Information geometry of Gaussian channels

Alex Monras[1,2] and Fabrizio Illuminati[1,2,3,*]

[1]*Dipartimento di Matematica e Informatica, Università degli Studi di Salerno, Via Ponte don Melillo, I-84084 Fisciano (SA), Italy*
[2]*CNR-INFM Coherentia, Napoli, Italy; CNISM Unità di Salerno; and INFN Sezione di Napoli,*
Gruppo collegato di Salerno, Baronissi (SA), Italy
[3]*ISI Foundation for Scientific Interchange, Villa Gualino, Viale Settimio Severo 65, I-10133 Torino, Italy*

We define a local Riemannian metric tensor in the manifold of Gaussian channels and the distance that it induces. We adopt an information-geometric approach and define a metric derived from the Bures-Fisher metric for quantum states. The resulting metric inherits several desirable properties from the Bures-Fisher metric and is operationally motivated by distinguishability considerations: It serves as an upper bound to the attainable quantum Fisher information for the channel parameters using Gaussian states, under generic constraints on the physically available resources. Our approach naturally includes the use of entangled Gaussian probe states. We prove that the metric enjoys some desirable properties like stability and covariance. As a by-product, we also obtain some general results in Gaussian channel estimation that are the continuous-variable analogs of previously known results in finite dimensions. We prove that optimal probe states are always pure and bounded in the number of ancillary modes, even in the presence of constraints on the reduced state input in the channel. This has experimental and computational implications. It limits the complexity of optimal experimental setups for channel estimation and reduces the computational requirements for the evaluation of the metric: Indeed, we construct a converging algorithm for its computation. We provide explicit formulas for computing the multiparametric quantum Fisher information for dissipative channels probed with arbitrary Gaussian states and provide the optimal observables for the estimation of the channel parameters (e.g., bath couplings, squeezing, and temperature).

I. INTRODUCTION

The theory of quantum channels provides a broad conceptual and mathematical framework to describe physical transformations on quantum states. Progress in quantum information technology is bringing long-standing questions related to quantum channels to the front line of research. Topics such as dissipation-assisted quantum computation [1], quantum teleportation [2], quantum memories [3], and quantum state engineering [4,5] all have in common that they deal with quantum channels in one way or another. One major question that has recently received a significant amount of attention is the definition of a distance among quantum channels [6–8]. The main motivation for such notion is the identification of a *gold standard* [7] against which all quantum processes could be compared, which would unify and systematize the way in which errors are treated and quantified. In [8], a reasonable, physically motivated set of requirements for such a distance was introduced and some particular cases fulfilling most of them were identified. However, despite recent progress in the field [9], many questions remain open.

A similarly motivated program, as developed between the late sixties and the early nineties, regarded the problem of defining distances among quantum states. The resulting theory is today well encompassed within the framework of information geometry [10–12]. In particular, this approach addresses questions regarding the distinguishability and the estimation of sets of quantum states. These questions led to the notions of quantum fidelity [13], Bures distance, and quantum Fisher information (QFI) [14–19]. As a result, the manifold

of quantum states is endowed with a local Riemannian metric tensor which is physically motivated and serves as the gold standard for comparing quantum states. Additionally, the metric allows one to define the Bures distribution [11,20,21], the quantum analog to the Jeffreys' prior [22], which, among other applications, provides an operationally motivated prior distribution for Bayesian tomography and estimation techniques [23].

There exists in the literature a number of proposals for defining a distance for quantum channels. Most notably, the *Jamiolkowski process distance* [8], based on the Choi-Jamiolkowski isomorphism [24], and the *completely bounded trace norm* [8,25] establish a distance between two channels $\mathcal{S}_1$ and $\mathcal{S}_2$ by considering the distance $d(\rho_1,\rho_2)$ between the respective images of an arbitrary probe state ρ under the action of the channels $\rho_i = (\mathcal{S}_i \otimes \mathcal{I})\rho$. Such distance is then maximized over all possible probe states ρ in order to obtain a fundamental measure of distinguishability between $\mathcal{S}_1$ and $\mathcal{S}_2$. The completely bounded trace norm is arguably the most appropriate choice for both physically motivated reasons and practical considerations: Indeed, it has been recently shown that it can be computed in polynomial time in the dimension of the system upon which the channels act [9]. Despite this and other significant contributions, existing distance measures are inadequate to address a number of relevant situations. In particular, defining a distance on a manifold instead of a metric tensor hinders the task of defining natural prior distributions over the set of channels. Moreover, most of the existing distances are hard, if not impossible, to compute in infinite-dimensional systems such as continuous variables [26,27]. Additionally, in a number of situations, arbitrarily good distinguishability between any two infinite-dimensional channels can be achieved if one allows for a large enough

―――――――
*Corresponding author: illuminati@sa.infn.it

amount of resources [28], thus rendering the direct approach of optimization [9] useless unless some restrictions on the resources (regularization) are enforced.

In the present work we address the problem of defining a metric tensor on the set of n-mode Gaussian channels $\mathfrak{C}$ [32], and we study the ensuing physical consequences. The set of Gaussian channels can be regarded as a manifold once a proper parametrization $X : \mathcal{O} \rightarrow \mathfrak{C}$ is established, where $\mathcal{O} \subseteq \mathbb{R}^d$ is an open subset in a d-dimensional real vector space. The set $\mathfrak{C} = \{\mathcal{S}(X)\}$ can be equipped with a Riemannian metric following the spirit of the statistical distance [33] and the Bures distance. Considering any two infinitesimally close quantum channels $\mathcal{S}_1 \equiv \mathcal{S}(X)$ and $\mathcal{S}_2 \equiv \mathcal{S}(X + dX)$, a metric tensor $\mathfrak{J}$ provides the infinitesimal distance between them as

$$d^2(\mathcal{S}_1, \mathcal{S}_2) = \mathfrak{J}_{\mu\nu}(X) dX^\mu dX^\nu, \qquad (1)$$

where $\mathfrak{J}(X) \geqslant 0$. We use Einstein's summation convention, that is, an index appearing once as a sub-index and once as a super-index is automatically contracted.

As pointed out in [8] some physically motivated requirements should be imposed. Namely, the metric should meet the following criteria.

(a) *Stability*. The metric should be invariant under the addition of ancillary modes, namely, $d^2(\mathcal{S}_1, \mathcal{S}_2) = d^2(\mathcal{S}_1 \otimes \mathcal{I}, \mathcal{S}_2 \otimes \mathcal{I})$, where $\mathcal{I}$ corresponds to the identity channel on an arbitrary number of ancillary modes.

(b) *Measurability*. This amounts to saying that, once the metric is defined, the channel parameters can be determined by experimental means.

(c) *Computability*. The metric $\mathfrak{J}(X)$ should be computable. This requirement is obviously unclear for the existing proposals when addressing infinite-dimensional channels.

(d) *Physical meaning*. The metric should have a clear operational meaning. This is certainly the case if the metric is derived from distinguishability considerations.

Other requirements (symmetry, nondegeneracy, and the triangle inequality) are guaranteed by any distance stemming from a metric tensor. The chaining condition imposed on [8] is not immediate to translate into the metric approach, and we do not address it in the present work. Finally, a purely formal requirement is in order for any well-defined metric tensor.

(e) *Covariance*. The metric tensor has to transform covariantly under a reparametrization. Namely, if we perform the change $X \rightarrow X'(X)$, with $dX'^\mu = \Delta^\mu_\nu(X) dX^\nu$, the metric tensor has to transform as $\mathfrak{J}_{\mu\nu}(X) = \mathfrak{J}'_{\lambda\sigma}(X') \Delta^\lambda_\mu(X) \Delta^\sigma_\nu(X)$ so that the distance $d(\mathcal{S}_1, \mathcal{S}_2)$ is invariant under reparametrization.

