



# Non-adaptive Heisenberg-limited metrology with multi-channel homodyne measurements

Danilo Triggiani<sup>1,a</sup> , Paolo Facchi<sup>2,3</sup> , Vincenzo Tamma<sup>1,4,b</sup> 

<sup>1</sup> School of Mathematics and Physics, University of Portsmouth, Portsmouth PO1 3QL, UK

<sup>2</sup> Dipartimento di Fisica and MECENAS, Università di Bari, 70126 Bari, Italy

<sup>3</sup> INFN, Sezione di Bari, 70126 Bari, Italy

<sup>4</sup> Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK

Received: 7 October 2021 / Accepted: 29 December 2021

© The Author(s) 2022

**Abstract** We show a protocol achieving the ultimate Heisenberg-scaling sensitivity in the estimation of a parameter encoded in a generic linear network, without employing any auxiliary networks, and without the need of any prior information on the parameter nor on the network structure. As a result, this protocol does not require a prior coarse estimation of the parameter, nor an adaptation of the network. The scheme we analyse consists of a single-mode squeezed state and homodyne detectors in each of the  $M$  output channels of the network encoding the parameter, making it feasible for experimental applications.

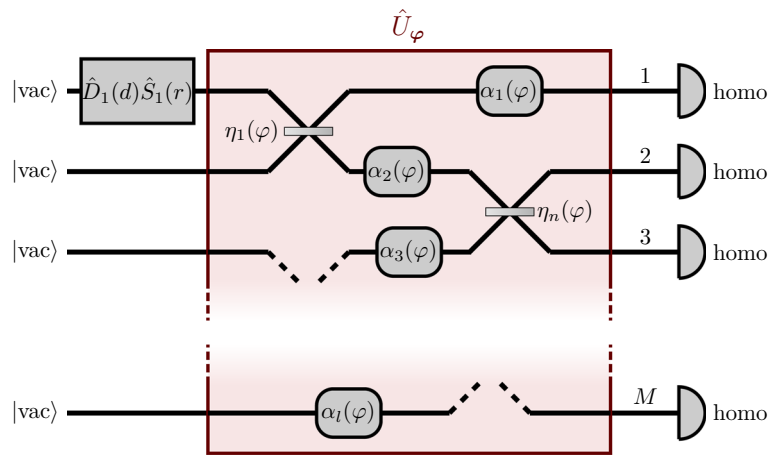
## 1 Introduction

Increasing the level of precision achievable in the estimation of physical properties of systems, such as temperatures, optical lengths and magnitude of external fields among others, is one of the multiple applications of quantum technologies that have been extensively studied in the recent years. In particular, the goal of quantum metrology—the field of science laying between quantum mechanics and estimation theory—is to propose and analyse estimation protocols that surpass the precision achievable by classical strategies by employing quantum probes and quantum measurement schemes. In fact it is well known that the classical limit on the precision achievable in the estimation of an unknown parameter when employing  $N$  probes, known as shot-noise limit, for which the error is of order  $1/\sqrt{N}$ , can be surpassed by quantum strategies achieving the ultimate Heisenberg limit, where the estimation scales as  $1/N$  [1–9].

The first proposed protocols reaching Heisenberg-scaling sensitivity heavily employed entanglement as a metrological resource [1–3], and several entanglement-based strategies have been recently studied with interesting results, especially in the cases of simultaneous estimation of multiple parameters with non-commuting generators [5–10]. Nonetheless, the entanglement fragility and the complicated procedures needed to generate entangled metrological probes, such as NOON or GHZ states [11, 12], are two of the challenges that stimulated the search for more feasible estimation schemes making use of protocols implementing metrological resources that are easier to generate and to manipulate. Squeezed light [13, 14] manifests useful properties (e.g. robustness to decoherence, relatively easy implementation, reduced noise below the vacuum shot-noise) which make it a perfect candidate as a feasible metrological resource [15–19]. Motivated by these favourable properties, many works have recently focused on the analysis and proposal of Gaussian metrological schemes, namely involving squeezed states as probes and homodyne detection as measurement, and the ultimate precision that these can achieve in the estimation of a single localised parameter [15–18, 20, 21], a function of parameters [22–24], or a single distributed parameter, such as the temperature or the electromagnetic field, affecting several component of the network [8, 25–30]. These schemes typically perform an analysis based on the study of the Cramér–Rao bound (CRB) [31], or its quantum counterpart [32], to assess the ultimate precision achievable in the estimation, and show whether the Heisenberg scaling can be achieved. Although the CRB analysis does not generally assure that the ultimate bound on the precision found can be achieved globally, namely with an estimation strategy which does not depend on the true value of the parameter to be estimated [33, 34], local estimations are relevant in the typical interferometric framework, in which small deviations of the parameters need to be measured. Interestingly, it has been recently found that it is always possible to reach Heisenberg-scaling sensitivity regardless of the structure of the network encoding the parameter, only employing a single squeezed vacuum state, a single-homodyne measurement, and an auxiliary network suitably engineered, whose preparation only requires a knowledge on the unknown parameter that can be obtained by a classical measurement [26, 29, 30]. The need for an auxiliary stage in such protocols arises from the fact that in general the probe is scattered by the network in all its output ports, while only a single port is eventually

<sup>a</sup> e-mail: [danilo.triggiani@port.ac.uk](mailto:danilo.triggiani@port.ac.uk) (corresponding author)

<sup>b</sup> e-mail: [vincenzo.tamma@port.ac.uk](mailto:vincenzo.tamma@port.ac.uk)



**Fig. 1** Optical set-up for the estimation at the Heisenberg-scaling precision of a single unknown parameter  $\varphi$  encoded arbitrarily in a generic passive  $M$ -channel linear network. The parameter can be either localised in a single element of the network, or represent a global property affecting several components, such as a temperature or a magnetic field. A single source of coherent squeezed states with  $N = N_D + N_S$  average photons, where  $N_D = d^2$  and  $N_S = \sinh^2 r$  are the number of displaced and squeezed photons, respectively, is employed in a single input channel (the first in figure), while homodyne measurements are taken in every output channel

measured through homodyne detection, so that an auxiliary network is required in order to refocus the probe on the only channel observed. A question that naturally arises is whether incrementing the number of observed channels would ease, if not completely lift, the requirement of an auxiliary stage and ultimately the requirement of a prior classical knowledge on the unknown parameter. Moreover, different Gaussian protocols rely on encoding the information about the unknown parameter on the displacement of the probe, requiring that a portion of the resources in the probe are employed in a non-vanishing displacement [28, 35]. Despite concentrating all the photons in the squeezing is known to be the optimal allocation of the resources in the probe [26], encoding the parameter into a non-vanishing displacement can reduce the estimation process into the relatively simple task of inferring the parameter from the expectation value of a Gaussian probability density function [28].

In this work we investigate the ultimate precision achievable in the estimation of a parameter encoded in a generic linear network, when employing a single-mode squeezed coherent Gaussian state and performing homodyne detection on all the output channels (see Fig. 1). We show that, without making any assumption on the structure of the linear network nor on the nature of the parameter, it is always possible to reach Heisenberg-scaling sensitivity with such set-up, without the use of any auxiliary network. This allows for estimation protocols not requiring a preparatory stage nor a prior coarse estimation of the parameter, as opposed to the single-channel homodyne protocols in Refs. [29, 30]. We also show that two independent contributions on the precision arise from our analysis: one originated from the presence of displaced photons in addition to squeezed photons, and the other from the squeezing of the probe. We find that both contributions can reach Heisenberg-scaling sensitivity independently, and this can be achieved expectedly when the local oscillator phases are chosen such that the noise in the outcome is reduced, namely when the squeezed quadratures are observed in each output channel. Thus, differently from the schemes in Refs. [29, 30], it is then possible to employ this set-up to retrieve information on the parameter through measurements of the average value (i.e. the displacement) of the signal observed with homodyne detection, as well as through the modulation of the noise. Although it is not required to reach the Heisenberg-scaling sensitivity, the presence of an auxiliary network in general affects the precision of the estimation through a pre-factor multiplying the scaling. This comes in useful in those cases where priority is given to increasing the precision, at the expenses of engineering an auxiliary network to be added before the estimation protocol is started.

### 2 Set-up

Let us consider a generic  $M \times M$  passive linear network whose action on any injected photon probe is given by the unitary operator  $\hat{U}_\varphi$ , in which the unknown parameter  $\varphi$  to be estimated is encoded in an arbitrary manner. The linearity and passivity of the network allow us to describe it with an  $M \times M$  unitary matrix  $U_\varphi$  related to the evolution operator  $\hat{U}_\varphi$  by

$$\hat{U}_\varphi^\dagger \hat{a}_i \hat{U}_\varphi = \sum_{j=1}^M (U_\varphi)_{ij} \hat{a}_j. \tag{1}$$

The input probe is prepared in a single-mode squeezed coherent state  $|\psi_{\text{in}}\rangle = \hat{D}_1(d)\hat{S}_1(r)|\text{vac}\rangle$  with  $N = \sinh^2 r + d^2/2 \equiv N_S + N_D$  average number photons, where  $\hat{S}_1(r) = \exp(r(\hat{a}_1^{\dagger 2} - \hat{a}_1^2)/2)$  is the squeezing operator with real squeezing parameter  $r$ , and

$\hat{D}_1(d) = \exp(d(\hat{a}_1^\dagger - \hat{a}_1)/\sqrt{2})$  is the displacement operator with real displacement  $d$ , and it is injected in one input channel, say the first, of the linear network. In this case, only the first row of  $U_\varphi$  is relevant in this protocol

$$(U_\varphi)_{1j} = \sqrt{P_j} e^{i\bar{\gamma}_j}, \tag{2}$$

where we have made explicit the probability  $P_j$  that each photon exits from the  $j$ th output port of the network, and the phase  $\bar{\gamma}_j$  acquired in the process, with  $j$  from 1 to  $M$ . The unitarity of  $U_\varphi$  assures that  $\sum_j P_j = 1$ .

