T \mathcal{F}(\boldsymbol{\theta}) \boldsymbol{\beta}^* = 0 \quad (\forall i \neq 1). \quad (21b)$$

Then, using $\boldsymbol{\alpha}^{(i)} \cdot \boldsymbol{\beta}^{(j)} = \delta_{ij}$, we obtain the condition

$$\mathcal{F}(\boldsymbol{\theta}) \boldsymbol{\beta}^* = 4t^2[\Delta(\boldsymbol{\beta}^* \cdot \hat{\boldsymbol{g}})_{\rho^*}]^2 \boldsymbol{\alpha}. \quad (22)$$

Matrix elements of $\mathcal{F}(\boldsymbol{\theta})$ for pure probe states and unitary evolution are given via

$$\mathcal{F}(\boldsymbol{\theta})_{ij} = 4\left[\frac{1}{2}\langle\{\mathcal{H}_i, \mathcal{H}_j\}\rangle - \langle\mathcal{H}_i\rangle\langle\mathcal{H}_j\rangle\right], \quad (23)$$

where $\mathcal{H}_i = -iU^\dagger \partial_i U$ with $\partial_i := \partial/\partial\theta_i$, U is the unitary generated by Eq. (1) and the expectation values are taken with respect to the initial probe state [36].

We refer to protocols that make use of probe states and controls so that Eq. (22) is satisfied as optimal. However, we caution that the existence of an optimal probe state does not imply the existence of measurements on this state that allow one to extract an estimate of the parameter q saturating the lower bounds we have derived. This issue of the optimal measurements to extract parameters is also discussed extensively in, e.g., Ref. [37], with some convenient, nearly optimal, protocols presented in Refs. [38–40]. Such methods are the origin of the "small multiplicative constants" that arise in the explicit protocols above. In fact, lower bounds derived via the quantum Cramér-Rao bound can be obtained only up to a constant $\geqslant \pi^2$ [41]. See Appendix G for a brief explanation of these ideas.

For the particular cases considered in this paper, $\boldsymbol{\beta}^*$ has been explicitly calculated (see Appendices A and B), so Eq. (22) can be expressed in a more meaningful form. For number operator coupling, we obtain the condition

$$\sum_{i \in \mathcal{P}} \mathcal{F}(\boldsymbol{\theta})_{ij} = \frac{N^2 t^2}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}} \alpha_j, \quad (24)$$

for all j. Similarly, for the quadrature coupling, an optimal protocol requires

$$\mathcal{F}(\boldsymbol{\theta}) \boldsymbol{\alpha} \sim 4\overline{N} t^2 \boldsymbol{\alpha}, \quad (25)$$

where $\sim$ denotes asymptotically in $\overline{N}$. Equations (24) and (25) provide a generic route to finding new protocols: consider a set of parameterized families of probe states $\mathcal{T}$ that one can coherently switch between using available controls $\hat{H}_c(t)$

(here a "family" of states refers to a particular superposition of Fock states with an arbitrary relative phase). One can then calculate $\mathcal{F}(\boldsymbol{\theta})$ via Eq. (23) and allocate the time spent in a particular family of states such that the associated quantum Fisher information condition is achieved. As a limiting case, one could consider $|\mathcal{T}| = 1$, removing the necessity of coherent control; the protocols considered in the previous section are of this sort (and, in Appendix D, we show that these protocols do, indeed, achieve the saturability conditions).

The possible choices for families of states $\mathcal{T}$ that allow for such a solution are actually quite limited, even given access to arbitrary control Hamiltonians and ancilla modes. In particular, we prove the following in the case where $\hat{g}_j := \hat{n}_j$:

Lemma 1. Any optimal protocol using N photons and M passes through interferometers with a coupling as in Eq. (1) with $\hat{g}_j = \hat{n}_j$ requires that, for every pass m, the probe state $|\psi_m\rangle$ be of the form

$$|\psi_m\rangle \propto |\boldsymbol{N}(m)\rangle_{\mathcal{P}}|\boldsymbol{0}\rangle_{\mathcal{N}\mathcal{R}} + e^{i\varphi_m}|\boldsymbol{0}\rangle_{\mathcal{P}}|\boldsymbol{N}'(m)\rangle_{\mathcal{N}\mathcal{R}}, \quad (26)$$

where $\mathcal{P}$, $\mathcal{N}$, and $\mathcal{R}$ represent the modes with $\alpha_j \geqslant 0, \alpha_j < 0$, and the (arbitrary number of) reference modes, respectively, $\boldsymbol{N}(m)$ and $\boldsymbol{N}'(m)$ are strings of occupation numbers such that $|\boldsymbol{N}(m)| = |\boldsymbol{N}'(m)| = N$ for all passes m. φ_m is an arbitrary phase.

The proof follows straightforwardly from an explicit calculation of the Fisher information matrix for $\hat{g}_j = \hat{n}_j$, but is somewhat algebraically tedious so we relegate it to Appendix E.

Lemma 1 suggests a particular choice of $\mathcal{T}$ from which we can pick an optimal protocol for function estimation in the $\hat{g}_j = \hat{n}_j$ case. In particular, define a set of vectors

$$\mathcal{W} := \left\{ \boldsymbol{\omega} \in \mathbb{Z}^d \mid \|\boldsymbol{\omega}\|_{1,\mathcal{P}} = N, \ \|\boldsymbol{\omega}\|_{1,\mathcal{N}} \leqslant N, \ \omega_j \alpha_j \geqslant 0 \ \forall j \right\}. \quad (27)$$

Further, consider the restriction $\boldsymbol{\omega}|_{\mathcal{P}} \in \mathbb{Z}^d$ with components

$$(\boldsymbol{\omega}|_{\mathcal{P}})_j = \begin{cases} \omega_j, & j \in \mathcal{P} \\ 0, & \text{otherwise,} \end{cases} \quad (28)$$

and the restriction $\boldsymbol{\omega}|_{\mathcal{N}}$, defined similarly. Armed with these vectors, we can define a particular choice $\mathcal{T}$ of one-parameter families of probe states in an occupation number basis where each $|\psi(\boldsymbol{\omega}; \varphi)\rangle \in \mathcal{T}$ is labeled by a particular choice of $\boldsymbol{\omega}$ such that

$$|\psi(\boldsymbol{\omega}; \varphi)\rangle \propto |\boldsymbol{\omega}|_{\mathcal{P}}\rangle|0\rangle + e^{i\varphi}|-\boldsymbol{\omega}|_{\mathcal{N}}\rangle|N - \|\boldsymbol{\omega}|_{\mathcal{N}}\|_1\rangle, \quad (29)$$

where $\varphi \in \mathbb{R}$ is an arbitrary parameter and the last mode is a reference mode. It should be clear that these families of states are of the form specified by Lemma 1. Furthermore, note that the proportionally weighted N00N state in Eq. (19) is also of this form.

Our protocols proceed as follows: starting in a state $|\psi(\boldsymbol{\omega}; 0)\rangle$, after any given pass through the interferometers we use control unitaries to coherently switch between families of probe states such that the relative phase between the branches is preserved (that is, we change $\boldsymbol{\omega}$, but not φ). The fact that an optimal protocol must coherently map between such states is proven in Lemma 5 in Appendix E. We stay in the family of states $|\psi(\boldsymbol{\omega}_n; \varphi)\rangle$ for a fraction p_n of the passes where $p_n = \frac{r_n}{M}$

for $r_n \in \{0, 1, \ldots, M\}$ such that $\sum_n p_n = 1$. Here n indexes some enumeration of the families of states in $\mathcal{T}$.

The value of the component ω_j in a given probe state determines the contribution of the parameter θ_j coupled to sensor j to the relative phase between the two branches of the probe state during a single pass. In particular, in a single pass with a probe state in the family $|\psi(\boldsymbol{\omega}; \varphi)\rangle$, the relative phase between the two branches of the probe state becomes $\boldsymbol{\omega} \cdot \boldsymbol{\theta} + \varphi$. Assuming an initial probe state with $\varphi = 0$ and summing over all passes we obtain a total relative phase

$$\varphi_{\text{tot}} = M \sum_n p_n (\boldsymbol{\omega}_n \cdot \boldsymbol{\theta}) \quad (30)$$

$$=: (W\boldsymbol{r}) \cdot \boldsymbol{\theta}. \quad (31)$$

In the second line, we implicitly defined W as a matrix whose columns are the vectors $\boldsymbol{\omega}_n \in \mathcal{W}$ and $\boldsymbol{r} := M\boldsymbol{p} \in \mathbb{Z}^{|\mathcal{T}|}$. Explicitly computing the Fisher information matrix for these states demonstrates that the optimality condition in Eq. (24) is satisfied if

$$W\boldsymbol{r} = NM \frac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}}; \quad (32)$$

see Appendix D for details. Consequently, any integer solution $\boldsymbol{r}$ to Eq. (32) such that

$$\|\boldsymbol{r}\|_1 = M,$$
$$\boldsymbol{r} \geqslant 0, \quad (33)$$

yields an optimal protocol. The protocols of Ref. [8], described above and generalized in Appendix D, are a particularly simple case within this class with $M = 1$ and $\boldsymbol{\omega} = \frac{N\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}}$, i.e., we select out only a single column of $\mathcal{W}$.

Solutions to Eqs. (32) and (33) are not guaranteed to exist for all N, M. In particular, we require that

$$NM \frac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}} \in \mathbb{Z}^d. \quad (34)$$

For $\boldsymbol{\alpha} \in \mathbb{Q}$ and sufficiently large N or M this hold true. Setting up the system of equations in Eqs. (32) and (33) that must be solved to pick out explicit protocols requires identifying the set of vectors $\mathcal{W}$ defined in Eq. (27). While computationally straightforward, if expensive, to construct and enumerate this set, the number of states is extremely large, yielding a correspondingly large set of linear Diophantine equations in Eq. (32). Consequently, it is reasonable to place further, experimentally motivated constraints to limit this set of states and pick out advantageous protocols. For instance, one such constraint is to limit the amount of entanglement between modes on any given pass. We consider this case in the following section.

