SAND2018-2493J

Optimal receivers for distributed quantum sensing

T.J. Volkoff[1,*] and Mohan Sarovar[2,†]

[1] Department of Physics, Konkuk University, Seoul 05029, Korea

[2] Extreme-scale Data Science and Analytics, Sandia National Laboratories, Livermore, California 94550, USA

We study the problem of estimating a function of many parameters acquired by sensors that are distributed in space, *e.g.,* the spatial gradient of a field. We restrict ourselves to a setting where the distributed sensors are probed with experimentally practical resources, namely, field modes in separable displaced thermal states, and focus on the optimal design of the optical receiver that measures the phase-shifted returning field modes. We prove that the locally optimal measurement strategy that achieves the standard quantum limit in this setting is a Gaussian measurement, and moreover, one that is separable. We also demonstrate the utility of adaptive phase measurements for making estimation performance robust in cases where one has little prior information on the unknown parameters. In this setting we identify a regime where it is beneficial to use structured optical receivers that entangle the received modes before measurement.

The technical maturity and low cost of a variety of sensors has made distributed sensor networks ubiquitous [1]. Such sensor networks are advantageous for extracting and processing a variety of spatially distributed information to achieve tasks such as boundary detection and precise estimation of spatially varying fields. With the rapid maturation and miniaturization of quantum sensing technologies, *distributed quantum sensing* is naturally emerging as a technological possibility. However, there are still open questions regarding the extent to which quantum sensors can provide performance improvements for distributed sensing problems.

In the distributed sensing context, one can have two types of quantum sensors. In the first type, each of the N sensing nodes in a network could operate quantum mechanically, but independently of all other nodes, while in the second type, all sensing nodes could be coherently linked, *e.g.,* by sharing an entangled state or by being jointly measured by an entangling measurement. For the first type, any quantum enhancement in performance is the same as in the non-distributed setting since one just has N independent sensors. For the second type, there is potential for a quantum-enhancement for sensing distributed properties due to shared quantum resources, and we will focus on this case here. In this context, Proctor *et al.* have recently shown that in a network where the quantum state of each sensing node is dependent on a separate parameter, whether there is a benefit to using quantum resources (such as entanglement across the nodes or an entangling measurement) depends on the form of the distributed quantity one is interested in sensing [2]. In particular, they show by computing the quantum Fisher information (QFI), that if the goal is to estimate all parameters, there is no benefit to using quantum resources, but that if the goal is to estimate a global (non-local) function of the parameters, then one can obtain a $1/N$ enhancement in precision by initializing all sensor nodes in a quantum entangled state.

While the QFI optimized over input states yields the ultimate bound on asymptotic estimation variance, it can be misleading if the measurements required to achieve this bound are not considered since these measurements may be unfeasible under practical constraints. Moreover, the QFI-optimal input states are usually non-classical (and sometimes entangled) states, and preparing many remote quantum sensors in non-classical states (or probing many sensors with entangled probe states) will be technically challenging in the near-term.

Motivated by these considerations, in this work we consider a practical variant of the distributed quantum sensing problem, and quantify the benefits of using entangling measurements to estimate functions of distributed parameters. In particular, we consider a scenario where N quantum sensors are interrogated by separable states that can be measured jointly after interacting with the sensors, see Fig. 1. Although such a setting is strictly less powerful than the more general one where one also allows for entangled probe states [3], it is more practical in the near-term, where constructing joint measurements is more technically feasible. We explicitly construct the ideal and resource-constrained (Gaussian, including adaptive Gaussian) measurement strategies for distributed sensing, and show that separable Gaussian measurements can achieve the standard quantum limit in this setting. This is surprising because in a similar physical setting, joint measurements enabled by structured optical receivers have been shown to be necessary for achieving classical communication capacities [4, 5]. Finally, we identify a special case where a mismatch in prior information about the distributed parameters yields a benefit to using a structured optical receiver that entangles the received light.

Setting: Consider N sensors that are individually probed by N optical probes, each of which is initially in a displaced thermal state and acquires a phase shift θ_i, see Fig. 1. The N modes are collected by a receiver, which also has a local phase reference, and the goal is to estimate a function $f(\theta_1, \theta_2, ..., \theta_N)$ of all the parameters. The classical strategy is to measure each mode separately and compute the function f from the measure-

* volkoff@konkuk.ac.kr
† mnsarov@sandia.gov

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

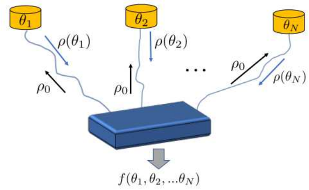

FIG. 1. Schematic of the distributed quantum sensing setting considered here.

ment results. We ask if performing a joint measurement on the N modes (plus the phase reference mode) is of any benefit. Such a setting is relevant to any experimental scenario where information is imprinted in the phase of optical probes. Two examples are: laser phase-shift based range finding [6] and off-resonant optical probing of an array of neutral atoms encoding sensed information in clock state populations [7].

Two-mode, noiseless case: We first consider the case $N = 2$ with no propagation loss or measurement noise in order to present the main concepts. The probe state is a two mode displaced thermal state, $\rho_{\text{in}} = D(\alpha_1, \alpha_2)\rho_{\beta_1} \otimes \rho_{\beta_2} D^\dagger(\alpha_1, \alpha_2)$, where $D(\alpha_1, \alpha_2) := e^{\sum_{j=1}^2 \alpha_j a_j^\dagger - \overline{\alpha_j} a_j}$ is the two-mode displacement operator, $\rho_\beta := (1 - e^{-\beta}) \sum_{n=0}^\infty e^{-\beta n} |n\rangle\langle n|$ is a centered, thermal state, and we take $\alpha_j \in \mathbb{R}$ for simplicity. The phase shifted state received by the receiver is then $\rho_{\vec{\theta}} = U_{\vec{\theta}} \rho_{\text{in}} U_{\vec{\theta}}^\dagger$, where $i \ln U_{\vec{\theta}} = \theta_1 a_1^\dagger a_1 + \theta_2 a_2^\dagger a_2 =: H(\vec{\theta})$. Note that ρ_{in} and $\rho_{\vec{\theta}}$ are both two-mode Gaussian states [8].