A homodyne detection is then performed at each of the output channels, and the quadratures  $\hat{x}_{i,\theta_i}$  are measured, where  $\theta_i$  is the  $i$ th local oscillator reference phase, from which we want to infer the value of  $\varphi$ . Due to the Gaussian nature of the scheme, the joint probability distribution  $p(\mathbf{x}|\varphi)$  associated with the  $M$ -mode homodyne measurement is Gaussian

$$p(\mathbf{x}|\varphi) = \frac{1}{\sqrt{(2\pi)^M |\Sigma|}} \exp\left[-\frac{(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})}{2}\right]. \tag{3}$$

Here,  $\Sigma$  is the  $M \times M$  covariance matrix with elements (see ‘‘Appendix A’’)

$$\Sigma_{ij} = \frac{\delta_{ij}}{2} + \sqrt{P_i P_j} (\cos(\gamma_i - \gamma_j) \sinh(r)^2 + \cos(\gamma_i + \gamma_j) \cosh(r) \sinh(r)), \tag{4}$$

where  $\delta_{ij}$  is the Kronecker delta,  $\gamma_i = \bar{\gamma}_i - \theta_i$  is the phase delay at the output of the  $i$ th channel relative to the correspondent local oscillator, and  $|\Sigma|$  is the determinant of  $\Sigma$ , which reads (see ‘‘Appendix A’’)

$$\begin{aligned} |\Sigma| = & \frac{1}{2^M} + \frac{\sinh(r)}{2^{M-1}} \sum_{i=1}^M P_i (\sinh(r) + \cos(2\gamma_i) \cosh(r)) \\ & - \frac{\sinh^2(r)}{2^{M-2}} \sum_{i=1}^M \sum_{j=i+1}^M P_i P_j \sin^2(\gamma_i - \gamma_j), \end{aligned} \tag{5}$$

and  $\boldsymbol{\mu}$  is the mean vector

$$\mu_i = d\sqrt{P_i} \cos \gamma_i. \tag{6}$$

For any given unbiased estimator  $\tilde{\varphi}$ , the statistical error in the estimation of  $\varphi$  after  $\nu$  iterations of the measurement is limited by the Cram er-Rao bound (CRB) [31]

$$\text{Var}[\tilde{\varphi}] \geq \frac{1}{\nu \mathcal{F}(\varphi)}, \tag{7}$$

where  $\mathcal{F}(\varphi)$  is the Fisher information

$$\mathcal{F}(\varphi) = \int d\mathbf{x} p(\mathbf{x}|\varphi) (\partial_\varphi \log p(\mathbf{x}|\varphi))^2, \tag{8}$$

associated with the Gaussian distribution (3), and reads (see ‘‘Appendix B’’)

$$\mathcal{F}(\varphi) = \frac{1}{|\Sigma|} \partial_\varphi \boldsymbol{\mu}^T C \partial_\varphi \boldsymbol{\mu} + \frac{1}{2} \left( \frac{\partial_\varphi |\Sigma|}{|\Sigma|} \right)^2 - \frac{1}{2|\Sigma|} \text{Tr}[(\partial_\varphi \Sigma)(\partial_\varphi C)], \tag{9}$$

where  $C = |\Sigma| \Sigma^{-1}$  is the cofactor matrix of  $\Sigma$  and  $\text{Tr}[\cdot]$  denotes the trace. In the following we will discuss in detail expression (9) in the asymptotic limit of large  $N$ , showing which condition must be met in order for this set-up to reach Heisenberg-scaling precision in the estimation of  $\varphi$ , and compare differences and advantages with respect to the Heisenberg-scaling single-homodyne schemes [29,30].

We conclude this section by remarking that, in the case of a single channel,  $M = 1$ , the last term in the right-hand side of (9) vanishes, and thus the only relevant terms are the first two, containing the derivative of the mean  $\boldsymbol{\mu}$  and of the determinant of  $\Sigma$ , which reduces to the variance of a single-homodyne measurement. Interestingly enough, we will show that also in the multi-homodyne case, only the first two terms are relevant for the Heisenberg scaling in the asymptotic regime.

### 3 Heisenberg scaling of the Fisher information

In order to investigate the asymptotic behaviour of the Fisher information (9), it is convenient to express the elements of the cofactor matrix  $C$  in terms of the squeezing factor  $r$  (see ‘‘Appendix B’’):

$$C_{ss} = \frac{1}{2^{M-1}} + \frac{1}{2^{M-2}} \sum_{\substack{i=1 \\ i \neq s}}^M \left( \Sigma_{ii} - \frac{1}{2} \right) - \frac{1}{2^{M-3}} \sum_{\substack{i=1 \\ i \neq s}}^M \sum_{\substack{j=i+1 \\ j \neq s}}^M S_{ij},$$

$$C_{st} = -\frac{1}{2^{M-2}} \Sigma_{st} + \frac{1}{2^{M-3}} \sum_{\substack{i=1 \\ i \neq s,t}}^M S_{sti}, \quad s \neq t, \quad (10)$$

where

$$S_{sti} = \sinh^2 r \sqrt{P_s P_t} P_i \sin(\gamma_s - \gamma_i) \sin(\gamma_t - \gamma_i). \quad (11)$$

Notice that every element of  $C$  in the previous expressions, and of  $\Sigma$  in (4), scale at most as quick as  $N$ , namely  $C_{st} = O(N_S)$  and  $\Sigma_{st} = O(N_S)$ , while the mean vector  $\boldsymbol{\mu}$  is of order  $O(\sqrt{N_D})$  (see ‘‘Appendix C’’), and the same asymptotic bounds hold for their derivatives with respect to  $\varphi$ , since neither  $P_i$  nor  $\bar{\gamma}_i$  depend on  $N$ . For this reason, in order for the Fisher information in (9) to asymptotically grow with Heisenberg scaling, it is essential to study the asymptotics of the determinant  $|\Sigma|$  in (5) and find the conditions for which it does not grow with  $N$ .

In fact, it is evident from Eq. (5) that in general  $|\Sigma| = O(N_S)$ , and we show in ‘‘Appendix C’’ that the necessary condition for it to scale slower than  $N_S$  is that the relative phases  $\gamma_i$  tend to  $\pm\pi/2$  for large  $N_S$ : in other words, the larger the number of photons employed in the squeezing of the probe to reach higher precisions, the closer the local oscillator phase needs to be tuned to the minimum-variance quadrature of each mode.

More precisely, as shown in ‘‘Appendix C’’, the conditions to reach Heisenberg scaling in the Fisher information (9), read

$$\gamma_i = \pm \frac{\pi}{2} + O(N_S^{-1}), \quad i = 1, \dots, M. \quad (12)$$

When these conditions hold, we can introduce the finite quantities  $k_i = \lim_{N_S \rightarrow \infty} N_S(\gamma_i \mp \pi/2)$ , and the determinant  $|\Sigma|$  reduces to

$$|\Sigma| = \frac{1}{2^{M-2} N_S} \left( \left( \sum_{i=1}^M P_i k_i \right)^2 + \frac{1}{16} \right), \quad (13)$$

while  $\partial_\varphi |\Sigma|$ ,  $\partial_\varphi \Sigma$ ,  $\partial_\varphi C$  and  $C$  tend to constant values, and  $\partial_\varphi \boldsymbol{\mu}$  scales as  $\sqrt{N_D}$ , thus making only the first two terms of the Fisher information dominant for large  $N$ .

As expected, the determinant of the covariance matrix  $\Sigma$  reaches its minimum value when  $\gamma_i = \pi/2$ , or  $k_i = 0$  for  $i = 1, \dots, M$ , namely when the squeezed quadratures are measured. When conditions (12) are met, we can neglect the trace term in Eq. (9), and we can write

$$\begin{aligned} \mathcal{F}(\varphi) &\simeq \frac{1}{|\Sigma|} \partial_\varphi \boldsymbol{\mu}^T C \partial_\varphi \boldsymbol{\mu} + \frac{1}{2} \left( \frac{\partial_\varphi |\Sigma|}{|\Sigma|} \right)^2 \\ &\simeq 8(\partial\gamma)_{\text{avg}}^2 (2\zeta(k_{\text{avg}}) N_D N_S + \varrho(k_{\text{avg}}) N_S^2), \end{aligned} \quad (14)$$

where  $k_{\text{avg}} \equiv \sum_{i=1}^M P_i k_i$ ,  $(\partial\gamma)_{\text{avg}} \equiv \sum_{i=1}^M P_i \partial_\varphi \gamma_i$ , and  $\varrho(x) = (8x)^2 / (16x^2 + 1)^2$  and  $\zeta(x) = (16x^2 + 1)^{-1}$  are positive, even function which reach their maxima at  $x = \pm 1/4$  and  $x = 0$ , respectively, namely  $\varrho(1/4) = 1$  and  $\zeta(0) = 1$ . The Cram er-Rao bound (7) with the Fisher information (14) is saturated for large  $\nu$  by the maximum-likelihood estimator [36–38], and thus, Heisenberg-scaling precision can be achieved. The expression of the maximum-likelihood estimator for this estimation scheme can be found in ‘‘Appendix D’’.