It is also important to note that integer linear programming is NP-hard [42], so finding a particular solution once we add additional constraints is not a computationally easy task. Regardless, in applications one can apply standard (possibly heuristic) algorithms for integer linear programming to seek solutions. If a solution is found, it is known to be optimal. Consequently, proving the existence or lack thereof of a solution with certain additional constraints may be intractable for large problem instances.

Similar arguments to those that go into proving Lemma 1 allow us to show that, for quadrature sensing, the condition in

TABLE I. Comparison of the lower bounds on the mean square error and entanglement requirements for an (asymptotically) optimal protocol obeying the corresponding conditions on the quantum Fisher information for the task of estimating a linear function $q = \boldsymbol{\alpha} \cdot \boldsymbol{\theta}$ with qubit, phase sensing, and displacement-sensing quantum sensor networks.

	Qubit phase sensing	Phase sensing	Displacement sensing
Parameter coupling	$\frac{1}{2}\hat{\sigma}_i^z \theta_i$	$\hat{n}_i \theta_i$	$\frac{i}{2}(\hat{a}_i^\dagger - \hat{a}_i)\theta_i$
Resources	Qubit number, d	Photon number, N	Avg. photon number,
	sensing time, t	sensing time, t	$\overline{N}$ sensing time, t
MSE (separable)	$\geq \frac{\|\boldsymbol{\alpha}\|_2^2}{t^2}$ [10]	$\geq \frac{\|\boldsymbol{\alpha}\|_{2/3}^2}{N^2 t^2}$ [8]	$\geq \frac{\|\boldsymbol{\alpha}\|_1^2}{4\overline{N}t^2}$
MSE (entangled)	$\geq \frac{\|\boldsymbol{\alpha}\|_\infty^2}{t^2}$ [10]	$\geq \frac{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}^2}{N^2 t^2}$	$\geq \frac{\|\boldsymbol{\alpha}\|_2^2}{4\overline{N}t^2}$ [13]
Entanglement needed (discrete controls)	$k \geq \max\{\lceil \frac{\|\boldsymbol{\alpha}\|_1}{\|\boldsymbol{\alpha}\|_\infty} \rceil, \lceil \frac{\|\boldsymbol{\alpha}\|_0}{M} \rceil\}$	$k > \lceil \frac{\|\boldsymbol{\alpha}\|_0}{M} \rceil$	$k \geq \lceil \frac{\|\boldsymbol{\alpha}\|_0}{M} \rceil$
Entanglement needed (arbitrary controls)	$\frac{\|\boldsymbol{\alpha}\|_1}{\|\boldsymbol{\alpha}\|_\infty} \in (k-1, k]$ [18]	$k = 2$	No entanglement
k-partite entanglement protocol always exists?	Yes [18]	No	Yes

Eq. (25) can be reduced to the condition that

$$\mathcal{F}(\boldsymbol{\theta})_{ij} \sim \frac{4\overline{N}t^2}{\|\boldsymbol{\alpha}\|_2^2} \alpha_i \alpha_j, \qquad (35)$$

which is proven in Appendix F. However, there is not a clearly interesting family of states that can be leveraged to achieve this quantum Fisher information, as in the case of number operator coupling or qubit sensors [18]. However, the existing optimal protocols described above do obey this condition asymptotically in average photon number $\overline{N}$.

V. ENTANGLEMENT REQUIREMENTS

The remaining flexibility in the choice of optimal probe states enabled by some control also allows us to impose further experimentally relevant constraints. One reasonable constraint is the amount of intermode entanglement required during the sensing process. This was considered in Ref. [18] for the case of qubit sensors.

The answer to the entanglement question in the current context depends crucially on the sorts of control operations we allow. In the number operator case, with arbitrary time-dependent control, only two-mode entanglement is needed at any given time, as one can simply prepare a N00N state between the reference and one of the sensing modes and coherently switch which sensing mode is entangled with the reference mode such that the time spent entangled with mode j is given by $t_j = |\alpha_j| t / \|\boldsymbol{\alpha}\|_1$. For similar reasons, no entanglement is needed for displacement sensing; here no reference mode is needed and one can simply sequentially apply displacement operators for a time $t_j = |\alpha_j| t / \|\boldsymbol{\alpha}\|_1$ on a single-mode squeezed state, followed by a homodyne measurement. When control operations to change the probe state are allowed only at M discrete time intervals, as described by Eq. (3), the problem becomes more interesting. For number operator coupling, subject to a fixed photon number constraint, any optimal protocol requires at least $(\lceil \|\boldsymbol{\alpha}\|_0/M \rceil + 1)$-mode entanglement. This bound is fairly trivial: it merely states that one must be entangled with each nontrivial mode for at least one pass. For displacement operator coupling, subject to a fixed average photon number constraint, an essentially identical argument allows us to prove that any optimal protocol requires at least $\lceil \|\boldsymbol{\alpha}\|_0/M \rceil$-mode entanglement. The difference of one is because, unlike displacement sensing, phase sensing generally

requires entanglement with a reference mode. In the $M \to \infty$ limit, we recover the continuous control case, so these trivial bounds can be tight. This triviality is in contrast to the qubit case, where results analogous to Lemma 1 lead to significantly tighter constraints on the minimum amount of necessary entanglement for optimal protocols [18]. This discrepancy arises due to the fact that, unlike with photonic resources which must be distributed in a zero-sum way between modes, for qubit sensors one can be maximally sensitive to all coupled parameters simultaneously.

VI. CONCLUSION AND OUTLOOK

We have determined the fundamental achievable performance limits for phase sensing and have extended proofs of lower bounds for displacement sensing beyond just an average to arbitrary functions. In the process, we proved a long-standing conjecture regarding function estimation with number operator coupling [8] and showed that some of the protocols that exist in the literature [8,11,12] are, in fact, optimal in the asymptotic limit. By considering different implementations of a quantum sensor network within a single framework, we reveal the role of entanglement and controls as they relate to the type of coupling and whether the relevant resource is "parallel" (as in qubit sensor networks, where all parameters can simultaneously be measured to maximal precision) or "sequential" (as in photonic sensor networks, where the photons must be optimally distributed between modes). Our approach to proving our bounds also enables an algebraic framework for developing further optimal protocols, subject to various constraints. Here we considered the particular case of entanglement-based constraints, enabling comparison to similar work in the case of qubit sensors [18]. These results, and how they fit into the landscape of known results for quantum sensor networks, are summarized in Table I. How other constraints impact the existence of and control requirements for optimal protocols remains an interesting open question deserving of further study.

APPENDIX A: BOUND FOR LOCAL PHASE SHIFTS

In this Appendix, we derive lower bounds for the mean square error of measuring a linear function $q(\boldsymbol{\theta}) = \boldsymbol{\alpha} \cdot \boldsymbol{\theta}$ of local phase shifts, generated via coupling to the number operator $\hat{n}_j$, as specified by the Hamiltonian in Eq. (1) and Eq. (2a).

In particular, we seek to solve the optimization problem in Eq. (9), restated here for convenience:

$$\min_{\boldsymbol{\beta}} \max_{\rho} [\Delta(\boldsymbol{\beta} \cdot \hat{\boldsymbol{g}})_\rho]^2, \quad \text{subject to } \boldsymbol{\alpha} \cdot \boldsymbol{\beta} = 1. \quad \text{(A1)}$$

Here $\hat{\boldsymbol{g}} = \hat{\boldsymbol{n}} = (\hat{n}_1, \hat{n}_2, \ldots, \hat{n}_d)^T$. For fixed particle number N, the Hilbert space on which possible probe states ρ are defined is finite dimensional, and it holds that [33]

$$[\Delta(\boldsymbol{\beta} \cdot \hat{\boldsymbol{n}})_\rho]^2 \leqslant \frac{\|\boldsymbol{\beta} \cdot \hat{\boldsymbol{n}}\|_{s,N}^2}{4}, \quad \text{(A2)}$$

where $\|\boldsymbol{\beta} \cdot \hat{\boldsymbol{n}}\|_{s,N}$ is the Fock-space-restricted seminorm of $\boldsymbol{\beta} \cdot \hat{\boldsymbol{n}}$ (defined as the difference between the maximum and minimum eigenvalues of $\boldsymbol{\beta} \cdot \hat{\boldsymbol{n}}$ restricted to the N-photon subspace). As we want to maximize the quantum Fisher information with respect to the choice of probe state ρ, and because Eq. (A2) is saturable when ρ is an equal superposition of the eigenstates of $\boldsymbol{\beta} \cdot \hat{\boldsymbol{n}}$ with maximum and minimum eigenvalues, we can consider the following optimization problem:

$$\text{minimize (w.r.t. } \boldsymbol{\beta}) \ \|\boldsymbol{\beta} \cdot \hat{\boldsymbol{n}}\|_{s,N},$$
$$\text{subject to } \boldsymbol{\alpha} \cdot \boldsymbol{\beta} = 1. \quad \text{(A3)}$$

To begin, note that the largest eigenvalue of $\boldsymbol{\beta} \cdot \hat{\boldsymbol{n}}$ in the N-particle subspace is given by

$$\lambda_{\max}(\boldsymbol{\beta} \cdot \hat{\boldsymbol{n}}) = N \max \left\{ \max_j \beta_j, 0 \right\} =: N\beta_{\max}, \quad \text{(A4)}$$

where we have implicitly defined $\beta_{\max}$. This largest eigenvalue corresponds to the eigenstate that consists of placing all photons in the mode corresponding to the largest positive β_j. If all $\beta_j \leqslant 0$, the largest eigenvalue is zero, obtained by any state with no particles in the sensor modes. Note that this requires the use of an extra mode (an ancilla or so-called "reference mode") to "store" these photons, as we fix the total photon number of our state to be N.

Similarly, the smallest eigenvalue of $\boldsymbol{\beta} \cdot \hat{\boldsymbol{n}}$ in the N-particle subspace is given by

$$\lambda_{\min}(\boldsymbol{\beta} \cdot \hat{\boldsymbol{n}}) = N \min \left\{ \min_j \beta_j, 0 \right\} =: N\beta_{\min}, \quad \text{(A5)}$$

where we have implicitly defined $\beta_{\min}$.