As mentioned earlier, some physical constraint on the probe states needs to be imposed in order to guarantee that the metric is well defined. We assume that this constraint is given in the form of a real-valued function $\phi(\rho) \leqslant \phi^\star$, where the specific form of ϕ is not of particular relevance. This function may have a practical motivation, and thus be chosen according to technical considerations, or may be used to regularize the divergencies that appear when resources are unlimited; we will refer to it as a *resource budget*. From now on we tacitly assume that all quantities we define depend implicitly on the given choice of ϕ.

The article is organized as follows. Section II introduces the Bures-Fisher metric, a concept central to our work, together with some of its properties. Section III defines our channel metric and derives some of its main properties, namely, covariance and stability, as well as showing that it can be computed with arbitrary precision, provided that one can compute QFI matrices for all Gaussian probe states. In doing so, we generalize some results known previously for finite-dimensional channels into the constrained Gaussian setup. In Sec. IV we derive explicit formulas for obtaining the QFI matrices for dissipative channels. Section V provides some remarks and stresses the main questions left open, as well as the near-future applications of our results to the estimation of relevant channel properties such as bath couplings, temperature, and squeezing. The mathematical details are reported in five technical appendices.

II. THE BURES-FISHER METRIC

The main motivation for defining a metric in the manifold of Gaussian channels $\mathfrak{C}$ is to obtain a parametrization-independent measure of distinguishability on $\mathfrak{C}$. By expressing distances by means of a metric tensor we also obtain a notion of *density* of channels. Such density thus provides a measure of *how many* different channels can be distinguished in a neighborhood of a point $\mathcal{S}(X)$ with a given amount of resources $\phi^\star$. Such distinguishability-derived density is captured by the Jeffrey's prior [22] in the case of probability distributions and the Bures prior [14,17] for the set of quantum states. These densities play a central role in the theory of Bayesian estimation [23].

Given two channels $\mathcal{S}_1$ and $\mathcal{S}_2$ and a specific probe state ρ_0, the maximal statistical distance [33] attainable by any quantum measurement between the states resulting from the action of the channels, $\rho = \mathcal{S}_1 \rho_0$ and $\sigma = \mathcal{S}_2 \rho_0$, is given by the Bures distance [13,14]

$$d_B^2(\rho, \sigma) = 2\big(1 - \sqrt{F(\rho, \sigma)}\big), \qquad (2)$$

where

$$F(\rho, \sigma) = \mathrm{tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}} \qquad (3)$$

is the quantum fidelity. For infinitesimally close channels, $\mathcal{S}_1 = \mathcal{S}(X)$ and $\mathcal{S}_2 = \mathcal{S}(X + dX)$, the Bures distance can be expressed as

$$d_B^2(\rho, \sigma) = \tfrac{1}{4} \mathcal{J}(X|\rho_0)_{\mu\nu} dX^\mu dX^\nu, \qquad (4)$$

where $\mathcal{J}(X|\rho_0)$ is the QFI matrix of the model $\{\rho(X) = \mathcal{S}(X)\rho_0\}$ [15,16,18,19]. The QFI matrix can be computed from the *symmetric logarithmic derivatives* (SLD) Λ_μ, which are the Hermitian operators that satisfy the equation

$$\partial_\mu \rho(X) = \Lambda_\mu(X) \circ \rho(X), \qquad (5)$$

where we have introduced the symmetric product for operators, $A \circ B = \frac{1}{2}(AB + BA)$. The QFI is then

$$\mathcal{J}_{\mu\nu}(X) = \mathrm{tr}[\rho \Lambda_\mu \circ \Lambda_\nu], \qquad (6)$$

where we have dropped the explicit X dependency in ρ and Λ_μ. The SLDs play an important role in the theory of quantum inference [19].

The QFI has been reviewed a number of times in the literature, from which we emphasize [11,12,34,35]. It enjoys several useful properties:

(i) The SLD has zero expectation,

$$\text{tr}[\rho\Lambda_\mu] = \text{tr}[\rho \circ \Lambda_\mu] = \partial_\mu\text{tr}\rho = 0. \qquad (7)$$

(ii) The QFI is real, symmetric, and positive semidefinite,

$$\theta^\mu\mathcal{J}_{\mu\nu}(X)\theta^\nu = \text{tr}[\rho(\theta^\mu\Lambda_\mu)^2] \geqslant 0, \qquad (8)$$

which follows from $\rho \geqslant 0$ and $(\theta^\mu\Lambda_\mu)^2 \geqslant 0$.

(iii) The SLD and the QFI are covariant quantities. Given a parameter $\theta(X)$ such that $\partial_\theta\rho(X) = \theta^\mu\partial_\mu\rho(X)$, we can define the SLD associated with θ as $\Lambda_\theta = \theta^\mu\Lambda_\mu$. We can equally define the QFI associated with θ as $\mathcal{J}_\theta = \theta^\mu\mathcal{J}_{\mu\nu}\theta^\nu$.

(iv) The QFI is monotonic under completely positive, trace-preserving (CPTP) maps. Writing $\mathcal{J}(\rho)_{\mu\nu} = \text{tr}[\rho\Lambda_\mu^\rho \circ \Lambda_\nu^\rho]$, we have

$$\mathcal{J}(\rho) \geqslant \mathcal{J}(\mathcal{E}\rho), \qquad (9)$$

where $\mathcal{E}$ is any CPTP map. In particular, if $\mathcal{E}$ is a unitary map, it holds that $\mathcal{J}(\rho) = \mathcal{J}(\mathcal{E}\rho)$.

Apart from these properties, the QFI plays a central role in the theory of quantum statistical inference by placing a lower bound on the attainable variance of any unbiased estimator of the parameters X [15,16,19,34–36].

III. A CHANNEL METRIC

In this section we define our proposed channel metric and study its properties. Analogously to Eq. (4), we expect that the distance between neighboring channels can be expressed as in Eq. (1). This is a requirement that is not met by all distance measures, and it is yet unclear that a straightforward extrapolation of the statistical distance applied to channels will fulfill such expectation. Thus, instead of focusing on maximizing the statistical distance, we search for a metric tensor that provides an upper bound to it. Namely,

(1) $\mathfrak{J}(X)$ should provide an upper bound to the Bures distance between states attainable within some resource budget $\phi^\star$.

Following these considerations, and sticking to our definition of ρ and σ, we impose the following condition on $\mathfrak{J}(X)$,

$$d_B^2(\rho,\sigma) \leqslant \tfrac{1}{4}\mathfrak{J}(X)_{\mu\nu}dX^\mu dX^\nu, \qquad (10)$$

for any possible dX and any chosen probe state ρ_0 fulfilling $\phi(\rho_0) \leqslant \phi^\star$. Imposing this for any dX implies that

$$\mathcal{J}(X|\rho_0) \leqslant \mathfrak{J}(X) \quad \forall\rho_0 \text{ s.t. } \phi(\rho_0) \leqslant \phi^\star. \qquad (11)$$

Notice, however, that this condition is not sufficient to uniquely specify $\mathfrak{J}(X)$. In general, one would expect that the distance established between two points is not unnecessarily large. However, a tight bound for $d_B(X, X + dX)$ for any dX may not be expressible in the form of Eq. (1). This is because minimizing the distance between a pair of points $\mathcal{S}(X)$ and $\mathcal{S}(X + dX)$ may not automatically minimize the distance between $\mathcal{S}(X)$ and another $\mathcal{S}(X + dX')$. This can be formalized as follows. Let C be the set of achievable QFI matrices for a given channel and resource budget $\phi^\star$. There may exist several matrices which are tight upper bounds

to C, namely, that they are upper bounds, and no smaller matrix exists which is also an upper bound. This is a direct consequence of the partial-ordered nature of matrices. There is, however, a natural way to precisely specify $\mathfrak{J}(X)$ while respecting all the above criteria. One expects that

(2) The volume element specified by $\mathfrak{J}(X)$ should be minimal.

This translates into imposing that $\sqrt{\det\mathfrak{J}(X)}$ is minimal.