Noticeably, both terms in the asymptotic Fisher information (14) give a Heisenberg-scaling precision in the estimation of the parameter  $\varphi$ , provided that both the average number of photons in the displacement  $N_D$  and in the squeezing  $N_S$  scale with the total average number of photons  $N = N_S + N_D$ , namely  $N_S = \beta N$  and  $N_D = (1 - \beta)N$ , for any value  $0 < \beta \leq 1$  independent of  $N$ .

Moreover, it is worth noticing that the first term in Eq. (9), and thus in Eq. (14), depends on the information encoded in the displacement of the probe, and thus it vanishes if  $\boldsymbol{\mu} = 0$ , namely if the probe is a squeezed vacuum and  $N_D = 0$ . The second term instead depends on the information on  $\varphi$  encoded in the variance of the measurement itself: it arises only from the interaction with the squeezed photons and vanishes if  $\partial_\varphi |\Sigma| = 0$ , namely when  $k_{\text{avg}} = 0$  in Eq. (14), and in particular when  $\gamma_i = \pm\pi/2$ , for  $i = 1, \dots, M$ , in Eq. (12), corresponding to quadratures with minimum squeezed variances, and thus locally insensible to the variations of the parameter.

Interestingly, this latter case is similar to the single squeezed vacuum and single-homodyne scenario found in the literature [29, 30]: in fact, the second term in (14) represents a generalisation of the single-homodyne Fisher information  $\mathcal{F}_1(\varphi) = 8\varrho(k)(\partial_\varphi \gamma)^2 N^2$ , and it can be obtained by substituting  $k$  and  $\partial_\varphi \gamma$  with their averages over the probabilities  $P_i$ , namely  $k \rightarrow \sum_i P_i k_i$  and  $\partial_\varphi \gamma \rightarrow \sum_i P_i \partial_\varphi \gamma_i$ .

We have then found that, also when employing multiple homodyne detections, one for each output port of the interferometer, the Heisenberg-scaling precision obtained through measurements of the squeezed noise (i.e.  $\partial_\varphi \boldsymbol{\mu} = 0$ ) is only reached when the quantum fluctuations of the observed quadratures are reduced to their quantum limit, i.e.  $|\Sigma| = O(N_S^{-1})$ , while the variations of the unknown parameter  $\varphi$  still yield a visible effect on the outcomes of the measurements, i.e.  $\partial_\varphi |\Sigma|$  is not vanishing.

However, at the expense of introducing a nonzero displacement in the probe, it is possible to relax the condition  $\partial_\varphi |\Sigma| \neq 0$ , thus allowing us to choose  $k_i = 0$  in Eq. (13), thus effectively measuring the maximally squeezed quadratures at  $\gamma_i = \pm\pi/2$ . Indeed in

such a case, even if the contribution to the Fisher information associated with only the squeezed photons in Eq. (14) is vanishing, it is still possible to reach Heisenberg-scaling precision through the information on the parameter encoded in the displacement of the probe.

An important feature of this protocol, which differentiate it from its single-homodyne counterpart, is that it does *not* require any adaptation of the network to the value of the unknown parameter, namely no auxiliary networks needs to be added at the input nor the output of  $\hat{U}_\varphi$  to reach Heisenberg-scaling precision. The only condition (12) can be thought as a minimum-resolution requirement on the local oscillators phases, which can thus be achieved without adding further auxiliary networks.

However, this does not mean that the form of the network  $U_\varphi$  does not affect the precision of our protocol in the estimation of  $\varphi$ : the terms  $k_{\text{avg}}$  and  $(\partial\gamma)_{\text{avg}}$  appearing in the constant factor in the Fisher information in Eq. (14) depend on the transition probabilities  $P_i$  and on the derivatives of the relative phases  $\partial_\varphi\gamma_i$ . In particular, an exceptionally poorly conceived network, e.g. one for which  $\gamma_i$  is independent on  $\varphi$  for every  $i$  such that  $P_i \neq 0$ , can be associated with a null factor  $(\partial\gamma)_{\text{avg}}$  that sets to zero the Fisher information. In this case, adding a  $\varphi$ -independent auxiliary network  $V$ , either at the input or at the output of  $U_\varphi$ , might modify both  $P_i$  and  $\gamma_i$ , and thus  $(\partial\gamma)_{\text{avg}}$ .

### 4 Conclusions

We have shown that performing homodyne measurements at each output channel of an arbitrary linear network encoding an unknown distributed parameter  $\varphi$  to be estimated allows us to reach Heisenberg-scaling precision for a single-mode squeezed probe with no prior information on  $\varphi$ . The information on  $\varphi$  is encoded both in the displacement and in the squeezing of the probe, leading to two independent contributions which can both provide Heisenberg-scaling sensitivity. We have shown that the determinant of the covariance matrix associated with the measurement outcomes plays an important role in the enhanced sensitivity: in particular, we demonstrated that the conditions to reach Heisenberg scaling in either of the two contributions, which can be met manipulating the phases of the local oscillators, correspond to imposing that the determinant of the covariance matrix is of order  $N^{-1}$  for large  $N$ . Differently from protocols involving only homodyne measurements at a single channel, here there is no need for a refocusing auxiliary stage: the procedure is independent of the network and of the value of the parameter. This allows us to safely entrust the measurement operation to an independent party without sharing any information on the structure of the network, possibly opening up a further path towards secure sensing and cryptographic quantum metrology [39–41]. On the other hand, we showed that, despite not required to achieve the Heisenberg limit, one can still employ an auxiliary stage to further enhance the estimation precision by a constant factor. As a future step, it would be interesting to extend the results hereby presented beyond the Cramér–Rao bound analysis, also in the framework of a parameter distributed arbitrarily in a network.

**Acknowledgements** This work was supported by the Office of Naval Research Global (N62909-18-1-2153). PF is partially supported by Istituto Nazionale di Fisica Nucleare (INFN) through the project QUANTUM, and by the Italian National Group of Mathematical Physics (GNFM-INdAM).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

### Appendix A: Joint Detection Probability

Here we will first obtain the expectation value  $\mu$  associated with the homodyne measurements on the probe after the interaction with the linear network shown in (6), then we derive the expression of the covariance matrix  $\Sigma$  shown in (4), and its determinant  $|\Sigma|$  in (5).

The initial phase-space displacement of the injected probe  $\alpha_0 = \langle \Psi_{\text{in}} | \hat{z} | \Psi_{\text{in}} \rangle$ , where  $\hat{z} = (\hat{x}_1, \dots, \hat{x}_M, \hat{p}_1, \dots, \hat{p}_M)$  and  $|\Psi_{\text{in}}\rangle = \hat{D}_1(d)\hat{S}_1(r)|\text{vac}\rangle$  is a  $2M$ -vector

$$\alpha_0 = \begin{pmatrix} d \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{15}$$

with  $d = \sqrt{N_D}$ . This vector is transformed by the linear network, and at the output reads

$$\alpha = R\alpha_0 = d \begin{pmatrix} \sqrt{P_1} \cos \bar{\gamma}_1 \\ \vdots \\ \sqrt{P_M} \cos \bar{\gamma}_M \\ \sqrt{P_1} \sin \bar{\gamma}_1 \\ \vdots \\ \sqrt{P_M} \sin \bar{\gamma}_M \end{pmatrix} \tag{16}$$

where  $R$  is the  $2M \times 2M$  orthogonal and symplectic matrix associated with the interferometer unitary matrix  $U_\varphi$

$$R = \begin{pmatrix} \text{Re}[U_\varphi] & -\text{Im}[U_\varphi] \\ \text{Im}[U_\varphi] & \text{Re}[U_\varphi] \end{pmatrix}. \tag{17}$$

The local oscillator phases in the homodyne measurements are described by the  $2M \times 2M$  orthogonal matrix

$$O_\theta = \begin{pmatrix} \cos(\Theta) & \sin(\Theta) \\ -\sin(\Theta) & \cos(\Theta) \end{pmatrix}, \tag{18}$$

with  $\Theta = \text{diag}(\theta) = \text{diag}(\theta_1, \dots, \theta_M)$ . This matrix represents a clockwise rotation in phase space for each of the  $M$  channels of the network, of angles  $\theta_i$  for the  $i$ th mode. The mean vector  $\mu$  in Eq. (6) is then given by the first  $M$  elements of  $O_\theta \alpha$ .