Using the facts above about the maximum and minimum eigenvalues of $\boldsymbol{\beta} \cdot \hat{\boldsymbol{n}}$ in the N-particle subspace we can rewrite the optimization problem in Eq. (9) as

$$\text{minimize } N(\beta_{\max} - \beta_{\min}),$$
$$\text{subject to } \boldsymbol{\alpha} \cdot \boldsymbol{\beta} = 1. \quad \text{(A6)}$$

As in the main text, define $\mathcal{P} := \{j \,|\, \alpha_j \geqslant 0\}$ and $\mathcal{N} := \{j \,|\, \alpha_j < 0\}$. We then have the following lemma.

Lemma 2. The solution $\boldsymbol{\beta}^*$ to Eq. (A6) is such that $\beta_j^* \geqslant 0$ for all $j \in \mathcal{P}$, and $\beta_j^* \leqslant 0$ for all $j \in \mathcal{N}$. That is, $\alpha_j \beta_j^* \geqslant 0$ for all j.

Proof. We proceed by contradiction. Let $\mathcal{J}_- = \{j \,|\, \alpha_j \beta_j^* < 0\}$ and $\mathcal{J}_+ = \{j \,|\, \alpha_j \beta_j^* \geqslant 0\}$. Suppose the solution vector $\boldsymbol{\beta}^*$ to Eq. (A6) has $\mathcal{J}_- \neq \emptyset$. We can construct an alternative candidate solution vector $\boldsymbol{\beta}'$ as follows: First, let $\boldsymbol{\beta}' = \boldsymbol{\beta}^*$. Then set $\beta_j' = 0$ for all $j \in \mathcal{J}_-$. In order to still satisfy the constraint $\boldsymbol{\alpha} \cdot \boldsymbol{\beta}' = 1$, we must reduce the values of some other components in $\boldsymbol{\beta}'$. In particular, it is simple to calculate that a valid solution is, for $j \in \mathcal{J}_+$,

$$\beta_j' = \frac{\beta_j^*}{\sum_{j \in \mathcal{J}_+} \alpha_j \beta_j^*} = \frac{\beta_j^*}{1 - \sum_{j \in \mathcal{J}_-} \alpha_j \beta_j^*}. \quad \text{(A7)}$$

Again, when $j \in \mathcal{J}_-$, $\beta_j' = 0$.

Let $\beta_{\max}' := \max\{\max_j \beta_j', 0\}$ and $\beta_{\min}' := \max\{\min_j \beta_j', 0\}$. By construction, $\beta_{\max}' \leqslant \beta_{\max}^*$ and $0 = \beta_{\min}' \geqslant \beta_{\min}^*$. Consequently, $\boldsymbol{\beta}'$ yields a smaller solution candidate than $\boldsymbol{\beta}^*$. This contradicts the fact that $\boldsymbol{\beta}^*$ is the optimal solution. The lemma statement follows as an immediate consequence. ∎

Lemma 2 allows us to rewrite the minimization problem in Eq. (A6) once again as

$$\text{minimize } N\left[\max_{j \in \mathcal{P}} \beta_j - \min_{j \in \mathcal{N}} \beta_j \right],$$
$$\text{where } \beta_j \geqslant 0 \ \forall \, j \in \mathcal{P},$$
$$\beta_j \leqslant 0 \ \forall \, j \in \mathcal{N},$$
$$\text{subject to } \boldsymbol{\alpha} \cdot \boldsymbol{\beta} = 1. \quad \text{(A8)}$$

In the above, we define $\max_{j \in \mathcal{P}} \beta_j$ ($\min_{j \in \mathcal{N}} \beta_j$) to be zero if $\mathcal{P} = \emptyset$ ($\mathcal{N} = \emptyset$). A further simplification is enabled by another lemma.

Lemma 3. The solution vector $\boldsymbol{\beta}^*$ to Eq. (A8) is such that $\beta_j^* = \beta_{\max}^*$ for all $j \in \mathcal{P}$ and $\beta_j^* = \beta_{\min}^*$ for all $j \in \mathcal{N}$.

Proof. We proceed by contradiction. Suppose the solution vector $\boldsymbol{\beta}^*$ is such that $\beta_i^* \neq \beta_j^*$ for some $i, j \in \mathcal{P}$. Then we could consider an alternative candidate solution vector $\boldsymbol{\beta}'$ where $\beta_k' = \frac{\sum_{l \in \mathcal{P}} \alpha_l \beta_l^*}{\sum_{l \in \mathcal{P}} \alpha_l}$ for all $k \in \mathcal{P}$. Similarly, if $\beta_i^* \neq \beta_j^*$ for some $i, j \in \mathcal{N}$ we could consider $\beta_k' = \frac{\sum_{l \in \mathcal{N}} \alpha_l \beta_l^*}{\sum_{l \in \mathcal{N}} \alpha_l}$ for all $k \in \mathcal{N}$. Clearly, $\boldsymbol{\beta}'$ still satisfies the constraint

$$\boldsymbol{\alpha} \cdot \boldsymbol{\beta}' = \sum_{m \in \mathcal{P}} \alpha_m \left(\frac{\sum_{l \in \mathcal{P}} \alpha_l \beta_l^*}{\sum_{l \in \mathcal{P}} \alpha_l} \right) + \sum_{m \in \mathcal{N}} \alpha_m \left(\frac{\sum_{l \in \mathcal{N}} \alpha_l \beta_l^*}{\sum_{l \in \mathcal{N}} \alpha_l} \right)$$
$$= \boldsymbol{\alpha} \cdot \boldsymbol{\beta}^* = 1. \quad \text{(A9)}$$

Additionally, $\boldsymbol{\beta}'$ also clearly still has $\beta_j' \geqslant 0$ when $j \in \mathcal{P}$ and $\beta_j' \leqslant 0$ when $j \in \mathcal{N}$. But, by construction (because the weighted average of a set is less than its maximum element),

$$N\left[\max_{j \in \mathcal{P}} \beta_j' - \min_{j \in \mathcal{N}} \beta_j' \right] < N\left[\max_{j \in \mathcal{P}} \beta_j^* - \min_{j \in \mathcal{N}} \beta_j^* \right]. \quad \text{(A10)}$$

So $\boldsymbol{\beta}^*$ is not the solution vector and we have arrived at a contradiction. ∎

As a direct consequence of Lemma 3 we can rewrite the optimization problem in Eq. (A8) one last time as

$$\text{minimize (w.r.t. } \beta_{\min}, \beta_{\max}) \, N[\beta_{\max} - \beta_{\min}], \tag{A11}$$

$$\text{subject to } \beta_{\max} \geqslant 0, \beta_{\min} \leqslant 0, \tag{A12}$$

$$\beta_{\max} \sum_{j \in \mathcal{P}} \alpha_j + \beta_{\min} \sum_{j \in \mathcal{N}} \alpha_j = 1. \tag{A13}$$

Because this is a linear objective function, the optimal solution will be one of the two boundary solutions: $\beta_{\max} = \frac{1}{\sum_{i \in \mathcal{P}} \alpha_i}, \beta_{\min} = 0$ or $\beta_{\min} = \frac{1}{\sum_{i \in \mathcal{N}} \alpha_i}, \beta_{\max} = 0$. Minimizing over these two candidate solutions, we obtain the final result

$$\|\hat{g}_q\|_{s,N}^2 = \frac{N^2}{\max\left(\sum_{i \in \mathcal{P}} \alpha_i, \sum_{i \in \mathcal{N}} \alpha_i\right)^2}. \tag{A14}$$

Consequently, via the quantum Cramér-Rao bound, Eq. (10),

$$\mathcal{M} \geqslant \frac{\max\left\{\sum_{i \in \mathcal{P}} \alpha_i, \sum_{i \in \mathcal{N}} \alpha_i\right\}^2}{N^2 t^2}$$

$$=: \frac{\max\left\{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}^2, \|\boldsymbol{\alpha}\|_{1,\mathcal{N}}^2\right\}}{N^2 t^2}, \tag{A15}$$

which is Eq. (12), and where $\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}$ and $\|\boldsymbol{\alpha}\|_{1,\mathcal{N}}$ are the one-norm restricted to positive and negative values, respectively, of $\boldsymbol{\alpha}$. In the special case of all positive coefficients (i.e., $\mathcal{N} = \emptyset$), this reduces to

$$\mathcal{M} \geqslant \frac{\|\boldsymbol{\alpha}\|_1^2}{N^2 t^2}, \tag{A16}$$

which, as described in the main text, proves a conjecture from Ref. [8] that this is the minimum attainable variance for $\boldsymbol{\alpha} \in \mathbb{Q}^d$ with $\boldsymbol{\alpha} \geqslant 0$.

APPENDIX B: BOUND FOR LOCAL DISPLACEMENTS

In this Appendix, we derive Eq. (15) for the mean square error attainable for measuring a linear function of local displacements, restricting to probe states with fixed average photon number $\overline{N}$.