To motivate the Gaussian measurements considered later, let us first derive the unconstrained optimal question (*i.e.*, two element projection-valued measurement) for estimation of the phase difference, $\varphi_1 := \frac{\theta_1 - \theta_2}{\sqrt{2}}$ between the two modes. We compute the QFI and optimal measurement that saturates it for this case by explicitly computing the symmetric logarithmic derivative (SLD) [9]:

$$L_{\varphi_1} = \frac{1}{\sqrt{2}} \sum_{j=1}^{2} (-1)^{j+1} \frac{\alpha_j}{N_j + \frac{1}{2}} \left(i e^{i\theta_j} a_j + h.c. \right) \quad (1)$$

where $N_j := \langle a_j^\dagger a_j \rangle_{\rho_{\vec{\theta}}} = (e^{\beta_j} - 1)^{-1}$. For background on the quantum Cramér-Rao bound, the symmetric logarithmic derivative (SLD), and estimation of bosonic Gaussian states, see Ref. [10]. The QFI is independent of θ_1 and θ_2 and has the value $\text{tr} L_{\varphi_1}^2 \rho_{\vec{\theta}} = \sum_{j=1}^2 \frac{\alpha_j^2}{N_j + \frac{1}{2}}$. When $N_1 = N_2 = 0$ this quantity is $2\alpha_1^2 + 2\alpha_2^2$, which is the standard quantum limit (SQL) for estimation of φ_1 with separable probe states having total intensity $\bar{n}_{\text{tot}} = \alpha_1^2 + \alpha_2^2$ [11] [12]. Furthermore, for $N_1 = N_2 = 0$ and $\alpha_1 = \alpha_2 =: \alpha$, the SLD L_{φ_1} can be replaced by a rank 2 self-adjoint operator $P L_{\varphi_1} P$ given by projecting L_{φ_1} on both sides, such that $P L_{\varphi_1} P = 2\partial_{\varphi_1} U_{\vec{\theta}} (|\alpha\rangle \otimes |\alpha\rangle\langle\alpha| \otimes \langle\alpha|) U_{\vec{\theta}}^\dagger = i[a_1^* a_1 - a_2^* a_2, |\alpha e^{-i\theta_1}\rangle \otimes |\alpha e^{-i\theta_2}\rangle\langle\alpha e^{-i\theta_1}| \otimes \langle\alpha e^{-i\theta_2}|]$ (see Supplemental Material for generic construction of P). At $\theta_1 = \theta_2 = 0$, the optimal question is given by the spectral projections $\{|\xi_\pm\rangle\langle\xi_\pm|\}$ of $P L_{\varphi_1} P|_{\vec{\theta}=0}$, where $|\xi_\pm\rangle := \frac{|e_1\rangle \pm i |e_2\rangle}{\sqrt{2}}$ and

$$|e_1\rangle := |\alpha\rangle \otimes |\alpha\rangle, \qquad |e_2\rangle := \frac{\left(a_2^\dagger a_2 - a_1^\dagger a_1\right) |\alpha\rangle \otimes |\alpha\rangle}{\alpha\sqrt{2}}$$

are orthogonal states. The state $|e_2\rangle$ is a superposition of photon-added coherent states. Therefore, the optimal question requires projection onto entangled non-Gaussian states, suggesting that highly non-trivial quantum resources are necessary to achieve the SQL. However, we proceed to show that a separable Gaussian measurement can approach the same performance.

We denote by $z := (x_1, y_1, x_2, y_2)^\mathsf{T}$ the column vector of coordinates on $\mathbb{R}^4$, $R := (q_1, p_1, q_2, p_2)$ the row vector of canonical observables that satisfy the Heisenberg uncertainty principle $[Rz, Rz'] = i z^\mathsf{T} \Delta z' \mathbb{I}_4$ for all z, $z' \in \mathbb{R}^4$ ($\Delta := \oplus_{j=1}^2 i\sigma_y$ is the standard symplectic form on $\mathbb{R}^4$ and we have taken $\hbar = 1$), and $W(z) := e^{iRz}$ is a unitary operator that defines the Weyl form of the canonical commutation relations via $W(z)W(z') = e^{-\frac{i}{2} z^\mathsf{T} \Delta z'} W(z + z')$. $W(z)$ is equal to the two-mode quantum optical displacement operator $D(\alpha_1, \alpha_2)$ if one takes $z = (\sqrt{2}\text{Im}\alpha_1, -\sqrt{2}\text{Re}\alpha_1, \sqrt{2}\text{Im}\alpha_2, -\sqrt{2}\text{Re}\alpha_2)^\mathsf{T}$. A Gaussian quantum state S on two modes of the electromagnetic field is associated with a mean vector $m_S := \text{tr} S R$ and a 4×4 covariance matrix $(\Sigma_S)_{i,j} := \text{tr} S((R_i - m_i) \circ (R_j - m_j))$, where $\circ$ denotes the Jordan product. An energy-constrained Gaussian measurement (ECGM) on two modes is defined by $E \geq 0$ and a positive operator-valued measure $M_S(d^4 z) := W(z) S W(-z) d^4 z$ with symplectic outcome space $\mathbb{R}^4$ such that S is a two-mode, centered Gaussian state (i.e., $m_S = \text{tr} S R = (0, 0, 0, 0)$) and $\text{tr} S \sum_{j=1}^2 a_j^\dagger a_j = E$. Due to the fact that S is centered, the energy constraint can be rewritten $\frac{1}{2} \text{Tr} \Sigma_S - 1 = E$ [13]. Note that when $E = 0$, this ECGM simply describes a heterodyne measurement, and similary, when $E \to \infty$, it describes a homodyne measurement. We will refer to these as the heterodyne and homodyne limits, respectively. For $0 < E < \infty$, the ECGM prescribes projection onto a state with finite squeezing along some quadrature of a mode, which is practically implemented as an adaptive phase measurement [14]. Hence, the parameter E enables us to consider the full class of Gaussian measurements, including adaptive strategies.

We consider the single-parameter estimation problem with Cramér-Rao bound defined by the Fisher information $\tilde{F}(\rho_{\vec{\theta}})_{1,1} := (J^\mathsf{T} F(\rho_{\vec{\theta}}) J)_{1,1}$, where $F(\rho_{\vec{\theta}})$ is the Fisher information metric on the two-dimensional tangent subspace spanned by $(\partial_{\theta_1}, \partial_{\theta_2})$ at the probability density $p_{\vec{\theta}}(z) := \text{tr} W(z) S W(-z) \rho_{\vec{\theta}}$, and J is the Jacobian matrix of the transformation from (θ_1, θ_2) to

$(g_1(\theta_1,\theta_2), g_2(\theta_1,\theta_2))$. Explicitly, $\tilde{F}_{1,1}$ is given by [9]

$$\tilde{F}_{1,1}(\rho_{\vec{\theta}}) = \frac{1}{4}\left(\mathrm{tr}\left(\left(\partial_{\theta_1} g_1 \partial_{\theta_1}\Sigma_{\rho_{\vec{\theta}}} + \partial_{\theta_1} g_2 \partial_{\theta_2}\Sigma_{\rho_{\vec{\theta}}}\right)\Sigma^{-1}\right)\right)^2 + w_{1,1}\Sigma^{-1}w_{1,1}^{\mathsf{T}}. \quad (2)$$

where $\Sigma := \Sigma_{\rho_{\vec{\theta}}} + \Sigma_S$, $w_{1,1} := (\partial_{\theta_1} g_1 \partial_{\theta_1} m_{\rho_{\vec{\theta}}} + \partial_{\theta_1} g_2 \partial_{\theta_2} m_{\rho_{\vec{\theta}}}) \in \mathbb{R}^4$, and $\Sigma_{\rho_{\vec{\theta}}}$ and $m_{\rho_{\vec{\theta}}}$ are the covariance matrix and mean vector defining the state $\rho_{\vec{\theta}}$.

