

PAPER • OPEN ACCESS

Quantum sensing networks for the estimation of linear functions

To cite this article: Jesús Rubio *et al* 2020 *J. Phys. A: Math. Theor.* **53** 344001

View the [article online](#) for updates and enhancements.

You may also like

- [Slower deviations of the branching Brownian motion and of branching random walks](#)
Bernard Derrida and Zhan Shi
- [Scalar product for the XXZ spin chain with general integrable boundaries](#)
Samuel Belliard, Rodrigo A Pimenta and Nikita A Slavnov
- [Selective probing of magnetic order on Tb and Ir sites in stuffed \$Tb_2Ir_2O_7\$ using resonant x-ray scattering](#)
C Donnerer, M C Rahn, E Schierle *et al.*

Quantum sensing networks for the estimation of linear functions

Jesús Rubio^{1,5} , Paul A Knott² , Timothy J Proctor³ 
and Jacob A Dunningham⁴ 

¹ Department of Physics and Astronomy, University of Exeter, Stocker Road, Exeter EX4 4QL, United Kingdom

² Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems (CQNE), School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom

³ Sandia National Laboratories, Livermore CA 94550, United States of America

⁴ Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, United Kingdom

E-mail: J.Rubio-Jimenez@exeter.ac.uk and J.Dunningham@sussex.ac.uk

Received 11 March 2020, revised 19 May 2020

Accepted for publication 16 June 2020

Published 3 August 2020



CrossMark

Abstract

The theoretical framework for networked quantum sensing has been developed to a great extent in the past few years, but there are still a number of open questions. Among these, a problem of great significance, both fundamentally and for constructing efficient sensing networks, is that of the role of inter-sensor correlations in the *simultaneous* estimation of multiple linear functions, where the latter are taken over a collection local parameters and can thus be seen as global properties. In this work we provide a solution to this when each node is a qubit and the state of the network is sensor-symmetric. First we derive a general expression linking the amount of inter-sensor correlations and the geometry of the vectors associated with the functions, such that the asymptotic error is optimal. Using this we show that if the vectors are clustered around two special subspaces, then the optimum is achieved when the correlation strength approaches its extreme values, while there is a monotonic transition between such extremes for any other geometry. Furthermore, we demonstrate that entanglement can be detrimental for estimating non-trivial global properties, and that sometimes it is in fact irrelevant. Finally, we perform a non-asymptotic analysis of these results using a Bayesian approach, finding that the amount of correlations needed to enhance the precision crucially depends on the number of

⁵ Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

measurement data. Our results will serve as a basis to investigate how to harness correlations in networks of quantum sensors operating both in and out of the asymptotic regime.

Keywords: quantum sensing networks, multi-parameter estimation, quantum metrology, quantum correlations

(Some figures may appear in colour only in the online journal)

1. Introduction

An important task in quantum information science is to devise protocols for multi-parameter metrology and estimation by exploiting the quantum properties of light and matter. This problem has been widely explored not only in a theoretical fashion [1–22], but also in applications [9, 15, 16, 23–39] and experiments [27, 40–42]. As a result, new practical ways of enhancing our estimation schemes have recently emerged [43–48]. These protocols are normally formulated on the basis of d unknown parameters $\theta = (\theta_1, \dots, \theta_d)$ that arise naturally in the description of the system at hand, and in many cases these are the quantities of interest. However, sometimes we may wish or need to find l new quantities that are functions of θ , that is, $f(\theta) = (f_1(\theta), \dots, f_l(\theta))$. This is the case, in particular, when we analyse global properties in a quantum sensing network [32, 33], which is a model for spatially distributed sensing [46] and the main focus of this work. Indeed, in [32, 33] this model is defined as an array of quantum sensors where one or several parameters are locally encoded in each of them, and while a property of the network is said to be *local* if it is represented by parameters at a single sensor, a *global* property is thought of as a non-trivial function of two or more parameters at different sensors. Here we consider that a single parameter θ_i is encoded in the i th sensor, so that θ is a collection of local properties, and we assume that both parameters and functions are real-valued quantities. See figure 1 for a schematic representation.

Networked scenarios where global properties are relevant provide a natural testbed to identify the potential usefulness of entanglement in a broad range of multi-parameter schemes [32, 37]. Within this context, the optimal estimation of a single function $f(\theta)$ has been extensively studied [32, 33, 37, 46, 49–58], and it has been established that one can find entangled states that beat the best separable probe when that function is linear [32, 33]. In addition, Eldredge *et al* [49] derived a bound on the error for this scenario that was later generalised to accommodate a single analytical function [52], which can also be estimated with an enhanced precision when there is entanglement, while Gross and Caves [59] have reexamined the linear case using an elegant geometric approach. On the opposite extreme, it has been shown that a collection of $l = d$ linear functions that generates an orthogonal transformation (i.e., $f(\theta) = V^T \theta$ with $VV^{-1} = \mathbb{I}$) can be estimated optimally with a local strategy [32, 37].

Beyond these two types of global properties, the simultaneous estimation of $l > 1$ linear but otherwise arbitrary real functions has been a less travelled path. There exist generic bounds for this problem (see, e.g., [32, 60]), which in practice may arise in scenarios such as the estimation of phase differences [29, 60]. However, how quantum correlations may help for linear functions with arbitrary geometry has not been examined in detail. Given that this represents a richer regime than the $l = 1$ and $l = d$ with orthogonal functions cases, it can be argued that answering this question is essential for further progress in networked quantum metrology.

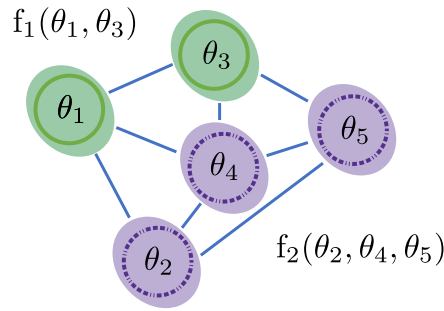


Figure 1. A network of $d = 5$ sensors. The parameters $\theta = (\theta_1, \dots, \theta_5)$ represent local properties, since each of them is locally encoded in a single sensor. On the contrary, $f_1(\theta_1, \theta_3)$ and $f_2(\theta_2, \theta_4, \theta_5)$ are global properties associated with sensors 1 and 3 (green solid lines) and sensors 2, 4 and 5 (purple dashed lines), respectively.

While a general answer is beyond the scope of our methods, here we obtain a definite solution for a subclass of schemes with sensor-symmetric pure qubit states, which we introduce in section 2.1. Using the Helstrom Cramér–Rao bound and the associated quantum Fisher information matrix, in section 3 we derive a general expression linking the geometry of the vector components associated with the functions and the strength of the inter-sensor correlations, such that the uncertainty in the asymptotic regime of many trials is optimal. Moreover, we show that there exists a physical state for many of the optimal configurations that our formula predicts. Equipped with this, we then derive a number of important results. First we find that the largest amounts of correlations are associated, for sensor-symmetric states, with two special subspaces: the direction of the vector of ones $\mathbf{1}^\top \equiv (1, 1, \dots)$, and the subspace orthogonal to it. This connection between entanglement in a pure state and how much the vectors are clustered around certain directions was precisely one of the open questions identified in [32], and our findings contribute towards its solution. In addition, we demonstrate that entanglement can be detrimental for estimating global properties other than those associated with orthogonal transformations, while a three-sensor network reveals that entanglement is sometimes irrelevant. This is consistent with the fact that the asymptotic uncertainty only depends on correlations of a pairwise nature, and thus other forms of entanglement do not affect the asymptotic error.

On the other hand, it is known that strategies with a good asymptotic precision found by optimising the Cramér–Rao bound sometimes have a particularly poor performance when the number of trials is very low (see, e.g., [61]). In fact, there is compelling evidence of the existence of a potential trade-off between the performances in the asymptotic and non-asymptotic regimes [62]. In view of this, a non-asymptotic analysis of our findings for sensing networks is in order. To do it, in section 2.2 we propose a multi-parameter Bayesian procedure that generalises its single-parameter counterpart in [61], and in section 4 we utilise it to examine the non-asymptotic properties of some of our results in section 3. Our central insight here is that trading a part of the asymptotic enhancement is sometimes associated with an improved performance in the non-asymptotic regime *also* in networked quantum metrology, and in general we find that the amount of correlations needed to enhance the precision crucially depends on the amount of data that has been collected. Due to the more complex (and often numerical) nature of Bayesian calculations, this study is restricted to the $d = 2$ case, although in section 5 we discuss some potential directions to overcome this limitation. To the best of our knowledge,

this work, together with [16, 54], constitutes one of the first Bayesian studies of a network of quantum sensors in this context.

Our approach to the simultaneous estimation of linear functions in a scheme for distributed quantum sensing will serve as a basis to investigate how to harness correlations in multi-parameter schemes, operating both in and out of the asymptotic regime. Since the construction of entangled networks is likely to be difficult in practice, these insights may prove to be crucial in the study and implementation of quantum sensing networks that operate with a realistic amount of data.

2. Formulation of the problem

2.1. Physical scheme and available information

Consider a network of d qubit sensors prepared in some initial state $\rho_0 = |\psi_0\rangle\langle\psi_0|$, with

$$|\psi_0\rangle = \sum_{i_1 \dots i_d=0}^1 a_{i_1 \dots i_d} |i_1 \dots i_d\rangle, \quad (1)$$

$\sum_{i_1 \dots i_d=0}^1 |a_{i_1 \dots i_d}|^2 = 1$, and the basis elements $\langle 0|_j = (1, 0)$ and $\langle 1|_j = (0, 1)$ for the j th sensor. In addition, suppose we encode d local parameters $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d)$, one per sensor, as $\rho(\boldsymbol{\theta}) = e^{-i\mathbf{K}\cdot\boldsymbol{\theta}} \rho_0 e^{i\mathbf{K}\cdot\boldsymbol{\theta}}$, where $\mathbf{K} = (K_1, \dots, K_d)$, each generator K_i has the form

$$2K_i = \mathbb{I}_1 \otimes \dots \otimes \mathbb{I}_{i-1} \otimes \sigma_z \otimes \mathbb{I}_{i+1} \otimes \dots \otimes \mathbb{I}_d \equiv \sigma_{z,i}, \quad (2)$$

and

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathbb{I}_i = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3)$$

This is an instance of the type of unitary encoding that arises in spatially distributed sensing [32, 33], and while it is separable, i.e.,

$$\exp(-i\mathbf{K}\cdot\boldsymbol{\theta}) = e^{-i\sigma_z\theta_1/2} \otimes \dots \otimes e^{-i\sigma_z\theta_d/2}, \quad (4)$$

in principle we allow for entangled pure states and any general measurement acting on all the sensors at once. When the state and the measurement present no quantum correlations, we say that the scheme implements a *local strategy*. Otherwise we have a *global strategy*. We also note that

$$[K_i, K_j] = [\sigma_{z,i}, \sigma_{z,j}]/4 = 0, \quad (5)$$

which is a useful feature of this system because it will allow us to saturate the asymptotic bound in section 2.2.

To introduce the subclass of *sensor-symmetric states* that we will exploit, first we recall that the strength of correlations between any pair of sensors, which we call *inter-sensor correlations*, may be quantified as [29, 32]

$$\mathcal{J}_{ij} = \frac{\langle K_i K_j \rangle - \langle K_i \rangle \langle K_j \rangle}{\Delta K_i \Delta K_j}, \quad (6)$$

for $i \neq j$, where $\Delta K_i^2 = \langle K_i^2 \rangle - \langle K_i \rangle^2$ and we use the notation $\langle \square \rangle \equiv \langle \psi_0 | \square | \psi_0 \rangle$. Furthermore, \mathcal{J}_{ij} in equation (6) is bounded as $-1 \leq \mathcal{J}_{ij} \leq 1$. Using this quantifier, we define sensor-symmetric states as those satisfying

$$v = \langle K_i^2 \rangle - \langle K_i \rangle^2, \quad c = \langle K_i K_j \rangle - \langle K_i \rangle \langle K_j \rangle \quad (7)$$

for all i, j , where c and v are fixed values that characterise the preparation of the network and the encoding of the parameters. In turn, equation (6) becomes $\mathcal{J}_{ij} = \mathcal{J} = c/v$, also for all $i \neq j$, and for our qubit model we see that

$$4v = \langle \sigma_{z,i}^2 \rangle - \langle \sigma_{z,i} \rangle^2 = 1 - \langle \sigma_{z,i} \rangle^2, \quad 4c = \langle \sigma_{z,i} \sigma_{z,j} \rangle - \langle \sigma_{z,i} \rangle \langle \sigma_{z,j} \rangle, \quad (8)$$

where $0 \leq 4v \leq 1$ due to the fact that the eigenvalues of σ_z are ± 1 and thus $|\langle \sigma_{z,i} \rangle| \leq 1$. This definition in terms of the conditions in equation (7) is a way of generalising the notion of path-symmetric states in optical interferometry [29, 63, 64], and it motivates our choice of initial probe.