Thus, our metric corresponds to an upper bound on the attainable QFI for any set of Gaussian quantum channels $\mathfrak{C}$, tested with Gaussian probe states fulfilling some regularization condition $\phi(\rho_0) \leqslant \phi^\star$ and providing the smallest possible volume element in the manifold of channels. Restricting to the Gaussian domain yields a manageable parametrization of the probe states and allows one to obtain explicit formulas for the QFI. In addition, the geometry induced by our metric has an immediate practical interpretation. It is a bound to the distinguishability attainable by Gaussian states. Moreover, recent analysis [37] suggests that entangled Gaussian states are as sensitive as single-mode de-Gaussified states [38] in quantum statistical inference. This claim has yet to be proven.

As a first approach to the problem of defining a channel metric it is relevant to discuss the Jamiolkowski distance introduced in [8]. It is well known that the Jamiolkowski isomorphism [24] can be used to encode a d-dimensional quantum channel $\mathcal{S}$ into a d^2-dimensional quantum state $\sigma_\mathcal{S} = (\mathcal{S} \otimes \mathbb{1})\Phi$, where Φ is a maximally entangled state. One may consider that distinguishability between $\sigma_{\mathcal{S}_1}$ and $\sigma_{\mathcal{S}_2}$ be related to the distinguishability between $\mathcal{S}_1$ and $\mathcal{S}_2$. The state Φ, however, is not likely to comply with any physically motivated resource restriction ϕ. Instead, one must consider the attainable distinguishability within the limits imposed by ϕ.

We proceed, in the next subsections, to formally define the metric tensor $\mathfrak{J}(X)$ based on the previously announced criteria (1) and (2) and prove the covariance, stability, and computability properties.

A. Defining $\mathfrak{J}(X)$

Given a parametrized family of n-mode channels $\mathfrak{C} = \{\mathcal{S}(X)\}$, let us define the *quantum model* $\mathcal{M}[\mathfrak{C}|\rho] = \{\rho(X) = \mathcal{S}(X)\rho_0\}$, that is, the parametrized set of all possible quantum states resulting from an initial probe ρ_0, under the action of the set of channels $\mathfrak{C}$. The associated QFI is

$$\mathcal{J}_{\mu\nu}(\mathcal{S}(X)|\rho_0) = \text{tr}[\rho(X)\Lambda_\mu \circ \Lambda_\nu], \qquad (12)$$

where Λ_μ are the associated SLDs. This corresponds to the situation depicted in Fig. 1(a).

Consider next the *m-completed channel* as $\mathcal{S}^{\star m}(X) = \mathcal{S}(X) \otimes \mathcal{I}_m$, where $\mathcal{I}_m$ represents the identity channel on m ancillary modes. We thus obtain the *m-completed set* $\mathfrak{C}^{\star m} = \{\mathcal{S}^{\star m}(X)\}$. Given an $(n + m)$-mode probe state ρ we can define a new quantum model, $\mathcal{M}[\mathfrak{C}^{\star m}|\rho_0] = \{\rho^{\star m}(X) = \mathcal{S}^{\star m}(X)\rho_0\}$. The QFI for the new model is

$$\mathcal{J}_{\mu\nu}(\mathcal{S}^{\star m}(X)|\rho_0) = \text{tr}[\rho^{\star m}(X)\Lambda_\mu^{\star m} \circ \Lambda_\nu^{\star m}], \qquad (13)$$

where $\rho^{\star m}(X) = \mathcal{S}^{\star m}(X)\rho_0$ is the probe state under the action of the m-completed channel, and the SLD's $\Lambda_\mu^{\star m}$ obviously
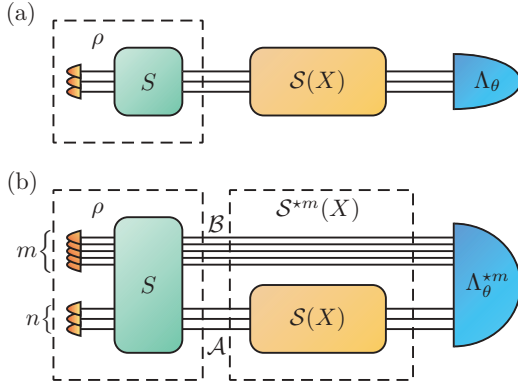
FIG. 1. (Color online) Measurement schemes (a) without ancillas and (b) with ancillas, combined with the most general possible measurement. Scheme (b) includes (a) and, hence, must be equally or more efficient than (a).

correspond to the model $\mathcal{M}[\mathfrak{C}^{\star m}|\rho_0]$. The extension of $\mathfrak{C}$ to $\mathfrak{C}^{\star m}$ is depicted in Fig. 1(b).

Now we turn to the constraint ϕ. Let us denote the n modes upon which the channels $\mathcal{S}(X)$ act as a_k, $k \in \{1,\dots,n\}$, and are collectively referred to as $\mathcal{A}$, while the ancillary modes are denoted by b_k, $k \in \{1,\dots,m\}$, collectively denoted as $\mathcal{B}$. We may find it useful to consider yet another set of modes, collectively denoted as $\mathcal{C}$. We define the *acceptable* constraints as those that only involve the reduced density operator in modes $\mathcal{A}$, $\rho_{\mathcal{A}} = \mathrm{tr}_{\bar{\mathcal{A}}}\rho$, namely,

$$\phi(\rho) \equiv \phi(\rho_{\mathcal{A}}). \tag{14}$$

This choice is physically motivated by the fact that relevant resources should only involve degrees of freedom accessible to $\mathcal{A}$. A natural choice for the function ϕ could be the average photon number in $\mathcal{A}$, $\phi(\rho) = \mathrm{tr}[\rho \sum_k a_k^\dagger a_k]$, which has been used previously as a comparison reference for several estimation problems. Nevertheless, our results are general and not restricted to this particular choice.

Finally let $C(\mathcal{S}(X)|\phi^\star) = \{\mathcal{J}(\mathcal{S}(X)|\rho)|\phi(\rho) \leqslant \phi^\star\}$ be the set of all QFI achievable by any Gaussian state fulfilling the constraint $\phi(\rho) \leqslant \phi^\star$, and let $C_{\mathrm{pure}}(\mathcal{S}(X)|\phi^\star) = \{\mathcal{J}(\mathcal{S}(X)|\rho)|\phi(\rho) \leqslant \phi^\star, \rho^2 = \rho\}$ be the corresponding set when restricting to pure states. Let C be any set of positive semidefinite bounded matrices, and $\mathcal{J} \in C$. Define the function $M(\mathcal{J}) = \{j \geqslant \mathcal{J}\}$ to be the set of all upper bounds to $\mathcal{J}$ and $M(C) = \cap_{\mathcal{J} \in C} M(\mathcal{J})$ to be the set of all upper bounds common to all matrices in C.

We are now in the position to introduce the metric for the set $\mathfrak{C}$. Let $\mathcal{S}$ be a channel in $\mathfrak{C}$. The metric at point $\mathcal{S}$ is

$$\mathfrak{J}(\mathcal{S}) = \underset{\{j \in M(C(\mathcal{S}^{\star n}|\phi^\star))\}}{\arg\inf} \det j, \tag{15}$$

that is, the matrix with the smallest determinant that is greater or equal to all possible QFI achievable by any $2n$-mode Gaussian probe state fulfilling the constraint $\phi(\rho) \leqslant \phi^\star$. This is a positive semidefinite matrix and certainly qualifies as a metric. It is measurable in the sense that it is a function of the parameters X which are themselves measurable [see Sec. IV A]. Notice that in very pathological cases, there may be more than one solution to the minimization problem. These are

highly symmetric and unlikely situations which we do not deal with. We dedicate the following subsections to discussing the most relevant properties of $\mathfrak{J}$ and providing an approximation method to compute it.