We now specialize to the case of phase-difference estimation, $g_j = \frac{1}{\sqrt{2}}(\theta_1 + (-1)^j\theta_2)$ for which Eq. (2) becomes $\tilde{F}_{1,1} = \frac{1}{4}(F_{1,1} + F_{2,2} - 2F_{1,2})$. When the probe state is a two-mode thermal state, *i.e.,* of the form ρ_{in}, and only when it is so, the covariance matrix $\Sigma_{\rho_{\vec{\theta}}}$ is independent of θ_1, θ_2. Explicitly, $\Sigma_{\rho_{\vec{\theta}}} = \oplus_{j=1}^2(N_j + \frac{1}{2})\mathbb{I}_2$, and $m_{\rho_{\vec{\theta}}} = (\sqrt{2}\mathrm{Re}\alpha, \sqrt{2}\mathrm{Im}\alpha, \sqrt{2}\mathrm{Re}\alpha, \sqrt{2}\mathrm{Im}\alpha)V_{\theta_1}\oplus V_{\theta_2}$, where $V_{\theta_j} = \begin{pmatrix}\cos\theta_j & \sin\theta_j \\ -\sin\theta_j & \cos\theta_j\end{pmatrix}$. In this case, $\tilde{F}_{1,1}$ simplifies to

$$\tilde{F}_{1,1} = \frac{1}{2}((\partial_{\theta_1} - \partial_{\theta_2})m_{\rho_{\vec{\theta}}})\Sigma^{-1}((\partial_{\theta_1} - \partial_{\theta_2})m_{\rho_{\vec{\theta}}})^{\mathsf{T}}. \quad (3)$$

We specialize to an isothermal ($\beta_1 = \beta_2 = \beta$), path-symmetric ($\alpha_1 = \alpha_2 = \alpha$) signal, *i.e.,* $\rho_{\vec{\theta}} = U_{\vec{\theta}} D(\alpha,\alpha)\rho_\beta \otimes \rho_\beta D(\alpha,\alpha)^\dagger U_{\vec{\theta}}^\dagger$, without sacrificing any important features of the problem. We seek to maximize $\tilde{F}_{1,1}$ over Σ_S in the case that the state S that defines the ECGM is a pure, two-mode Gaussian state, i.e., $\Sigma_S = \frac{1}{2}T^{\mathsf{T}}T$ for $T \in Sp(4,\mathbb{R})$. Under these assumptions, it follows that $\Sigma = (N_0 + \frac{1}{2})\mathbb{I}_4 + \frac{1}{2}T^{\mathsf{T}}T$, where $N_0 := (e^\beta - 1)^{-1}$. Because $\Sigma^{-1} > 0$, there exists an orthogonal matrix O that takes the eigenvector corresponding to the maximal eigenvalue of Σ^{-1} to the direction $(\partial_{\theta_1} - \partial_{\theta_2})m_{\rho_{\vec{\theta}}}$. Because $[O, c\mathbb{I}_4] = 0$ for any constant c, we may conjugate Σ by the adjoint action of O to achieve the maximum value of $\tilde{F}_{1,1}$, *i.e.,*

$$\begin{aligned}
&\max_{\substack{T\in Sp(4,\mathbb{R}) \\ \frac{1}{4}\mathrm{tr}T^{\mathsf{T}}T - 1 = E}} \tilde{F}_{1,1} \\
&= \max_{\substack{T\in Sp(4,\mathbb{R}) \\ \frac{1}{4}\mathrm{tr}T^{\mathsf{T}}T - 1 = E}} \frac{1}{2}\|(\partial_{\theta_1} - \partial_{\theta_2})m_{\rho_{\vec{\theta}}}\|^2\,\|\Sigma^{-1}\| \\
&= \max_{\substack{T\in Sp(4,\mathbb{R}) \\ \frac{1}{4}\mathrm{tr}T^{\mathsf{T}}T - 1 = E}} 2\alpha^2 \left\| \left(\left(N_0 + \frac{1}{2}\right)\mathbb{I} + \frac{1}{2}T^{\mathsf{T}}T\right)^{-1} \right\|. \quad (4)
\end{aligned}$$

We refer to this quantity, the Fisher information maximized over all Gaussian measurements, as the *Gaussian Fisher information* (GFI), and it is obviously upper bounded by the QFI. It follows from the Euler decomposition of $Sp(4,\mathbb{R})$ [15] and the fact that $\|O^{\mathsf{T}}\Sigma^{-1}O\| = \|\Sigma^{-1}\|$ that we may restrict attention to $\Sigma_S = \mathrm{diag}(e^{-2r_1}/2, e^{2r_1}/2, e^{-2r_2}/2, e^{2r_2}/2)$, $r_j \in \mathbb{R}$, such that $\sum_{j=1}^2 \sinh^2 r_j = E$. We then have that $\|\Sigma^{-1}\| = (N_0 + \frac{1}{2} + \frac{1}{2}e^{-\max\{r_1,r_2\}})^{-1}$, from which it follows that the constrained maximum of $\tilde{F}_{1,1}$ occurs when all the energy is invested into a single mode. The resulting maximum Fisher information is given by

$$\max_{\substack{T\in Sp(4,\mathbb{R}) \\ \frac{1}{4}\mathrm{tr}T^{\mathsf{T}}T - 1 = E}} \tilde{F}_{1,1} = \frac{2\alpha^2}{N_0 + 1 + E - \sqrt{E^2 + E}} \quad (5)$$

This is the GFI for the phase difference parameter. Note that in the homodyne limit (*i.e.,* $E \to \infty$), this quantity limits to $4\alpha^2/(2N_0+1)$, which coincides with the QFI, see discussion after Eq. (1). Hence, the optimal estimation strategy is achievable by a Gaussian measurement.

It remains to identify the ECGM that achieves the optimal value in Eq. (5). An arbitrary pure, centered, two mode Gaussian state S can be written as $S = |\Xi\rangle\langle\Xi|$ with $|\Xi\rangle := e^{i\sum_{j=1}^2 \phi_j a_j^\dagger a_j} U_\zeta e^{\sum_{j=1}^2 \frac{r_j}{2}(a_j^2 - a_j^{\dagger 2})}|0\rangle \otimes |0\rangle$, with $r_j \in \mathbb{R}$, and $U_\zeta := e^{\zeta a_1^\dagger a_2 - \bar{\zeta} a_2^\dagger a_1}$ being a beam-splitter (ζ is an angle in the closed complex disk with center 0 and radius $\pi/2$) [16]. We set $\phi_j = \mathrm{Arg}\zeta = 0$ because these parameters do not impact the GFI and hence can be set arbitrarily when defining the optimal measurement. Utilizing this explicit form for the ECGM, the energy constrained maximization of Eq.(3) at the parameter values $\theta_1 = \theta_2 = 0$ [17] reduces to maximization of the expression

$$2\alpha^2 \sum_{j=1}^2 \frac{1 + (-1)^{j+1}\sin 2|\zeta|}{2N_0 + 1 + e^{-2r_j}} \quad (6)$$

subject to $\sum_{j=1}^2 \sinh^2(r_j) = E$. Eq. (6) achieves the value in Eq. (5) when $|\zeta| = \pi/4$ and when all the energy is invested in squeezing a single mode, *i.e.,* $r_1 = \sinh^{-1}\sqrt{E}$ and $r_2 = 0$. In the homodyne limit, this corresponds to a homodyne measurement of $a_1 - a_2$, which is obviously an entangling measurement of the two received modes. In fact, the entanglement entropy in $S = |\Xi\rangle\langle\Xi|$, takes the value $H(\mathrm{tr}_2 S) = g(\frac{1}{2}(\sqrt{E+1} - 1))$, where $g(x) := (x+1)\log_2(x+1) - x\log_2 x$.