The  $2M \times 2M$  symplectic covariance matrix  $\Gamma_0$  of the squeezed state  $\hat{S}_1(r)|\text{vac}\rangle$  reads

$$\Gamma_0 = \frac{1}{2} \begin{pmatrix} e^{2\mathcal{R}} & 0 \\ 0 & e^{-2\mathcal{R}} \end{pmatrix}, \tag{19}$$

where  $\mathcal{R}$  is the  $M \times M$  diagonal matrix  $\mathcal{R} = \text{diag}(r, 0, \dots, 0)$ . Once again, the action of the linear network  $U_\varphi$  is represented by the orthogonal and symplectic matrix  $R$  in Eq. (17), so that the covariance matrix of the probe at the output is

$$\Gamma = R\Gamma_0R^T = \begin{pmatrix} \Sigma_x & \Sigma_{xp} \\ \Sigma_{xp}^T & \Sigma_p \end{pmatrix}, \tag{20}$$

where, by direct calculation,

$$\begin{aligned} \Sigma_x &\equiv \frac{1}{2} [\text{Re}[U_\varphi]e^{2\mathcal{R}}\text{Re}[U_\varphi^\dagger] - \text{Im}[U_\varphi]e^{-2\mathcal{R}}\text{Im}[U_\varphi^\dagger]] \\ &= \frac{1}{2} [\Re[U_\varphi \cosh(2\mathcal{R})U_\varphi^\dagger] + \Re[U_\varphi \sinh(2\mathcal{R})U_\varphi^T]], \end{aligned} \tag{21}$$

$$\begin{aligned} \Sigma_p &\equiv \frac{1}{2} [-\text{Im}[U_\varphi]e^{2\mathcal{R}}\text{Im}[U_\varphi^\dagger] + \text{Re}[U_\varphi]e^{-2\mathcal{R}}\text{Re}[U_\varphi^\dagger]] \\ &= \frac{1}{2} [\Re[U_\varphi \cosh(2\mathcal{R})U_\varphi^\dagger] - \Re[U_\varphi \sinh(2\mathcal{R})U_\varphi^T]], \end{aligned} \tag{22}$$

$$\begin{aligned} \Sigma_{xp} &\equiv \frac{1}{2} [-\text{Re}[U_\varphi]e^{2\mathcal{R}}\text{Im}[U_\varphi^\dagger] - \text{Im}[U_\varphi]e^{-2\mathcal{R}}\text{Re}[U_\varphi^\dagger]] \\ &= \frac{1}{2} [-\Im[U_\varphi \cosh(2\mathcal{R})U_\varphi^\dagger] + \Im[U_\varphi \sinh(2\mathcal{R})U_\varphi^T]]. \end{aligned} \tag{23}$$

The covariance matrix  $\Sigma = \Sigma_x$  at the detection stage is then obtained by extracting the first  $M$  rows and columns from the matrix  $O_\theta \Gamma O_\theta^T$ , and thus its elements as shown in (4) can be easily obtained.

To evaluate the determinant  $|\Sigma|$ , we first notice that  $\Gamma_0 \equiv \mathbb{I}/2 + K_0$ , where  $\mathbb{I}$  is the identity matrix and  $K_0 = \Gamma_0 - \mathbb{I}/2$  is a diagonal matrix of rank 2. Being the rank invariant under orthogonal rotations, the same holds true for  $O_\theta \Gamma O_\theta^T \equiv \mathbb{I}/2 + K$ , with  $\text{rank}(K) = 2$ . By definition of rank, none of the sub-matrices of  $K$  can have rank greater than 2; hence, we can write

$$\Sigma = \mathbb{I}/2 + (\Sigma - \mathbb{I}/2) \equiv \mathbb{I}/2 + A, \tag{24}$$

with  $\text{rank}(A) \leq 2$ , being  $A$  a sub-matrix of  $K$ . We can then apply the result presented in ‘‘Appendix E’’ to  $\Sigma$  and write  $|\Sigma|$  as a sum of determinants of the matrices obtained replacing any number of columns of  $\mathbb{I}/2$ , with the respective columns of  $A$

$$|\Sigma| = \frac{1}{2^M} + \frac{1}{2^{M-1}} \sum_{i=1}^M A_{ii} + \frac{1}{2^{M-2}} \sum_{i=1}^M \sum_{j=i+1}^M A_{ii} A_{jj} - A_{ij}^2, \tag{25}$$

where the first term is the determinant of  $\mathbb{I}/2$ , the terms in the first summation are the contributions that arise substituting the  $i$ th column of  $\mathbb{I}/2$  with the  $i$ th column of  $A$ , and the terms in the last summations from substituting the  $i$ th and  $j$ th columns, and we



also exploited the symmetry of  $A$ . Noticeably, since  $\text{rank } A \leq 2$ , all the contributions involving the replacement of three or more columns of  $A$  are vanishing. By direct calculation, the expression in (5) can then be easily obtained.

**Appendix B: Fisher information**

In this Appendix we will obtain the expression for the Fisher information shown in Eq. (9), and the expression for the cofactor matrix in Eq. (10).

By plugging the probability density function in Eq. (3) into the definition of the Fisher information (8), one can easily obtain

$$\begin{aligned} \mathcal{F}(\varphi) &= \partial_\varphi \boldsymbol{\mu}^T \Sigma^{-1} \partial_\varphi \boldsymbol{\mu} + \frac{1}{2} \text{Tr}[(\Sigma^{-1} \partial_\varphi \Sigma)^2] \\ &= \partial_\varphi \boldsymbol{\mu}^T \Sigma^{-1} \partial_\varphi \boldsymbol{\mu} - \frac{1}{2} \text{Tr}[(\partial_\varphi \Sigma^{-1})(\partial_\varphi \Sigma)], \end{aligned} \tag{26}$$

where we used the matrix identity  $\partial_\varphi \Sigma^{-1} = -\Sigma^{-1}(\partial_\varphi \Sigma)\Sigma^{-1}$ . We now can express the inverse of the covariance matrix in terms of its cofactor matrix  $C$  and its determinant, namely  $\Sigma^{-1} = C/|\Sigma|$ , where the symmetry of the covariance matrix allows us to consider directly the cofactor matrix, and not its transpose. The second term reads

$$-\frac{1}{2} \text{Tr}[(\partial_\varphi \Sigma^{-1})(\partial_\varphi \Sigma)] = \frac{1}{2} \frac{\partial_\varphi |\Sigma|}{|\Sigma|^2} \text{Tr}[C \partial_\varphi \Sigma] - \frac{1}{2|\Sigma|} \text{Tr}[(\partial_\varphi C)(\partial_\varphi \Sigma)]. \tag{27}$$

We recognise in the first term of Eq. (27) Jacobi’s formula for the derivative of the determinant,  $\partial_\varphi |\Sigma| = \text{Tr}[C \partial_\varphi \Sigma]$ , which allows us to obtain the expression shown in Eq. (9)

$$\mathcal{F}(\varphi) = \frac{1}{|\Sigma|} \partial_\varphi \boldsymbol{\mu}^T C \partial_\varphi \boldsymbol{\mu} + \frac{1}{2} \left( \frac{\partial_\varphi |\Sigma|}{|\Sigma|} \right)^2 - \frac{1}{2|\Sigma|} \text{Tr}[(\partial_\varphi \Sigma)(\partial_\varphi C)]. \tag{28}$$

In order to explicit the cofactor matrix  $C$  in terms of the elements of  $\Sigma$ , and thus in terms of squeezing parameter  $r$ , transition probabilities  $P_i$ , and relative phases acquired  $\gamma_i, i = 1, \dots, M$ , as displayed in (10), we first need to make some observations. First, the  $(s, t)$ -cofactor  $C_{st}$ , which is defined as the determinant of the  $L - 1 \times L - 1$  sub-matrix of  $\Sigma$  obtained deleting the  $s$ th row and  $t$ th column, then multiplied by  $(-1)^{s+t}$ , can also be thought as the determinant of the  $L \times L$  matrix  $\Sigma^{[s,t],1}$ , where we denote with  $X^{[s,t],n}$  the matrix obtained from the matrix  $X$  replacing all the elements in the the  $s$ th row and in the  $p$ th column with zeros, except the element  $(s, t)$  which is replaced by  $n$ , namely

$$C_{st} = \begin{vmatrix} \Sigma_{11} & \dots & \Sigma_{1t-1} & 0 & \Sigma_{1t+1} & \dots & \Sigma_{1L} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \Sigma_{s-11} & \dots & \Sigma_{s-1t-1} & 0 & \Sigma_{s-1t+1} & \dots & \Sigma_{s-1L} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \Sigma_{s+11} & \dots & \Sigma_{s+1t-1} & 0 & \Sigma_{s+1t+1} & \dots & \Sigma_{s+1L} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \Sigma_{L1} & \dots & \Sigma_{Lt-1} & 0 & \Sigma_{Lt+1} & \dots & \Sigma_{LL} \end{vmatrix}. \tag{29}$$

Second, as discussed in ‘‘Appendix A’’,  $\Sigma = \mathbb{I}/2 + A$ , where  $A$  is a symmetric matrix with  $\text{rank}(A) = \rho \leq 2$ . Thus, we can write the  $(s, t)$ -cofactor as  $C_{st} = |(\mathbb{I}/2)^{[s,t],0} + A^{[s,t],1}|$ , and evaluate this determinant as a sum of determinants of matrices obtained swapping columns of  $(\mathbb{I}/2)^{[s,t],0}$  and  $A^{[s,t],1}$ , as discussed in detail in ‘‘Appendix E’’. Noticeably, by replacing a row and a column of  $A$  may increase its rank by one, so that  $\text{rank}(A^{[s,t],1}) \leq 3$ . It is convenient now to consider separately the simpler case  $s = t$  first, and then  $s \neq t$ .