The idea behind this definition is to provide an upper bound to the achievable Fisher information for any parameter of interest, under the constraint $\phi(\rho) \leqslant \phi^\star$. While any matrix greater than those in $C(\mathcal{S}^{\star n}|\phi^\star)$ would certainly qualify as an upper bound, the partial-ordered nature of matrices prevents one from having, in the general case, a well-defined supremum. Instead, the minimization of the determinant in Eq. (15) aims at reducing the *volume element* to a minimum while still providing an upper bound to the attainable Fisher information.

B. Covariance

Notice that under a reparametrization $X \to X'(X)$ we have

$$\Lambda_\mu \to \Lambda_{\mu'} = \Lambda_\mu \Delta^\mu_{\mu'}, \tag{16}$$

where $\Delta^\mu_{\mu'} \equiv \partial X^\mu / \partial X'^{\mu'}$, where primed indices refer to the new coordinates and unprimed indices correspond to the old coordinates, as is customary in the notation of general relativity. Equation (16) in turn implies that the QFI is a covariant quantity,

$$\mathcal{J}_{\mu\nu} \to \mathcal{J}'_{\mu'\nu'} = \mathcal{J}_{\mu\nu} \Delta^\mu_{\mu'} \Delta^\nu_{\nu'}. \tag{17}$$

Let us denote this transformation law in shorthand notation $\mathcal{J} \to \Delta^T \mathcal{J} \Delta$. All elements in $C(\mathcal{S}^{\star n}(X)|\phi^\star)$ transform in the same way under a given reparametrization. Accordingly, the set $M(C(\mathcal{S}^{\star n}(X)|\phi^\star))$ also transforms covariantly. Namely, for any matrix $j \in M(C(\mathcal{S}^{\star n}(X)|\phi^\star))$ we have

$$\Delta^T j \Delta \in M(C(\mathcal{S}^{\star n}(X')|\phi^\star)), \tag{18}$$

and vice versa.

Finally the determinants of all elements in $M(C(\mathcal{S}(X')|\phi^\star))$ are just those of $M(C(\mathcal{S}(X)|\phi^\star))$ multiplied by a factor $\det \Delta^T \Delta$. Therefore, assuming that under parametrization X the minimum was achieved for $\mathfrak{J}$, the new minimum under parametrization X' will be achieved by $\mathfrak{J}' = \Delta^T \mathfrak{J} \Delta$ [$\mathfrak{J}'_{\mu'\nu'} = \mathfrak{J}_{\mu\nu} \Delta^\mu_{\mu'} \Delta^\nu_{\nu'}$], as required.

C. Stability

The stability requirement has been discussed previously in the literature [6,8]. We dedicate this section to proving that $\mathfrak{J}(\mathcal{S})$ is invariant under the addition of ancillas, namely,

$$\mathfrak{J}(\mathcal{S} \otimes \mathcal{I}) = \mathfrak{J}(\mathcal{S}), \tag{19}$$

where $\mathcal{I}$ is the identity channel on an arbitrary number n' of modes. We remark that, while adding ancillary modes, we still require the constraints ϕ to act on the original set of modes. Observe that

$$\mathfrak{J}(\mathcal{S} \otimes \mathcal{I}) = \underset{\{j \in M(C(\mathcal{S}^{\star(n+2n')}|\phi^\star))\}}{\arg\inf} \det j. \tag{20}$$

Observe that if $X \subseteq Y$ then $M(Y) \subseteq M(X)$. Given two sets of positive semidefinite matrices X and Y, we say $X \succeq Y$ if and only if for any $\forall y \in Y \exists x \in X$ such that $x \geqslant y$. This binary relation is reflexive ($X \succeq X$) and transitive ($X \succeq Y$
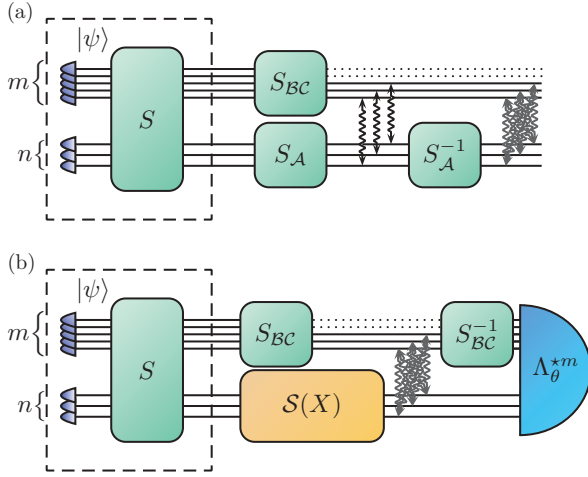
FIG. 2. (Color online) (a) The Schmidt decomposition theorem ensures that, in an $(n + m)$-mode state $(m > n)$, one can locally bring $m - n$ modes to the vacuum. Therefore, for any pair $(|\psi\rangle, \Lambda_\theta)$ of $n + m$ modes there exists a pair $(S_{\mathcal{BC}}|\psi\rangle, S_{\mathcal{BC}}^\dagger \Lambda_\theta S_{\mathcal{BC}})$ of $2n$-modes which performs equally well. The remaining modes can be traced out of $S_{\mathcal{BC}}|\psi\rangle$ leading to a generalized measurement. The optimal generalized measurement on the $2n$-mode state is given by the corresponding SLD, $\Lambda_\theta^{\star n}$.

and $Y \succeq Z \Rightarrow X \succeq Z$). Also $X \succeq Y \Rightarrow M(X) \subseteq M(Y)$. We will say that two sets are *equivalent* ($X \simeq Y$) if and only if $M(X) = M(Y)$.

Let n be the number of modes upon which $\mathcal{S}$ acts. The proof consists of two main steps.

1. We first show that the set of QFI matrices attainable using *pure* probe states with $m > n$ ancillary modes is equivalent to those attainable with *pure* probes using only n ancillary modes.

2. We then show that the set of QFI matrices attainable with $(n + m)$-mode states $(m > n)$ is equivalent to those attainable with *pure* $(n + m)$-mode states.

Proof of 1. Consider an $(n + m)$-mode setup such as the one in Fig. 1(b) $(m > n)$. The Schmidt decomposition-like theorem [39–42] states that, given an $(n + m)$-mode pure Gaussian state $|\psi\rangle_{\mathcal{ABC}}$ separated into parties $\mathcal{A}$ (n modes through the channel), $\mathcal{B}$ (n modes), and $\mathcal{C}$ ($m - n$ modes), one can always reduce the system to n pairwise squeezed states between $\mathcal{A}$ and $\mathcal{B}$, and $m - n$ vacuum states in $\mathcal{C}$ by means of local symplectic transformations on $\mathcal{A}$ ($S_{\mathcal{A}}$) and joint symplectic transformations on $\mathcal{B}$ and $\mathcal{C}$ ($S_{\mathcal{BC}}$) [Fig. 2(a)]. Therefore, for any pure Gaussian state $|\psi\rangle_{\mathcal{ABC}}$ there is another $2n$-mode state $|\psi_{2n}\rangle_{\mathcal{AB}}$ such that

$$S_{\mathcal{BC}}|\psi\rangle_{\mathcal{ABC}} = |\psi_{2n}\rangle_{\mathcal{AB}}|0\rangle_{\mathcal{C}}^{\otimes(n-m)}. \quad (21)$$

Defining $\mathcal{E}$ as $\mathcal{E}\rho = S_{\mathcal{BC}}\rho S_{\mathcal{BC}}^\dagger$ and noting that $\mathcal{S}^{\star(n+m)}\mathcal{E}\rho = \mathcal{E}\mathcal{S}^{\star(n+m)}\rho$ we have, following from the unitary invariance of the QFI,

$$\mathcal{J}(\mathcal{S}^{\star m}||\psi\rangle_{\mathcal{ABC}}) = \mathcal{J}(\mathcal{S}^{\star m}|\mathcal{E}|\psi\rangle_{\mathcal{ABC}}). \quad (22)$$