1. Separable bound

To begin, it is helpful to present the bound for the more restricted case where we use separable input states. Begin by considering the lower bound on the variance of measuring a displacement φ coupled to a single mode via $H = \varphi \hat{p}$, following the proof sketched in Ref. [7]. The quantum Fisher information is given by

$$\mathcal{F}(\varphi) = 4[\Delta(\hat{p})_\rho]^2, \tag{B1}$$

where ρ is the probe state, which is restricted to have an average photon number $\overline{N}$. An initial displacement does not enhance precision [7], so we can consider zero-mean displacement input states. For such probe states,

$$(\Delta \hat{p})^2 = -\tfrac{1}{4}\langle(\hat{a}^\dagger - \hat{a})^2\rangle$$
$$= -\tfrac{1}{4}(\langle\hat{a}^\dagger\hat{a}^\dagger\rangle - \langle\hat{a}^\dagger\hat{a}\rangle - \langle\hat{a}\hat{a}^\dagger\rangle + \langle\hat{a}\hat{a}\rangle), \tag{B2}$$

$$(\Delta \hat{x})^2 = \tfrac{1}{4}\langle(\hat{a}^\dagger + \hat{a})^2\rangle$$
$$= \tfrac{1}{4}(\langle\hat{a}^\dagger\hat{a}^\dagger\rangle + \langle\hat{a}^\dagger\hat{a}\rangle + \langle\hat{a}\hat{a}^\dagger\rangle + \langle\hat{a}\hat{a}\rangle), \tag{B3}$$

so that

$$\overline{N} = \langle\hat{a}^\dagger\hat{a}\rangle = (\Delta\hat{p})^2 + (\Delta\hat{x})^2 - \tfrac{1}{2}, \tag{B4}$$

where we used that $\hat{a}\hat{a}^\dagger = \hat{a}^\dagger\hat{a} + 1$. We can then use the uncertainty principle

$$(\Delta\hat{p})^2(\Delta\hat{x})^2 \geqslant \tfrac{1}{16}, \tag{B5}$$

which follows from our definition of the quadrature operators as $\hat{x} = (\hat{a}^\dagger + \hat{a})/2$ and $\hat{p} = i(\hat{a}^\dagger - \hat{a})/2$. Therefore,

$$\xi\left(\overline{N} - \xi + \tfrac{1}{2}\right) \geqslant \tfrac{1}{16}, \tag{B6}$$

where we let $\xi := (\Delta\hat{p})^2$. Then

$$-16\xi^2 + (16\overline{N} + 8)\xi - 1 \geqslant 0. \tag{B7}$$

To maximize ξ, this inequality must be saturated, so we can solve the corresponding quadratic to obtain the solution

$$\xi = \frac{-8(2\overline{N} + 1) + \sqrt{64(2\overline{N} + 1)^2 - 64}}{-32}$$

$$\Rightarrow 4\xi = (\sqrt{\overline{N}} + \sqrt{\overline{N} + 1})^2 \sim 4\overline{N}. \tag{B8}$$

It is worth noting that the $O(\overline{N})$ asymptotic behavior of the maximum variance of $\hat{p}$ could have been obtained with no calculation just from examining the constraint in Eq. (B4) under the assumption that $(\Delta\hat{x})^2$ can be made negligibly small.

Putting everything back together, we have found that, optimizing over states with fixed average photon number $\overline{N}$, the following holds:

$$[\Delta(\tilde{\varphi})]^2 \geqslant \frac{1}{\mathcal{F}} \geqslant \frac{1}{t^2(\sqrt{\overline{N}} + \sqrt{\overline{N} + 1})^2} = \frac{1}{4t^2\overline{N}} + O\left(\frac{1}{t^2\overline{N}^2}\right). \tag{B9}$$

Working in the asymptotic in $\overline{N}$ limit, we can use Eq. (B9) to obtain a bound on performance for estimating a linear function $q(\boldsymbol{\theta}) = \boldsymbol{\alpha} \cdot \boldsymbol{\theta}$ with an unentangled protocol as

$$(\Delta\tilde{q})^2 \geqslant \frac{1}{t^2} \min_{\{\overline{N}_j\}} \sum_{j=1}^d \frac{|\alpha_j|^2}{4\overline{N}_j} + O\left(\frac{1}{\overline{N}_j^2}\right), \tag{B10}$$

where $\overline{N}_j = \langle\hat{a}_j^\dagger\hat{a}_j\rangle$ is the average number of photons used in mode j and $\sum_j \overline{N}_j = \overline{N}$. Assume without loss of generality that $|\alpha_j| > 0$ for all j (that is, no $\alpha_j = 0$) and independent of $\overline{N}$. Then we can optimize (at leading order in $\frac{1}{\overline{N}}$) the distribution of photons amongst the modes using the Lagrangian

$$\mathcal{L} = \sum_{j=1}^d \frac{|\alpha_j|^2}{4\overline{N}_j} + \gamma\left(\sum_{j=1}^d \overline{N}_j - \overline{N}\right), \tag{B11}$$

where γ is a Lagrange multiplier. A bit of algebra yields that

$$\frac{\partial \mathcal{L}}{\partial \overline{N}_j} = 0 \Rightarrow \overline{N}_j = \frac{|\alpha_j|}{2\sqrt{\gamma}}. \tag{B12}$$

This further implies that

$$\overline{N} = \sum_{j=1}^d \overline{N}_j = \frac{\|\boldsymbol{\alpha}\|_1}{2\sqrt{\gamma}}, \tag{B13}$$

allowing us to obtain the optimal division of photons as

$$\overline{N}_j = \frac{|\alpha_j|}{\|\boldsymbol{\alpha}\|_1} \overline{N}. \tag{B14}$$

We note that this solution is clearly the desired minimum of the Lagrangian, as maximizing the objective would lead to setting any $\overline{N}_j$ to 0. Plugging this back into Eq. (B10) we obtain the (asymptotic in $\overline{N}$) separable bound

$$[\Delta \tilde{q}]^2 \geqslant \frac{\|\boldsymbol{\alpha}\|_1^2}{4\overline{N}t^2} + O\left(\frac{1}{\overline{N}^2}\right). \tag{B15}$$

This bound can be achieved by using the single-mode protocols in Ref. [7] for each mode and then computing the function of interest classically as a linear combination of the individual estimators.

2. General function estimation bound

In this subsection, we turn to our primary task: deriving Eq. (15) for the mean square error attainable for measuring a linear function of local displacements, restricting to probe states with fixed average photon number $\overline{N}$.

To derive this bound, we must solve the optimization problem in Eq. (9) for $\hat{g}_j = \hat{p}_j$:

$$\min_{\boldsymbol{\beta}} \max_{\rho} [\Delta(\boldsymbol{\beta} \cdot \hat{\boldsymbol{p}})_\rho]^2, \quad \text{subject to } \boldsymbol{\alpha} \cdot \boldsymbol{\beta} = 1. \tag{B16}$$

We can write

$$
\begin{aligned}
[\Delta(\boldsymbol{\beta} \cdot \hat{\boldsymbol{p}})]^2 &= \sum_{i,j=1}^d \beta_i \beta_j \mathrm{Cov}(\hat{p}_i, \hat{p}_j) \\
&\leqslant \sum_{i,j=1}^d \beta_i \beta_j \sqrt{(\Delta \hat{p}_i)^2 (\Delta \hat{p}_j)^2} \\
&= \left[\sum_{j=1}^d \beta_j \Delta \hat{p}_j \right]^2 \\
&\leqslant \|\boldsymbol{\beta}\|_2^2 \sum_{j=1}^d (\Delta \hat{p}_j)^2,
\end{aligned}
\tag{B17}
$$

where we applied the Cauchy-Schwarz inequality twice. Using the same assumption of zero-displacement states we made in the previous section, we can further bound $\sum_j (\Delta \hat{p}_j)^2$ using the constraint on average photon number

$$\sum_{j=1}^d [(\Delta \hat{p}_j)^2 + (\Delta \hat{x}_j)^2] - \frac{d}{2} = \sum_{j=1}^d \langle a_j^\dagger a_j \rangle = \overline{N}, \tag{B18}$$

implying that

$$\sum_{j=1}^d (\Delta \hat{p}_j)^2 \leqslant \overline{N} + \frac{d}{2}. \tag{B19}$$

Equation (B19) is tight when $(\Delta \hat{x}_j)^2 = 0$ for all j. This is, of course, impossible to achieve, but can be approached asymptotically with increasing $\overline{N}$ ($\overline{N} \gg d$). Furthermore, using the fact that $\boldsymbol{\alpha}$ is dual to $\boldsymbol{\beta}$ and the Cauchy-Schwarz inequality, it holds that

$$1 = \boldsymbol{\alpha} \cdot \boldsymbol{\beta} \leqslant \|\boldsymbol{\beta}\|_2 \|\boldsymbol{\alpha}\|_2. \tag{B20}$$

As we want to minimize with respect to $\boldsymbol{\beta}$, we consider the case where this inequality is saturated (i.e. $\boldsymbol{\beta}^* = \frac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|_2^2}$). Therefore, $\|\boldsymbol{\beta}^*\|_2 = \frac{1}{\|\boldsymbol{\alpha}\|_2}$, and we obtain

$$[\Delta(\boldsymbol{\beta} \cdot \hat{\boldsymbol{p}})]^2 \leqslant \frac{\overline{N}}{\|\boldsymbol{\alpha}\|_2^2} + O\left(\frac{d}{\|\boldsymbol{\alpha}\|_2^2}\right). \tag{B21}$$

This yields the final bound

$$\mathcal{M} \geqslant \frac{\|\boldsymbol{\alpha}\|_2^2}{4\overline{N}t^2} - O\left(\frac{d\|\boldsymbol{\alpha}\|_2^2}{\overline{N}^2 t^2}\right). \tag{B22}$$

From the derivation alone, it is not obvious that this bound can be saturated, but the existence of protocols that achieve it [12] indicate that this bound is, indeed, tight asymptotically in $\overline{N}$.

APPENDIX C: QUANTUM FISHER INFORMATION MATRIX ELEMENTS

In this Appendix, we derive the matrix elements of the quantum Fisher information matrix for generators $\hat{n}_j$ and $\hat{p}_j$ under the unitary evolution Eq. (4). For number operator coupling $\hat{g}_j = \hat{n}_j$,

$$
\begin{aligned}
\mathcal{H}_j &= -iU^\dagger \partial_j U = -\sum_{m=1}^M \left(\prod_{l=1}^{m-1} U^{(l)} V \right)^\dagger \hat{n}_j \left(\prod_{l=1}^{m-1} U^{(l)} V \right) \\
&=: -\sum_{m=1}^M \hat{n}_j(m),
\end{aligned}
\tag{C1}
$$

where in the second line we implicitly defined $\hat{n}_j(m)$. Consequently, we can compute the quantum Fisher information matrix elements via Eq. (23) to be

$$
\begin{aligned}
\mathcal{F}(\boldsymbol{\theta})_{ij} = 4 \Bigg[&\sum_{l=1}^M \sum_{m=1}^M \frac{1}{2} \langle \{\hat{n}_i(l), \hat{n}_j(m)\} \rangle \\
&- \left(\sum_{m=1}^M \langle \hat{n}_i(m) \rangle \right) \left(\sum_{m=1}^M \langle \hat{n}_j(m) \rangle \right) \Bigg].
\end{aligned}
\tag{C2}
$$

When $\hat{U}^{(j)} = I$ for all j, this reduces to

$$\mathcal{F}(\boldsymbol{\theta})_{ij} = 4M^2 [\langle \hat{n}_i \hat{n}_j \rangle - \langle \hat{n}_i \rangle \langle \hat{n}_j \rangle]. \tag{C3}$$

For quadrature operator coupling $\hat{g}_j = \hat{p}_j$, essentially identical manipulations yield

$$\mathcal{F}(\boldsymbol{\theta})_{ij} = 4 \left[\sum_{l=1}^{M} \sum_{m=1}^{M} \frac{1}{2} \langle \{\hat{p}_i(l), \hat{p}_j(m)\} \rangle \right.$$
$$\left. - \left(\sum_{m=1}^{M} \langle \hat{p}_i(m) \rangle \right) \left(\sum_{m=1}^{M} \langle \hat{p}_j(m) \rangle \right) \right], \quad \text{(C4)}$$

where $\hat{p}_j(l)$ is defined as in Eq. (C1) with $\hat{n}_j \to \hat{p}_j$.