Comparison to separable strategy: Having identified the optimal ECGM, we now compare this to the best separable Gaussian strategy, where each received mode is measured separately subject to a total energy constraint. We maximize Eq. (3) over separable, pure, centered S, *i.e.,* $S = |\Phi\rangle\langle\Phi|$ with $|\Phi\rangle := e^{\sum_{j=1}^2 \frac{r_j}{2}(a_j^2 - a_j^{\dagger 2})}|0\rangle \otimes |0\rangle$, $r_j \in \mathbb{R}$, and $\sum_{j=1}^2 \sinh^2 r_j = E$. The state S is a tensor product of single-mode squeezed states and the restriction to real r_j is possible because a local rotation of S only decreases the maximum constrained value of $\tilde{F}_{1,1}$. Utilizing this explicit form for the separable ECGM, the maximization of Eq.(3) reduces to maximization of $2\alpha^2\sum_{j=1}^2(2N_0 + 1 + e^{-2r_j})^{-1}$ subject to $\sum_{j=1}^2 \sinh^2 r_j = E$. While for finite E this quantity is always less than the QFI, $4\alpha^2/(2N_0+1)$, in the homodyne limit it asymptotes to the QFI. Hence, the SQL is achievable by *separable* homodyne measurements on the two modes. This is surprising for two reasons: (i) the analysis

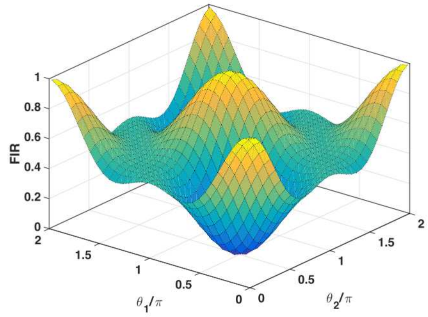

FIG. 2. Ratio between the achievable and maximal Fisher information (the FIR) when the measurement used is formulated assuming parameter values $\theta_1 = \theta_2 = 0$, while the true parameter values are indicated on the axes. $\alpha = 1, N_0 = 0$.

based on the SLD suggested that a non-Gaussian, entangling measurement is necessary to achieve the SQL, and (ii) the proven necessity of entangling measurements to achieve classical communication capacities over bosonic channels [4].

Generalizations: In [9] we generalize the above calculations to the case of N probe modes and estimation of arbitrary linear functions of the parameters θ_i. In such general settings we also find that the optimal GFI can be achieved by an entangling Gaussian measurement in the homodyne limit, *and* a separable Gaussian measurement in the homodyne limit.

Finally, we note that the effects of common imperfections in the transmission channel can be easily incorporated into the above analysis. Transmission through common media such a fibers and free-space is modeled well by a linear bosonic channel that results in loss and injection of thermal noise. These effects simply rescale the amplitude and effective temperature of the received state, $\rho_{\vec{\theta}}$, respectively; *i.e.,* $\alpha \to \eta\alpha$, where $0 \leq \eta \leq 1$, and $N_0 \to N_0 + N_{\text{channel}}$.

Local optimality versus robustness: So far we have shown that separable homodyne measurements achieve the optimized Cramer-Rao bound for distributed sensing with displaced thermal state probes; in essence, the best thing to do is the classical strategy of estimating each parameter separately and then computing the function $f(\theta_1, ...\theta_N)$. However, it is important to note that the QFI analysis results in locally optimal strategies [18]. In particular, the form of the optimal measurement is dependent on the values of the parameters θ_i. In the two mode example, the values of θ_1, θ_2 dictate the local phase parameters ϕ_1, ϕ_2 in the state $|\Xi\rangle$ that determines the optimal measurement. This is not a practical issue if one has a prior distribution over the parameters that is narrow. However, in cases where this is unavailable, or the prior distribution has broad support (*e.g.,* a uniform or maximally uninformative prior) the locally optimal estimation strategy can fail spectacularly. To illustrate this, in Fig. 2 we plot the ratio between (i) the actual Fisher information achieved when applying the optimal measurement (with $E = 10^8$, close to the homodyne limit) formulated for $\theta_1 = \theta_2 = 0$ to a returning state imprinted with different values of θ_i, and (ii) the maximal GFI ($4\alpha^2/(2N_0+1)$). If the actual values of the parameters are significantly different from the assumed ones this Fisher information ratio (FIR) can be less that one, and in some cases go to zero.

We note that a Gaussian strategy that does not suffer from this sensitivity to prior information employs heterodyne measurements for all modes. The Fisher information for heterodyne measurement ($E = 0$) is $4\alpha^2/(2N_0 + 2)$, regardless of whether we allow for entangling, or only separable, measurements. Since this measurement has no dependence on the actual value of the parameters (*i.e.,* $|\Xi\rangle = |\Phi\rangle = |0\rangle \otimes |0\rangle$) the Fisher information remains constant regardless of the actual value of the parameters. However, this lack of sensitivity comes at the cost of a smaller value of Fisher information.

One way to negotiate this trade-off between estimation precision and robustness is to use an adaptive measurements ($0 < E < \infty$) that smoothly interpolate between heterodyne (which prefers no quadrature) and homodyne (which prefers one particular quadrature). In this sense, E can be considered a parameter that quantifies the degree of confidence in the prior information on the parameters. This also suggests a scenario where there is a benefit to using a structured optical receiver. Namely, consider a setting where one is very uncertain about the distribution of the individual parameters θ_i, but has a narrow prior on the collective parameter $f(\theta_1, ...\theta_N)$. If one is concerned with minimizing uncertainty in estimation precision (*e.g.,* quantified by the variance in Fisher information) then the best separable strategy is to use heterodyne measurements on all modes, in which case the Fisher information is $4\alpha^2/(2N_0 + 2)$. However, if one employs an entangling measurement that concentrates the collective parameter into a single mode, one can exploit the narrow prior on this parameter and apply a homodyne measurement on this mode to attain the optimized Fisher information for the estimation problem $4\alpha^2/(2N_0 + 1)$. Although this is only a constant gain in estimation precision it could be beneficial in extremely low-power, low-noise applications where $\alpha^2 \ll 1$ and $N_0 \ll 1$.