We notice that the matrix  $(\mathbb{I}/2)^{[s,s],0}$  has a single zero eigenvalue, and thus each contribution to  $C_{ss}$  is non-vanishing only if the  $s$ th columns of  $(\mathbb{I}/2)^{[s,s],0}$  is replaced. We thus obtain

$$C_{ss} = \frac{1}{2^{M-1}} + \frac{1}{2^{M-2}} \sum_{\substack{i=1 \\ i \neq s}}^M A_{ii} - \frac{1}{2^{M-3}} \sum_{\substack{i=1 \\ i \neq s}}^M \sum_{\substack{j=i+1 \\ j \neq s}}^M A_{ii} A_{jj} - A_{ij}^2, \tag{30}$$

which is the sum of terms obtained substituting the  $s$ th, the  $s$ th and  $i$ th, and the  $s$ th,  $i$ th and  $j$ th columns, respectively. Noticeably, replacing more than 3 columns yields vanishing contributions, since  $\text{rank}(A^{[s,t],1}) \leq 3$ . When  $s \neq t$ ,  $(\mathbb{I}/2)^{[s,t],0}$  has two null eigenvalues; hence, all the non-vanishing contributions must replace the  $s$ th and  $t$ th columns. For example, the only contribution obtained swapping the  $s$ th and  $t$ th columns is of the type

$$\frac{1}{2^{M-2}} \left| \begin{pmatrix} 0 & 1 \\ A_{ts} & 0 \end{pmatrix} \right| = -\frac{1}{2^{M-2}} A_{st}, \tag{31}$$

where we also exploited the symmetry of  $A$ ; the contribution obtained swapping the  $s$ th,  $t$ th and  $i$ th columns, with  $i \neq s, t$  are of the type

$$\frac{1}{2^{M-3}} \left| \begin{pmatrix} 0 & 1 & 0 \\ A_{ts} & 0 & A_{ti} \\ A_{is} & 0 & A_{ii} \end{pmatrix} \right| = \frac{1}{2^{M-3}} (A_{si} A_{ti} - A_{st} A_{ii}), \quad (32)$$

where once again, we exploited the symmetry of  $A$ . Replacing more than 3 columns once again yields no contributions since  $\text{rank}(A^{[s,t,1]}) \leq 3$ . The final expression for  $C_{st}$ ,  $s \neq t$  thus reads

$$C_{st} = -\frac{1}{2^{M-2}} A_{st} + \frac{1}{2^{M-3}} \sum_{\substack{i=1 \\ i \neq s,t}}^M (A_{si} A_{ti} - A_{st} A_{ii}). \quad (33)$$

Replacing in (30) and (33) the definition of  $A = \Sigma - \mathbb{I}/2$ , it is straightforward to obtain Eqs. (10).

### Appendix C: Asymptotics

In this appendix we will study the asymptotic regime of the Fisher information (9),

$$\begin{aligned} \mathcal{F}(\varphi) &= \frac{1}{|\Sigma|} \partial_\varphi \boldsymbol{\mu}^T C \partial_\varphi \boldsymbol{\mu} + \frac{1}{2} \left( \frac{\partial_\varphi |\Sigma|}{|\Sigma|} \right)^2 \\ &\quad - \frac{1}{2|\Sigma|} \text{Tr}[(\partial_\varphi \Sigma)(\partial_\varphi C)], \end{aligned} \quad (34)$$

for large  $N = N_S + N_D = \sinh^2 r + d^2$ . We will show that the only conditions needed to reach Heisenberg scaling are the ones shown in (12), and that in this case the asymptotic expression for the Fisher information is (14).

First, from the explicit expressions of  $\Sigma$  in (4) and  $C$  in (10), we notice that each of their matrix elements is at most of order of  $N_S$ , since  $\sinh r = \sqrt{N_S}$ , the single-photon probabilities  $P_i$  are independent of  $N_S$ , and the cosine and sine functions are limited.

The same holds true for their derivatives: in particular  $\partial_\varphi P_i$  and  $\partial_\varphi \gamma_i = \partial_\varphi \bar{\gamma}_i$  do not depend on  $N_S$ , where  $\gamma_i = \bar{\gamma}_i - \theta_i$  is the phase acquired by the signal through the  $i$ th output port relatively to the local oscillators reference phases. Moreover, similar considerations can be applied for  $\boldsymbol{\mu} = O(\sqrt{N_D})$ , and specifically for its derivative  $\partial_\varphi \boldsymbol{\mu}$ . This implies that, in order to reach a Heisenberg-scaling sensitivity, namely a scaling of order of  $N^2$  in the Fisher information, the determinant  $|\Sigma|$  cannot be of any order higher than  $N^0$ .

We thus first focus our attention on  $|\Sigma|$

$$\begin{aligned} |\Sigma| &= \frac{1}{2^M} + \frac{\sinh(r)}{2^{M-1}} \sum_{i=1}^M P_i (\sinh(r) + \cos(2\gamma_i) \cosh(r)) \\ &\quad - \frac{\sinh^2(r)}{2^{M-2}} \sum_{i=1}^M \sum_{j=i+1}^M P_i P_j \sin^2(\gamma_i - \gamma_j), \end{aligned} \quad (35)$$

shown in Eq. (5). In particular, we will suppose that, for large  $N$ ,  $\gamma_i$  tend to finite values  $\gamma_{0i}$ , namely that  $\gamma_i = \gamma_{0i} + k_i N^{-\alpha}$ , with  $k_i \in \mathbb{R}$  of order 1 and  $\alpha > 0$ , since if  $\gamma_i$  were to grow with  $N$  (i.e.  $\alpha < 0$ ), it would give rise to an oscillating asymptotic behaviour to  $|\Sigma|$ .

By expanding the squeezing parameter  $r$  in powers of  $N_S$  in (35), we obtain

$$|\Sigma| = D_1 N_S + D_2 + D_3 \frac{1}{N_S} + O\left(\frac{1}{N_S^2}\right), \quad (36)$$

where

$$\begin{aligned} D_1 &= \frac{1}{2^{M-1}} \left( 1 + \sum_{i=1}^M P_i \cos(2\gamma_i) \right) \\ &\quad - \frac{1}{2^{M-2}} \left( \sum_{i=1}^M \sum_{j=i+1}^M P_i P_j \sin(\gamma_i - \gamma_j)^2 \right), \end{aligned} \quad (37)$$

$$D_2 = \frac{1}{2^M} \left( 1 + \sum_{i=1}^M P_i \cos(2\gamma_i) \right), \quad (38)$$



$$D_3 = -\frac{1}{2^{M+2}} \sum_{i=1}^M P_i \cos(2\gamma_i). \tag{39}$$

In order to cancel the scaling with  $N_S$ ,  $D_1$  must be equal to, or tend to, zero.

After some trigonometry, and exploiting the passivity of the linear network  $U_\varphi$  which sets  $\sum_i P_i = 1$ , we can rewrite

$$D_1 = \frac{1}{2^{M-2}} \left( \left( \sum_{i=1}^M P_i \cos^2 \gamma_i \right)^2 + \left( \sum_{i=1}^M P_i \sin \gamma_i \cos \gamma_i \right)^2 \right), \tag{40}$$

which tends to zero iff  $\gamma_{0i} = \pi/2 + n\pi$ , with  $n \in \mathbb{Z}$ .

In particular, the scaling of  $|\Sigma|$  for large  $N_S$  will be of order  $N_S^0$  or lower only if  $\gamma_i = \pi/2 + n\pi + k_i N_S^{-\alpha}$ , with  $\alpha \geq 1/2$ . To see that, we notice that, for  $\gamma_{0i} = \pi/2$ ,  $D_1$  and  $D_2$  scale with  $N_S^{-2\alpha}$ , while  $D_3$  scale with  $N_S^0$ . Thus,  $|\Sigma|$  scales with  $N_S^{1-2\alpha}$  for  $\alpha \leq 1$  (and in particular with  $N_S^0$  for  $\alpha = 1/2$ ), and with  $N_S^{-1}$  for  $\alpha > 1$ . Noticeably, also  $D_2$  tends to zero iff  $\gamma_{0i} = \pi/2 + n\pi$ .