APPENDIX D: PROTOCOLS FOR LOCAL PHASE SHIFTS

In this Appendix, we elaborate on the families of optimal protocols for measuring a linear function of phase shifts that we described in Sec. IV.

1. An optimal protocol for functions with positive coefficients

We begin by reviewing a protocol from Ref. [8] for the special case of a linear function with positive coefficients (i.e., $\boldsymbol{\alpha} \geqslant 0$). Our results in Appendix A show that, as those authors conjectured, this protocol is optimal. In particular, consider using as the probe state a so-called proportionally weighted N00N state over $d+1$ modes:

$$|\psi\rangle \propto \left| N \frac{\alpha_1}{\|\boldsymbol{\alpha}\|_1}, \dots, N \frac{\alpha_d}{\|\boldsymbol{\alpha}\|_1}, 0 \right\rangle + |0, \dots, 0, N\rangle, \quad \text{(D1)}$$

where we have expressed the state in an occupation number basis over $d+1$ modes and have dropped the normalization for concision. The last mode serves as a reference mode. Observe that, for this state to be well defined, it is essential that $\frac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|_1} \in \mathbb{Q}^d$ and that N is such that the resulting occupation numbers are integers, which may require that N be large.

Following imprinting of the parameters $\boldsymbol{\theta}$ onto the probe state via M passes through the interferometers, one obtains

$$|\psi_M\rangle = e^{-iM\hat{\boldsymbol{n}}\cdot\boldsymbol{\theta}} |\psi\rangle \propto \left| N \frac{\alpha_1}{\|\boldsymbol{\alpha}\|_1}, \dots, N \frac{\alpha_d}{\|\boldsymbol{\alpha}\|_1}, 0 \right\rangle$$
$$+ e^{i\boldsymbol{\alpha}\cdot\boldsymbol{\theta}\frac{NM}{\|\boldsymbol{\alpha}\|_1}} |0, \dots, 0, N\rangle. \quad \text{(D2)}$$

This process allows us to saturate the bound in Eq. (14). In particular, using Eq. (C2) [which reduces to Eq. (C3) because there is no control required], it is straightforward to calculate that the quantum Fisher information matrix for the probe state is

$$\mathcal{F}(\boldsymbol{\theta}) = \frac{(MN)^2}{\|\boldsymbol{\alpha}\|_1^2} \boldsymbol{\alpha}\boldsymbol{\alpha}^T, \quad \text{(D3)}$$

which clearly satisfies the condition in Eq. (24) (recalling that $\|\boldsymbol{\alpha}\|_1 = \|\boldsymbol{\alpha}\|_{1,\mathcal{P}}$ here because we have assumed all coefficients are non-negative, and also recalling that $\Delta t = 1$ such that $M = t$).

While the conditions on the quantum Fisher information matrix for an optimal protocol are met, a full protocol requires a description of the measurements used to extract the quantity of interest from the relative phase between the branches of $|\psi_M\rangle$. As described in the main text, this can be done via the robust phase estimation protocols of Refs. [38–40] with a small multiplicative constant overhead relative to the quantum

Cramér-Rao bound (we also briefly discuss the idea behind robust phase estimation in Appendix G). The details of implementing the necessary parity measurements for N00N-like states are discussed in detail in Appendix A of Ref. [40] and Ref. [43].

2. Extending the optimal protocol to negative coefficients

While not explicitly considered in Ref. [8], it is straightforward to extend the above protocol to the situation where $\mathcal{N} \neq \emptyset$, which we do here. Without loss of generality, assume the coefficients are ordered so that $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_d$. Using our standard assumption that $\|\boldsymbol{\alpha}\|_{1,\mathcal{P}} \geqslant \|\boldsymbol{\alpha}\|_{1,\mathcal{N}}$, we claim that the following probe state is optimal:

$$|\psi\rangle \propto \bigotimes_{j\in\mathcal{P}} \left| N \frac{\alpha_j}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}} \right\rangle |0\rangle^{\otimes|\mathcal{N}|} |0\rangle$$
$$+ |0\rangle^{\otimes|\mathcal{P}|} \bigotimes_{j\in\mathcal{N}} \left| N \frac{|\alpha_j|}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}} \right\rangle \left| N - N \frac{\|\boldsymbol{\alpha}\|_{1,\mathcal{N}}}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}} \right\rangle, \quad \text{(D4)}$$

where, again, the last mode is a reference mode, and we have dropped the normalization of the state. Interestingly, observe that, if $\|\boldsymbol{\alpha}\|_{1,\mathcal{P}} = \|\boldsymbol{\alpha}\|_{1,\mathcal{N}}$, the reference mode factors out and is unnecessary. Similar to the $\boldsymbol{\alpha} \geqslant 0$ case, for this state to be well defined, we require that $N|\alpha_j|/\|\boldsymbol{\alpha}\|_{1,\mathcal{P}} \in \mathbb{N}$ for all j, which is always true for some sufficiently large N provided $\boldsymbol{\alpha} \in \mathbb{Q}^d$.

Consider applying the encoding unitary for M passes through the interferometers. For $\|\boldsymbol{\alpha}\|_{1,\mathcal{P}} \geqslant \|\boldsymbol{\alpha}\|_{1,\mathcal{N}}$, this yields

$$|\psi_M\rangle \propto \bigotimes_{j\in\mathcal{P}} \left| N \frac{\alpha_j}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}} \right\rangle |0\rangle^{\otimes|\mathcal{N}|} |0\rangle$$
$$+ e^{i\boldsymbol{\alpha}\cdot\boldsymbol{\theta}\frac{NM}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}}} |0\rangle^{\otimes|\mathcal{P}|} \bigotimes_{j\in\mathcal{N}} \left| N \frac{|\alpha_j|}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}} \right\rangle \left| N - N \frac{\|\boldsymbol{\alpha}\|_{1,\mathcal{N}}}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}} \right\rangle. \quad \text{(D5)}$$

This probe state is optimal in the sense of satisfying the Fisher information condition in Eq. (24). In the main text, we described an even more general family of protocols. Within this more general framework, we will prove this optimality.

3. A family of optimal protocols

Finally, we describe a family of optimal protocols that satisfy the conditions on the quantum Fisher information matrix given in Eq. (24). In the main text, we defined a family of optimal protocols in terms of vectors from the set

$$\mathcal{W} := \{\boldsymbol{\omega} \in \mathbb{Z}^d \mid \|\boldsymbol{\omega}\|_{1,\mathcal{P}} = N, \; \|\boldsymbol{\omega}\|_{1,\mathcal{N}} \leqslant N, \; \omega_j\alpha_j \geqslant 0 \; \forall j\}. \quad \text{(D6)}$$

In particular, from these vectors, we defined a set $\mathcal{T}$ of one-parameter families of probe states in an occupation number basis where each $|\psi(\boldsymbol{\omega}; \varphi)\rangle \in \mathcal{T}$ is labeled by a particular choice of $\boldsymbol{\omega}$ such that

$$|\psi(\boldsymbol{\omega}; \varphi)\rangle \propto |\boldsymbol{\omega}|_{\mathcal{P}}\rangle|0\rangle + e^{i\varphi}|-\boldsymbol{\omega}|_{\mathcal{N}}\rangle|N - \|\boldsymbol{\omega}|_{\mathcal{N}}\|_1\rangle, \quad \text{(D7)}$$

where $\varphi \in \mathbb{R}$ is an arbitrary parameter and the last mode is a reference mode. Recall also that $\boldsymbol{\omega}_{\mathcal{P}}$ and $\boldsymbol{\omega}_{\mathcal{N}}$ are defined in

Eq. (28) as the restriction of $\boldsymbol{\omega}$ to $j \in \mathcal{P}$ and $\mathcal{N}$, respectively (for j not in the correct set, the value is set to 0). Note that such states are of the form of those in Lemma 1. We claimed that, by explicitly computing the Fisher information matrix for these states, one could demonstrate that the optimality condition in Eq. (24) is satisfied for a protocol such that

$$W\boldsymbol{r} = NM\frac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}}, \tag{D8}$$

where $\boldsymbol{r} \in \mathbb{Z}^{|\mathcal{T}|}$ is as defined in the main text and must obey the conditions

$$\|\boldsymbol{r}\|_1 = M,$$
$$\boldsymbol{r} \geqslant 0. \tag{D9}$$

Recall that W is a matrix whose columns are the vectors $\boldsymbol{\omega}_n \in \mathcal{W}$.