Conclusions: We have analyzed distributed quantum sensing applications where thermal probe fields are imprinted with phase shifts proportional to distributed parameters, and one is interested in a global function of the parameters. We proved that a separable, Gaussian measurement strategy can achieve the SQL in this setting. In addition, we showed that this is a locally optimal strategy that can result in large estimation performance variation unless one has narrow prior distributions over the parameters. Finally, we highlighted a scenario where one has a mismatch between prior information about the individual parameters and the global function of the parameters, where an entangling measurement can yield some benefit.

Therefore, even when using probe states that are almost classical (separable displaced thermal states), there are situations where one can gain an estimation advantage by utilizing a structured receiver that interferes/entangles the returning light.

ACKNOWLEDGMENTS

T.J.V. thanks Yongkyung Kwon for hosting at Konkuk University during the completion of this work. M.S. thanks Grant Biedermann and Tim Proctor for useful discussions about distributed sensing, and Howard Wiseman for discussions about adaptive measurement POVMs. T.J.V. acknowledges support from the Korea Research Fellowship Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (2016H1D3A1908876) and from the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2015R1D1A1A09056745). Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

SUPPLEMENTAL MATERIAL

Symmetric logarithmic derivative for the phase difference parameter for $N = 2$

Here we present the detailed derivation of SLD for the phase difference parameter in the $N = 2$ case analyzed in the main text. By explicit calculation, one finds that for $\rho_{\vec{\theta}} = D(\alpha_1, \alpha_2)\rho_{\beta_1} \otimes \rho_{\beta_2} D(\alpha_1, \alpha_2)^\dagger$, where $\alpha_j \in \mathbb{R}$, $N_j \geq 0$,

$$\partial_{\theta_j} \rho_{\vec{\theta}} = ie^{i\theta_j} \frac{\alpha_j}{N_j} a_j \rho_{\vec{\theta}} + h.c. \,. \tag{7}$$

However, using the identities

$$\begin{aligned} a_j \rho_{\vec{\theta}} &= \rho_{\vec{\theta}} \left(e^{-\beta_j}(a_j - \alpha_j e^{-i\theta_j}) + \alpha_j e^{-i\theta_j}\right) \\ \rho_{\vec{\theta}} a_j^\dagger &= \left(e^{-\beta_j}(a_j^\dagger - \alpha_j e^{i\theta_j}) + \alpha_j e^{i\theta_j}\right) \rho_{\vec{\theta}} \end{aligned} \tag{8}$$

it follows that $\partial_{\theta_j} \rho_{\vec{\theta}} = \rho_{\vec{\theta}} \circ L_{\theta_j}$, where

$$L_{\theta_j} = \frac{\alpha_j}{N_j + \frac{1}{2}} \left(ie^{i\theta_j} a_j + h.c.\right) = L_{\theta_j}^\dagger \tag{9}$$

are the SLD operators in the ∂_{θ_j} directions. Using the Jacobian to transform two-dimensional tangent subspace $\text{span}\{L_{\theta_j}\}_{j=1,2}$ at $\rho_{\vec{\theta}}$ to the basis $\{L_{\varphi_j}\}_{j=1,2}$ gives Eq.(1).

For a pure state $|\psi\rangle$, the defining equation of the SLD (in the direction ∂_{φ_1}) $\partial_{\varphi_1}|\psi\rangle\langle\psi| = \frac{1}{2}|\psi\rangle\langle\psi|L_{\varphi_1} + \frac{1}{2}L_{\varphi_1}|\psi\rangle\langle\psi|$ combined with the fact that $\langle\psi|L_{\varphi_1}|\psi\rangle = 0$ implies that $P(\partial_{\varphi_1}|\psi\rangle\langle\psi|)P = \partial_{\varphi_1}|\psi\rangle\langle\psi|$, where P is the projection to the two-dimensional complex Hilbert space $\overline{\text{span}\{|\psi\rangle, L_{\varphi_1}|\psi\rangle\}}$ (clearly, P is dependent on $|\psi\rangle$; here we omit the subscripts on P, shown explicitly in the main text, that indicate this dependence). Then, since $[P, |\psi\rangle\langle\psi|] = 0$, it follows that $\partial_{\varphi_1}|\psi\rangle\langle\psi| = \frac{1}{2}|\psi\rangle\langle\psi|PL_{\varphi_1}P + \frac{1}{2}PL_{\varphi_1}P|\psi\rangle\langle\psi|$. Calculation of the spectral projections of $PL_{\varphi_1}P$ amounts to diagonalization of a 2×2 matrix.

Explicit form of the Fisher information in the Gaussian setting

Here we present the detailed derivation of Eq. (2) in the main text. The measure $\mu_{\rho_{\vec{\theta}}}\left(d^4z\right) := \text{tr}\left(\rho_{\vec{\theta}} W(z) S W(-z)\right) \frac{d^4z}{(2\pi)^2}$ on $\mathbb{R}^4$ has density $p_{\vec{\theta}}(z) := \frac{1}{(2\pi)^2}\text{tr}\left(\rho_{\vec{\theta}} W(z) S W(-z)\right)$ with respect to Lebesgue measure. The quantity of interest is the $2{\times}2$ Fisher information metric

$$\begin{aligned} F_{i,j} &:= \int d^4z \, p_{\vec{\theta}}(z) \partial_{\theta_i} \log p_{\vec{\theta}}(z) \partial_{\theta_j} \log p_{\vec{\theta}}(z) \\ &= \int d^4z \, p_{\vec{\theta}}(z)^{-1} \partial_{\theta_i} p_{\vec{\theta}}(z) \partial_{\theta_j} p_{\vec{\theta}}(z) \end{aligned} \tag{10}$$

Actually, for phase difference estimation, we really only care about $\tilde{F}_{1,1}$, where $\tilde{F} = J^\intercal F J$ is the Fisher information metric in terms of φ_1, φ_2 after the change of parameters $\theta_1 = \frac{\varphi_1 + \varphi_2}{\sqrt{2}}$, $\theta_2 = \frac{\varphi_2 - \varphi_1}{\sqrt{2}}$ and $J = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. To calculate $p_{\vec{\theta}}(z)$, we use the expansion of the states over the CCR C*-algebra, e.g., $\rho_{\vec{\theta}} = \int \frac{d^4z_1}{(2\pi)^2} \chi_{\rho_{\vec{\theta}}}(z_1) W(-z_1)$ where $\chi_{\rho_{\vec{\theta}}}(z_1) := e^{-\frac{1}{2}z_1^\intercal \Sigma_{\rho_{\vec{\theta}}} z_1 + i m_{\rho_{\vec{\theta}}}^\intercal z_1}$ is the characteristic function of $\rho_{\vec{\theta}}$ defined by the covariance matrix $\Sigma_{\rho_{\vec{\theta}}}$ and the mean vector $m_{\rho_{\vec{\theta}}}$.