The final piece required before we can formulate the estimation problem of interest is to establish what prior information is available. The properties of the network that we wish to estimate are those that can be modelled linearly as

$$\mathbf{f}(\boldsymbol{\theta}) = (f_1(\boldsymbol{\theta}), \dots, f_l(\boldsymbol{\theta})) = \mathbf{V}^\top \boldsymbol{\theta} + \mathbf{a}, \quad (9)$$

where \mathbf{V} is a $(d \times l)$ matrix and \mathbf{a} is a column vector with l components. We consider that the form of these functions is known and so there is no uncertainty associated with the matrix \mathbf{V} or the vector \mathbf{a} . Furthermore, we assume that the unknown parameters $\boldsymbol{\theta}$ can be initially thought of as independent in the statistical sense, such that there are no prior correlations between them, and we suppose that the magnitude of the i th parameter can be found somewhere within an interval of width $W_{0,i}$ centred around $\bar{\theta}_i$, which is a moderate amount of prior knowledge [45, 62, 65]. This state of information can be represented by the separable prior probability

$$p(\boldsymbol{\theta}) = 1 / \left(\prod_{i=1}^d W_{0,i} \right), \quad (10)$$

for $\boldsymbol{\theta} \in [\bar{\theta}_1 - W_{0,1}/2, \bar{\theta}_1 + W_{0,1}/2] \times \dots \times [\bar{\theta}_d - W_{0,d}/2, \bar{\theta}_d + W_{0,d}/2]$, and zero otherwise. Equivalently, equation (10) may also be written as $p(\boldsymbol{\theta}) = 1/\Delta_0$, with hypervolume $\Delta_0 = \prod_{i=1}^d W_{0,i}$ centred around $\bar{\boldsymbol{\theta}} = (\bar{\theta}_1, \dots, \bar{\theta}_d)$. The interested reader will find in appendix A a way of justifying this prior from the perspective of the so-called *objective* version of the Bayesian framework.

2.2. Estimation method: a hybrid approach

Starting with the transformed network state $\rho(\boldsymbol{\theta})$ in section 2.1, the next step is to consider μ identical and independent measurements on this system, which we see as *trials* or *repetitions*. In particular, the i th measurement is represented by a POVM $E(m_i)$ with outcome m_i , and the probability of this process generating the outcomes $\mathbf{m} = (m_1, \dots, m_\mu)$ is given by the likelihood function

$$p(\mathbf{m}|\boldsymbol{\theta}) = \prod_{i=1}^{\mu} p(m_i|\boldsymbol{\theta}) = \prod_{i=1}^{\mu} \text{Tr}[E(m_i)\rho(\boldsymbol{\theta})]. \quad (11)$$

Since the form of the functions $\mathbf{f}(\cdot)$ has been assumed to be known, it is appropriate to construct their estimators as

$$\tilde{\mathbf{f}}(\mathbf{m}) = \mathbf{f}[\tilde{\boldsymbol{\theta}}(\mathbf{m})] = (f_1[\tilde{\boldsymbol{\theta}}(\mathbf{m})], \dots, f_l[\tilde{\boldsymbol{\theta}}(\mathbf{m})]) = \mathbf{V}^\top \tilde{\boldsymbol{\theta}}(\mathbf{m}) + \mathbf{a}, \quad (12)$$

where $\tilde{\boldsymbol{\theta}}(\mathbf{m}) = (\tilde{\theta}_1(\mathbf{m}), \dots, \tilde{\theta}_d(\mathbf{m}))$ are the estimators for the parameters $\boldsymbol{\theta}$, and we evaluate the uncertainty of our estimates $\tilde{\mathbf{f}}(\mathbf{m})$ as

$$\bar{\epsilon}_{\text{mse}} = \int d\boldsymbol{\theta} d\mathbf{m} p(\boldsymbol{\theta}) p(\mathbf{m}|\boldsymbol{\theta}) \text{Tr}\{\mathcal{W}[\tilde{\mathbf{f}}(\mathbf{m}) - \mathbf{f}(\boldsymbol{\theta})][\tilde{\mathbf{f}}(\mathbf{m}) - \mathbf{f}(\boldsymbol{\theta})]^\top\}, \quad (13)$$

where $p(\boldsymbol{\theta})$ is the prior, $\mathcal{W} = \text{diag}(w_1, \dots, w_l)$ is a weighting matrix, $w_i \geq 0$ represents the relative importance of estimating the i th parameter, and $\text{Tr}(\mathcal{W}) = 1$. Importantly, although a square error is generally not suitable for quantities associated with topologies other than that for the real line, it can still be a good approximation to the uncertainty for other topologies when the prior knowledge about $\boldsymbol{\theta}$ is moderate or high (see, e.g., [45, 47, 61, 62, 66, 67]), which is our case.

By using equations (10)–(12) and the network configuration in section 2.1, equation (13) becomes

$$\begin{aligned} \bar{\epsilon}_{\text{mse}} = & \int \frac{d\boldsymbol{\theta} d\mathbf{m}}{\Delta_0} \prod_{i=1}^{\mu} \text{Tr} [E(m_i) e^{-i\mathbf{K}\cdot\boldsymbol{\theta}} \rho_0 e^{i\mathbf{K}\cdot\boldsymbol{\theta}}] \\ & \times \text{Tr}\{\mathcal{W}V^\top [\tilde{\boldsymbol{\theta}}(\mathbf{m}) - \boldsymbol{\theta}][\tilde{\boldsymbol{\theta}}(\mathbf{m}) - \boldsymbol{\theta}]^\top V\} \end{aligned} \quad (14)$$

for our system. We note that this error does not depend on \mathbf{a} , so that we can set $\mathbf{a} = \mathbf{0}$ without loss of generality. Hence, from now on the functions are $\mathbf{f}(\boldsymbol{\theta}) = V^\top \boldsymbol{\theta}$ and the coefficients are encoded in the columns of V .

Ideally, we would like to minimise the error in equation (14) with respect to the estimators $\tilde{\boldsymbol{\theta}}(\mathbf{m})$, the measurement scheme $E(m_i)$ and the initial sensor-symmetric state ρ_0 , so that we can find the optimal configuration of the network and study its properties. Since, in general, this is a very challenging problem, in this work we follow an approximate procedure that combines asymptotic and non-asymptotic optimisations. We now describe this *hybrid* approach and how to use it for our analysis of sensing networks (a discussion of other methods in the literature can be found in appendix B).

On the one hand, equation (14) can be minimised with respect to $\tilde{\boldsymbol{\theta}}(\mathbf{m})$ in a straightforward way (e.g., using calculus of variations; see [16, 68]). This provides the familiar result that

$$\tilde{\boldsymbol{\theta}}(\mathbf{m}) = \int d\boldsymbol{\theta} p(\boldsymbol{\theta}|\mathbf{m}) \boldsymbol{\theta} \quad (15)$$

are the optimal estimators [68, 69], where $p(\boldsymbol{\theta}|\mathbf{m}) = p(\mathbf{m}|\boldsymbol{\theta})/[\Delta_0 p(\mathbf{m})]$ is the posterior probability and $p(\mathbf{m}) = \int d\boldsymbol{\theta} p(\mathbf{m}|\boldsymbol{\theta})/\Delta_0$. As a consequence, inserting equation (15) in equation (14) we have that

$$\bar{\epsilon}_{\text{mse}} \geq \sum_{i=1}^l w_i \int d\mathbf{m} p(\mathbf{m}) \left\{ \int d\boldsymbol{\theta} p(\boldsymbol{\theta}|\mathbf{m}) f_i^2(\boldsymbol{\theta}) - \left[\int d\boldsymbol{\theta} p(\boldsymbol{\theta}|\mathbf{m}) f_i(\boldsymbol{\theta}) \right]^2 \right\} \equiv \epsilon_{\text{opt}}^c, \quad (16)$$

where $f_i(\boldsymbol{\theta}) = \sum_{j=1}^d V_{ji} \theta_j$. This is the optimal uncertainty based on the probabilities that emerge from the measurements in a given quantum strategy ($E(m_i)$ plus ρ_0), and is valid and exact for any number of trials μ .

On the other hand, we may select the quantum strategy such that it is optimal in the asymptotic regime of many trials, where $\mu \gg 1$. First we recall that, if the true values $\boldsymbol{\theta}'$ lie within the prior hypervolume Δ_0 , and the likelihood $p(\mathbf{m}|\boldsymbol{\theta})$, which we assume to be sufficiently regular, becomes concentrated around $\boldsymbol{\theta}'$ as μ grows, then the posterior probability $p(\boldsymbol{\theta}|\mathbf{m})$ can

be approximated as a multivariate Gaussian density, and the uncertainty ϵ_{opt}^c in equation (16) satisfies [68, 70, 71]

$$\epsilon_{\text{opt}}^c \approx \int \frac{d\boldsymbol{\theta}'}{\mu \Delta_0} \text{Tr} [\mathcal{W}V^\top F(\boldsymbol{\theta}')^{-1}V] \equiv \epsilon_{\text{asym}}^c, \quad (17)$$

where

$$F(\boldsymbol{\theta}) = \int \frac{dm}{p(m|\boldsymbol{\theta})} \left[\frac{\partial p(m|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] \left[\frac{\partial p(m|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]^\top \quad (18)$$

is the Fisher information matrix for a single trial with outcome m (for a derivation of this approximation, see, e.g., [68, 70, 71] and section 6.2.2 of [45], and [8, 72, 73] for a rigorous treatment). At the same time, given that the form of the unitary encoding is $\exp(-i\mathbf{K} \cdot \boldsymbol{\theta})$ and the state $\rho_0 = |\psi_0\rangle\langle\psi_0|$ is pure, the Helstrom Cramér–Rao bound establishes the matrix inequality [43, 44, 46, 47]

$$F(\boldsymbol{\theta})^{-1} \geq F_q^{-1}, \quad \text{with } (F_q)_{ij} = 4 (\langle\psi_0|K_i K_j|\psi_0\rangle - \langle\psi_0|K_i|\psi_0\rangle\langle\psi_0|K_j|\psi_0\rangle), \quad (19)$$

F_q being the quantum counterpart of the information matrix. Then, the combination of equations (16), (17) and (19) implies that, in the asymptotic regime,

$$\bar{\epsilon}_{\text{mse}} \geq \epsilon_{\text{opt}}^c \approx \epsilon_{\text{asym}}^c \geq \frac{1}{\mu} \text{Tr} (\mathcal{W}V^\top F_q^{-1}V) \equiv \bar{\epsilon}_{\text{cr}}. \quad (20)$$

The quantum Cramér–Rao bound $\bar{\epsilon}_{\text{cr}}$ in equation (20) is a function of ρ_0 only, since \mathbf{K} , V , \mathcal{W} and μ are fixed, and it does not depend on the measurement. As such, if we choose the POVM $E(m_i)$ for the i th repetition such that $\epsilon_{\text{asym}}^c = \bar{\epsilon}_{\text{cr}}$, then that measurement will be asymptotically optimal. It can be shown that a measurement such that $F(\boldsymbol{\theta}) = F_q$ (and thus $\epsilon_{\text{asym}}^c = \bar{\epsilon}_{\text{cr}}$) always exists when the generators \mathbf{K} commute with each other [12, 13], and equation (5) demonstrates that this is indeed satisfied by our qubit network. Hence, we will use this criterion to construct the POVM. Regarding the optimisation of the state, we will proceed by first calculating $\bar{\epsilon}_{\text{cr}}$ as a function of the properties that characterise the sensor-symmetric state ρ_0 , which, as we will see, are the variance v and the correlation strength \mathcal{J} , and then minimising the resulting bound with respect to the pair (v, \mathcal{J}) . Once we know the optimal estimators

$$\tilde{\mathbf{f}}(\mathbf{m}) = V^\top \int d\boldsymbol{\theta} p(\boldsymbol{\theta}|\mathbf{m}) \boldsymbol{\theta} \quad (21)$$

and the asymptotically optimal state and measurement as prescribed above, we can complete the estimation by inserting these in the Bayesian uncertainty for μ repetitions in equation (14), which here will be calculated numerically with the algorithm in section 6.2.3 of [45] (the reader interested in reproducing our numerical results will find the associated MATLAB code in appendix C of the same work).

It is important to realise that our approach can fail when the asymptotic approximation is not valid. This could happen, for example, if the prior information provided within the hyper-volume Δ_0 is not sufficient to distinguish a single point [61, 68], or if the Fisher information matrix (classical or quantum) is singular. Therefore, we will concern ourselves with schemes where the information matrix is invertible, and, once we have found the asymptotically optimal quantum strategy, we will also check that the likelihood $p(\mathbf{m}|\boldsymbol{\theta})$ associated with it does not

present ambiguities in the relevant portion of the parameter space. Nevertheless, note that, in general, a potentially ambiguous likelihood function or a singular $F(\theta)$ do not introduce any fundamental difficulty for Bayesian estimation itself (this will be demonstrated in section 4 with an example).

In summary, the estimation method that emerges from the previous discussion requires that we:

- (a) calculate the quantum Cramér–Rao bound $\bar{\epsilon}_{\text{cr}}$ and find the sensor-symmetric state that makes it minimal,
- (b) search for a POVM such that $\epsilon_{\text{asym}}^c = \bar{\epsilon}_{\text{cr}}$,
- (c) verify that the quantum strategy (state plus POVM) allows for unambiguous estimation given the prior information represented in equation (10),
- (d) calculate the optimal estimators for the linear functions in equation (21), and
- (e) calculate the μ -trial Bayesian uncertainty in equation (14).

While the protocols constructed in this way may not be optimal for low μ , [61] demonstrated that this technique can provide important information about the non-asymptotic regime in optical interferometry, and here we will show that this is also true for networked quantum sensing. Moreover, a very useful feature of our approach is that the analysis of the role of inter-sensor correlations emerging from (a) and (b) will be relevant for researchers interested only in the Cramér–Rao bound, while those that also require an analysis based on a finite number of repetitions will benefit from the insights arising from (c)–(e). The next section is dedicated to the former.