Moreover, using $\mathcal{S}^{\star m}(\rho \otimes |0\rangle\langle 0|^{\otimes(m-n)}) = (\mathcal{S}^{\star n}\rho) \otimes |0\rangle\langle 0|^{\otimes(m-n)}$ together with Eq. (21) and the additivity of

the QFI [obviously $\mathcal{J}(\mathcal{I}||0\rangle^{\otimes(m-n)}) = 0$] we have

$$\mathcal{J}(\mathcal{S}^{\star m}|\mathcal{E}|\psi\rangle_{\mathcal{ABC}}) = \mathcal{J}(\mathcal{S}^{\star m}||\psi_{2n}\rangle \otimes |0\rangle^{\otimes(n-m)})$$
$$= \mathcal{J}(\mathcal{S}^{\star n}||\psi_{2n}\rangle_{\mathcal{AB}}), \quad (23)$$

while $\phi(|\psi\rangle) = \phi(|\psi_{2n}\rangle)$. This shows that $C_{\text{pure}}(\mathcal{S}^{\star m}|\phi^\star) \subseteq C_{\text{pure}}(\mathcal{S}^{\star n}|\phi^\star)$, which combined with the trivial inverse inclusion yields

$$C_{\text{pure}}(\mathcal{S}^{\star m}|\phi^\star) = C_{\text{pure}}(\mathcal{S}^{\star n}|\phi^\star), \quad \forall m > n, \quad (24)$$

and therefore

$$M(C_{\text{pure}}(\mathcal{S}^{\star n}|\phi^\star)) = M(C_{\text{pure}}(\mathcal{S}^{\star m}|\phi^\star)), \quad \forall m > n. \quad (25)$$

This makes it clear that all $C_{\text{pure}}(\mathcal{S}^{\star m}|\phi^\star)$ with $m \geqslant n$ are equivalent. This result has a similar flavor to the well-known fact that maximally entangled states are optimal for estimation of d-dimensional unitary operations and several other finite-dimensional channels [43], which means that a d-dimensional ancilla is sufficient. In our case, maximal entanglement will in general be forbidden by the resource constraint $\phi^\star$, but nevertheless the ancilla need not be larger than the system itself.

Proof of 2. Notice that for any $(n + m)$-mode QFI $\mathcal{J}(\mathcal{S}^{\star m}|\rho)$ $(m \geqslant n)$ there exists an $(n + m + m')$-mode pure state $|\psi\rangle$ such that $\mathcal{J}(\mathcal{S}^{\star(m+m')}||\psi\rangle) \geqslant \mathcal{J}(\mathcal{S}^{\star m}|\rho)$. To see this, one simply needs to construct a pure $(n + m + m')$-mode Gaussian state $|\psi\rangle$ such that $\text{tr}_{\mathcal{C}}|\psi\rangle\langle\psi| = \rho$, with $\mathcal{C}$ denoting the *purifying* extra (m') modes [Fig. 3(b)]. Since $\mathcal{S}^{\star m}\rho = \text{tr}_{\mathcal{C}}\mathcal{S}^{\star(m+m')}|\psi\rangle\langle\psi|$ the monotonicity of the QFI under CPTP maps [44] guarantees that

$$\mathcal{J}(\mathcal{S}^{\star(m+m')}||\psi\rangle) \geqslant \mathcal{J}(\mathcal{S}^{\star m}|\rho). \quad (26)$$

Notice that the constraint $\phi(\rho)$ is not affected by this construction, as long as it depends only on the reduced state $\rho_{\mathcal{A}} = \text{tr}_{\bar{\mathcal{A}}}\rho$, since $\phi(\rho) = \phi(|\psi\rangle)$. This shows that the optimal probe states can always be taken to be pure, provided that one enlarges sufficiently the set of ancillary modes. Therefore $C_{\text{pure}}(\mathcal{S}^{\star(m+m')}|\phi^\star) \succeq C(\mathcal{S}^{\star m}|\phi^\star)$ yielding $M(C_{\text{pure}}(\mathcal{S}^{\star(m+m')}|\phi^\star)) \subseteq M(C(\mathcal{S}^{\star m}|\phi^\star))$.

Using Eq. (25) we can, furthermore, say that $M(C_{\text{pure}}(\mathcal{S}^{\star m}|\phi^\star)) \subseteq M(C(\mathcal{S}^{\star m}|\phi^\star))$. On the other hand, it is trivial that $C_{\text{pure}}(\mathcal{S}^{\star m}|\phi^\star) \subseteq C(\mathcal{S}^{\star m}|\phi^\star)$, yielding $M(C(\mathcal{S}^{\star m}|\phi^\star)) \subseteq M(C_{\text{pure}}(\mathcal{S}^{\star m}|\phi^\star))$. Consequently,

$$M(C_{\text{pure}}(\mathcal{S}^{\star m}|\phi^\star)) = M(C(\mathcal{S}^{\star m}|\phi^\star)). \quad (27)$$

This concludes the proof. ∎

This result is also a Gaussian version of previously known results for the estimation of finite-dimensional channels [45]. The added value of this result is that the restriction to pure states can be made while preserving the Gaussian character of the probe states as well as respecting the resource budget $\phi^\star$.

All these results imply that $M(C(\mathcal{S}^{\star(n+2n')}|\phi^\star)) = M(C_{\text{pure}}(\mathcal{S}^{\star(n+2n')}|\phi^\star)) = M(C_{\text{pure}}(\mathcal{S}^{\star n}|\phi^\star))$, so that Eq. (20) reduces to

$$\mathfrak{J}(\mathcal{S} \otimes \mathcal{I}) = \arg\inf_{\{j \in M(C_{\text{pure}}(\mathcal{S}^{\star n}|\phi^\star))\}} \det j = \mathfrak{J}(\mathcal{S}). \quad (28)$$

Summarizing, we have shown that any Gaussian state of $2n$ modes (or more) can be reduced to a $2n$-mode Gaussian *pure*
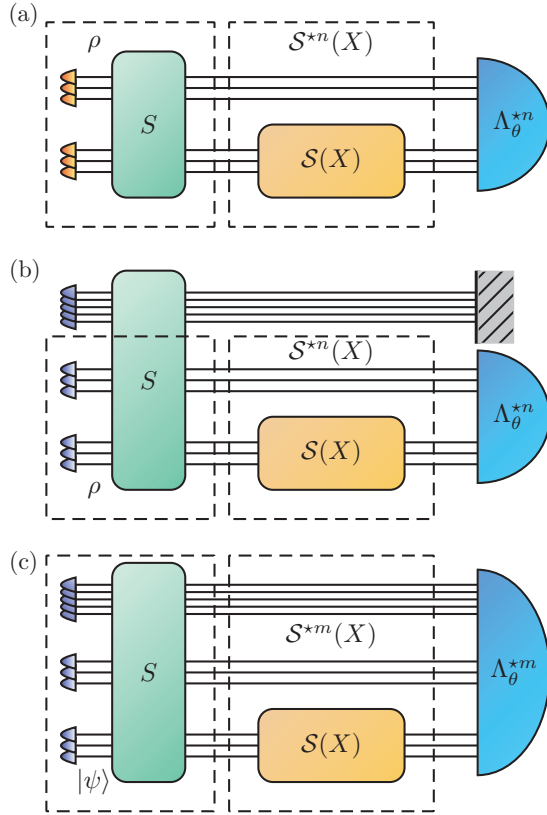
FIG. 3. (Color online) Measurement schemes (a) with a thermal state in a $2n$-mode setup. (b) Equivalent setup with a pure state containing ancillary modes that are not measured. (c) The optimal measurement including the ancillary modes. Obviously, the setups in (a) and (b) are equivalent, while the setup in (c) has more freedom and includes (b). Consequently $J_a = J_b \leqslant J_c$.

state which performs equally well or better than the original. The stability property thus follows.