$$\begin{aligned} p_{\vec{\theta}}(z) &= \frac{1}{(2\pi)^2}\mathrm{tr}\int \frac{d^4z_1}{(2\pi)^2}\frac{d^4z_2}{(2\pi)^2}\chi_{\rho_{\vec{\theta}}}(z_1)\chi_S(z_2)W(-z_1)W(z)W(-z_2)W(-z) \\ &= \frac{1}{(2\pi)^2}\mathrm{tr}\int \frac{d^4z_1}{(2\pi)^2}\frac{d^4z_2}{(2\pi)^2}\chi_{\rho_{\vec{\theta}}}(z_1)\chi_S(z_2)e^{-i\Delta(z_2,z)}e^{-\frac{i}{2}\Delta(z_1,z_2)}W(-z_1-z_2) \\ &= \frac{1}{(2\pi)^2}\int \frac{d^4z_1}{(2\pi)^2}\chi_{\rho_{\vec{\theta}}}(z_1)\chi_S(-z_1)e^{i\Delta(z_1,z)} \\ &= \frac{1}{(2\pi)^2}\left(\det\left(\Sigma_{\rho_{\vec{\theta}}}+\Sigma_S\right)\right)^{-1/2}e^{-\frac{1}{2}\left(m_{\rho_{\vec{\theta}}}-m_S-z^\mathsf{T}\Delta\right)\left(\Sigma_{\rho_{\vec{\theta}}}+\Sigma_S\right)^{-1}\left(m_{\rho_{\vec{\theta}}}-m_S+\Delta z\right)^\mathsf{T}} \end{aligned} \tag{11}$$

where, in the third line, we have used $\mathrm{tr}W(z)=(2\pi)^2\delta(z)$ for a two mode system. Now, we calculate $\partial_{\theta_1}p_{\vec{\theta}}(z)$ by using the third line Eq.(11) and a generating function.

$$\begin{aligned} \partial_{\theta_1}p_{\vec{\theta}}(z) = \frac{1}{(2\pi)^2}\int\frac{d^4z_1}{(2\pi)^2}&\Bigg[\Bigg[-\frac{1}{2}\sum_{m,n=1}^{2}(-i\partial_{j_m})[\partial_{\theta_1}\Sigma_{\rho_{\vec{\theta}}}]_{m,n}(-i\partial_{j_n}) \\ &+\sum_{n=1}^{2}i[\partial_{\theta_1}m_{\rho_{\vec{\theta}}}]_n(-i\partial_{j_n})\Bigg] \\ &e^{-\frac{1}{2}z_1^\mathsf{T}\Sigma z_1+i(m_{\rho_{\vec{\theta}}}-m_S-z^\mathsf{T}\Delta)z_1}e^{ij^\mathsf{T}z_1}\Big|_{j=0}\Bigg] \\ &= \frac{1}{(2\pi)^2\sqrt{\det\Sigma}}\left[-\frac{1}{2}\mathrm{tr}\left((\partial_{\theta_1}\Sigma_{\rho_{\vec{\theta}}})\Sigma^{-1}\right)-(\partial_{\theta_1}m_{\rho_{\vec{\theta}}})\Sigma^{-1}(m_{\rho_{\vec{\theta}}}-m_S-z^\mathsf{T}\Delta)^\mathsf{T}\right]p_{\vec{\theta}}(z) \end{aligned} \tag{12}$$

where the Gaussian integral version of Wick's theorem has been used to get the last line. Now, we perform a final integration over z to get the Fisher metric. We will show the off-diagonal element $F_{1,2}$ so that it will be clear how the other elements go. Finally, we will present the result for $\tilde{F}_{1,1}$.

$$\begin{aligned} F_{1,2} &= \int d^4z\, p_{\vec{\theta}}(z)^{-1}\partial_{\theta_1}p_{\vec{\theta}}(z)\partial_{\theta_2}p_{\vec{\theta}}(z) \\ &= \frac{1}{(2\pi)^2}\left(\det\Sigma\right)^{-1/2} \\ &\quad\int d^4z\left[\frac{1}{2}\mathrm{tr}\left(\left(\partial_{\theta_1}\Sigma_{\rho_{\vec{\theta}}}\right)\Sigma^{-1}\right)+(\partial_{\theta_1}m_{\rho_{\vec{\theta}}})\Sigma^{-1}(m_{\rho_{\vec{\theta}}}-m_S-z^\mathsf{T}\Delta)\right] \\ &\quad\left[\frac{1}{2}\mathrm{tr}\left(\left(\partial_{\theta_2}\Sigma_{\rho_{\vec{\theta}}}\right)\Sigma^{-1}\right)+(\partial_{\theta_2}m_{\rho_{\vec{\theta}}})\Sigma^{-1}(m_{\rho_{\vec{\theta}}}-m_S-z^\mathsf{T}\Delta)\right] \\ &\quad e^{-\frac{1}{2}\left(m_{\rho_{\vec{\theta}}}-m_S-z^\mathsf{T}\Delta\right)\Sigma^{-1}\left(m_{\rho_{\vec{\theta}}}-m_S+\Delta z\right)^\mathsf{T}} \end{aligned} \tag{13}$$

Expanding the brackets and noting that: 1) for any $v\in\mathbb{R}^4$, and positive $A\in M_4(\mathbb{R})$, $\int d^4z\, v^\mathsf{T}u\, e^{-\frac{1}{2}u^\mathsf{T}Au}=0$, 2) taking $u=m_{\rho_{\vec{\theta}}}-m_S+\Delta z$ gives

$$\begin{aligned} F_{1,2} &= \left(\frac{1}{4}\mathrm{tr}\left(\left(\partial_{\theta_1}\Sigma_{\rho_{\vec{\theta}}}\right)\Sigma^{-1}\right)\,\mathrm{tr}\left(\left(\partial_{\theta_2}\Sigma_{\rho_{\vec{\theta}}}\right)\Sigma^{-1}\right)\right) \\ &\quad+\frac{(\det\Sigma)^{-1/2}}{(2\pi)^2}\int d^4u\left((\partial_{\theta_1}m_{\rho_{\vec{\theta}}})\Sigma^{-1}u^\mathsf{T}\right)\left((\partial_{\theta_2}m_{\rho_{\vec{\theta}}})\Sigma^{-1}u^\mathsf{T}\right)e^{-\frac{1}{2}u\Sigma^{-1}u^\mathsf{T}}. \end{aligned} \tag{14}$$

Using the identity

$$\frac{(\det\Sigma)^{-1/2}}{(2\pi)^2}\left[\sum_{n_1,n_2=1}^{2}(L)_{n_1}(-i\partial_{j_{n_1}})(Q)_{n_2}(-i\partial_{j_{n_2}})\right]\int d^4u e^{-\frac{1}{2}u^\mathsf{T}\Sigma^{-1}u}e^{iu^\mathsf{T}j}\Big|_{j=0}=L\Sigma Q^\mathsf{T} \tag{15}$$

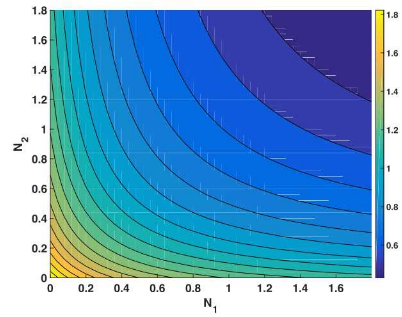

FIG. 3. Maximal value of $\tilde{F}_{1,1}$ for probe state $\rho_{\text{in}} = D(\alpha,\alpha)\rho_{\beta_1} \otimes \rho_{\beta_2} D^\dagger(\alpha,\alpha)$, with $\alpha = 1$. The average thermal photon numbers $N_i = (e^{\beta_i} - 1)^{-1}$ are swept across the two axes. The optimal ECGM is obtained for the energy constraint $E = 4$. $\tilde{F}_{1,1}$ is symmetric about $N_1 = N_2$.