3. Asymptotic estimation of global properties

3.1. Estimation of arbitrary linear functions

Our first step is to examine the quantum strategies that are optimal in the regime where the square error $\bar{\epsilon}_{\text{mse}}$ converges to the quantum Cramér–Rao bound $\bar{\epsilon}_{\text{cr}} = \text{Tr}(\mathcal{W}V^\top F_q^{-1}V)/\mu$ as μ grows. If we denote by $\{\mathbf{e}_i\}$ the basis components of the real space where \mathcal{W} , V and F_q are defined, with $\mathbf{e}_i^\top \mathbf{e}_j = \delta_{ij}$, then from equations (8) and (19) we have that

$$\begin{aligned} F_q &= \sum_{i,j=1}^d (\langle \sigma_{z,i} \sigma_{z,j} \rangle - \langle \sigma_{z,i} \rangle \langle \sigma_{z,j} \rangle) \mathbf{e}_i \mathbf{e}_j^\top = 4 \left(v \sum_{i=1}^d \mathbf{e}_i \mathbf{e}_i^\top + c \sum_{\substack{i,j=1 \\ i \neq j}}^d \mathbf{e}_i \mathbf{e}_j^\top \right) \\ &= 4[(v-c)\mathbb{I} + c\mathcal{I}] = 4v[(1-\mathcal{J})\mathbb{I} + \mathcal{J}\mathcal{I}], \end{aligned} \quad (22)$$

where \mathcal{I} is a $(d \times d)$ matrix of ones and \mathbb{I} the $(d \times d)$ identity matrix. This is the quantum Fisher information matrix for sensor-symmetric states.

To invert F_q , we need to impose the condition of positive definiteness, which is equivalent to requiring that its eigenvalues are strictly positive. Expressing \mathcal{I} as $\mathcal{I} = \mathbf{1}\mathbf{1}^\top$, where we recall that $\mathbf{1}^\top = (1, 1, \dots)$ is the vector of ones, the information matrix becomes $F_q = 4v[(1-\mathcal{J})\mathbb{I} + \mathcal{J}\mathbf{1}\mathbf{1}^\top]$. In that case, the characteristic equation for the eigenvalues $\{\lambda\}$ is

$$\det \left\{ 4v \left[\left(1 - \mathcal{J} - \frac{\lambda}{4v} \right) \mathbb{I} + \mathcal{J}\mathbf{1}\mathbf{1}^\top \right] \right\} = 0, \quad (23)$$

which upon using the identity $\det(X + \mathbf{y}\mathbf{z}^\top) = (1 + \mathbf{z}^\top X^{-1}\mathbf{y})\det(X)$, with $X = [4v(1 - \mathcal{J}) - \lambda]\mathbb{I}$, $\mathbf{y} = 4v\mathcal{J}\mathbf{1}$ and $\mathbf{z} = \mathbf{1}$, implies that

$$\{4v[1 + (d - 1)\mathcal{J}] - \lambda\}[4v(1 - \mathcal{J}) - \lambda]^{d-1} = 0. \quad (24)$$

As a result, the eigenvalues of F_q are $\lambda_1 = 4v[1 + (d - 1)\mathcal{J}]$, with multiplicity 1, and $\lambda_2 = 4v(1 - \mathcal{J})$, with multiplicity $d - 1$, and by imposing that they are positive we conclude that F_q is invertible when $1/(1 - d) < \mathcal{J} < 1$. The rest of our calculations assume that \mathcal{J} lies in such open interval under this assumption.

We can now calculate the inverse of F_q in equation (22), which is [32]

$$F_q^{-1} = \frac{[1 + (d - 1)\mathcal{J}]\mathbb{I} - \mathcal{J}\mathcal{I}}{4v(1 - \mathcal{J})[1 + (d - 1)\mathcal{J}]}. \quad (25)$$

Utilising this result we find that the asymptotic uncertainty for the estimation of linear functions is given by

$$\bar{\epsilon}_{\text{cr}} = \frac{[1 + (d - 2)\mathcal{J}]\text{Tr}(\mathcal{W}\mathcal{V}^\top\mathcal{V}) - \mathcal{J}\text{Tr}(\mathcal{W}\mathcal{V}^\top\mathcal{X}\mathcal{V})}{4\mu v(1 - \mathcal{J})[1 + (d - 1)\mathcal{J}]}, \quad (26)$$

where we have introduced the $(d \times d)$ matrix $\mathcal{X} \equiv \mathcal{I} - \mathbb{I}$ to separate the contribution to the uncertainty due to the diagonal elements of F_q^{-1} , which are the errors for each of the parameters, from that of the rest of the matrix.

The expression in equation (26) shows that the uncertainty depends on three types of quantities: (i) the number of repetitions μ and the number of parameters d , (ii) the combined properties of state and generators through the correlation strength \mathcal{J} and the variance v , and (iii) two quantities, $\text{Tr}(\mathcal{W}\mathcal{V}^\top\mathcal{V})$ and $\text{Tr}(\mathcal{W}\mathcal{V}^\top\mathcal{X}\mathcal{V})$, that are defined in terms of the functions encoded in \mathcal{V} and the weighting matrix \mathcal{W} . The next step is to investigate the physical meaning of these two quantities in (iii).

By relabelling the vector formed by the components of the j th linear function as \mathbf{f}_j (i.e., $f_j(\boldsymbol{\theta}) = \sum_{i=1}^d V_{ij}\theta_i \equiv \mathbf{f}_j^\top\boldsymbol{\theta}$), we can rewrite the first quantity in a more suggestive form as

$$\begin{aligned} \text{Tr}(\mathcal{W}\mathcal{V}^\top\mathcal{V}) &= \sum_{i,j=1}^l \sum_{k=1}^d (\mathcal{W})_{ij} V_{kj} V_{ki} = \sum_{j=1}^l w_j \sum_{k=1}^d V_{kj} V_{kj} \\ &= \sum_{j=1}^l w_j \mathbf{f}_j^\top \mathbf{f}_j = \sum_{j=1}^l w_j |\mathbf{f}_j|^2. \end{aligned} \quad (27)$$

where the norm in the last term is defined as $|\mathbf{v}|^2 = \sum_k v_k^2$ for a real vector \mathbf{v} . This is the weighted sum of the squared magnitudes of the vectors associated with the linear functions. Since $\mathcal{V}\mathcal{W}\mathcal{V}^\top$ is positive semi-definitive, and excluding the degenerate case where all the coefficients vanish, we have that $\text{Tr}(\mathcal{W}\mathcal{V}^\top\mathcal{V}) = \text{Tr}(\mathcal{V}\mathcal{W}\mathcal{V}^\top) > 0$. In addition, when the functions are normalised, that is, $|\mathbf{f}_i| = 1$ for $1 \leq i \leq l$, and recalling that $\text{Tr}(\mathcal{W}) = \sum_{i=1}^l w_i = 1$, we have that $\text{Tr}(\mathcal{W}\mathcal{V}^\top\mathcal{V}) = 1$. Hence, we define the *normalisation term*

$$\mathcal{N} \equiv \text{Tr}(\mathcal{W}\mathcal{V}^\top\mathcal{V}) = \sum_{j=1}^l w_j |\mathbf{f}_j|^2 \quad (28)$$

satisfying that $\mathcal{N} > 0$, with $\mathcal{N} = 1$ for normalised linear functions.

As for the second quantity, we can rewrite it as

$$\begin{aligned}
 \text{Tr}(\mathcal{W}\mathcal{V}^T\mathcal{X}\mathcal{V}) &= \text{Tr}[\mathcal{W}\mathcal{V}^T(\mathcal{I} - \mathbb{1})\mathcal{V}] = -\mathcal{N} + \sum_{i,j=1}^l \sum_{k,m=1}^d (\mathcal{W})_{ij} V_{kj} \mathcal{I}_{km} V_{mi} \\
 &= -\mathcal{N} + \sum_{j=1}^l w_j \sum_{k,m=1}^d V_{kj} \mathbf{1}_k \mathbf{1}_m V_{mj} = -\mathcal{N} + \sum_{j=1}^l w_j \left(\sum_{k=1}^d V_{kj} \mathbf{1}_k \right)^2 \\
 &= -\mathcal{N} + \sum_{j=1}^l w_j (\mathbf{f}_j^T \mathbf{1})^2 = -\mathcal{N} + d \sum_{j=1}^l w_j |\mathbf{f}_j|^2 \cos^2(\varphi_{\mathbf{1},j}) \\
 &= \sum_{j=1}^l w_j |\mathbf{f}_j|^2 [d \cos^2(\varphi_{\mathbf{1},j}) - 1], \tag{29}
 \end{aligned}$$

where $\varphi_{\mathbf{1},j}$ is the angle between the vector associated with the j th function and the direction defined by the vector of ones $\mathbf{1}$, and we have used the fact that $|\mathbf{1}| = \sqrt{d}$.

Recalling that $|\cos(\varphi_{\mathbf{1},j})| \leq 1$ and using equation (29), we see that $\text{Tr}(\mathcal{W}\mathcal{V}^T\mathcal{X}\mathcal{V})$ is bounded as

$$-\mathcal{N} \leq \text{Tr}(\mathcal{W}\mathcal{V}^T\mathcal{X}\mathcal{V}) \leq \mathcal{N}(d - 1), \tag{30}$$

and that the extremes are realised when either the functions are aligned with the direction of the vector of ones $\mathbf{1}$, or they lie in a subspace orthogonal to it and of dimension $(l - 1)$. So, for sensor-symmetric networks with properties modelled by linear functions, there are two kinds of global properties that play a special role: the sum of all the natural parameters with equal weights, and any linear combination of them such that the sum of its coefficients vanishes. Any other set of global properties will produce some value for $\text{Tr}(\mathcal{W}\mathcal{V}^T\mathcal{X}\mathcal{V})$ lying within the interval in equation (30), and this will be given by the geometry of the transformation defined by $V\mathcal{W}\mathcal{V}^T$. This motivates the introduction of the *geometry parameter*

$$\mathcal{G} \equiv \frac{1}{\mathcal{N}} \text{Tr}(\mathcal{W}\mathcal{V}^T\mathcal{X}\mathcal{V}) = \frac{1}{\mathcal{N}} \sum_{j=1}^l w_j |\mathbf{f}_j|^2 [d \cos^2(\varphi_{\mathbf{1},j}) - 1], \tag{31}$$

which satisfies that $-1 \leq \mathcal{G} \leq (d - 1)$.

Inserting equations (28) and (31) in equation (26), the asymptotic uncertainty finally becomes

$$\bar{\epsilon}_{\text{cr}} = \frac{\mathcal{N}}{4\mu\nu} h(\mathcal{J}, \mathcal{G}, d), \tag{32}$$

where

$$h(\mathcal{J}, \mathcal{G}, d) = \frac{[1 + (d - 2 - \mathcal{G})\mathcal{J}]}{(1 - \mathcal{J})[1 + (d - 1)\mathcal{J}]}. \tag{33}$$

Given a sensor-symmetric network with d local properties, the factor $h(\mathcal{J}, \mathcal{G}, d)$ in equation (33) codifies the interplay between the inter-sensor correlations of strength \mathcal{J} and the geometry parameter \mathcal{G} for any linear property, which may be local or global. A representation of this interplay can be found in figure 2. The formulas in equations (32) and (33) have been obtained without imposing further restrictions on the functions, and this implies that this formalism can be applied to any number of linear functions whose coefficients generate vectors that can form any angle and have any length.

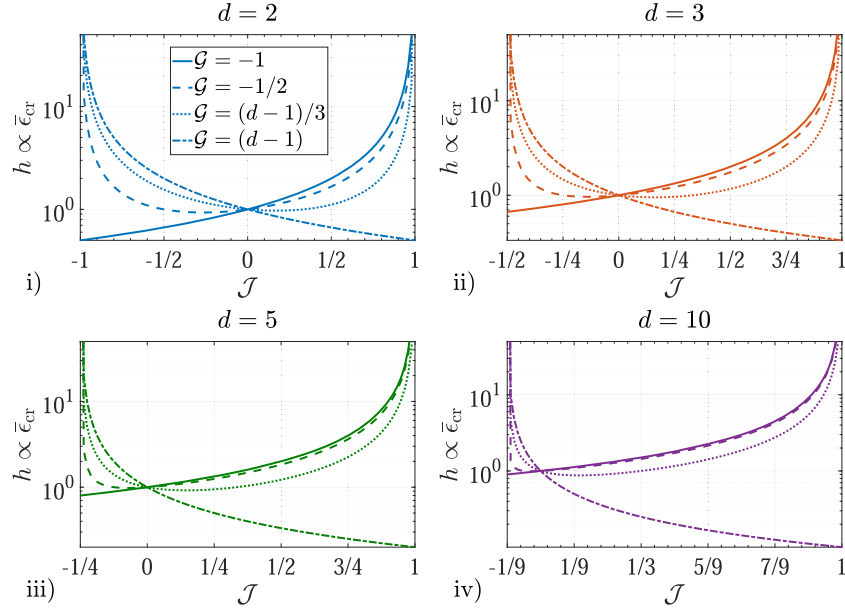


Figure 2. Representation of the interplay between the correlation strength \mathcal{J} and the geometry parameter \mathcal{G} in equation (33) for a quantum sensing network with (i) $d = 2$, (ii) $d = 3$, (iii) $d = 5$ and (iv) $d = 10$ parameters. We observe that, given $\mathcal{G} \in (-1, (d - 1))$, the minimum asymptotic uncertainty is achieved using a scheme with inter-sensor correlations of strength $\mathcal{J} \in (1/(1 - d), 1)$. The quantitative characterisation of these minima is provided in section 3.2.