Here we explicitly demonstrate this. We can easily evaluate

$$\langle \hat{n}_j(m) \rangle = \langle \psi(\boldsymbol{\omega}^{(m)}; \varphi) | \hat{n}_j | \psi(\boldsymbol{\omega}^{(m)}; \varphi) \rangle = \frac{\left| \omega_j^{(m)} \right|}{2} \tag{D10}$$

and

$$
\begin{aligned}
&\langle \hat{n}_i(l)\hat{n}_j(m) \rangle \\
&= \langle \psi(\boldsymbol{\omega}^{(l)}; \varphi) | \hat{n}_i U(m \leftrightarrow l) \hat{n}_j | \psi(\boldsymbol{\omega}^{(m)}; \varphi) \rangle \\
&= \frac{\left| \omega_i^{(l)} \omega_j^{(m)} \right|}{2} \langle \psi_l(\boldsymbol{\omega}^{(l)}; \varphi) | U(m \leftrightarrow l) | \psi_m(\boldsymbol{\omega}^{(m)}; \varphi) \rangle,
\end{aligned}
\tag{D11}
$$

where $\hat{n}_j(m)$ are defined as in Eq. (C1), and

$$U(m \leftrightarrow l) = \begin{cases} \prod_{k=m}^{l-1} U^{(k)}V, & \text{if } l \geqslant m \\ \prod_{k=l}^{m-1} (U^{(k)}V)^{\dagger}, & \text{otherwise,} \end{cases} \tag{D12}$$

i.e., it is the unitary that converts between the mth and lth probe states. Additionally, $\boldsymbol{\omega}^{(m)}$ refers to the vector associated to the mth probe state; correspondingly $|\psi_l(\boldsymbol{\omega}^{(l)}; \varphi)\rangle$ is the branch of $|\psi(\boldsymbol{\omega}^{(l)}; \varphi)\rangle$ with nonzero occupation number on mode l and $|\psi_m(\boldsymbol{\omega}^{(m)}; \varphi)\rangle$ is the branch of $|\psi(\boldsymbol{\omega}^{(m)}; \varphi)\rangle$ with nonzero occupation number on mode m. For an optimal protocol, $U(m \leftrightarrow l)$ coherently maps the first (second) branch of $|\psi(\boldsymbol{\omega}^{(l)}; \varphi)\rangle$ to the first (second) branch of $|\psi(\boldsymbol{\omega}^{(m)}; \varphi)\rangle$; therefore, we have that the matrix element $\langle \psi_l(\boldsymbol{\omega}^{(l)}; \varphi) | U(m \leftrightarrow l) | \psi_m(\boldsymbol{\omega}^{(m)}; \varphi) \rangle$ is nonzero if and only if the branches with nonzero occupation on modes l and m are the same. So we have that

$$\langle \hat{n}_i(l)\hat{n}_j(m) \rangle = \frac{\left| \omega_i^{(l)} \omega_j^{(m)} \right|}{2} \xi_{ij}, \tag{D13}$$

where

$$\xi_{ij} := \begin{cases} 1, & \text{if } i, j \in \mathcal{P} \text{ or } i, j \in \mathcal{N} \\ 0, & \text{otherwise.} \end{cases} \tag{D14}$$

Putting everything together we obtain that

$$\mathcal{F}(\boldsymbol{\theta})_{ij} = (-1)^{\xi_{ij}+1} \left(\sum_{m=1}^{M} \left| \omega_i^{(m)} \right| \right) \left(\sum_{m=1}^{M} \left| \omega_j^{(m)} \right| \right). \tag{D15}$$

To prove the protocols work, we need to show that this Fisher information matrix obeys the condition in Eq. (24).

Without loss of generality, consider the case that $\|\boldsymbol{\alpha}\|_{1,\mathcal{P}} \geqslant \|\boldsymbol{\alpha}\|_{1,\mathcal{N}}$. We have that

$$\sum_{j \in \mathcal{P}} \mathcal{F}(\boldsymbol{\theta})_{ij} = \text{sgn}(\alpha_i) \left(\sum_{m=1}^{M} \left| \omega_i^{(m)} \right| \right) MN, \tag{D16}$$

where we used that $\|\boldsymbol{\omega}\|_{1,\mathcal{P}} = N$. So, to obey the condition in Eq. (24), we require that

$$\sum_{m=1}^{M} \left| \omega_i^{(m)} \right| = MN\frac{|\alpha_i|}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}}. \tag{D17}$$

Or, in vector form:

$$\sum_{m=1}^{M} \left| \boldsymbol{\omega}^{(m)} \right| = MN\frac{|\boldsymbol{\alpha}|}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}}. \tag{D18}$$

Protocols in our family satisfy this condition by construction as, for any valid protocol,

$$\sum_{m=1}^{M} \left| \boldsymbol{\omega}^{(m)} \right| = |W|\boldsymbol{r}, \tag{D19}$$

where $|W|$ denotes taking the element-wise absolute value of the elements of W. Consequently, noting that $\text{sgn}(\omega_j^{(m)}) = \text{sgn}(\alpha_j)$ for all m, we require

$$W\boldsymbol{r} = MN\frac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}}, \tag{D20}$$

which is Eq. (D8).

APPENDIX E: PROOF OF LEMMA 1

Here we provide a proof of Lemma 1 in the main text, restated here for convenience.

Lemma 4. Any optimal protocol using N photons and M passes through interferometers with a coupling as in Eq. (1) with $\hat{g}_j = \hat{n}_j$ requires that, for every pass m, the probe state $|\psi_m\rangle$ be of the form

$$|\psi_m\rangle \propto |N(m)\rangle_{\mathcal{P}} |\boldsymbol{0}\rangle_{\mathcal{N}\mathcal{R}} + e^{i\varphi_m} |\boldsymbol{0}\rangle_{\mathcal{P}} |N'(m)\rangle_{\mathcal{N}\mathcal{R}}, \tag{E1}$$

where $\mathcal{P}$, $\mathcal{N}$, and $\mathcal{R}$ represent the modes with $\alpha_j \geqslant 0$, $\alpha_j < 0$, and the (arbitrary number of) reference modes, respectively, $N(m)$ and $N'(m)$ are strings of occupation numbers such that $|N(m)| = |N'(m)| = N$ for all passes m. φ_m is an arbitrary phase.

Proof. The quantum Fisher information matrix elements for any protocol with $\hat{g}_j = \hat{n}_j$ are given by

$$
\begin{aligned}
\mathcal{F}(\boldsymbol{\theta})_{ij} &= 4\Bigg[\sum_{l=1}^{M} \sum_{m=1}^{M} \frac{1}{2} \langle \{\hat{n}_i(l), \hat{n}_j(m)\} \rangle \\
&\quad - \left(\sum_{m=1}^{M} \langle \hat{n}_i(m) \rangle \right) \left(\sum_{m=1}^{M} \langle \hat{n}_j(m) \rangle \right) \Bigg] \\
&= 4 \sum_{l=1}^{M} \sum_{m=1}^{M} \text{Cov}(\hat{n}_i(l), \hat{n}_j(m)), \tag{E2}
\end{aligned}
$$

where the expectation values are taken with respect to the initial probe state, and $\hat{n}_j(m)$ are the number operators on the jth mode in the Heisenberg picture prior to the mth pass, as

specified in Eq. (C1). Without loss of generality, we make the assumption that $\|\boldsymbol{\alpha}\|_{1,\mathcal{P}} \geqslant \|\boldsymbol{\alpha}\|_{1,\mathcal{N}}$. Summing over $i, j \in \mathcal{P}$, we have that, for an optimal protocol,

$$\sum_{i\in\mathcal{P}}\sum_{j\in\mathcal{P}}\mathcal{F}(\boldsymbol{\theta})_{ij} = \sum_{j\in\mathcal{P}}\frac{(MN)^2}{\|\boldsymbol{\alpha}\|_{1,\mathcal{P}}}\alpha_j = (MN)^2, \quad (\text{E3})$$

where we used the condition in Eq. (24) for an optimal protocol, and we recall that, for $j \in \mathcal{P}$, all $\alpha_j > 0$. For convenience, define

$$\hat{P}(m) := \sum_{j\in\mathcal{P}}\hat{n}_j(m). \quad (\text{E4})$$

Armed with this definition, we can upper bound the sum over $i, j \in \mathcal{P}$ in the explicit expression from Eq. (E2) as

$$\sum_{i\in\mathcal{P}}\sum_{j\in\mathcal{P}}\mathcal{F}(\boldsymbol{\theta})_{ij} = 4\sum_{l=1}^{M}\sum_{m=1}^{M}\text{Cov}(\hat{P}(l), \hat{P}(m))$$

$$\leqslant 4\sum_{l=1}^{M}\sum_{m=1}^{M}\sqrt{\text{Var}(\hat{P}(l))\text{Var}(\hat{P}(m))}$$

$$= 4\left(\sum_{l=1}^{M}\sqrt{\text{Var}(\hat{P}(l))}\right)^2$$

$$\leqslant 4\left(\sum_{l=1}^{M}\frac{\|\hat{P}(l)\|_{s,N}}{2}\right)^2$$

$$\leqslant (NM)^2, \quad (\text{E5})$$

where in the first line we use the Cauchy-Schwarz inequality, in the second line we use that once restricted to the N-particle subspace $\text{Var}(A) \leqslant \|A\|_{s,N}^2/4$ (where, again, $\|A\|_{s,N}$ is the seminorm restricted to the N-particle subspace) for any Hermitian operator A, and in the final line we use that $\|\hat{P}(l)\|_{s,N} \leqslant N$. Comparing Eq. (E5) with Eq. (E3), we find that, for any optimal protocol, all inequalities in Eq. (E5) must be saturated. Specifically,

$$\text{Cov}(\hat{P}(l), \hat{P}(m))^2 = \text{Var}(\hat{P}(l))\text{Var}(\hat{P}(m)), \quad (\text{E6})$$

$$\text{Var}(\hat{P}(l)) = \frac{N^2}{4}. \quad (\text{E7})$$

The second condition, Eq. (E7), means that, at all times, the state of our system must be of the form

$$\frac{|\boldsymbol{N}(l)\rangle_{\mathcal{P}}|\boldsymbol{0}\rangle_{\mathcal{N}\mathcal{R}} + e^{i\varphi_l}|\boldsymbol{0}\rangle_{\mathcal{P}}|\boldsymbol{N}'(l)\rangle_{\mathcal{N}\mathcal{R}}}{\sqrt{2}}, \quad (\text{E8})$$

where we are using the simplifying notation from the statement of the lemma. In particular, the subscripts $\mathcal{P}, \mathcal{N}, \mathcal{R}$ refer to the collection of all modes associated with $\alpha_j \geqslant 0, \alpha_j < 0$, and the reference modes, respectively. Therefore, the state $|\boldsymbol{N}\rangle_{\mathcal{P}}|\boldsymbol{0}\rangle_{\mathcal{N}\mathcal{R}}$ means that all photons are distributed (in some potentially arbitrary way) among the modes with non-negative α_j, and there are no photons in the modes with negative α_j or in the reference modes. Contrastingly, $|\boldsymbol{0}\rangle_{\mathcal{P}}|\boldsymbol{N}'(l)\rangle_{\mathcal{N}\mathcal{R}}$ refers to a state where there are N photons in the negative and reference modes, and there are no photons in the non-negative modes. We have also shifted to the Schrödinger picture where we move the time dependence onto the state as opposed to the

operators. It is simple to verify that this state satisfies Eq. (E7), and it is also simple to verify these are the most general states that achieve this. Intuitively, $|\psi_m\rangle$ is a generalized N00N state between the positive and negative/reference modes. ∎

In addition, we have the following useful characterization of optimal protocols:

Lemma 5. Let $|\psi_i\rangle$ be a state of the form in Lemma 1. Refer to the first and second parts of its superposition as, respectively, the first and second or positive and nonpositive branches. Let U_m be the unitary that maps the initial state $|\psi_1\rangle$ to the state just before the mth pass, $|\psi_m\rangle$, given by

$$U_m = \begin{cases} \prod_{i=1}^{m-1} U^{(i)}V, & M+1 \geqslant m \geqslant 2 \\ I, & m = 1. \end{cases} \quad (\text{E9})$$

in agreement with Eq. (4). Then, if U_m is part of an optimal protocol, it coherently maps the first (second) branch of $|\psi_1\rangle$ to the first (second) branch of $|\psi_m\rangle$.