D. Computing $\mathfrak{J}(X)$

We now turn to the problem of computing the metric $\mathfrak{J}(X)$. Take $\mathcal{S}$ as the channel for which the metric needs to be computed, under the constraint $\phi \leqslant \phi^\star$. Let $C^\star = C(\mathcal{S}^{\star n}|\phi^\star)$. The problem can be written as

$$\text{minimize} \quad \det j \tag{29}$$

$$\text{subject to} \quad j \geqslant \mathcal{J} \quad \forall \mathcal{J} \in C^\star. \tag{30}$$

This can be cast as a semi-infinite programming problem [46], where there are finitely many variables but infinitely many inequalities to satisfy. In the following we provide an approximation method by discretizing the problem into a convergent sequence of convex programming problems. Despite the apparently untractable nature of the problem, consistent approximations can be computed from the following prescription: Generate a finite subset of constraints by sampling a subset $C_n \subset C^\star$ of n random QFI matrices from $C^\star$. Let $\mathfrak{J}_n$ be the solution to the discretized problem with matrices in C_n. Then $\mathfrak{J}_n$ can be computed following standard convex optimization methods. In fact, the problem of computing $\mathfrak{J}_n$ can be recast as the problem of finding the Löwner-Jones
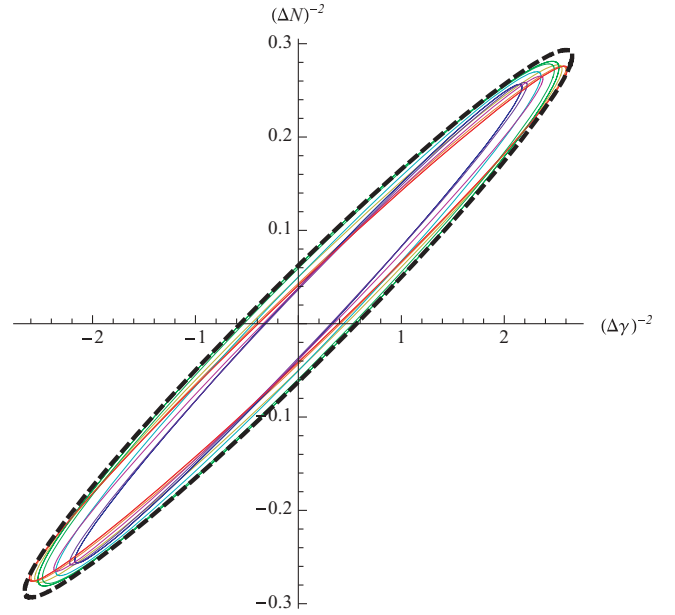

FIG. 4. (Color online) Illustration of the problem for the model $X = (\gamma, N, 0, 0)$ at the point $(\gamma, N) = (0.1, 1)$, tested with Gaussian states fulfilling $\phi(\rho) = \text{tr}[a^\dagger a \, \rho] = 0.2$. The thick-dashed (black) outermost ellipse that bounds all the inner ones corresponds to the metric $\mathfrak{J}$. The inner ellipses correspond to the points $X \mathcal{J}^{-1} X = 1$, $\forall \mathcal{J} \in C^\star$. The various inner ellipses correspond to the different distribution of the available energy among single mode squeezing (red), two-mode squeezing (green) and displacement (blue). Ellipsoids corresponding to the QFI's for different probe states are drawn with matching Red-Green-Blue color scheme.

ellipsoid of a union of ellipsoids [47] (see Fig. 4), for which efficient methods exist.

We show in Appendix A that this approximation method converges to the true value $\mathfrak{J}$. More precisely, we show that for any $\epsilon > 0$ there exists a sufficiently large n such that $\Pr(\|\mathfrak{J}_n - \mathfrak{J}\| > \epsilon) < k \exp(-n)$, where $\|\cdot\|$ is the operator norm.

This iterative method assumes that one can generate any QFI matrix in $C^\star$. This is indeed a nontrivial task. In principle, QFI matrices can always be computed numerically for finite-dimensional quantum systems, but in the case of infinite dimensions, a general method does not exist. In the following section we concentrate on a particular class of channels and provide closed analytic formulas for the QFI matrix.

IV. EXPLICIT FORMULAS FOR DISSIPATIVE CHANNELS

We dedicate this section to the issue of measurability and computability stated in the Introduction. In principle the SLD is defined for any possible model $\mathcal{S}(X)\rho$, and thus the parameters X are always measurable. Similarly, the QFI is always defined and can be obtained from the second derivatives of the quantum fidelity $F(\rho(X), \rho(X'))$ with respect to X' at $X' = X$ [17]. Therefore, the matrix $\mathcal{J}(X)$ is always defined. However, computing the QFI in the general case can be extremely challenging. Generic formulas can be found in the literature, but obtaining explicit expressions often requires knowledge of the eigenbasis of $\mathcal{S}(X)\rho$. On the other hand, brute force evaluation of the SLD amounts to solving a Sylvester equation

[48]. Several algorithms exist for performing such a task numerically. Unfortunately, this is of little use in the case of infinite dimensions.

In this section we develop explicit formulas for the computation of the SLD and the QFI of dissipative channels probed by arbitrary Gaussian probe states. Dissipative Gaussian channels are of the form

$$\mathcal{S}(X) = \exp\mathcal{G}(X),\qquad(31)$$

where

$$\mathcal{G}(X) = \sum_{k=1}^{n}\frac{\gamma_k}{2}(N_k L[a_k^\dagger] + (N_k+1)L[a_k]$$
$$+ M_k^* D[a_k] + M_k D[a_k^\dagger]),\qquad(32)$$

with $L[o]\rho = 2o\rho o^\dagger - o^\dagger o\rho - \rho o^\dagger o$ and $D[o] = 2o\rho o - o^2\rho - \rho o^2$, and $a_k^\dagger$ and a_k are creation and annihilation operators in mode k, satisfying the commutation relations $[a_i, a_j^\dagger] = \delta_{ij}$. We have used the shorthand notation $X \equiv \bigoplus_{k=1}^{n} x_k = (\gamma_1, N_1, \mathrm{Re}M_1, \mathrm{Im}M_1, \gamma_2, \dots)$ while x_k stands for $x_k = (0, \dots, 0, \gamma_k, N_k, \mathrm{Re}M_k, \mathrm{Im}M_k, 0, \dots, 0)$, with nonzero parameter values only in mode k. We have chosen to work with $(\mathrm{Re}M, \mathrm{Im}M)$ rather than (M, M^*) to ensure hermiticity of the SLD. These are the most general Gaussian dissipative channels, where the $\{\gamma, \dots, \gamma_k\}$ parameters can be regarded as coupling strengths, N corresponds to the mean photon number of the reservoir modes, and the M's are their squeezing parameters. Notice that this parametrization can be taken beyond the natural interpretation of the parameters, for example, by setting $\gamma = \mathcal{C} - \mathcal{A}$ and $N = \mathcal{A}/(\mathcal{C} - \mathcal{A})$ one can account for amplification processes where $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are the gain, saturation, and decay rate, respectively, in the linear regime ($\mathcal{B} = 0$) [49]. Moreover, the squeezing parameters M can accommodate phase sensitivity in the amplification process.

The dynamics of an arbitrary state ρ undergoing the most general Gaussian dissipative evolution, in the interaction picture and within the Markovian approximation, can be described by the master equation [50]

$$\frac{d\rho}{dt} = \mathcal{G}(X)\rho,\qquad(33)$$

and evolves, after a time t, from state ρ_0 to $\rho = \exp[t\mathcal{G}(X)]\rho_0$. For all practical purposes the coupling constants always appear as $\gamma_k t$, thus the time t can be absorbed in the coupling constants γ_k, allowing one to take $t = 1$ without loss of generality.

There are several ways to compute the QFI. As stated previously, one consists of computing the Hessian of the fidelity. This, however, does not provide information about the optimal observables, namely, the SLD's Λ_μ's. Hence, we take here the longer route by first computing the SLDs. This settles the issue about measurability, as we explicitly provide the optimal observables to estimate the value of X^μ. Then the QFI is immediately given by their covariance matrix. In order to preserve the clarity, we introduce notation by reviewing some well-known facts about Gaussian states.