3.2. The role of inter-sensor correlations I

Let us exploit the previous result to address the problem of selecting a sensor-symmetric network state that is optimal to estimate a given set of linear functions. This amounts to finding the values for v and \mathcal{J} that are optimal for a given \mathcal{G} . One approach is to use the fact that, for qubits, $0 \leq 4v \leq 1$, which allows us to lower bound equation (32) as

$$\bar{\epsilon}_{\text{cr}} \geq \frac{\mathcal{N}}{\mu} h(\mathcal{J}, \mathcal{G}, d) \equiv \bar{\epsilon}_{\text{f}}. \quad (34)$$

We then search for the \mathcal{J} that minimises this bound after having fixed \mathcal{G} , d and μ . In principle, there is no guarantee that the pairs of values $(4v = 1, \mathcal{J})$ generated by this method will correspond to any physical state, although the bounds on the asymptotic error constructed in this way would still be valid. Nevertheless, later we will study an example that realises a large portion of the pairs $(4v = 1, \mathcal{J})$ that we will predict.

By minimising $\bar{\epsilon}_{\text{f}}$ (see appendix C) we find that, if $4v = 1$, and restricting our attention to the range $1/(1 - d) < \mathcal{J} < 1$ where the information matrix is invertible, the optimal strength for the inter-sensor correlations of the network is

$$\mathcal{J}_{\text{opt}} = \frac{1}{\mathcal{G} + 2 - d} \left[1 - \sqrt{\frac{(\mathcal{G} + 1)(d - 1 - \mathcal{G})}{d - 1}} \right], \quad (35)$$

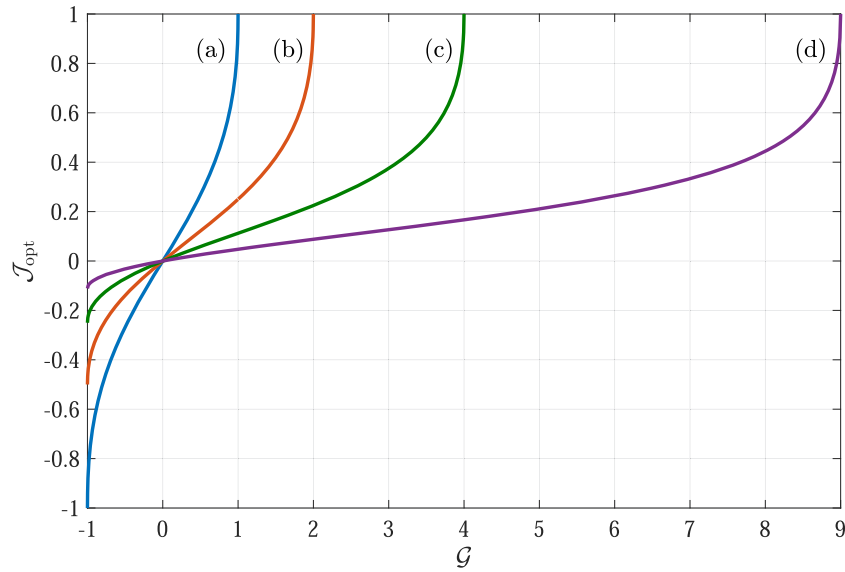


Figure 3. Optimal inter-sensor correlation strength \mathcal{J}_{opt} versus the geometry \mathcal{G} of a set of arbitrary linear functions, for $d = 2, 3, 5$ and 10 parameters (lines (a)–(d) respectively). These monotonic curves provide a quantitative representation for the uncertainty minima identified in figure 2, and the associated analytical formula is in equation (35). This result shows that, the more a collection of functions is clustered around the direction of $\mathbf{1}$, so that $\mathcal{G} = d - 1$, the larger the amount of correlations must be in order to perform the estimation optimally (provided that $4v = 1$), while the opposite is true if the functions are instead clustered around the subspace orthogonal to $\mathbf{1}$, for which $\mathcal{G} = -1$. Remarkably, any amount of correlations is detrimental when $\mathcal{G} = 0$, even though a vanishing geometry parameter can *also* be obtained for properties of the network that are global.

for $-1 < \mathcal{G} < d - 1$, which is determined by the structure of the functions alone via \mathcal{G} (once d has been fixed). This provides a map between correlation strength and geometry with one-to-one correspondence (note that $\mathcal{J}_{\text{opt}} \rightarrow (d - 2)/[2(d - 1)]$ when $\mathcal{G} \rightarrow d - 2$), as is illustrated in figure 3, and this is the central result of our asymptotic analysis.

The expression in equation (35) reveals that, the more a collection of functions is clustered around the vector of ones $\mathbf{1}$, the larger the amount of positive correlations is required to be in order to perform the estimation optimally (provided that $4v = 1$). Similarly, the amount of correlations with negative strength needs to be large if the functions are instead clustered around the subspace orthogonal to $\mathbf{1}$. The potential existence of this type of connection between geometry and quantum correlations was precisely one of the general open questions identified in [32].

Furthermore, equation (35) (and figure 3) shows that any non-zero pairwise correlation strength is detrimental whenever the geometry parameter vanishes. It is therefore interesting to investigate which linear functions imply that $\mathcal{G} = 0$, as well as the form of the associated optimal strategy. To achieve this, let us recall the original definition for \mathcal{G} in equation (31), that is, $\mathcal{G} = \text{Tr}(\mathcal{W}V^T \mathcal{X}V)/\mathcal{N}$. If we choose the uniform weighting matrix $\mathcal{W} = \mathbb{I}/l$, and if V is an orthogonal transformation (i.e., $VV^T = V^T V = \mathbb{I}$), then

$$\mathcal{G} = \frac{1}{\mathcal{N}l} \text{Tr}(VV^T \mathcal{X}) = \frac{1}{\mathcal{N}l} \text{Tr}(\mathcal{X}) = \frac{1}{\mathcal{N}l} \text{Tr}(\mathcal{I} - \mathbb{I}) = 0. \tag{36}$$

Now we observe that $\mathcal{J} = 0$, which is the optimal choice for the previous scenario, is always achieved by a separable qubit state $|\psi_0\rangle = (\sqrt{a}|0\rangle + \sqrt{1-a}|1\rangle)^{\otimes d}$, and by selecting $a = 1/2$ we have that $4v = 1$. Thus we can say that the estimation of a set of $l = d$ linear functions that are equally relevant and orthogonal can be carried out optimally by preparing our scheme with separable states. Moreover, since the estimation of the parameters θ is equivalent to choosing $V = \mathbb{I}$, our result implies that separable states are also optimal in that case. So, our present formalism is consistent with previous results [32, 33, 37, 74].

The above conclusion is sufficient to affirm that while entangled pure states are generally useful for the optimal estimation of global properties, it is not true that we *always* need entangled probes in such case. However, a transformation that is orthogonal preserves angles and lengths, and thus one may argue that, in a sense, the information encoded by a set of functions that gives rise to an orthogonal transformation is equivalent to the information content of the original parameters, provided that the weighting matrices are uniform. Hence, it is perhaps not surprising that a local estimation strategy is preferred here, since [32, 33] had already shown that the estimation of local properties associated with commuting generators can be performed optimally with a local scheme. In view of this, it is important to establish whether there are other global properties with $\mathcal{G} = 0$ that instead select information that is not equivalent to estimating all the original parameters. First we observe that the eigendecomposition of \mathcal{X} , which is a symmetric matrix, is (see appendix D)

$$\mathcal{X}_D = U_X^\top \mathcal{X} U_X = \text{diag} [(d-1), -1, \dots, -1], \quad (37)$$

where the eigenvector for the first eigenvalue is $\mathbf{1}$ and those for the other eigenvalues belong to the orthogonal subspace. That implies that if we choose a single linear function as $V = f = U_X \mathbf{1}$, then we will have that $\mathcal{G} = \mathbf{1}^\top U_X^\top \mathcal{X} U_X \mathbf{1} / d = \mathbf{1}^\top \mathcal{X}_D \mathbf{1} / d = 0$. Now consider a three-parameter network, so that

$$f = U_X \mathbf{1} = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & \sqrt{3} & 1 \\ \sqrt{2} & -\sqrt{3} & 1 \\ \sqrt{2} & 0 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} + \sqrt{3} + 1 \\ \sqrt{2} - \sqrt{3} + 1 \\ \sqrt{2} - 2 \end{pmatrix}. \quad (38)$$

Clearly, this gives rise to a global property, as these are the coefficients of a non-trivial function of three local parameters. Yet, $\mathcal{G} = 0$, and so, according to equation (35), pairwise correlations are detrimental. Therefore, entanglement is sometimes not needed in scenarios where we are estimating non-trivial global properties. Interestingly, the same argument fails for $d = 2$, since in that case

$$f = U_X \mathbf{1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix}, \quad (39)$$

and this is associated with a local property because it simply rescales the first parameter. Nonetheless, our conclusion above is still valid in general.

For the link between geometry and correlations in equation (35) to be truly relevant, it is necessary that there are physical states with the properties that such a link predicts as optimal. In [32] we studied the estimation of $1 \leq l \leq d = 2$ linear and normalised but otherwise arbitrary functions using the sensor-symmetric state

$$|\psi_0\rangle = \frac{1}{\sqrt{2(1+\gamma^2)}} [|00\rangle + \gamma (|01\rangle + |10\rangle) + |11\rangle], \quad (40)$$

with $-\infty < \gamma < \infty$, and we provided a complete solution to this two-parameter estimation problem. The fact that this is a particular case of the more general formalism that we develop in this work suggests that, for the $d = 2$ case, it may be possible to use the state in equation (40) to realise all the pairs ($4v = 1, \mathcal{J}$) that are optimal according to our results. We will now show that this is the case.

Recalling that $\sigma_z |i\rangle = (-1)^i |i\rangle$, we see that, for the state in equation (40), $\langle \sigma_{z,1} \rangle = \langle \sigma_{z,2} \rangle = 0$ and $\langle \sigma_{z,1} \sigma_{z,2} \rangle = \langle \sigma_{z,1} \sigma_{z,2} \rangle = (1 - \gamma^2)/(1 + \gamma^2)$, so that the variance is $4v = 4v_1 = 4v_2 = 1$ and the quantifier for the inter-sensor correlations can be written as a function of γ as $\mathcal{J} = (1 - \gamma^2)/(1 + \gamma^2)$. This function reaches the maximum $\mathcal{J} = 1$ at $\gamma = 0$, while it tends monotonically from such point to $\mathcal{J} = -1$ when $\gamma \rightarrow \pm\infty$. In other words, for $d = 2$ there is always a physical state that satisfies the condition imposed in equation (35) when $4v = 1$.

It is interesting to observe that γ splits the state into a part where the sum of the parameters is encoded and a part that encodes the difference. More concretely,

$$e^{-\frac{i}{2}(\sigma_{z,1}\theta_1 + \sigma_{z,2}\theta_2)} |\psi_0\rangle = \frac{1}{\sqrt{2(1 + \gamma^2)}} \left[e^{-\frac{i}{2}(\theta_1 + \theta_2)} |00\rangle + e^{\frac{i}{2}(\theta_1 + \theta_2)} |11\rangle \right] + \frac{\gamma}{\sqrt{2(1 + \gamma^2)}} \left[e^{-\frac{i}{2}(\theta_1 - \theta_2)} |01\rangle + e^{\frac{i}{2}(\theta_1 - \theta_2)} |10\rangle \right]. \quad (41)$$

A partial extension of this idea to the d -parameter case can be achieved by constructing a state where the part that encodes functions aligned with the direction of $\mathbf{1}$ is isolated in an analogous fashion, i.e.,

$$|\psi_0\rangle = \frac{1}{\sqrt{2[1 + (2^{d-1} - 1)\gamma^2]}} \left[|00\dots 0\rangle + |11\dots 1\rangle + \gamma(\text{all other terms}) \right] = \frac{1}{\sqrt{2[1 + (2^{d-1} - 1)\gamma^2]}} \left[(1 - \gamma) \left(|0\rangle^{\otimes d} + |1\rangle^{\otimes d} \right) + \gamma \left(|0\rangle + |1\rangle \right)^{\otimes d} \right]. \quad (42)$$

For this probe, $4v_i = 1 - \langle \sigma_{z,i} \rangle^2 = 1 = 4v$ for all i , and $4c_{ij} = \langle \sigma_{z,i} \sigma_{z,j} \rangle - \langle \sigma_{z,i} \rangle \langle \sigma_{z,j} \rangle = \langle \sigma_{z,i} \sigma_{z,j} \rangle = (1 - \gamma^2)/[1 + (2^{d-1} - 1)\gamma^2] = 4c$ for all $i \neq j$, which verifies that the state in equation (42) is also sensor symmetric. As a result, we can see that its inter-sensor correlations are given by

$$\mathcal{J} = \frac{1 - \gamma^2}{1 + (2^{d-1} - 1)\gamma^2}. \quad (43)$$

If $0 \leq |\gamma| \leq 1$, then we have that $1 \geq \mathcal{J} \geq 0$. This implies that there always exists a physical state associated with all the results in this section that require either positive inter-sensor correlations, or the absence of them. On the other hand, the amount of negative correlations that this state can cover lies in $0 > \mathcal{J} > -1/(2^{d-1} - 1)$, which corresponds to $1 < |\gamma| < \infty$. Unfortunately, the amount of negative correlations that equation (35) might predict can lie in $0 > \mathcal{J} > 1/(1 - d)$, where $1/(1 - d) \leq -1/(2^{d-1} - 1)$ for $d \geq 2$ and the inequality is only saturated when $d = 2$. Thus there is a subinterval not covered by equation (42). Whether there are other physical states that may realise the missing values is an open question.