Proof. We use the covariance equality in Eq. (E6). To proceed, we evaluate the expectation value of $\hat{P}$ in the initial state. Here we will again use the Schrödinger picture.

$$\langle\psi_1|\hat{P}(l)|\psi_1\rangle = \langle\psi_l|\hat{P}|\psi_l\rangle \quad (\text{E10})$$

$$= \frac{1}{2}(\langle\boldsymbol{N}(l)|_{\mathcal{P}}\langle\boldsymbol{0}|_{\mathcal{N}\mathcal{R}} + e^{-i\varphi_l}\langle\boldsymbol{0}|_{\mathcal{P}}\langle\boldsymbol{N}'(l)|_{\mathcal{N}\mathcal{R}})$$

$$\times \hat{P}(|\boldsymbol{N}(l)\rangle_{\mathcal{P}}|\boldsymbol{0}\rangle_{\mathcal{N}\mathcal{R}} + e^{i\varphi_l}|\boldsymbol{0}\rangle_{\mathcal{P}}|\boldsymbol{N}'(l)\rangle_{\mathcal{N}\mathcal{R}}) \quad (\text{E11})$$

$$= \frac{1}{2}(\langle\boldsymbol{N}(l)|_{\mathcal{P}}\langle\boldsymbol{0}|_{\mathcal{N}\mathcal{R}} + e^{-i\varphi_l}\langle\boldsymbol{0}|_{\mathcal{P}}\langle\boldsymbol{N}'(l)|_{\mathcal{N}\mathcal{R}})$$

$$\times N(|\boldsymbol{N}(l)\rangle_{\mathcal{P}}|\boldsymbol{0}\rangle_{\mathcal{N}\mathcal{R}}) \quad (\text{E12})$$

$$= \frac{N}{2}. \quad (\text{E13})$$

We next evaluate the covariance:

$$\text{Cov}(\hat{P}(l), \hat{P}(m))$$

$$= \langle\psi_1|\hat{P}(l)\hat{P}(m)|\psi_1\rangle$$

$$- \langle\psi_1|\hat{P}(l)|\psi_1\rangle\langle\psi_1|\hat{P}(m)|\psi_1\rangle \quad (\text{E14})$$

$$= \langle\psi_l|\hat{P}U_lU_m^\dagger\hat{P}|\psi_m\rangle - \langle\psi_l|\hat{P}|\psi_l\rangle\langle\psi_m|\hat{P}|\psi_m\rangle \quad (\text{E15})$$

$$= \frac{N^2}{2}\langle\boldsymbol{N}(l)|_{\mathcal{P}}\langle\boldsymbol{0}|_{\mathcal{N}\mathcal{R}}U_lU_m^\dagger|\boldsymbol{N}(m)\rangle_{\mathcal{P}}|\boldsymbol{0}\rangle_{\mathcal{N}\mathcal{R}} - \frac{N^2}{4}, \quad (\text{E16})$$

where in the last line we have used the fact that $\hat{P}$ gives a factor of N when acting on the first branch of states $|\psi_l\rangle$ and $|\psi_m\rangle$, but it annihilates the second branch that has zero photons in the positive modes.

In order for Eq. (E6) to be satisfied, and using Eq. (E7), we therefore require that, for all pairs of passes l, m,

$$\langle\boldsymbol{N}(l)|_{\mathcal{P}}\langle\boldsymbol{0}|_{\mathcal{N}\mathcal{R}}U_lU_m^\dagger|\boldsymbol{N}(m)\rangle_{\mathcal{P}}|\boldsymbol{0}\rangle_{\mathcal{N}\mathcal{R}} = 1. \quad (\text{E17})$$

Choosing $l = 1$, this implies that we require that

$$U_m^\dagger|\boldsymbol{N}(m)\rangle_{\mathcal{P}}|\boldsymbol{0}\rangle_{\mathcal{N}\mathcal{R}} = |\boldsymbol{N}(0)\rangle_{\mathcal{P}}|\boldsymbol{0}\rangle_{\mathcal{N}\mathcal{R}} \equiv |\psi_1\rangle_{\mathcal{P}}, \quad (\text{E18})$$

where we are defining $|\psi_1\rangle_{\mathcal{P}}, |\psi_1\rangle_{\mathcal{N}\mathcal{R}}$ such that $|\psi_0\rangle \propto |\psi_1\rangle_{\mathcal{P}} + |\psi_1\rangle_{\mathcal{N}\mathcal{R}}$ in the obvious way. Moving the unitary onto the right hand side of the equation yields

$$|\psi_m\rangle_{\mathcal{P}} = U_m|\psi_1\rangle_{\mathcal{P}}, \quad (\text{E19})$$

which of course implies the corresponding equation for the second branch by linearity. ∎

APPENDIX F: FISHER INFORMATION MATRIX CONDITIONS FOR QUADRATURE DISPLACEMENTS

In this Appendix, we provide conditions on the quantum Fisher information matrix for an optimal protocol in the case of quadrature generators. This result yields a simpler form of the saturability condition of Eq. (25), although the set of states that it picks out is less clear than in the number operator case. This issue is compounded by the fact that the bound is not actually saturable (it can only be approached asymptotically as $\overline{N} \to \infty$). Regardless, it allows us to bring quadrature displacements into our general formalism and suggests a route towards designing additional optimal protocols beyond those already in the literature.

In particular, starting with the definition of $\hat{p}_i(l)$ from Eq. (C4), we can bound the sum over the quantum Fisher information matrix elements as

$$\sum_{i=1, j=1}^{d} \mathcal{F}(\boldsymbol{\theta})_{ij}$$

$$= \sum_{i=1, j=1}^{d} 4 \sum_{l=1}^{M} \sum_{m=1}^{M} \mathrm{Cov}(\hat{p}_i(l), \hat{p}_j(m)) \tag{F1}$$

$$\leqslant 4 \sum_{l=1}^{M} \sum_{m=1}^{M} \sqrt{\mathrm{Var}\left(\sum_{i=1}^{d} \hat{p}_i(l)\right) \mathrm{Var}\left(\sum_{i=1}^{d} \hat{p}_j(m)\right)} \tag{F2}$$

$$= 4 \left(\sum_{l=1}^{M} \sqrt{\mathrm{Var}\left(\sum_{i=1}^{d} \hat{p}_i(l)\right)}\right)^2 \tag{F3}$$

$$\leqslant 4 \left(\sum_{l=1}^{M} \sqrt{\overline{N} - \frac{d}{2}}\right)^2 = 4M^2 \left(\overline{N} - \frac{d}{2}\right) \sim 4M^2 \overline{N}. \tag{F4}$$

Above, in Eq. (F2), we used the Cauchy-Schwarz inequality; in Eq. (F4), we used the uncertainty relation in Eq. (B19). Consistent with the rest of the paper, the $\sim$ symbol denotes asymptotically in $\overline{N}$ (for $\overline{N} \gg d$).

The saturability condition in Eq. (25) states that, for an optimal protocol (asymptotically in $\overline{N}$), it must hold that $\boldsymbol{\alpha}$ is an eigenvector of $\mathcal{F}(\boldsymbol{\theta})$ with eigenvalue $4M^2 \overline{N}$. Thus, for an optimal protocol,

$$\mathrm{Tr}(\mathcal{F}) = \sum_{j=1}^{d} \lambda_j \gtrsim 4M^2 \overline{N}, \tag{F5}$$

where λ_j are the eigenvalues of $\mathcal{F}$. This implies that the chain of inequalities leading to Eq. (F4) must be saturated (asymptotically in $\overline{N}$) for an optimal protocol and that the largest eigenvalue of $\mathcal{F}$ must be $\lambda_1 \sim 4m^2 \overline{N}$ with all other eigenvalues zero. It immediately follows that the saturability condition for quadrature displacements can be written as

$$\mathcal{F}(\boldsymbol{\theta})_{ij} \sim \frac{4M^2 \overline{N}}{\|\boldsymbol{\alpha}\|_2^2} \alpha_i \alpha_j. \tag{F6}$$

APPENDIX G: APPROACHING THE SINGLE-SHOT LIMIT AND ROBUST PHASE ESTIMATION

As pointed out in the footnote preceding Eq. (8) and in the discussion of what defines an information-theoretically optimal protocol in Sec. IV B, it is not, in practice, possible to construct an unbiased estimator achieving the single shot ($\mu = 1$) quantum Cramér-Rao bound that we analyze in this paper, as the quantum Cramér-Rao bound is only guaranteed to be achievable in the limit of asymptotically large amounts of data ($\mu \to \infty$). Resolving this tension while still achieving asymptotic Heisenberg scaling in the total amount of resources (here, μN photons) requires carefully designed protocols. In particular, extracting a relative phase from the probe states considered in the protocols in this paper requires a proper division of resources so that, asymptotically, the single-shot bound is achieved up to a small constant.