We now study the asymptotics of the numerators appearing in the Fisher information, when the condition  $\gamma_i = \pi/2 + k_i N_S^{-\alpha}$ , with  $\alpha \geq 1/2$ , is true. We first obtain the derivative of  $\Sigma$  from Eq. (4), and substitute  $\gamma_i = \pi/2 + k_i/N_S^\alpha$ , with  $\alpha \geq 1/2$ ,

$$\begin{aligned} \partial_\varphi \Sigma_{ij} = & -\frac{1}{2} \partial_\varphi(\sqrt{P_i P_j}) + \sqrt{P_i P_j} \left( \partial_\varphi(\gamma_i + \gamma_j)(k_i + k_j) - \partial_\varphi(\gamma_i - \gamma_j)(k_i - k_j) \right) N_S^{1-\alpha} \\ & + O(N_S^{1-2\alpha}) + O(N_S^{-1}), \end{aligned} \tag{41}$$

and we notice that it scales at most with  $N^{1-\alpha}$  for  $1/2 \leq \alpha < 1$  and at most with  $N^0$  for  $\alpha \geq 1$ . We then analyse the auxiliary term (11)

$$\begin{aligned} S_{sti} = & \sinh^2(r) \sqrt{P_s P_t} P_i \sin(\gamma_s - \gamma_i) \sin(\gamma_t - \gamma_i) \\ = & O(N_S^{1-2\alpha}), \end{aligned} \tag{42}$$

of which we evaluate the derivative when  $\gamma_i = \pi/2 + k_i/N_S^\alpha$ , with  $\alpha \geq 1/2, i = 1, \dots, M$

$$\partial_\varphi S_{sti} = \sqrt{P_s P_t} P_i \left( (\partial_\varphi(\gamma_s - \gamma_i))(k_t - k_i) + (\partial_\varphi(\gamma_t - \gamma_i))(k_s - k_i) \right) N_S^{1-\alpha} + O(N^{1-2\alpha}), \tag{43}$$

which scales at most with  $N_S^{1-\alpha}$  for large  $N_S$ .

Moreover, the covariance matrix  $\Sigma$  in (4) asymptotically reads

$$\Sigma_{ij} = \frac{\delta_{ij} - \sqrt{P_i P_j}}{2} + O(N_S^{1-2\alpha}) + O(N_S^{-1}), \tag{44}$$

and thus, inserting (42) and (44) in the cofactor matrix (10), we obtain

$$C_{st} = \frac{1}{2^{M-1}} \sqrt{P_s P_t} + O(N_S^{1-2\alpha}) + O(N_S^{-1}). \tag{45}$$

Finally, we can write

$$\partial_\varphi |\Sigma| = \frac{1}{2^{M-1}} \sum_{i=1}^M \partial_\varphi \Sigma_{ii} - \frac{1}{2^{M-2}} \sum_{i=1}^M \sum_{j=i+1}^M \partial_\varphi S_{ijj}, \tag{46}$$

$$\partial_\varphi C_{ss} = \frac{1}{2^{M-2}} \sum_{\substack{i=1 \\ i \neq s}}^M \partial_\varphi \Sigma_{ii} - \frac{1}{2^{M-3}} \sum_{\substack{i=1 \\ i \neq s}}^M \sum_{\substack{j=i+1 \\ j \neq s}}^M \partial_\varphi S_{ijj}, \tag{47}$$

$$\partial_\varphi C_{st} = -\frac{1}{2^{M-2}} \partial_\varphi \Sigma_{st} + \frac{1}{2^{M-3}} \sum_{\substack{i=1 \\ i \neq s,t}}^M \partial_\varphi S_{sti}. \tag{48}$$

It is easy to see that  $\partial_\varphi |\Sigma|$  scales at most with  $N_S^{1-\alpha}$ , since  $\sum_i^M \partial_\varphi P_i = 0$ , while the derivatives of the elements of  $C$  scale at most with  $N_S^{1-\alpha}$  for  $1/2 \leq \alpha < 1$ , and at most with  $N_S^0$  for  $\alpha \geq 1$ .

Our last step is to evaluate the asymptotics for  $\partial_\varphi \mu$ , which can be easily evaluated differentiating equation (6)

$$\begin{aligned} \partial_\varphi \mu_i = & \sqrt{2N_D} \left( \frac{\partial_\varphi P_i}{2\sqrt{P_i}} \cos \gamma_i - \sqrt{P_i} \partial_\varphi \gamma_i \sin \gamma_i \right) \\ = & -\sqrt{2N_D} P_i \partial_\varphi \gamma_i + O(\sqrt{N_D} N_S^{-\alpha}). \end{aligned} \tag{49}$$

Now we can finally draw our conclusions and obtain the scaling of the Fisher information (9) by putting together all the asymptotic regimes found. First, we notice that, independently of  $\alpha \geq 1/2$ , the last term in Eq. (9) always scales with  $N_S$ , hence only reaching shot-noise precision. The only terms which allow the Heisenberg scaling are then the first two: in fact, the first term scales with  $N_D N_S^{2\alpha-1}$  for  $1/2 \leq \alpha \leq 1$ , and with  $N_D N_S$  for  $\alpha > 1$ , and thus reaching Heisenberg scaling for  $\alpha \geq 1$  (condition  $\alpha > 1$  includes the case  $\gamma_i = \pi/2$ ), while the second term reaches sub-shot-noise scaling  $N_S^{2-2|\alpha-1|}$  for  $1/2 < \alpha < 3/2$ , with Heisenberg scaling for  $\alpha = 1$ .

This shows that the only condition to reach Heisenberg scaling is that

$$\gamma_i = \pi/2 + O(N_S^{-1}), \tag{50}$$

for  $i = 1, \dots, M$ , as shown in (12), or equivalently that asymptotically  $\gamma_i \simeq \pi/2 + k_i/N_S$  with  $k_i \in \mathbb{R}$  of order 1, as well as that the only relevant terms under this condition are the first two in expression (9), and that only for  $\alpha = 1$  the first term is non-vanishing.

We can now finally prove (14). Substituting condition (50) inside Eqs. (41) and (43), we obtain from (46) the asymptotics

$$\begin{aligned} \partial_\varphi |\Sigma| &\simeq \frac{1}{2^{M-1}} \sum_{i=1}^M 4P_i k_i (\partial_\varphi \gamma_i) \\ &\quad - \frac{1}{2^{M-2}} \sum_{i=1}^M \sum_{j=i+1}^M 2P_i P_j (k_i - k_j) (\partial_\varphi \gamma_i - \partial_\varphi \gamma_j) \\ &\simeq \frac{1}{2^{M-3}} \sum_{i=1}^M P_i k_i \sum_{j=1}^M P_j \partial_\varphi \gamma_j \end{aligned} \tag{51}$$

where we once again exploited the passivity of the interferometer, so that  $\sum_{i=1}^M P_i = 1$ , while we obtain from (36)

$$|\Sigma| \simeq \frac{1}{2^{M-2} N} \left( \left( \sum_{i=1}^M P_i k_i \right)^2 + \frac{1}{16} \right), \tag{52}$$

which is the expression shown in (13). Lastly, from Eq. (49) when conditions (50) hold, we obtain the asymptotics

$$\partial_\varphi \mu_i \simeq -\sqrt{2N_D P_i} \partial_\varphi \gamma_i. \tag{53}$$

Eqs. (51), (52) and (53) yield the asymptotic expression for the Fisher information shown in (14).

### Appendix D: Maximum-likelihood estimator

We will now obtain the implicit equation that defines the maximum-likelihood estimator  $\tilde{\varphi}_{\text{MLE}}$ , which saturates the Heisenberg-scaling Cramér–Rao bound (7), with the Fisher information given in Eq. (14), in the asymptotic regime of large  $\nu$ .

Let us then imagine that, after  $\nu$  iterations, we collect  $\nu$  sets of outcomes  $\mathbf{x}_i$ ,  $i = 1, \dots, \nu$  from the  $M$  homodyne measurements at the output of the linear network  $\hat{U}_\varphi$ . The likelihood that these outcomes are observed for a given value  $\varphi$  of the unknown parameter is given by

$$\mathcal{L}(\varphi|\mathbf{x}_1, \dots, \mathbf{x}_\nu) = \prod_{j=1}^\nu p(\mathbf{x}_j|\varphi), \tag{54}$$

where  $p(\mathbf{x}_j|\varphi)$  is the probability density function given in (3). The maximum-Likelihood estimator  $\tilde{\varphi} \equiv \tilde{\varphi}(\mathbf{x}_1, \dots, \mathbf{x}_\nu)$  is defined as the value of  $\varphi$  which maximises the likelihood  $\mathcal{L}$  that the outcomes  $\mathbf{x}_1, \dots, \mathbf{x}_\nu$  are observed, and it is usually found by maximising the log-likelihood function

$$\begin{aligned} 0 &= \partial_\varphi \log \mathcal{L}(\varphi|\mathbf{x}_1, \dots, \mathbf{x}_\nu) \Big|_{\varphi=\tilde{\varphi}_{\text{MLE}}} \\ &= \partial_\varphi \sum_{j=1}^\nu \log p(\mathbf{x}_j|\varphi) \Big|_{\varphi=\tilde{\varphi}_{\text{MLE}}} \\ &= \left[ -\frac{\nu}{2} \partial_\varphi \log(|\Sigma|) - \frac{1}{2} \partial_\varphi \sum_{j=1}^\nu (\mathbf{x}_j - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x}_j - \boldsymbol{\mu}) \right] \Big|_{\varphi=\tilde{\varphi}_{\text{MLE}}} \\ &= \left[ -\frac{\nu}{2} \text{Tr}[\Sigma^{-1} \partial_\varphi \Sigma] - \frac{1}{2} \partial_\varphi \sum_{j=1}^\nu \text{Tr}[\Sigma^{-1} (\mathbf{x}_j - \boldsymbol{\mu})(\mathbf{x}_j - \boldsymbol{\mu})^T] \right] \Big|_{\varphi=\tilde{\varphi}_{\text{MLE}}} \end{aligned}$$