Finally, we note that the only entangled pure probes that may be asymptotically relevant for sensor-symmetric networks are those that give rise to inter-sensor correlations, while any other form of entanglement will be irrelevant in this type of scenario. To illustrate this idea, let

we consider the state in equation (42) for $d = 3$, and suppose that the functions to be estimated give rise to the geometry parameter $\mathcal{G} = 0$. We have seen that, in that case, no inter-sensor correlations are needed to perform the estimation optimally, which implies that, according to equation (43), $\gamma = \pm 1$. By inserting these parameters in equation (42) we find that the optimal states are

$$|\psi_+\rangle = \frac{1}{2\sqrt{2}}(|0\rangle + |1\rangle)^{\otimes 3} \quad (44)$$

and

$$|\psi_-\rangle = \frac{1}{2\sqrt{2}} \left[2(|0\rangle^{\otimes 3} + |1\rangle^{\otimes 3}) - (|0\rangle + |1\rangle)^{\otimes 3} \right]. \quad (45)$$

The first state is separable, but $|\psi_-\rangle$ is not. More concretely, if we tried to write the latter as $|\psi_-\rangle = (x_0|0\rangle + x_1|1\rangle)(y_0|0\rangle + y_1|1\rangle)(z_0|0\rangle + z_1|1\rangle)$, with $|x_0|^2 + |x_1|^2 = |y_0|^2 + |y_1|^2 = |z_0|^2 + |z_1|^2 = 1$, we would find contradictions such as

$$[(x_0 = x_1) \wedge (x_0 = -x_1)] \wedge (|x_0|^2 + |x_1|^2 = 1), \quad (46)$$

which by *reductio ad absurdum* allows us to conclude that the state with $\gamma = -1$ and $d = 3$ is entangled. Hence, while here entanglement is not required to reach the asymptotic optimum, neither is it necessarily detrimental. The only requirement imposed by our formalism is the absence of pairwise correlations, and the presence or absence of any other kind of correlation does not affect the asymptotic uncertainty.

3.3. Optimal POVM in the asymptotic regime

The final step of the asymptotic analysis is to find some POVM that is optimal in the large- μ regime, in the sense that it saturates the quantum Cramér–Rao bound as $\epsilon_{\text{asym}}^c = \bar{\epsilon}_{\text{cr}}$, and we can achieve this by requiring that $F(\theta) = F_q$ [12, 13]. That the latter condition refers to the parameters but not to the functions, together with the fact that the former can be estimated optimally using a local strategy [32, 33] (see also section 3.2), suggests that a local POVM might be sufficient to make the classical and quantum information matrices equal. In fact, this would be very useful, since then we could associate any enhancement derived from the presence of correlations with the initial state alone. In the following we demonstrate this for a network with $d = 2$ parameters.

Consider a local POVM with elements

$$|n, k\rangle = [|0\rangle + (-1)^n |1\rangle] \otimes [|0\rangle + (-1)^k |1\rangle] / 2, \quad (47)$$

where $n, k = 0, 1$. Furthermore, we have seen that, if $d = 2$, then the state in equation (40) is general enough to realise all the asymptotic results predicted by our theory. As such, this is the probe that we will use in this calculation. Combining this POVM with the transformed state $|\psi(\theta_1, \theta_2)\rangle = e^{-\frac{i}{2}(\sigma_{z,1}\theta_1 + \sigma_{z,2}\theta_2)} |\psi_0\rangle$ in equation (41), we find the amplitude

$$\begin{aligned} \langle n, k | \psi(\theta_1, \theta_2) \rangle &\propto e^{-\frac{i}{2}(\theta_1 + \theta_2)} + (-1)^{n+k} e^{\frac{i}{2}(\theta_1 + \theta_2)} \\ &\quad + \gamma \left[(-1)^k e^{-\frac{i}{2}(\theta_1 - \theta_2)} + (-1)^n e^{\frac{i}{2}(\theta_1 - \theta_2)} \right] \\ &\propto \cos \{ [\theta_1 + \theta_2 + \pi(k + n)] / 2 \} + \gamma \cos \{ [\theta_1 - \theta_2 - \pi(k - n)] / 2 \}, \end{aligned} \quad (48)$$

the modulus of the proportionality factor being $1/\sqrt{2(1+\gamma^2)}$. This allows us to arrive at the likelihood function

$$p(n, k|\theta_1, \theta_2) = \|\langle n, k|\psi(\theta_1, \theta_2)\rangle\|^2 = [\cos(x_+) + \gamma \cos(x_-)]^2/[2(1+\gamma^2)], \quad (49)$$

where we have introduced the notation $x_{\pm} \equiv [\theta_1 \pm \theta_2 \pm \pi(k \pm n)]/2$.

The elements of the classical Fisher information matrix in equation (18) for the quantum probability in equation (49) are

$$\begin{aligned} [F(\theta)]_{11} &= \sum_{n,k=0}^1 \frac{1}{p(n, k|\theta_1, \theta_2)} \left[\frac{\partial p(n, k|\theta_1, \theta_2)}{\partial \theta_1} \right]^2 \\ &= \frac{1}{2(1+\gamma^2)} \sum_{n,k=0}^1 [\sin(x_+) + \gamma \sin(x_-)]^2 = 1, \end{aligned} \quad (50)$$

$$\begin{aligned} [F(\theta)]_{22} &= \sum_{n,k=0}^1 \frac{1}{p(n, k|\theta_1, \theta_2)} \left[\frac{\partial p(n, k|\theta_1, \theta_2)}{\partial \theta_2} \right]^2 \\ &= \frac{1}{2(1+\gamma^2)} \sum_{n,k=0}^1 [\sin(x_+) - \gamma \sin(x_-)]^2 = 1, \end{aligned} \quad (51)$$

and

$$\begin{aligned} [F(\theta)]_{12} &= \sum_{n,k=0}^1 \frac{1}{p(n, k|\theta_1, \theta_2)} \frac{\partial p(n, k|\theta_1, \theta_2)}{\partial \theta_1} \frac{\partial p(n, k|\theta_1, \theta_2)}{\partial \theta_2} \\ &= \frac{1}{2(1+\gamma^2)} \sum_{n,k=0}^1 [\sin^2(x_+) - \gamma^2 \sin^2(x_-)] = \frac{1-\gamma^2}{1+\gamma^2}, \end{aligned} \quad (52)$$

with $[F(\theta)]_{21} = [F(\theta)]_{12}$. Additionally, in sections 3.1 and 3.2 we have seen that, for this configuration,

$$F_q = \begin{pmatrix} 1 & \mathcal{J} \\ \mathcal{J} & 1 \end{pmatrix} = \begin{pmatrix} 1 & (1-\gamma^2)/(1+\gamma^2) \\ (1-\gamma^2)/(1+\gamma^2) & 1 \end{pmatrix}, \quad (53)$$

which is identical to the classical Fisher information matrix in equations (50)–(52). We thus conclude that the quantum strategy formed by the local POVM in equation (47) and the state in equation (40) is asymptotically optimal. This completes our solution for the asymptotic estimation of linear functions in a two-parameter network, and will be our starting point to perform a Bayesian analysis.

4. Bayesian analysis of non-asymptotic quantum sensing networks

Now we turn to the more general problem of estimating linear functions when different amounts of data are available, which may include cases with a low number of trials. Thanks to the simplicity of the asymptotic approach, in section 3 we were able to discuss examples where $d = 2, 3, 5$ and 10 , and many of the results there were valid for any d . However, due to the more challenging nature of the numerical calculations associated with Bayesian estimation, in the remainder of this work we will focus on *two-parameter* sensor-symmetric qubit networks.

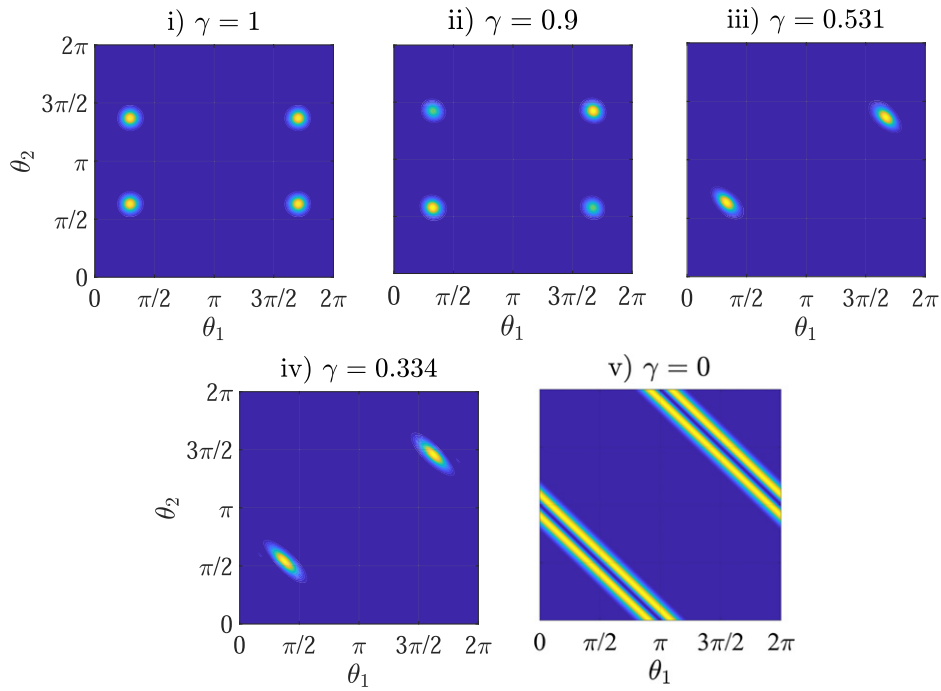


Figure 4. Posterior density functions for random simulations of $\mu = 100$ trials, a flat prior and the quantum strategy represented by the likelihood in equation (49), with (i) $\gamma = 1$, (ii) $\gamma = 0.9$, (iii) $\gamma = 0.531$, (iv) $\gamma = 0.334$ and (v) $\gamma = 0$. The simulated true values of the parameters are $\theta'_1 = 1$ and $\theta'_2 = 2$. This figure shows that the potential ambiguities in the estimation associated with scenarios (i)–(iv) can be generally avoided if the prior area satisfies $\Delta_0 \leq \pi^2$. On the contrary, while the scheme (v) can be exploited to estimate the sum of the parameters, in general it cannot provide good estimates for other linear functions, independently of the value for Δ_0 . We draw attention to the fact that a similar pattern emerges as $\gamma \rightarrow \infty$, but with the posterior peaks tending to the direction orthogonal to that in (v).

4.1. Regions of unambiguous information

Our aim is to use the asymptotically optimal strategy in equations (41), (47) and (49) as a guide to perform a non-asymptotic analysis. Following our discussion in section 2.2, this approach is best justified when, as μ grows, the likelihood function

$$p(\mathbf{n}, \mathbf{k} | \theta_1, \theta_2) = \prod_{i=1}^{\mu} p(n_i, k_i | \theta_1, \theta_2), \tag{54}$$

with each $p(n_i, k_i | \theta_1, \theta_2)$ given by equation (49), becomes concentrated around a *unique* absolute maximum within the prior area Δ_0 . Indeed, this condition helps to prevent the estimation process from giving ambiguous answers [68]. Hence, before we proceed we need to find how large Δ_0 can be such that the above requirement is satisfied.

One way of estimating this size is to first represent the posterior probability $p(\theta_1, \theta_2 | \mathbf{n}, \mathbf{k}) \propto p(\mathbf{n}, \mathbf{k} | \theta_1, \theta_2)$ as a function of (θ_1, θ_2) , where the outcomes (\mathbf{n}, \mathbf{k}) come from a simulation with true values (θ'_1, θ'_2) , and then visualise the regions with an asymptotically unique absolute maximum in a direct fashion (see [61]).

The previous method generates the results shown in figure 4 for several values of γ . First we note that the simulations in figure 4 have been restricted to the area $(\theta_1, \theta_2) \in [0, 2\pi] \times [0, 2\pi]$ because the single-shot likelihood in equation (49) is invariant under $\theta_i \rightarrow \theta_i + 2\pi m$, with $m = 0, \pm 1, \pm 2, \dots$ and $i = 1, 2$, and thus it suffices to examine its symmetries within one period. Depending on the value for γ , we see that the posterior probability in figures 4(i)–(iv) develops either two or four identical absolute maxima as μ grows, and that each of these peaks is located within an extension of area π^2 . Therefore, in the presence of complete ignorance, i.e., $\Delta_0 = 4\pi^2$, the quantum strategy under analysis cannot select a unique answer, a phenomenon already encountered in single-parameter metrology [45, 61, 62, 75]. In view of this, to avoid the ambiguities in figures 4(i)–(iv) we shall impose that the prior area satisfies the condition $\Delta_0 \leq \pi^2$.