At best, this constant can be reduced to π^2 [41], but the nonadaptive robust phase estimation scheme of Refs. [38–40] provides a relatively simple-to-implement approach with a multiplicative overhead of $(24.26\pi)^2$. In brief, these protocols work by dividing the protocol into K stages where in stage j one uses N_j photons (or $\overline{N}_j$ average photons for displacement sensing). In each stage, one imprints the unknown function into the phase between two branches of a catlike state of N_j photons and then performs a measurement, as described in the main text. The experiment is performed ν_j times, allowing one to obtain an estimate of the unknown phase. This estimate is refined over the course of the K stages, with more photons used in each additional stage such that the total photon resources are

$$N = \sum_{j=1}^{K} \nu_j N_j. \tag{G1}$$

An optimal choice of ν_j and N_j ensures that, asymptotically, $N_K = \Theta(N)$ and $\nu_K = O(1)$, and, thus, the asymptotic scaling of the single-shot bound is obtained up to a multiplicative constant that depends on the details of the optimization. The proof of this and the associated optimization are detailed in Refs. [38–40].

[1] M. Piani, M. Cianciaruso, T. R. Bromley, C. Napoli, N. Johnston, and G. Adesso, Robustness of asymmetry and coherence of quantum states, Phys. Rev. A **93**, 042107 (2016).

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

[3] M. M. Taddei, B. M. Escher, L. Davidovich, and R. L. de Matos Filho, Quantum speed limit for physical processes, Phys. Rev. Lett. **110**, 050402 (2013).

[4] S. Deffner and S. Campbell, Quantum speed limits: From Heisenberg's uncertainty principle to optimal quantum control, J. Phys. A: Math. Theor. **50**, 453001 (2017).

[5] L. P. García-Pintos, S. B. Nicholson, J. R. Green, A. del Campo, and A. V. Gorshkov, Unifying quantum and classical speed limits on observables, Phys. Rev. X **12**, 011038 (2022).

[6] 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).

[7] Z. Zhang and Q. Zhuang, Distributed quantum sensing, Quantum Sci. Technol. **6**, 043001 (2021).

[8] T. J. Proctor, P. A. Knott, and J. A. Dunningham, Networked quantum sensing, arXiv:1702.04271.

[9] W. Ge, K. Jacobs, Z. Eldredge, A. V. Gorshkov, and M. Foss-Feig, Distributed quantum metrology with linear networks and separable inputs, Phys. Rev. Lett. **121**, 043604 (2018).

[10] Z. Eldredge, M. Foss-Feig, J. A. Gross, S. L. Rolston, and A. V. Gorshkov, Optimal and secure measurement protocols for quantum sensor networks, Phys. Rev. A **97**, 042337 (2018).

[11] T. J. Proctor, P. A. Knott, and J. A. Dunningham, Multiparameter estimation in networked quantum sensors, Phys. Rev. Lett. **120**, 080501 (2018).

[12] Q. Zhuang, Z. Zhang, and J. H. Shapiro, Distributed quantum sensing using continuous-variable multipartite entanglement, Phys. Rev. A **97**, 032329 (2018).

[13] Y. Xia, Q. Zhuang, W. Clark, and Z. Zhang, Repeater-enhanced distributed quantum sensing based on continuous-variable multipartite entanglement, Phys. Rev. A **99**, 012328 (2019).

[14] J. Gross and C. Caves, One from many: Estimating a function of many parameters, J. Phys. A: Math. Theor. **54**, 014001 (2021).

[15] J. Rubio, P. A. Knott, T. J. Proctor, and J. A. Dunningham, Quantum sensing networks for the estimation of linear functions, J. Phys. A: Math. Theor. **53**, 344001 (2020).

[16] C. Oh, C. Lee, S. H. Lie, and H. Jeong, Optimal distributed quantum sensing using Gaussian states, Phys. Rev. Res. **2**, 023030 (2020).

[17] D. Triggiani, P. Facchi, and V. Tamma, Heisenberg scaling precision in the estimation of functions of parameters in linear optical networks, Phys. Rev. A **104**, 062603 (2021).

[18] A. Ehrenberg, J. Bringewatt, and A. V. Gorshkov, Minimum entanglement protocols for function estimation, Phys. Rev. Res. **5**, 033228 (2023).

[19] C. Oh, L. Jiang, and C. Lee, Distributed quantum phase sensing for arbitrary positive and negative weights, Phys. Rev. Res. **4**, 023164 (2022).

[20] M. Malitesta, A. Smerzi, and L. Pezzè, Distributed quantum sensing with squeezed-vacuum light in a configurable array of Mach-Zehnder interferometers, Phys. Rev. A **108**, 032621 (2023).

[21] Y. Yang, B. Yadin, and Z.-P. Xu, Quantum-enhanced metrology with network states, arXiv:2307.07758.

[22] X. Guo, C. R. Breum, J. Borregaard, S. Izumi, M. V. Larsen, T. Gehring, M. Christandl, J. S. Neergaard-Nielsen, and U. L. Andersen, Distributed quantum sensing in a continuous-variable entangled network, Nat. Phys. **16**, 281 (2020).

[23] Y. Xia, W. Li, W. Clark, D. Hart, Q. Zhuang, and Z. Zhang, Demonstration of a reconfigurable entangled radio-frequency photonic sensor network, Phys. Rev. Lett. **124**, 150502 (2020).

[24] L.-Z. Liu, Y.-Z. Zhang, Z.-D. Li, R. Zhang, X.-F. Yin, Y.-Y. Fei, L. Li, N.-L. Liu, F. Xu, Y.-A. Chen *et al.*, Distributed quantum phase estimation with entangled photons, Nat. Photon. **15**, 137 (2021).

[25] S.-R. Zhao, Y.-Z. Zhang, W.-Z. Liu, J.-Y. Guan, W. Zhang, C.-L. Li, B. Bai, M.-H. Li, Y. Liu, L. You *et al.*, Field demonstration of distributed quantum sensing without post-selection, Phys. Rev. X **11**, 031009 (2021).

[26] K. Qian, Z. Eldredge, W. Ge, G. Pagano, C. Monroe, J. V. Porto, and A. V. Gorshkov, Heisenberg-scaling measurement protocol for analytic functions with quantum sensor networks, Phys. Rev. A **100**, 042304 (2019).

[27] T. Qian, J. Bringewatt, I. Boettcher, P. Bienias, and A. V. Gorshkov, Optimal measurement of field properties with quantum sensor networks, Phys. Rev. A **103**, L030601 (2021).

[28] A. Hamann, P. Sekatski, and W. Dür, Approximate decoherence free subspaces for distributed sensing, Quantum Sci. Technol. **7**, 025003 (2022).

[29] J. Rubio and J. Dunningham, Bayesian multiparameter quantum metrology with limited data, Phys. Rev. A **101**, 032114 (2020).

[30] J. Bringewatt, I. Boettcher, P. Niroula, P. Bienias, and A. V. Gorshkov, Protocols for estimating multiple functions with quantum sensor networks: Geometry and performance, Phys. Rev. Res. **3**, 033011 (2021).

[31] E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino, Photonic quantum metrology, AVS Quantum Sci. **2**, 024703 (2020).

[32] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum metrology, Phys. Rev. Lett. **96**, 010401 (2006).

[33] S. Boixo, S. T. Flammia, C. M. Caves, and J. M. Geremia, Generalized limits for single-parameter quantum estimation, Phys. Rev. Lett. **98**, 090401 (2007).

[34] Clearly, with $\mu = 1$, we are not guaranteed the existence of an unbiased estimator, so there is some subtlety in this restriction. The choice is sufficient for determining bounds and optimal probe states, but, when considering measurements to extract the quantity of interest, realistic protocols must use more than one shot. For instance, robust phase estimation allows for $\mu = \mathcal{O}(1)$, while still allowing us to obtain an unbiased estimator that achieves the quantum Cramér-Rao bound up to a multiplicative constant [38–40]. In Appendix G, for completeness, we briefly summarize this approach. See also Refs. [44,45] and Ref. [18] for further discussion of these issues.

[35] Note the use of a minimax as opposed to a maximin in Eq. (9). This follows from the fact that the minimax of some objective function is always greater than or equal to the maximin and we seek to maximize the quantum Fisher information.

[36] J. Liu, H. Yuan, X. Lu, and X. Wang, Quantum Fisher information matrix and multiparameter estimation, J. Phys. A: Math. Theor. **53**, 023001 (2020).

[37] M. Hayashi, S. Vinjanampathy, and L. Kwek, Resolving unattainable Cramér–Rao bounds for quantum sensors, J. Phys. B: At. Mol. Opt. Phys. **52**, 015503 (2019).

[38] S. Kimmel, G. H. Low, and T. J. Yoder, Robust calibration of a universal single-qubit gate set via robust phase estimation, Phys. Rev. A **92**, 062315 (2015).

[39] S. Kimmel, G. H. Low, and T. J. Yoder, Erratum: Robust calibration of a universal single-qubit gate set via robust phase estimation [Phys. Rev. A **92**, 062315 (2015)], Phys. Rev. A **104**, 069901(E) (2021).

[40] F. Belliardo and V. Giovannetti, Achieving Heisenberg scaling with maximally entangled states: An analytic upper bound for

the attainable root-mean-square error, Phys. Rev. A **102**, 042613 (2020).

[41] W. Górecki, R. Demkowicz-Dobrzański, H. M. Wiseman, and D. W. Berry, π-corrected Heisenberg limit, Phys. Rev. Lett. **124**, 030501 (2020).

[42] S. Arora and B. Barak, *Computational Complexity: A Modern Approach* (Cambridge University Press, Cambridge, 2009).

[43] A. Chiruvelli and H. Lee, Parity measurements in quantum optical metrology, J. Mod. Opt. **58**, 945 (2011).

[44] Y. Yang, G. Chiribella, and M. Hayashi, Attaining the ultimate precision limit in quantum state estimation, Commun. Math. Phys. **368**, 223 (2019).

[45] J. Suzuki, Y. Yang, and M. Hayashi, Quantum state estimation with nuisance parameters, J. Phys. A: Math. Theor. **53**, 453001 (2020).