for row vectors L, $Q \in \mathbb{R}^4$, allows one to simplify the last line of Eq.(14).

$$F_{1,2} = \left(\frac{1}{4}\text{tr}\left(\left(\partial_{\theta_1}\Sigma_{\rho_{\vec{\theta}}}\right)\Sigma^{-1}\right)\,\text{tr}\left(\left(\partial_{\theta_2}\Sigma_{\rho_{\vec{\theta}}}\right)\Sigma^{-1}\right)\right) + (\partial_{\theta_1} m_{\rho_{\vec{\theta}}})\Sigma^{-1}(\partial_{\theta_2} m_{\rho_{\vec{\theta}}})^\intercal \tag{16}$$

For $F_{1,1}$, $F_{2,2}$, just use the appropriate ∂_{θ_j}. From the transformation $\tilde{F}_{i,j} = (J^\intercal F J)_{i,j}$ that arises from an arbitrary diffeomorphism $\vec{\theta} = (\theta_1, \theta_2) \mapsto (g_1(\theta_1,\theta_2), g_2(\theta_1,\theta_2))$, one finds that

$$\tilde{F}_{i,j} = F_{1,1}(\partial_{\theta_i} g_1)(\partial_{\theta_j} g_1) + F_{2,2}(\partial_{\theta_i} g_2)(\partial_{\theta_j} g_2) + F_{1,2}(\partial_{\theta_i} g_1 \partial_{\theta_j} g_2 + \partial_{\theta_i} g_2 \partial_{\theta_j} g_1). \tag{17}$$

Eq.(2) in the main text follows immediately from Eq.(16) and Eq.(17).

Non-isothermal probe states

In the main text the probe state is assumed to be isothermal (*i.e.,* the temperature of all modes is the same). In this section we compute the GFI for non-isothermal states for completeness.

Consider the $N = 2$ case, and path-symmetric, non-isothermal ($\beta_1 \neq \beta_2$) probes for phase difference estimation. Because $\Sigma_{\rho_{\vec{\theta}}}$ is no longer a constant multiple of the identity matrix, the arguments leading to Eq.(4) in the main text cannot be applied. In this case, it is most convenient to carry out constrained numerical optimization of Eq.(3) over states S defined by $S = |\Xi\rangle\langle\Xi|$ with $|\Xi\rangle$ as defined in the main text, and Fig. 3 presents the results of this calculation. It is clear from this data that the maximal value of $\tilde{F}_{1,1}$ decreases most rapidly for uniform probe state noise. The entanglement entropy of the optimal S (not shown) satisfies the following properties: 1) it is constant along the $N_1 = N_2$ line and in agreement with Eq.(**??**) for all values of E, and 2) it monotonically decreases from the $N_1 = N_2$ value along the quarter circle of radius $\sqrt{N_1^2 + N_2^2}$.

Scaling analysis and entanglement gain

We consider the general setting where one has N sensors, each probed by a displaced thermal state that picks up a phase shift θ_i. We restrict to computing linear functions of the phases, and define the quantity of interest as the first component of the function $g(\vec{\theta}) = (\vec{v}_1 \cdot \vec{\theta}, \dots, \vec{v}_N \cdot \vec{\theta})$ where $\{\vec{v}_k\}_{k=1,\dots,N}$ is an orthonormal set in $\mathbb{R}^N$. The Jacobian matrix of the map $\vec{\theta} \mapsto [g_1(\vec{\theta}), \dots, g_N(\vec{\theta})]$ is given by $J := [\vec{v}_1, \dots, \vec{v}_N]$. Then, from $\tilde{F} = J^\intercal F J$, we get

$$\tilde{F}_{1,1} = (\vec{v}_1 \cdot \nabla_\theta) m_{\rho_{\vec{\theta}}} \left(\Sigma_{\rho_{\vec{\theta}}} + \Sigma_S\right)^{-1} \left((\vec{v}_1 \cdot \nabla_\theta) m_{\rho_{\vec{\theta}}}\right)^\intercal.$$

We seek to maximize this quantity, subject to the energy constraint $\langle \sum_{j=1}^{N} a_j^\dagger a_j \rangle_S = E$; *i.e.,*

$$\max_{\substack{T\in Sp(2N,\mathbb{R})\\ \frac{1}{4}\mathrm{tr}T^\mathsf{T}T-\frac{N}{2}=E}} \|(\vec{v}_1\cdot\nabla_\theta)m_{\rho_{\vec{\theta}}}\|^2 \|(\Sigma_{\rho_{\vec{\theta}}}+\frac{1}{2}T^\mathsf{T}T)^{-1}\|. \tag{18}$$

This can be solved in the same way as the two mode case if we take the isothermal, path symmetric probe state $\rho_{\text{in}} = D(\alpha,\ldots,\alpha)\rho_\beta^{\otimes N}D(\alpha,\ldots,\alpha)^\dagger$, since in this case $\|(\Sigma_{\rho_{\vec{\theta}}}+\frac{1}{2}T^\mathsf{T}T)^{-1}\| = \|((N_0+(1/2))\mathbb{I}_{2N}+\Sigma_S)^{-1}\|$, and we can assume that Σ_S is diagonal, *i.e.,* is a tensor product of squeezed states with squeezing of the q or p quadrature only.. Clearly, the matrix norm will be maximized if all the squeezing is in one mode (*i.e.,* all the energy is used for squeezing), and we have $\|((N_0+(1/2))\mathbb{I}_{2N}+\Sigma_S)^{-1}\| = ((N_0+(1/2)+\frac{e^{-2r}}{2})^{-1}$, where $\sinh^2 r = E$. Rewriting, and using the fact that $m_{\rho_{\vec{\theta}}} = m_{\rho_{\vec{\theta}=0}} \bigoplus_{j=1}^{N} V_{\theta_j}$, gives:

$$\begin{aligned}\max_{\substack{T\in Sp(2N,\mathbb{R})\\ \frac{1}{4}\mathrm{tr}T^\mathsf{T}T-\frac{N}{2}=E}} \tilde{F}_{1,1} &= \frac{\|(\vec{v}_1\cdot\nabla_\theta)m_{\rho_{\vec{\theta}}}\|^2}{\left(N_0+1+E-\sqrt{E^2+E}\right)} \\ &= \frac{2\alpha^2}{\left(N_0+1+E-\sqrt{E^2+E}\right)}\end{aligned} \tag{19}$$

This quantity has no dependence on the number of modes, N. To understand the import of this we must compare this quantity to the GFI when one is limited to separable measurements. To do so, we maximize $\tilde{F}_{1,1}$ under the restriction of energy constrained separable measurements, see [19] for details. We define the ratio of Eq. (19) to the maximum achieved by separable strategies, the *entanglement gain* (EG). The best separable strategy actually depends on the structure of $\vec{v}_1$; if $\vec{v}_1$ is dominated by one entry (the *unbalanced* case), say $(\vec{v}_1)_1$, then its preferable to invest most of the energy available for measurement into measuring the first mode. In contrast, if $\vec{v}_1$ contains entries of almost equal magnitude (the *balanced* case), then the best separable strategy distributes the energy available for measurement among all N modes. As the homodyne limit is taken ($E\to\infty$), the entanglement gain asymptotes to $EG^{\text{unbal}} \to \frac{2}{(\vec{v}_1)_1^2+1}$ for the unbalanced case, and $EG^{\text{bal}} \to 1$ for the balanced case. We see that in the general N case also, that separable and entangling Gaussian measurements can achieve the same estimation performance in the homodyne limit.