The situation for $\gamma = 0$ in figure 4(v) is, however, different. In that case, no single peak can be selected even if $\mu \gg 1$, which implies that such scheme does not have an asymptotic approximation in the sense of section 2.2. This is consistent with the fact that, if $\gamma = 0$, then $\mathcal{J} = 1$, and this case must be excluded for F_q to be invertible (see section 3.1). Moreover, the same type of behaviour would have been observed if we had examined the limit $|\gamma| \rightarrow \infty$, for which $\mathcal{J} \rightarrow -1$. Hence, we only need to impose the existence of a unique absolute maximum for $0 < |\gamma| < \infty$. Crucially, this does not imply that the scheme with $\gamma = 0$ is useless. Figure 4(v) shows that this scheme is giving information about the combination $\theta_1 + \theta_2 = \pi m$, with $m = 0, \pm 1, \pm 2, \dots$, that is, about the sum of the parameters. In fact, this can be readily seen by inserting $\gamma = 0$ in equation (49), since then the likelihood for a single shot is only sensitive to the equally weighted sum of the parameters. The calculations in the next section will reveal that while the asymptotic performance of this scheme is poor, it can be useful when μ is low.

4.2. The role of inter-sensor correlations II

Given the quantum strategy in equations (41) and (47) for a two-parameter qubit network, we wish to estimate two global properties of such network when the experiment operates both in and out of the regime of limited data. In particular, consider the linear functions $f_1(\theta) = (2\theta_1 + \pi\theta_2)/\sqrt{4 + \pi^2}$ and $f_2(\theta) = (2\theta_1 + \theta_2)/\sqrt{5}$, which can be encoded in the columns of V as

$$V = \frac{1}{\sqrt{20 + 5\pi^2}} \begin{pmatrix} 2\sqrt{5} & 2\sqrt{4 + \pi^2} \\ \pi\sqrt{5} & \sqrt{4 + \pi^2} \end{pmatrix}. \tag{55}$$

We assume that both functions are equally relevant, so that $\mathcal{W} = \mathbb{I}/2$, and that our prior knowledge is represented by the prior probability $p(\theta_1, \theta_2) = 4/\pi^2$, when $(\theta_1, \theta_2) \in [0, \pi/2] \times [0, \pi/2]$, and zero otherwise. The area associated with this prior assignment is sufficiently small for the square error to be a suitable figure of merit in phase estimation [62, 67], and, thanks to our analysis in section 4.1, we know that it will allow us to perform the estimation unambiguously when the asymptotically optimal strategies are employed, since $\Delta_0 = \pi^2/4 < \pi^2$.

Let us start by comparing a local strategy with an entangled scheme that is asymptotically optimal. The former assumes that the experiment is arranged such that $\gamma = 1$, $\mathcal{J} = 0$, while to find the properties of the latter we need to recall our results in section 3.2 for the asymptotic role of inter-sensor correlations. Equation (35) indicates that, for $d = 2$,

$$\mathcal{J}_{\text{opt}} = \left(1 - \sqrt{1 - \mathcal{G}^2}\right) / \mathcal{G}, \tag{56}$$

when $\mathcal{G} \neq 0$, and $\mathcal{J}_{\text{opt}} = 0$ if $\mathcal{G} = 0$. In addition, $\mathcal{J} = (1 - \gamma^2)/(1 + \gamma^2)$, and by combining the latter expression with equation (56) we find that

$$\gamma_{\text{opt}} = \pm \left(\frac{\mathcal{G} - 1 + \sqrt{1 - \mathcal{G}^2}}{\mathcal{G} + 1 - \sqrt{1 - \mathcal{G}^2}} \right)^{\frac{1}{2}}, \quad (57)$$

when $\mathcal{G} \neq 0$, and $\gamma_{\text{opt}} = 1$ if $\mathcal{G} = 0$. The normalisation term for the functions in equation (55) is simply $\mathcal{N} = \text{Tr}(\mathcal{W}\mathcal{V}^\top \mathcal{V}) = 1$, while the geometry parameter is

$$\mathcal{G} = \frac{1}{\mathcal{N}} \text{Tr}(\mathcal{W}\mathcal{V}^\top \mathcal{K}\mathcal{V}) = \frac{8 + 10\pi + 2\pi^2}{20 + 5\pi^2} \approx 0.853. \quad (58)$$

By inserting this result in equations (56) and (57) we have that $\gamma_{\text{opt}} \approx \pm 0.531$ (we can choose the positive solution without loss of generality) and $\mathcal{J} = 0.561$, where the latter verifies that this state is indeed entangled (note that the two-sensor state in equation (40) is only separable when $\gamma^2 = 1$).

Next we perform the numerical calculation of the Bayesian uncertainty $\bar{\epsilon}_{\text{mse}}$ in equation (14) for these two sensor-symmetric states, whose form as a function of γ is in equation (40); the measurement $E(n_i, k_i) = |n_i, k_i\rangle\langle n_i, k_i|$ in equation (47) for the i th repetition in a sequence of μ trials; and the optimal estimators

$$\begin{aligned} \begin{pmatrix} \tilde{f}_1(\mathbf{n}, \mathbf{k}) \\ \tilde{f}_2(\mathbf{n}, \mathbf{k}) \end{pmatrix} &= \frac{4}{\pi^2 \sqrt{20 + 5\pi^2}} \begin{pmatrix} 2\sqrt{5} & \pi\sqrt{5} \\ 2\sqrt{4 + \pi^2} & \sqrt{4 + \pi^2} \end{pmatrix} \\ &\times \int_0^{\pi/2} d\theta_1 \int_0^{\pi/2} d\theta_2 p(\mathbf{n}, \mathbf{k} | \theta_1, \theta_2) \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, \end{aligned} \quad (59)$$

which arise from equation (21) after inserting equation (55). The results have been represented in figure 5(i) as graphs (a) for the local scheme and (b) for the optimal entangled strategy. We can observe that the local strategy performs worse than the entangled one for any number of repetitions. Therefore, in this case we have that the prediction made by the asymptotic theory is qualitatively preserved in the non-asymptotic regime. However, a closer analysis reveals that the distance between the two lines is considerably less when $1 \leq \mu \lesssim 20$ than when $\mu \gg 1$, and this behaviour is reminiscent of that of a Mach–Zehnder interferometer [62]. Indeed, optical probes with a large Fisher information (and thus a good asymptotic performance) have sometimes an error very close to that of a coherent laser beam in the regime of limited data, and coherent probes can be seen as an optical analogue of the notion of local strategy in this work. Moreover, the optical study in [62] also demonstrated that a better asymptotic error is sometimes associated with a worse performance in the regime of low μ . As a consequence, a natural question is whether a similar phenomenon can be exploited here, so that we can obtain an uncertainty that is lower than the error for the asymptotically optimal entangled state when the network operates in the non-asymptotic regime.

To test this idea, let us select a third arrangement with an asymptotic error that lies between those of the local scheme and the asymptotically optimal strategy. The asymptotic error for our network can be written in terms of γ as (see equations (32) and (33))

$$\bar{\epsilon}_{\text{cr}} = \frac{(1 + \gamma^2) [(1 - \mathcal{G}) + (1 + \mathcal{G})\gamma^2]}{4\mu\gamma^2} \equiv \bar{\epsilon}_{\text{qbit}}(\gamma). \quad (60)$$

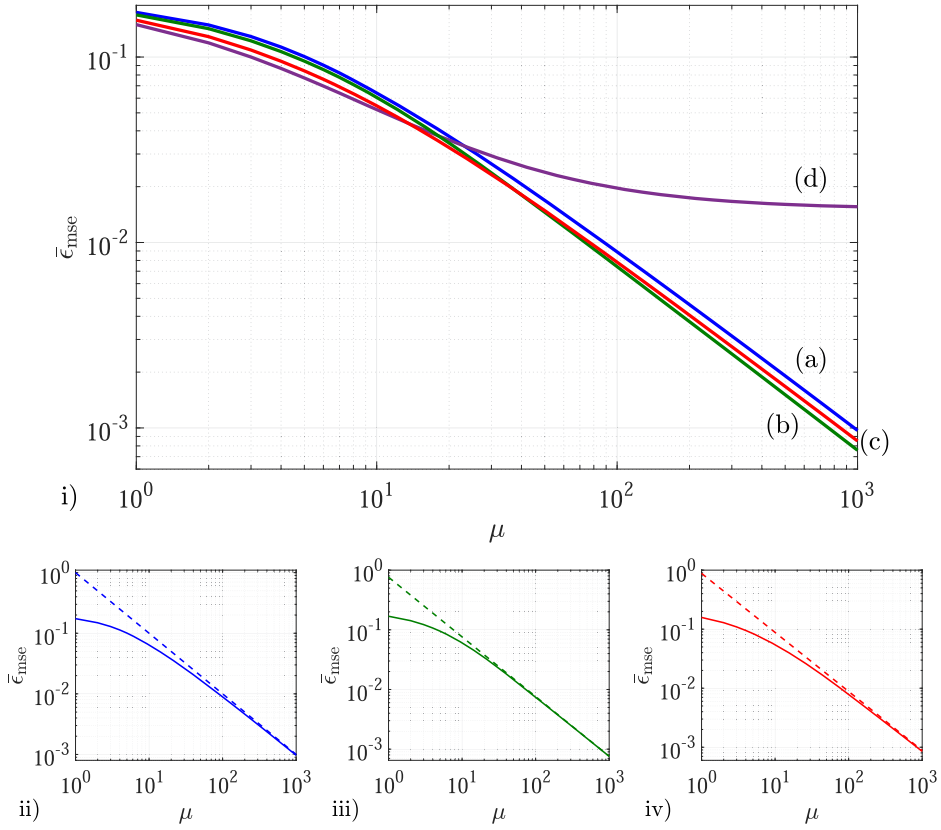


Figure 5. (i) Mean square error for the estimation of the linear functions $f_1(\boldsymbol{\theta}) = (2\theta_1 + \pi\theta_2)/\sqrt{4 + \pi^2}$ and $f_2(\boldsymbol{\theta}) = (2\theta_1 + \theta_2)/\sqrt{5}$ by means of the two-sensor qubit network introduced in the main text, where ((a), blue line) is a local strategy, with $\gamma = 1$, $\mathcal{J} = 0$; ((b), green line) is the asymptotically optimal entangled strategy, with $\gamma = 0.531$, $\mathcal{J} = 0.561$; ((c), red line) is a strategy whose enhancement has been balanced between the asymptotic and non-asymptotic regimes, with $\gamma = 0.334$, $\mathcal{J} = 0.799$; and ((d), purple line) is a maximally entangled state, with $\gamma = 0$, $\mathcal{J} = 1$. Figures (ii)–(iv) compare the mean square error (solid lines) and the multi-parameter quantum Cramér–Rao bound (dashed lines) for the strategies in (a)–(c), respectively, verifying that the latter is recovered asymptotically. All the calculations assume the weighting matrix $\mathcal{W} = \mathbb{I}/2$ and a flat prior of area $\Delta_0 = \pi^2/4$ centred around $(\pi/4, \pi/4)$.

Using this we can find the value of γ for the strategy satisfying our desideratum above by imposing that

$$\bar{\epsilon}_{\text{qbit}}(\gamma) = \frac{1}{2} [\bar{\epsilon}_{\text{qbit}}(\gamma_{\text{loc}} = 0) + \bar{\epsilon}_{\text{qbit}}(\gamma_{\text{ent}} = 0.531)], \quad (61)$$

and the solutions are $\gamma \approx \pm 0.334, \pm 0.842$. So we take our third strategy to be the state in equation (40) with $\gamma = 0.334$ (and thus $\mathcal{J} = 0.799$), a choice motivated by the fact that this is the option with the lower uncertainty for a single shot (in particular, $\bar{\epsilon}_{\text{mse}}(\mu = 1, \gamma = 0.334) \approx 0.158$ and $\bar{\epsilon}_{\text{mse}}(\mu = 1, \gamma = 0.842) \approx 0.173$).

The uncertainty $\bar{\epsilon}_{\text{mse}}$ for the third scheme has been represented as a function of the number of trials in figure 5(i), where it is labelled as (c). As expected, this error lies equidistantly

between the local and the asymptotically optimal strategies when $\mu \gg 1$, but this is no longer the case in the regime of limited data. More concretely, the graphs for the asymptotically optimal strategy and the new scheme cross each other when $\mu \approx 40$, so that the former is optimal when $\mu > 40$ and the latter is the preferred choice if $1 \leq \mu \lesssim 40$. Consequently, we may say that trading a part of the asymptotic enhancement is sometimes associated with an improved performance in the non-asymptotic regime, which constitutes a multi-parameter generalisation of the analogous phenomenon in [62] for a Mach–Zehnder interferometer.

Interestingly, the balanced strategy ($\gamma = 0.334$, $\mathcal{J} = 0.799$), which provides a better precision in the non-asymptotic regime, is associated with larger inter-sensor correlations, and in what follows we propose a potential explanation for this. Let us first recall that, when μ is large, the information about the global properties is essentially provided by the measurement outcomes that we accumulate as μ grows, which contrasts with the non-asymptotic regime where the information is a mixture of prior knowledge and experimental data. This implies that the optimal correlation strength predicted by the asymptotic theory is implicitly assuming a large amount of information, while the information available in the non-asymptotic regime is poorer because μ is low and the prior in equation (10) is only moderately informative. It is thus reasonable to expect that the asymptotically optimal amount of entanglement is *generally inappropriate* in the non-asymptotic regime. One can then try to compensate the low amount of information in the regime with limited data by choosing \mathcal{J} judiciously. In our case, we observe that our functions are clustered around the equally weighted sum of the parameters, since the geometry parameter of the former is $\mathcal{G} \approx 0.853$ and this is relatively close to the geometry parameter of the latter, $\mathcal{G} = 1$. In turn, this motivates choosing a \mathcal{J} that is closer to that associated with $\mathbf{1}$, which is $\mathcal{J} = 1$, in order to enhance the precision when μ is low, and this is what (b) and (c) in figure 5(i) show.