At variance with the previous sections, where we have made explicit the distinction between the channel modes $\mathcal{A}$ and the ancillary modes $\mathcal{B}$ and $\mathcal{C}$, we now treat all modes indistinctly. Considering channels with ancillas only amounts to considering parameter spaces of the form

$$X = \bigoplus_{k=1}^{n-m} x_k \oplus \vec{0},\qquad(34)$$

where $\vec{0} = \bigoplus_{k=1}^{m}(0,0,0,0)$. We thus have n bosonic modes with annihilation operators a_k fulfilling the canonical commutation relations $[a_i, a_j^\dagger] = \delta_{ij}$, all other commutators being zero. We arrange all operators into a vector $\chi = (a_1, a_1^\dagger, a_2, a_2^\dagger, \dots)$. The commutation relations are expressed as

$$[\chi^i, \chi^j] = \Omega^{ij},\qquad(35)$$

where $\Omega = \bigoplus_{k=1}^{n}\omega$, and ω is the symplectic matrix

$$\omega = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \omega^T = -\omega, \quad \omega^2 = -\mathbb{1}.\qquad(36)$$

Equivalently, the canonical operators Q_k and P_k are arranged in the vector $R = (Q_1, P_1, Q_2, P_2, \dots)$, related to χ by $R^i = H_j^i\chi^j$, where H is the unitary transformation

$$H = \frac{1}{\sqrt{2}}\bigoplus_{k=1}^{n}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix},\qquad(37)$$

and with commutation relations $[R^i, R^j] = i\Omega^{ij} = H_{i'}^i H_{j'}^j\Omega^{i'j'}$.

A Gaussian state is defined as a state with Gaussian characteristic function [51], which is fully described by the first moments $\langle R^i\rangle$ and by the covariance matrix (CM) Σ^{ij}:

$$\langle R^i\rangle = \mathrm{tr}[\rho R^i],\qquad(38a)$$
$$\Sigma^{ij} = \mathrm{tr}[\rho(R^i - \langle R^i\rangle)\circ(R^j - \langle R^j\rangle)].\qquad(38b)$$

Williamson's theorem [52] ensures that a Gaussian state can always be expressed as a thermal state under the action of a symplectic transformation S and a displacement operator D,

$$\rho = DS\left(\bigotimes_{k=1}^{n}\rho_{\nu_k}\right)S^\dagger D^\dagger,\qquad(39)$$

with

$$\rho_{\nu_k} = Z_k^{-1}\exp(-\beta_k a_k^\dagger a_k),\qquad(40)$$

where Z_k is a normalization factor and the β_k are inverse temperatures. The thermal state has zero first moments $\langle R^i\rangle = 0$ and a covariance matrix $\Sigma_{\mathrm{th}}^{ij} = \mathrm{tr}[(\bigotimes_{k=1}^{n}\rho_{\nu_k})R^i\circ R^j]$,

$$\Sigma_{\mathrm{th}} = \frac{1}{2}\bigoplus_{k=1}^{n}\nu_k\mathbb{1}_2, \quad \nu_k = \coth(\beta_k/2).\qquad(41)$$

In order to deal with states with nonzero first moments, it is convenient to define the displaced bosonic operators $\tilde{\chi}^i = D\chi^i D^\dagger = \chi^i - \langle\chi^i\rangle$. The displaced canonical operators $\tilde{R}^i$ are defined likewise. With this, the CM reads

$$\Sigma^{ij} = \mathrm{tr}[\rho\tilde{R}^i\circ\tilde{R}^j].\qquad(42)$$

Symplectic transformations are linear in the creation and annihilation operators and preserve the commutation relations

$$S^\dagger\chi^i S = u_j^i\chi^j,\qquad(43a)$$
$$u_{i'}^i u_{j'}^j\Omega^{i'j'} = \Omega^{ij},\qquad(43b)$$

where u is a matrix representation of the symplectic transformation S. The latter equation is often written as

$$u\Omega u^T = \Omega, \tag{44}$$

and we will use the fact that

$$u^T \Omega = \Omega u^{-1}. \tag{45}$$

In analogy with (43a), the canonical operators transform according to

$$S^\dagger R^i S = H^i_j u^j_k [H^\dagger]^k_l R^l \equiv s^i_j R^j, \tag{46}$$

where we have defined $s = HuH^\dagger$, fulfilling the same relations (44) and (45). From (42) and (46), the transformation rule for the CM under a symplectic transformation $\rho \to U\rho U^\dagger$ takes the explicit form

$$\Sigma \xrightarrow{U} s\Sigma s^T, \tag{47a}$$

$$\Sigma^{ij} \xrightarrow{U} s^i_{i'} s^j_{j'} \Sigma^{i'j'}, \tag{47b}$$

and from Eqs. (39) and (47a) we have $\Sigma = s\Sigma_{\text{th}} s^T$.

It is worth stressing at this point the doubly contravariant character of the covariance matrix Σ. Although the matrix equation (47a) suggests otherwise, it is clear from Eq. (47b) that both indices transform in the same way. In order to construct functions of the covariance matrices with proper transformation rules, we find it suitable to define the covariant-contravariant version by lowering one index with the symplectic matrix [39]. Let us introduce the symplectic matrix with lower indices $\Omega_{ij} \equiv \Omega^{ij}$, so that $\Omega^{ij}\Omega_{jk} = -\delta^i_k$. Define the 1-1 *covariance tensor* $\tilde\Sigma^i_j$ as

$$\tilde\Sigma = \Sigma\Omega \quad \text{or} \quad \tilde\Sigma^i_j = \Sigma^{ik}\Omega_{kj}. \tag{48}$$

With this definition we see that, at variance with Eqs. (47a) and (47b), under a symplectic transformation

$$\tilde\Sigma \xrightarrow{U} s\tilde\Sigma s^{-1}. \tag{49}$$

Products of 1-1 tensors enjoy proper transformation rules:

$$[AB]^i_k = A^i_j B^j_k, \tag{50}$$

$$s(AB)s^{-1} = (sAs^{-1})(sBs^{-1}). \tag{51}$$

Tensor products $A \otimes B$ are 2-2 tensors with indices

$$[A \otimes B]^{ij}_{kl} = A^i_k B^j_l. \tag{52}$$

Moreover powers of a k-k tensor transform in the same way as T,

$$[T^n]^{i_1...i_k}_{j_1...j_k} = T^{i_1...i_k}_{r'_1...r'_k} T^{r'_1...r'_k}_{r''_1...r''_k} \cdots T^{r^{(n-1)}_1...r^{(n-1)}_k}_{j_1...j_k} \tag{53}$$

$$(s^{\otimes k} T (s^{-1})^{\otimes k})^n = s^{\otimes k} T^n (s^{-1})^{\otimes k}. \tag{54}$$

The identity k-k tensor is $\mathbb{1}^{\otimes k}$, with $[\mathbb{1}]^i_j = \delta^i_j$. Finally the inverse of a tensor T is defined as

$$T^{-1} \equiv \sum_m (\mathbb{1}^{\otimes k} - T)^m \tag{55}$$

Due to the antisymmetry of Ω it is important to be consistent in the way that indices are lowered and raised. Indices are lowered by contracting with the first index of Ω_{ij} and raised

by contracting with the second index of Ω^{ij}, so that $\Sigma^{ij} = \tilde\Sigma^i_k \Omega^{jk} = \Sigma^{il}\Omega_{lk}\Omega^{jk} = \Sigma^{il}\delta^j_l$.

Finally, given a probe state ρ_0, the action of the channel is

$$\rho(X) = \mathcal{S}(X)\rho_0, \quad \mathcal{S}(X) = \exp\mathcal{G}(X), \tag{56}$$

and the derivatives of ρ with respect to the channel parameters X^μ can be neatly expressed as

$$\partial_\mu \rho(X) = \mathcal{D}_\mu \rho(X), \tag{57}$$

where $\mathcal{D}_\mu$ are superoperators whose expressions are given in Appendix B.