For S a separable ECGM, from Eq. (18), we have that $\tilde{F}_{1,1} = \sum_{j=1}^{N} 2(\vec{v}_1)_j^2\alpha^2\left(N_0+(1/2)+\frac{e^{-2r_j}}{2}\right)^{-1}$, where $\sum_{j=1}^{2}\sinh^2 r_j = E$. When attempting to maximize this, care must be taken in consideration of the vector $\vec{v}_1$. If the vector $\vec{v}_1$ has a unique entry of maximum magnitude (the unbalanced case), we should put all the energy into the corresponding r_j. Without loss of generality, let $(\vec{v}_1)_1^2 > (\vec{v}_1)_j^2$, $j\neq 1$, and put all the constraint energy into measuring that mode. In that case,

$$\max_{\substack{S\text{ separable, Gaussian}\\ \langle\sum_{j=1}^{N}a_j^\dagger a_j\rangle_S=E}} \tilde{F}_{1,1} = 2\alpha^2\left(\frac{(\vec{v}_1)_1^2}{N_0+1+E-\sqrt{E^2+E}}+\sum_{j=2}^{N}\frac{(\vec{v}_1)_j^2}{N_0+1}\right) \qquad \text{(unbalanced)} \tag{20}$$

In the opposite extreme, let us consider $(\vec{v}_1)_j^2 = 1/N$ (the balanced case), which encompasses the case of two-mode phase difference sensing that is considered in the main text. In this case, the maximum is achieved when the constraint energy is distributed equally for squeezing each mode of the state S that defines the ECGM, and we can show,

$$\max_{\substack{S\text{ separable, Gaussian}\\ \langle\sum_{j=1}^{N}a_j^\dagger a_j\rangle_S=E}} \tilde{F}_{1,1} = \frac{2\alpha^2}{N_0+1+\frac{E}{N}-\sqrt{\left(\frac{E}{N}\right)^2+\frac{E}{N}}} \qquad \text{(balanced)} \tag{21}$$

In the main text, we presented the finite-energy EG for the balanced case, and here we examine the same quantity for the unbalanced case. Note that Eq. (19) in the main text nor Eq. (20) have an explicit dependence on the number of modes, N. However, there is an implicit dependence on this quantity through $(\vec{v}_1)_1$; namely, since this is assumed to be the largest element of the normalized vector $\vec{v}_1$, its magnitude bounds the number of modes, *i.e.,* $N > (1-(\vec{v}_1)_1^2)/(\vec{v}_1)_1^2$. In Fig. 4 we plot the EG as a function of E and $(\vec{v}_1)_1^2$. As seen in this figure, the benefit of an entangling measurement decreases with increasing $(\vec{v}_1)_1^2$; as the function becomes more unbalanced the estimation precision can be minimized by performing separable measurements, and minimizing the variance of the generator of the dominating parameter, θ_1.

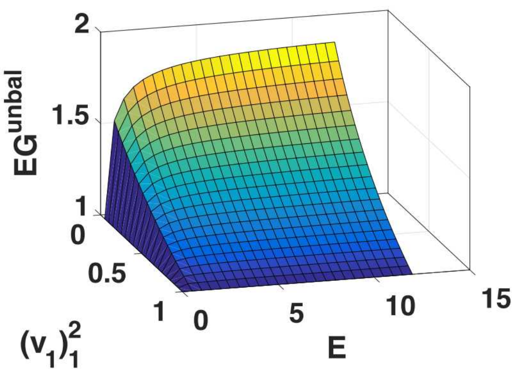

FIG. 4. Entanglement gain for the unbalanced case as a function of the largest element in $\vec{v}_1$, $(\vec{v}_1)_1^2$, and the energy constraint E.

[1] Y.-M. Huang, M.-Y. Hsieh, and F. E. Sandnes, in *Sensors* (Springer, Berlin, 2008) pp. 199–219.

[2] T. J. Proctor, P. A. Knott, and J. A. Dunningham, arXiv.org (2017), 1707.06252v1.

[3] V. Giovannetti, S. Lloyd, and L. Maccone, Nature Photonics **5**, 222 (2011).

[4] S. Guha, Physical Review Letters **106**, 240502 (2011).

[5] K. Banaszek and M. Jachura, arXiv (2017), 1711.02521.

[6] M. H. Conde, in *Compressive Sensing for the Photonic Mixer Device* (Springer Vieweg, Wiesbaden, Wiesbaden, 2017) pp. 11–88.

[7] J. Lodewyck, P. G. Westergaard, and P. Lemonde, Phys. Rev. A **79**, 061401 (2009).

[8] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Reviews of Modern Physics **84**, 621 (2012).

[9] See Supplemental Material for details.

[10] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).

[11] M. Jarzyna and R. Demkowicz-Dobrzański, Phys. Rev. A **85**, 011801 (2012).

[12] As Ref. [11] points out, the SQL for phase difference estimation depends on whether we assume access to a phase reference or not. In our setting, the receiver has a local oscillator that provides a phase reference (*e.g.,* has a fixed phase relationship to ρ_{in}) and hence the SQL in this setting is $\alpha_1^2 + \alpha_2^2$.

[13] We note that in order to construct a measurement $\tilde{M}_S(d\theta)$ with phase-valued outcomes (i.e., outcomes being measurable subsets of $[0, 2\pi)$) which is directly useful for estimation of a relative phase at a certain point in quantum state space, one must push forward the Gaussian measurement $M_S(d^4z)$ via post-processing of the phase space measurement outcome. However, the Fisher information and optimal measurement depend only on the probe state and the Gaussian state S that defines the ECGM.

[14] H. M. Wiseman, Quantum Semiclass. Opt. **8**, 205 (1996).

[15] D. Šafránek and I. Fuentes, Phys. Rev. A **94**, 062313 (2016).

[16] A. B. Dutta, N. Mukunda, and R. Simon, Pramana **45**, 471 (1995).

[17] While the GFI does note depend on the values of the parameters, the form of the optimal measurement does.

[18] R. Demkowicz-Dobrzański, Phys. Rev. A **83**, 061802 (2011).

[19] See Supplemental Material for details.