We may push this intuition further and consider a network with $\gamma = 0$, $\mathcal{J} = 1$, which makes the state in equation (40) maximally entangled. Its graph has been labelled as (d) in figure 5(i), and upon comparing it with the three previous strategies we see that the maximally entangled state is the best option when $1 \leq \mu \lesssim 10$. The price that we pay for this low- μ enhancement is that the scheme ceases to be useful after $\mu \approx 20$ trials, and it is asymptotically beaten by the rest of schemes, including the local strategy. We notice that this result is consistent with our analysis in section 4.1, where we established that this probe is only sensitive to the equally weighted sum of the parameters.

The maximally entangled state also illustrates how, despite the lack of an asymptotic approximation in the sense of section 2.2, we can still perform a Bayesian estimation using such strategy, even when it has limited usefulness. On the contrary, for the local, asymptotically optimal and balanced strategies we have that the Bayesian mean square errors converge to their respective Cramér–Rao bounds, as it may be verified by observing figures 5(ii)–(iv). The number of repetitions required for the relative error between these Bayesian uncertainties and the asymptotic bounds to be equal to or less than 5% runs from $\mu \sim 10$ to $\mu \sim 10^2$ (see table 1 for more details).

In summary, in this section we have demonstrated that the strength of the inter-sensor correlations that is useful to estimate a given collection of global properties changes substantially for different amounts of data, i.e., for different values of μ . Since this is the same type of behaviour that we had established for single-parameter schemes in [62], we *conjecture* that the novel effects associated with a limited number of trials, which here have been uncovered using specific examples, are a general feature of non-asymptotic quantum metrology, and that they are generally present in a wide range of experiments operating in the regime of limited data.

Table 1. Properties of different strategies based on a two-parameter qubit network, where γ selects the state and \mathcal{J} is the amount of inter-sensor correlations. The POVM is separable for all four schemes, but only the local strategy is based on a separable state. The asymptotically optimal strategy minimises the quantum Cramér–Rao bound. The balanced strategy has also been enhanced via quantum correlations, but it is not asymptotically optimal because part of this enhancement has been traded to instead enhance its non-asymptotic performance. Finally, the fourth strategy uses a maximally entangled state. We note that the fourth column provides the number of repetitions μ_τ needed such that the relative error between the Bayesian uncertainty and the Cramér–Rao bound is equal to or less than a 5% threshold (see [61]), and in general it depends on the available prior information. Importantly, this calculation does not apply to the strategy with a maximally entangled state, since the estimation uncertainty for the latter does not have an asymptotic limit in the sense of section 2.2. These results demonstrate the state-dependent nature of the conditions required to approach the Cramér–Rao in multi-parameter systems.

Strategy	γ	\mathcal{J}	$\mu_\tau(\Delta_0 = \pi^2/4)$
Local	1	0	4.58×10^2
Asymptotically optimal	0.531	0.561	4.3×10
Balanced enhancement	0.334	0.799	5.37×10^2
Maximally entangled	0	1	—

5. Summary and outlook

The central question addressed in this work has been that of the role of inter-sensor correlations in the estimation of linear functions with arbitrary geometry, having exploited a sensor-symmetric qubit network in the presence of different amounts of data. First we focused on the asymptotic part of the problem, and by optimising the class of sensor-symmetric states, we have established an optimal link between correlation strength and the geometry of the linear functions. Thanks to this we have been able to demonstrate that, while entanglement is useful for many geometrical configurations, it is sometimes detrimental even with functions that are non-trivial global properties. Furthermore, we have found that forms of entanglement other than those of a pairwise nature are in fact irrelevant in this regime. Hence, our approach significantly extends previous studies in networked quantum sensing that had only considered the estimation of a single function or a collection of $l = d$ orthogonal ones.

Given that, in practice, the number of trials μ is always finite and possibly small, we have also performed a non-asymptotic analysis of sensing networks. To this end we have introduced a *hybrid* estimation technique combining asymptotic and non-asymptotic optimisations in Bayesian estimation. This approximate but powerful approach has revealed that the correlation strength that is optimal for sensor-symmetric networks crucially depends on the number of times that we repeat the experiment. Additionally, we have demonstrated how the non-asymptotic precision may be enhanced by trading precision enhancements associated with the asymptotic regime.

Admittedly, while many of our asymptotic results are valid for d parameters, our Bayesian analysis has been restricted to the $d = 2$ case due to numerical complexity. Hence, developing methods to overcome this limitation may have a major impact in the long run. For instance, it would be interesting to examine whether the irrelevancy of forms of entanglement other than those that generate pairwise correlations is also true for a low number of trials, which is a question that requires simulations where $d \geq 3$. One possibility is to modify the multi-parameter algorithm in [45] that we have exploited in section 4, such that the integrals associated with the parameters θ are performed with Monte Carlo techniques. Alternatively, we could employ

some other quantum bound whose calculation is simple enough to study cases where both μ and d are unrestricted. One potential candidate fulfilling the latter is the multi-parameter quantum Ziv–Zakai bound in [9], although, according to our findings in [16], we cannot expect the results derived using this type of tool to be tight in general.

Another important direction for future work is to extend our analysis to include the potential effect that decoherence may have in our conclusions. For example, it would be desirable to establish whether, in such case, inter-sensor correlations are still generally detrimental for the estimation of linear functions whose geometry parameter vanishes, i.e., when $\mathcal{G} = 0$. Note that our hybrid estimation technique can still be employed here, but replacing the ideal quantum Cramér–Rao bound by its version for mixed states when such bound is applied to equation (17).

Notwithstanding these limitations, our methodology has revealed new important aspects of the role of entanglement in the simultaneous estimation of linear functions with networked schemes, and these results could contribute decisively towards a powerful theoretical framework for networked quantum sensing.

Acknowledgments

JR performed and interpreted the calculations of his contribution to this work during his time at the University of Sussex. We thank Jasminder Sidhu, Matthew Thornton and George Knee for helpful discussions. JR and JAD acknowledge support from the South East Physics Network (SEPnet) and the United Kingdom EPSRC through the Quantum Technology Hub: Networked Quantum Information Technology (Grant reference EP/M013243/1), and JR also from Engineering and Physical Sciences Research Council (EPSRC) Grant EP/T002875/1. PAK acknowledges support from the Royal Commission for the Exhibition of 1851. All statements of fact, subjective views, opinions, or conclusions expressed herein are strictly those of the authors; they do not represent the official views or policies of the Department of Energy or the US Government. 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 US Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525.

Appendix A. Constructing the multi-parameter prior probability

Suppose that, according to our prior information about the network, we know that: (a) *a priori* there is no reason to expect that the parameters θ are correlated in any way with each other and (b) we are ignorant of the magnitudes of the parameters, although (c) only within a hypervolume Δ_0 that is centred around $\bar{\theta} = (\bar{\theta}_1, \dots, \bar{\theta}_d)$. The purpose of this appendix is to construct a prior density that codifies this state of information.

Given (a), the parameters are initially thought of as independent in the statistical sense, which in turn allows us to formalise (b) as the assertion that a displacement by an arbitrary real vector c does not change our state of information. That is, θ and $\theta' = \theta + c$ generate equivalent estimation problems.

At the same time, this invariance in our state of information is equivalent to imposing that $p(\theta)d\theta = p(\theta')d\theta' = p(\theta + c)d\theta$, which gives rise to the functional equation $p(\theta) = p(\theta + c)$, and the latter can be satisfied with $p(\theta) \propto 1$.

Finally, (c) indicates that the argument in the previous two paragraphs can only be approximately fulfilled in a portion of the parameter domain with hypervolume Δ_0 centred around $\bar{\theta} = (\bar{\theta}_1, \dots, \bar{\theta}_d)$. Since *a priori* the parameters are thought of as independent, we may express the hypervolume Δ_0 as $\Delta_0 = \prod_{i=1}^d W_{0,i}$, where $W_{0,i}$ is the prior width for the i th parameter.

Therefore, our multi-parameter prior will be

$$p(\boldsymbol{\theta}) = 1/\Delta_0 = 1/\left(\prod_{i=1}^d W_{0,i}\right), \quad (\text{A.1})$$

for $\boldsymbol{\theta} \in [\bar{\theta}_1 - W_{0,1}/2, \bar{\theta}_1 + W_{0,1}/2] \times \cdots \times [\bar{\theta}_d - W_{0,d}/2, \bar{\theta}_d + W_{0,d}/2]$, and zero otherwise, which is the prior introduced in section 2.1 and employed in the main text. We notice that this is a multi-parameter application of a method proposed by Jaynes to construct objective prior probabilities [68, 76]. Other methods can be found in [77, 78].

Appendix B. Optimising the multi-parameter Bayesian uncertainty: review of techniques

There are several ways of addressing the problem of optimising the uncertainty in equation (14) with respect to the estimators $\hat{\boldsymbol{\theta}}(\mathbf{m})$, the measurement scheme $E(m_i)$ and the initial sensor-symmetric state ρ_0 . One option is to perform a direct minimisation [2–5], which is sometimes possible in covariant estimation [7, 23, 47, 79] but generally intractable. Alternatively, one can bound the estimation error and search for the strategy that better approaches that bound, which may be attempted with tools such as the Yuen–Lax bound [1], the quantum Weiss–Weinstein bound [11], or some multi-parameter version of the quantum Ziv–Zakai bound [9, 10], among others [8, 47]. This method usually suffers from the lack of tightness of the bounds, although this can be partially overcome with the Bayesian analogue of the Helstrom Cramér–Rao bound that we recently constructed in [16] (see also [46, 47]), since it can be saturated in certain cases and we showed how to exploit it for the estimation of local parameters (i.e., $\boldsymbol{\theta}$). Nevertheless, we have followed the weaker but computationally simpler hybrid approach in section 2.2 because the theory of estimating global properties of a network is more challenging, and we leave the application of more sophisticated methods to the estimation of linear functions for future work.

Appendix C. Minimisation of the asymptotic uncertainty for linear functions

The optimal strength for the inter-sensor correlations in equation (35) can be found as follows. If we look at $\bar{\epsilon}_f = \mathcal{N}h(\mathcal{J}, \mathcal{G}, d)/\mu$ as a function of \mathcal{J} , where we recall that, according to the discussion in section 3.1,

$$h(\mathcal{J}, \mathcal{G}, d) = \frac{[1 + (d - 2 - \mathcal{G})\mathcal{J}]}{(1 - \mathcal{J})[1 + (d - 1)\mathcal{J}]}, \quad (\text{C.1})$$

then the equation for its extrema is

$$\frac{\mathcal{N}}{\mu} \frac{\partial h(\mathcal{J}, \mathcal{G}, d)}{\partial \mathcal{J}} = \frac{\mathcal{N}}{\mu} \frac{(d - 1)(d - 2 - \mathcal{G})\mathcal{J}^2 + 2(d - 1)\mathcal{J} - \mathcal{G}}{(1 - \mathcal{J})^2[1 + (d - 1)\mathcal{J}]^2} = 0, \quad (\text{C.2})$$

whose solutions are

$$\mathcal{J}_{\pm} = \frac{1}{\mathcal{G} + 2 - d} \left[1 \mp \sqrt{\frac{(\mathcal{G} + 1)(d - 1 - \mathcal{G})}{d - 1}} \right]. \quad (\text{C.3})$$

Since we need to restrict our study to the range $1/(1 - d) < \mathcal{J} < 1$ for F_q to be invertible, only \mathcal{J}_+ is a valid candidate to find a minimum. Next we examine the sign of the slope in the

left-hand side of equation (C.2) for some values of \mathcal{J} around \mathcal{J}_+ . By noticing that $\mathcal{N}/\mu > 0$ and using the endpoints of the domain for \mathcal{J} we find that

$$\frac{\partial h(1 - \varepsilon, \mathcal{G}, d)}{\partial \mathcal{J}} > 0, \quad \frac{\partial h(1/(1 - d) + \varepsilon, \mathcal{G}, d)}{\partial \mathcal{J}} < 0 \quad (\text{C.4})$$

for an arbitrarily small $\varepsilon > 0$ when $\mathcal{G} \neq -1$, $\mathcal{G} \neq d - 1$, which we exclude to guarantee that $\mathcal{J} \neq 1/(1 - d)$, $\mathcal{J} \neq 1$. Consequently, \mathcal{J}_+ gives rise to the minimum that we were looking for.





Appendix D. Eigendecomposition of \mathcal{X}

The characteristic equation for $\mathcal{X} = \mathcal{I} - \mathbb{I}$ is

$$\det(\mathcal{X} - \lambda \mathbb{I}) = \det[\mathbf{1}\mathbf{1}^\top - (1 + \lambda)\mathbb{I}] \propto (1 - d + \lambda)(1 + \lambda)^{d-1} = 0, \quad (\text{D.1})$$

giving the eigenvalues $\lambda_1 = d - 1$, with multiplicity 1, and $\lambda_2 = -1$, with multiplicity $d - 1$ (see the calculations associated with equation (23), whose eigenvalues are obtained in the same way). By inspection we see that $\mathbf{1}$ is one of the eigenvectors. Since the latter satisfies that $\mathcal{X}\mathbf{1} = (\mathbf{1}\mathbf{1}^\top - \mathbb{I})\mathbf{1} = (d - 1)\mathbf{1}$, the rest of the eigenvalues must be associated with the subspace orthogonal to $\mathbf{1}$, and this concludes the eigendecomposition of \mathcal{X} .