$$= \frac{1}{2} \text{Tr} \left[ \partial_\varphi \Sigma^{-1} \left( \nu \Sigma - \sum_{j=1}^\nu (\mathbf{x}_j - \boldsymbol{\mu})(\mathbf{x}_j - \boldsymbol{\mu})^T \right) \right]_{\varphi=\tilde{\varphi}_{\text{MLE}}} + \left[ (\partial_\varphi \boldsymbol{\mu})^T \Sigma^{-1} \left( \nu \boldsymbol{\mu} - \sum_{j=1}^\nu \mathbf{x}_j \right) \right]_{\varphi=\tilde{\varphi}_{\text{MLE}}}, \tag{55}$$

where we exploited Jacobi’s formula for the derivative of the determinant of a matrix, the identity  $\text{Tr}(\Sigma^{-1} \partial_\varphi \Sigma) = -\text{Tr}(\partial_\varphi \Sigma^{-1} \Sigma)$ , and the symmetry of  $\Sigma$ . Once inserting the expressions for the covariance matrix  $\Sigma$  from (4) and for the mean vector  $\boldsymbol{\mu}$  from (6) into the previous equation, one is able to obtain with numerical methods the maximum-likelihood estimator as the  $\varphi$  that solves (55).

Equation (55) largely simplifies when the probe is a squeezed vacuum state—i.e.  $\boldsymbol{\mu} = 0$ —or when the maximally squeezed quadratures are measured—i.e.  $\partial_\varphi |\Sigma| = 0$ . In the first case, it becomes

$$0 = \frac{1}{2} \text{Tr} \left[ \partial_\varphi \Sigma^{-1} \left( \Sigma - \frac{1}{\nu} \sum_{j=1}^\nu \mathbf{x}_j \mathbf{x}_j^T \right) \right]_{\varphi=\tilde{\varphi}_{\text{MLE}}}, \tag{56}$$

where it is possible to recognise the usual mean squared error estimator  $\tilde{\Sigma} = \sum_{j=1}^\nu \mathbf{x}_j \mathbf{x}_j^T / \nu$  for the variance matrix  $\Sigma$ , so that the solution of Eq. (56) can be seen as the value of  $\varphi$  that sets equal to zero a weighted mean of the elements of covariance matrix estimator. In the latter case, Eq. (55) becomes

$$0 = \left[ (\partial_\varphi \boldsymbol{\mu})^T \Sigma^{-1} \left( \boldsymbol{\mu} - \frac{1}{\nu} \sum_{j=1}^\nu \mathbf{x}_j \right) \right]_{\varphi=\tilde{\varphi}_{\text{MLE}}}, \tag{57}$$

which can be seen as a weighted mean of the estimators  $\tilde{\boldsymbol{\mu}} = \sum_{j=1}^\nu \mathbf{x}_j / \nu$  of the mean  $\boldsymbol{\mu}$ .

**Appendix E: Useful formulas for the determinant of a sum of two matrices**

Let us consider an  $L \times L$  matrix  $Z$  which can be written as  $Z = D + W$ , where  $D = \text{diag}(d_1, \dots, d_L)$  is a real diagonal matrix, and  $\text{rank}(W) = \rho \leq L$ . In this appendix, we will show a way to write the determinant  $|Z|$  in terms of the elements of  $W$ , which is convenient for our purposes.

We exploit the identity [42]

$$|Z| = |D + W| = \sum_{\alpha=0}^L \Theta_\alpha(D, W), \tag{58}$$

where  $\Theta_\alpha(X, Y)$  is the sum of the determinants of the matrices obtained by replacing any set of  $\alpha$  columns (rows) of  $X$  with the corresponding  $\alpha$  columns (rows) of  $Y$ . Since  $\text{rank}(W) = \rho$ , the rank of any  $\alpha \times \alpha$  sub-matrix of  $W$  is zero if  $\alpha > \rho$ , so we can write

$$|Z| = \sum_{\alpha=0}^\rho \Theta_\alpha(D, W). \tag{59}$$

Let us make explicit the first, easier terms of the summation with the purpose to grasp the gist of this expression. For  $\alpha = 0$ , no columns are replaced from  $D$ , hence  $\Theta_0(D, W) = |D| = \prod_k d_k$ . We notice that, if at least one of the eigenvalues of  $D$  is zero, this term vanishes. For  $\alpha = 1$ ,  $\Theta_1(D, W)$  is the sum of determinants of matrices of the form

$$\begin{pmatrix} d_1 & 0 & \dots & 0 & W_{1i} & 0 & \dots & 0 & 0 \\ 0 & d_2 & \dots & 0 & W_{2i} & 0 & \dots & 0 & 0 \\ & & \ddots & & \vdots & & \ddots & & \\ 0 & 0 & \dots & d_{i-1} & W_{i-1i} & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & W_{ii} & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & W_{i+1i} & d_{i+1} & \dots & 0 & 0 \\ & & \vdots & & \vdots & & \ddots & & \\ 0 & 0 & \dots & 0 & W_{L-1i} & 0 & \dots & d_{L-1} & 0 \\ 0 & 0 & \dots & 0 & W_{Li} & 0 & \dots & 0 & d_L \end{pmatrix} \tag{60}$$

with  $i = 1, \dots, L$ . Due to the structure of this matrices, their determinant is straightforward and reduces to  $W_{ii} \times \prod_{k \neq i} d_k$ . This means that, in general,  $\Theta_1(D, W) = \sum_i W_{ii} \times \prod_{k \neq i} d_k$ . A key observation to make is that if two or more eigenvalues of  $D$  are zero, all these determinants are zero, and thus  $\Theta_1(D, W) = 0$ ; if instead a single eigenvalue is zero, say  $d_j = 0$ , with  $j = 1, \dots, L$ , then only one of these determinants is non-vanishing, i.e. the determinant of the matrix obtained by replacing the  $j$ th column of  $D$ . In this case then we have  $\Theta_1(D, W) = W_{jj} \times \prod_{k \neq j} d_k$ .

Let us now consider lastly the case  $\alpha = 2$ . The matrices whose determinants contribute to  $\Theta_2(D, W)$  are of the form

$$\begin{pmatrix} d_1 & \dots & 0 & W_{1i} & 0 & \dots & 0 & W_{1i'} & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots & & \vdots & & \vdots & \\ 0 & \dots & d_{i-1} & W_{i-1i} & 0 & \dots & 0 & W_{i-1i'} & 0 & \dots & 0 \\ 0 & \dots & 0 & W_{ii} & 0 & \dots & 0 & W_{ii'} & 0 & \dots & 0 \\ 0 & \dots & 0 & W_{i+1i} & d_{i+1} & \dots & 0 & W_{i+1i'} & 0 & \dots & 0 \\ & \vdots & & \vdots & & \ddots & & \vdots & & \vdots & \\ 0 & \dots & 0 & W_{i'-1i} & 0 & \dots & d_{i'-1} & W_{i'-1i'} & 0 & \dots & 0 \\ 0 & \dots & 0 & W_{i'i} & 0 & \dots & 0 & W_{i'i'} & 0 & \dots & 0 \\ 0 & \dots & 0 & W_{i'+1i} & 0 & \dots & 0 & W_{i'+1i'} & d_{i'+1} & \dots & 0 \\ & \vdots & & \vdots & & \vdots & & \vdots & & \ddots & \\ 0 & \dots & 0 & W_{Li} & 0 & \dots & 0 & W_{Li'} & 0 & \dots & d_L \end{pmatrix}, \quad (61)$$

with  $i < i' = 1, \dots, L$ . Once again, the determinants of these type of matrices are easy to be evaluated and read  $|W^{(i,i')}| \prod_{k \neq i, i'} d_k$ , where

$$W^{(i,i')} = \begin{pmatrix} W_{ii} & W_{ii'} \\ W_{i'i} & W_{i'i'} \end{pmatrix}. \quad (62)$$

We notice that  $|W^{(i,i')}| = |W^{(i',i)}|$ . Thus, in general,  $\Theta_2(D, W) = \sum_i \sum_{j>1} |W^{(i,i')}| \prod_{k \neq i, i'} d_k$ . Once again, key observations can be made: if  $D$  has at least three null eigenvalues, then  $\Theta_2(D, W)$  is vanishing. If only two eigenvalues are zero, e.g.  $d_j = d_{j'} = 0$ , then there is only one contribution to  $\Theta_2(D, W)$ , given by the matrix obtained substituting the  $j$ th and  $j'$ th columns of  $D$ , and in this case we have  $\Theta_2(D, W) = |W^{(j,j')}| \prod_{k \neq j, j'} d_k$ . If only one eigenvalue is zero, namely  $d_j = 0$ , then  $\Theta_2(D, W)$  is given by the sum of all the determinants of the matrices where the  $j$ th column has been replaced, namely  $\Theta_2(D, W) = \sum_{i \neq j} |W^{(i,j)}| \prod_{k \neq i, j} d_k$ .