ORCID iDs

Jesús Rubio  <https://orcid.org/0000-0002-8193-8273>
 Paul A Knott  <https://orcid.org/0000-0002-3809-3578>
 Timothy J Proctor  <https://orcid.org/0000-0003-0219-8930>
 Jacob A Dunningham  <https://orcid.org/0000-0001-7976-3278>

References

- [1] Yuen H and Lax M 1973 Multiple-parameter quantum estimation and measurement of nonselfadjoint observables *IEEE Trans. Inf. Theory* **19** 740–50
- [2] Holevo A S 1973 Statistical problems in quantum physics *Proceedings of the Second Japan-USSR Symposium on Probability Theory* ed G Maruyama and Y V Prokhorov (Berlin: Springer) pp 104–19
- [3] Holevo A S 1973 Statistical decision theory for quantum systems *J. Multivariate Anal.* **3** 337–94
- [4] Helstrom C and Kennedy R 1974 Noncommuting observables in quantum detection and estimation theory *IEEE Trans. Inf. Theory* **20** 16–24
- [5] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
- [6] Paris M G A 2009 Quantum estimation for quantum metrology *Int. J. Quantum Inf.* **07** 125–37
- [7] Holevo A S 2011 *Probabilistic and Statistical Aspects of Quantum Theory* (Berlin: Springer)
- [8] Gill R D and Guţă M I 2013 *On Asymptotic Quantum Statistical Inference (Collections vol 9)* (Beachwood: Institute of Mathematical Statistics) pp 105–27
- [9] Zhang Y-R and Fan H 2014 Quantum metrological bounds for vector parameters *Phys. Rev. A* **90** 043818
- [10] Berry D W, Tsang M, Hall M J W and Wiseman H M 2015 Quantum Bell–Ziv–Zakai bounds and Heisenberg limits for waveform estimation *Phys. Rev. X* **5** 031018
- [11] Lu X-M and Tsang M 2016 Quantum Weiss–Weinstein bounds for quantum metrology *Quantum Sci. Technol.* **1** 015002
- [12] Ragy S, Jarzyna M and Demkowicz-Dobrzański R 2016 Compatibility in multiparameter quantum metrology *Phys. Rev. A* **94** 052108

- [13] Pezzè L, Ciampini M A, Spagnolo N, Humphreys P C, Datta A, Walmsley I A, Barbieri M, Sciarrino F and Smerzi A 2017 Optimal measurements for simultaneous quantum estimation of multiple phases *Phys. Rev. Lett.* **119** 130504
- [14] Hall M J W 2018 Entropic Heisenberg limits and uncertainty relations from the Holevo information bound *J. Phys. A: Math. Theor.* **51** 364001
- [15] Gessner M, Pezzè L and Smerzi A 2018 Sensitivity bounds for multiparameter quantum metrology *Phys. Rev. Lett.* **121** 130503
- [16] Rubio J and Dunningham J 2020 Bayesian multiparameter quantum metrology with limited data *Phys. Rev. A* **101** 032114
- [17] Yang Y, Chiribella G and Hayashi M 2019 Attaining the ultimate precision limit in quantum state estimation *Commun. Math. Phys.* **368** 223–93
- [18] Carollo A, Spagnolo B, Dubkov A A and Valenti D 2019 On quantumness in multi-parameter quantum estimation *J. Stat. Mech.* **094010**
- [19] Albarelli F, Friel J F and Datta A 2019 Evaluating the Holevo Cramér–Rao bound for multiparameter quantum metrology *Phys. Rev. Lett.* **123** 200503
- [20] Tsang M 2019 The Holevo Cramér–Rao bound is at most thrice the Helstrom version (arXiv:1911.08359)
- [21] Sidhu J S, Ouyang Y, Campbell E T and Kok P 2019 Tight bounds on the simultaneous estimation of incompatible parameters (arXiv:1912.09218)
- [22] Albarelli F, Tsang M and Datta A 2019 Upper bounds to the Holevo Cramér–Rao bound for multiparameter parametric and semiparametric estimation (arXiv:1911.11036)
- [23] Macchiavello C 2003 Optimal estimation of multiple phases *Phys. Rev. A* **67** 062302
- [24] Spagnolo N, Aparo L, Vitelli C, Crespi A, Ramponi R, Osellame R, Mataloni P and Sciarrino F 2012 Quantum interferometry with three-dimensional geometry *Sci. Rep.* **2** 862
- [25] Chiribella G 2012 Optimal networks for quantum metrology: semidefinite programs and product rules *New J. Phys.* **14** 125008
- [26] Humphreys P C, Barbieri M, Datta A and Walmsley I A 2013 Quantum enhanced multiple phase estimation *Phys. Rev. Lett.* **111** 070403
- [27] Vidrighin M D, Donati G, Genoni M G, Jin X-M, Kolthammer W S, Kim M S, Datta A, Barbieri M and Walmsley I A 2014 Joint estimation of phase and phase diffusion for quantum metrology *Nat. Commun.* **5** 3532
- [28] Baumgratz T and Datta A 2016 Quantum enhanced estimation of a multidimensional field *Phys. Rev. Lett.* **116** 030801
- [29] Knott P A, Proctor T J, Hayes A J, Ralph J F, Kok P and Dunningham J A 2016 Local versus global strategies in multiparameter estimation *Phys. Rev. A* **94** 062312
- [30] Gagatsos C N, Branford D and Datta A 2016 Gaussian systems for quantum-enhanced multiple phase estimation *Phys. Rev. A* **94** 042342
- [31] Sidhu J S and Kok P 2017 Quantum metrology of spatial deformation using arrays of classical and quantum light emitters *Phys. Rev. A* **95** 063829
- [32] Proctor T J, Knott P A and Dunningham J A 2017 Networked quantum sensing (arXiv:1702.04271)
- [33] Proctor T J, Knott P A and Dunningham J A 2018 Multiparameter estimation in networked quantum sensors *Phys. Rev. Lett.* **120** 080501
- [34] Szczykulska M, Baumgratz T and Datta A 2017 Reaching for the quantum limits in the simultaneous estimation of phase and phase diffusion *Quantum Sci. Technol.* **2** 044004
- [35] Zhang L and Chan K W C 2017 Quantum multiparameter estimation with generalized balanced multimode NOON-like states *Phys. Rev. A* **95** 032321
- [36] Zhuang Q, Zhang Z and Shapiro J H 2017 Entanglement-enhanced lidars for simultaneous range and velocity measurements *Phys. Rev. A* **96** 040304
- [37] Altenburg S and Wölk S 2018 Multi-parameter estimation: global, local and sequential strategies *Phys. Scr.* **94** 014001
- [38] Sidhu J S and Kok P 2018 Quantum Fisher information for general spatial deformations of quantum emitters (arXiv:1802.01601)
- [39] Albarelli F, Barbieri M, Genoni M G and Gianani I 2020 A perspective on multiparameter quantum metrology: from theoretical tools to applications in quantum imaging *Phys. Lett. A* **384** 126311
- [40] Roccia E, Cimini V, Sbroscia M, Gianani I, Ruggiero L, Mancino L, Genoni M G, Ricci M A and Barbieri M 2018 Multiparameter approach to quantum phase estimation with limited visibility *Optica* **5** 1171–6

- [41] Polino E, Riva M, Valeri M, Silvestri R, Corrielli G, Crespi A, Spagnolo N, Osellame R and Sciarrino F 2019 Experimental multiphase estimation on a chip *Optica* **6** 288–95
- [42] Valeri M, Polino E, Poderini D, Gianani I, Corrielli G, Crespi A, Osellame R, Spagnolo N and Sciarrino F 2020 Experimental adaptive Bayesian estimation of multiple phases with limited data (arXiv:2002.01232)
- [43] Szczykulska M, Baumgratz T and Datta A 2016 Multi-parameter quantum metrology *Adv. Phys. X* **1** 621–39
- [44] Liu J, Yuan H, Lu X-M and Wang W 2019 Quantum Fisher information matrix and multiparameter estimation *J. Phys. A: Math. Theor.* **53** 023001
- [45] Rubio Jiménez J 2020 Non-asymptotic quantum metrology: extracting maximum information from limited data *PhD Thesis* University of Sussex
- [46] Sidhu J S and Kok P 2020 Geometric perspective on quantum parameter estimation *AVS Quantum Sci.* **2** 014701
- [47] Demkowicz-Dobrzański R, Gorecki W and Guta M 2020 Multi-parameter estimation beyond quantum Fisher information *J. Phys. A: Math. Theor.* (accepted) <https://doi.org/10.1088/1751-8121/ab8ef3>
- [48] Polino E, Valeri M, Spagnolo N and Sciarrino F 2020 Photonic quantum metrology (arXiv:2003.05821)
- [49] Eldredge Z, Foss-Feig M, Gross J A, Rolston S L and Gorshkov A V 2018 Optimal and secure measurement protocols for quantum sensor networks *Phys. Rev. A* **97** 042337
- [50] Ge W, Jacobs K, Eldredge Z, Gorshkov A V and Foss-Feig M 2018 Distributed quantum metrology with linear networks and separable inputs *Phys. Rev. Lett.* **121** 043604
- [51] Zhuang Q, Zhang Z and Shapiro J H 2018 Distributed quantum sensing using continuous-variable multipartite entanglement *Phys. Rev. A* **97** 032329
- [52] Qian K, Eldredge Z, Ge W, Pagano G, Monroe C, Porto J V and Gorshkov A V 2019 Heisenberg-scaling measurement protocol for analytic functions with quantum sensor networks *Phys. Rev. A* **100** 042304
- [53] Eldredge Z 2019 Generation and uses of distributed entanglement in quantum information *PhD Thesis* University of Maryland, ProQuest Dissertations Publishing
- [54] Sekatski P, Wölk S and Dür W 2020 Optimal distributed sensing in noisy environments *Phys. Rev. Res.* **2** 023052
- [55] Gatto D, Facchi P, Narducci F A and Tamma V 2019 Distributed quantum metrology with a single squeezed-vacuum source *Phys. Rev. Res.* **1** 032024
- [56] Guo X, Breum C R, Borregaard J, Izumi S, Larsen M V, Gehring T, Christandl M, Neergaard-Nielsen J S and Andersen U L 2020 Distributed quantum sensing in a continuous-variable entangled network *Nat. Phys.* **16** 281–4
- [57] Oh C, Lee C, Lie S H and Jeong H 2020 Optimal distributed quantum sensing using Gaussian states *Phys. Rev. Res.* **2** 023030
- [58] Zhuang Q, Preskill J and Jiang L 2020 Distributed quantum sensing enhanced by continuous-variable error correction *New J. Phys.* **22** 022001
- [59] Gross J A and Caves C M 2020 One from many: estimating a function of many parameters (arXiv:2002.02898)
- [60] Li X, Cao J-H, Liu Q, Tey M K and You L 2020 Multi-parameter estimation with multi-mode Ramsey interferometry *New J. Phys.* **22** 043005
- [61] Rubio J, Knott P A and Dunningham J A 2018 Non-asymptotic analysis of quantum metrology protocols beyond the Cramér–Rao bound *J. Phys. Commun.* **2** 015027
- [62] Rubio J and Dunningham J A 2019 Quantum metrology in the presence of limited data *New J. Phys.* **21** 043037
- [63] Hofmann H F 2009 All path-symmetric pure states achieve their maximal phase sensitivity in conventional two-path interferometry *Phys. Rev. A* **79** 033822
- [64] Sahota J and Quesada N 2015 Quantum correlations in optical metrology: Heisenberg-limited phase estimation without mode entanglement *Phys. Rev. A* **91** 013808
- [65] Demkowicz-Dobrzański R 2011 Optimal phase estimation with arbitrary *a priori* knowledge *Phys. Rev. A* **83** 061802
- [66] Demkowicz-Dobrzański R, Jarzyna M and Kołodyński J 2015 Quantum limits in optical interferometry *Prog. Opt.* **60** 345–435
- [67] Friis N, Orsucci D, Skotiniotis M, Sekatski P, Dunjko V, Briegel H J and Dür W 2017 Flexible resources for quantum metrology *New J. Phys.* **19** 063044

- [68] Jaynes E T 2003 *Probability Theory: The Logic of Science* (Cambridge: Cambridge University Press)
- [69] Kay S M 1993 *Fundamentals of Statistical Signal Processing: Estimation Theory* (Englewood Cliffs, NJ: Prentice Hall)
- [70] Cox D R and Hinkley D V 2000 *Theoretical Statistics* (London: Chapman & Hall)
- [71] Bernardo J M 1994 *Bayesian Theory* (Chichester: Wiley)
- [72] Le Cam L M 1986 *Asymptotic Methods in Statistical Decision Theory* (Berlin: Springer)
- [73] van der Vaart A W 1998 *Asymptotic Statistics* (Cambridge: Cambridge University Press)
- [74] Kok P, Dunningham J and Ralph J F 2017 Role of entanglement in calibrating optical quantum gyroscopes *Phys. Rev. A* **95** 012326
- [75] Kołodyński J 2014 Precision bounds in noisy quantum metrology *PhD Thesis* University of Warsaw (arXiv:1409.0535)
- [76] Jaynes E T 1968 Prior probabilities *IEEE Trans. Syst. Cybern.* **4** 227–41
- [77] Kass R E and Wasserman L 1996 The selection of prior distributions by formal rules *J. Am. Stat. Assoc.* **91** 1343–70
- [78] von Toussaint U 2011 Bayesian inference in physics *Rev. Mod. Phys.* **83** 943–99
- [79] Chiribella G, D’Ariano G M and Sacchi M F 2005 Optimal estimation of group transformations using entanglement *Phys. Rev. A* **72** 042338