Similarly, it is possible to extend these considerations to every value of  $\alpha$  and finally obtain the compact form

$$|Z| = \sum_{\alpha=0}^{\rho} \sum_{\gamma \in C_{\alpha}^L} |W^{(\gamma)}| \prod_{k \notin \gamma} d_k, \quad (63)$$

where  $C_{\alpha}^L$  is the set of all the combinations of  $\alpha$  items in a set of  $L$ , and  $W^{(\gamma)}$  denotes the  $\alpha \times \alpha$  sub-matrix of  $W$  obtained by selecting the rows and columns with indices  $\gamma_1, \dots, \gamma_{\alpha}$ .

## References

1. V. Giovannetti, S. Lloyd, L. Maccone, Quantum-enhanced measurements: beating the standard quantum limit. *Science* **306**(5700), 1330–1336 (2004)
2. V. Giovannetti, S. Lloyd, L. Maccone, Quantum metrology. *Phys. Rev. Lett.* **96**, 010401 (2006)
3. J.P. Dowling, Quantum optical metrology—the lowdown on high-n00n states. *Contemp. Phys.* **49**(2), 125–143 (2008)
4. V. Giovannetti, S. Lloyd, L. Maccone, Advances in quantum metrology. *Nat. Photonics* **5**, 010401 (2011)
5. J.P. Dowling, K.P. Seshadreesan, Quantum optical technologies for metrology, sensing, and imaging. *J. Lightw. Technol.* **33**, 2359–2370 (2015)
6. A. De Pasquale, P. Facchi, G. Florio, V. Giovannetti, K. Matsuoka, K. Yuasa, Two-mode bosonic quantum metrology with number fluctuations. *Phys. Rev. A* **92**, 042115 (2015)
7. S. Zhou, M. Zhang, J. Preskill, L. Jiang, Achieving the Heisenberg limit in quantum metrology using quantum error correction. *Nat. Commun.* **9**, 1–11 (2018)
8. W. Ge, K. Jacobs, Z. Eldredge, A.V. Gorshkov, M. Foss-Feig, Distributed quantum metrology with linear networks and separable inputs. *Phys. Rev. Lett.* **121**, 043604 (2018)
9. K. Qian, Z. Eldredge, W. Ge, G. Pagano, C. Monroe, J.V. Porto, A.V. Gorshkov, Heisenberg-scaling measurement protocol for analytic functions with quantum sensor networks. *Phys. Rev. A* **100**, 042304 (2019)
10. T.J. Proctor, P.A. Knott, J.A. Dunningham, Multiparameter estimation in networked quantum sensors. *Phys. Rev. Lett.* **120**, 080501 (2018)
11. D.M. Greenberger, M.A. Horne, A. Zeilinger, *Going Beyond Bell's Theorem* (Springer Netherlands, Dordrecht, 1989), pp. 69–72
12. P. Kok, H. Lee, J.P. Dowling, Creation of large-photon-number path entanglement conditioned on photodetection. *Phys. Rev. A* **65**, 052104 (2002)
13. M.O. Scully, M.S. Zubairy, *Quantum Optics* (Cambridge University Press, Cambridge, 1997)
14. W. Schleich, *Quantum Optics in Phase Space* (Wiley, New York, 2011)
15. A. Monras, Optimal phase measurements with pure gaussian states. *Phys. Rev. A* **73**, 033821 (2006)
16. L. Pezzé, A. Smerzi, Mach-Zehnder interferometry at the Heisenberg limit with coherent and squeezed-vacuum light. *Phys. Rev. Lett.* **100**, 073601 (2008)
17. M.D. Lang, C.M. Caves, Optimal quantum-enhanced interferometry using a laser power source. *Phys. Rev. Lett.* **111**, 173601 (2013)
18. J. Aasi, J. Abadie, B.P. Abbott, R. Abbott, T.D. Abbott, M.R. Abernathy, C. Adams, T. Adams, P. Addesso, R.X. Adhikari et al., Enhanced sensitivity of the ligo gravitational wave detector by using squeezed states of light. *Nat. Photonics* **7**, 613–619 (2013)
19. L. Maccone, A. Ricciardi, Squeezing metrology: a unified framework. *Quantum* **4**, 292 (2020)

20. M. Aspachs, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, Phase estimation for thermal gaussian states. *Phys. Rev. A* **79**, 033834 (2009)
21. C. Oh, C. Lee, C. Rockstuhl, H. Jeong, J. Kim, H. Nha, S.-Y. Lee, Optimal Gaussian measurements for phase estimation in single-mode gaussian metrology. *NPJ Quantum Inf.* **5**(1), 10 (2019)
22. D. Gatto, P. Facchi, V. Tamma, Phase space Heisenberg-limited estimation of the average phase shift in a Mach-Zehnder interferometer. *Int. J. Quantum Inf.* **2020**, 1941019 (2020)
23. Y. Xia, W. Li, W. Clark, D. Hart, Q. Zhuang, Z. Zhang, Demonstration of a reconfigurable entangled radio-frequency photonic sensor network. *Phys. Rev. Lett.* **124**, 150502 (2020)
24. D. Triggiani, P. Facchi, V. Tamma, Heisenberg scaling precision in the estimation of functions of parameters in linear optical networks. *Phys. Rev. A* **104**, 062603 (2021)
25. Q. Zhuang, Z. Zhang, J.H. Shapiro, Distributed quantum sensing using continuous-variable multipartite entanglement. *Phys. Rev. A* **97**, 032329 (2018)
26. T. Matsubara, P. Facchi, V. Giovannetti, K. Yuasa, Optimal gaussian metrology for generic multimode interferometric circuit. *New J. Phys.* **21**, 033014 (2019)
27. D. Gatto, P. Facchi, F.A. Narducci, V. Tamma, Distributed quantum metrology with a single squeezed-vacuum source. *Phys. Rev. Res.* **1**, 032024 (2019)
28. X. Guo, C.R. Breum, J. Borregaard, S. Izumi, M.V. Larsen, T. Gehring, M. Christandl, J.S. Neergaard-Nielsen, U.L. Andersen, Distributed quantum sensing in a continuous-variable entangled network. *Nat. Phys.* **16**, 281–284 (2020)
29. G. Gramegna, D. Triggiani, P. Facchi, F.A. Narducci, V. Tamma, Heisenberg scaling precision in multi-mode distributed quantum metrology. *New J. Phys.* **23**, 053002 (2021)
30. G. Gramegna, D. Triggiani, P. Facchi, F.A. Narducci, V. Tamma, Typicality of Heisenberg scaling precision in multimode quantum metrology. *Phys. Rev. Res.* **3**, 013152 (2021)
31. H. Cramér, *Mathematical Methods of Statistics*, vol. 9 (Princeton University Press, Princeton, 1999)
32. A. Holevo, Statistical decision theory for quantum systems. *J. Multivar. Anal.* **3**(4), 337–394 (1973)
33. M. Hayashi, Comparison between the Cramer–Rao and the mini-max approaches in quantum channel estimation. *Commun. Math. Phys.* **304**, 689–709 (2011)
34. M. Hayashi, S. Vinjanampathy, L.C. Kwek, Resolving unattainable Cramer–Rao bounds for quantum sensors. *J. Phys. B: At. Mol. Opt. Phys.* **52**, 015503 (2018)
35. M.R. Grace, C.N. Gagatsos, S. Guha, Entanglement-enhanced estimation of a parameter embedded in multiple phases. *Phys. Rev. Res.* **3**, 033114 (2021)
36. E.L. Lehmann, *Elements of Large-Sample Theory* (Springer, Berlin, 1999)
37. V.K. Rohatgi, A.M.E. Saleh, *An Introduction to Probability and Statistics* (Wiley, New York, 2000)
38. E. Polino, M. Valeri, N. Spagnolo, F. Sciarrino, Photonic quantum metrology. *AVS Quantum Sci.* **2**(2), 024703 (2020)
39. Z. Huang, C. Macchiavello, L. Maccone, Cryptographic quantum metrology. *Phys. Rev. A* **99**, 022314 (2019)
40. P. Yin, Y. Takeuchi, W.-H. Zhang, Z.-Q. Yin, Y. Matsuzaki, X.-X. Peng, X.-Y. Xu, J.-S. Xu, J.-S. Tang, Z.-Q. Zhou, G. Chen, C.-F. Li, G.-C. Guo, Experimental demonstration of secure quantum remote sensing. *Phys. Rev. Appl.* **14**, 014065 (2020)
41. N. Shettell, D. Markham, E. Kashefi, A cryptographic approach to quantum metrology (2021)
42. S. Xu, M. Darouach, J. Schaeffers, Expansion of  $\det(a+b)$  and robustness analysis of uncertain state space systems. *IEEE Trans. Autom. Control* **38**(11), 1671–1675 (1993)