A. Symmetric logarithmic derivatives

From now on we consistently drop the dependency on X and assume throughout that we are considering a particular point $\mathcal{S}(X) \in \mathfrak{C}$. The common structure of the $\mathcal{D}$ superoperators,

$$\mathcal{D}_\mu \rho = \alpha_{\mu,ij}[\chi^i \rho \chi^j - (\chi^j \chi^i) \circ \rho], \tag{58}$$

comes in very handy for computing the SLDs in a general manner. From Eqs. (5) and (57), these satisfy the equation,

$$\mathcal{D}_\mu \rho = \Lambda_\mu \circ \rho. \tag{59}$$

Using $\chi^j \chi^i = \chi^i \circ \chi^j - \frac{1}{2}\Omega^{ij}$ we have

$$\mathcal{D}_\mu \rho = \alpha_{\mu,ij} \left[\chi^i \rho \chi^j - \left(\chi^i \circ \chi^j - \frac{1}{2}\Omega^{ij} \right) \circ \rho \right]. \tag{60}$$

Combining this with Eq. (59) we get

$$\alpha_{\mu,ij} \chi^i \rho \chi^j = \left[\Lambda_\mu + \alpha_{\mu,ij} \left(\chi^i \circ \chi^j - \frac{1}{2}\Omega^{ij} \right) \right] \circ \rho. \tag{61}$$

This is a particular form of the Sylvester equation [48] $Y = Z \circ \rho$, which, for $\rho > 0$, has the formal solution

$$Z = 2 \int_0^\infty e^{-v\rho} Y e^{-v\rho} dv. \tag{62}$$

Therefore,

$$\Lambda_\mu = \alpha_{\mu,ij} \left[2 \int_0^\infty e^{-v\rho} \chi^i \rho \chi^j e^{-v\rho} dv - \left(\chi^i \circ \chi^j - \frac{1}{2}\Omega^{ij} \right) \right]. \tag{63}$$

We show in Appendix C that

$$
\int_0^\infty e^{-u\rho} \chi^i \rho \chi^j e^{-u\rho} \, du
$$
$$
= [H^\dagger \otimes H^\dagger]^{ij}_{i'j'} \bigg(\{[f(\tilde\Sigma) \otimes \mathbb{1} + \mathbb{1} \otimes f(\tilde\Sigma)^{-1}]^{-1}\}^{i'j'}_{kl}
$$
$$
\times \left(\tilde R^k \circ \tilde R^l + \frac{i}{2}\Omega^{kl} \right) + \{[\mathbb{1} + f(\tilde\Sigma)]^{-1}\}^{i'}_{i''} \tilde R^{i''} \langle R^{j'} \rangle
$$
$$
+ \{[\mathbb{1} + f(\tilde\Sigma)^{-1}]^{-1}\}^{j'}_{j''} \langle R^{i'} \rangle \tilde R^{j''} + \frac{1}{2}\langle R^{i'} \rangle \langle R^{j'} \rangle \bigg), \tag{64}
$$

where $\tilde\Sigma = \Sigma\Omega$ is the 1-1 covariance tensor of ρ and

$$f(x) = \frac{x - i/2}{x + i/2}. \tag{65}$$

Finally, defining

$$\tilde\alpha_{\mu,ij} = \alpha_{\mu,i'j'}[H^\dagger \otimes H^\dagger]^{i'j'}_{ij} \tag{66}$$

allows one to express the SLD as

$$\Lambda_\mu = \tilde{\alpha}_{\mu,ij}\big([L^{(0)}]\,^{ij}_{\ kl}\Omega^{kl} + [L^{(1)}]^{ij}_k\tilde{R}^k + [L^{(2)}]^{ij}_{kl}\tilde{R}^k \circ \tilde{R}^l\big),$$
(67)

where

$$L^{(0)} = i\big\{[f(\tilde{\Sigma}) \otimes \mathbb{1} + \mathbb{1} \otimes f(-\tilde{\Sigma})]^{-1} + \tfrac{1}{2}\mathbb{1} \otimes \mathbb{1}\big\}, \quad (68a)$$

$$[L^{(1)}]^{ij}_k = \{2[\mathbb{1} + f(\tilde{\Sigma})]^{-1} - \mathbb{1}\}^i_k\langle R^j\rangle$$
$$+ \{2[\mathbb{1} + f(\tilde{\Sigma})^{-1}]^{-1} - \mathbb{1}\}^j_k\langle R^i\rangle, \quad (68b)$$

$$L^{(2)} = 2\big\{[f(\tilde{\Sigma}) \otimes \mathbb{1} + \mathbb{1} \otimes f(-\tilde{\Sigma})]^{-1} - \tfrac{1}{2}\mathbb{1} \otimes \mathbb{1}\big\}. \quad (68c)$$

Here we have used the fact that $f(x)^{-1} = f(-x)$.

Notice that this expression overcomes the main difficulty in evaluating the QFI for Gaussian channels of continuous variable systems, namely, expressing the SLD in a manageable form. In our case, the problem is reduced from working in an infinite-dimensional Hilbert space to dealing with finite-dimensional vector spaces, where the tensors $L^{(0)}$, $L^{(1)}$, and $L^{(2)}$ are defined. More importantly, these tensors depend exclusively on the covariance matrix and the first moments of the quantum state output from the channel, which have a simple relation to the channel parameters [53]. This will not only provide a means for evaluating the channel metric $\mathfrak{J}$, but it will also allow the evaluation of the performance of several channel measurement schemes [37].

We conclude this subsection by commenting on possible difficulties when computing SLDs for singular states. This situation can arise when the CM of ρ in Williamson form contains a vacuum mode. This is the case, for instance, when probing a zero-temperature channel ($N = 0$) with a two-mode squeezed vacuum. In general the SLD is not defined on the kernel of ρ, that is, $\overline{P}\Lambda\overline{P}$ is undetermined by Eq. (5), where $\overline{P}$ is the projector on ker ρ. Let $\mathcal{P}[\Lambda] = \Lambda - \overline{P}\Lambda\overline{P}$. By measuring Λ_ν on the state $\rho(X + dX) = \rho(X) + \rho(X) \circ \Lambda_\mu dX^\mu$ it is easy to see that $\text{tr}[\rho(X + dX)\Lambda_\nu] = \text{tr}[\rho(X + dX)\mathcal{P}[\Lambda_\nu]]$. This is the freedom available in defining SLDs for singular ρ's. However, Eqs. (67) and (68) were derived from Eq. (59) by means of expression (62) which assumes that the density operator is nonsingular. Thus, these expressions can yield to divergencies for $\overline{P}\Lambda_\mu\overline{P}$. These are not observable divergencies and can be regularized by introducing a small temperature ϵ in the probe state ρ_0, projecting Λ with $\mathcal{P}$, and finally taking the limit ϵ to zero. There may be, on the other hand, observable divergencies. This situation often arises when one explores the vicinity of the boundary of the manifold. The classical Fisher information is well known to diverge in many statistical models, the most prominent case being the binomial distribution when $p \to 0$ or 1, giving rise to Poissonian statistics. Analogously, the Bures distance for mixed qubit states is well known to diverge in the limit of pure states, giving rise to interesting and counterintuitive effects in Bayesian qubit estimation [54]. A similar effect in quantum statistics has been reported in a particular Gaussian channel [$X = (\gamma,0,0,0)$], when the parameter $\gamma \to 0$ [55].

B. The quantum Fisher information matrix

The expression of the SLD, Eq. (67), allows for the computation of the QFI, Eq. (6). The full derivation of the most

general QFI is given in Appendix E, Eq. (E13). We reproduce here the resulting expression:

$$\mathcal{J}_{\mu\nu}(X) = \tilde{\alpha}_{\mu,i'j'}\tilde{\alpha}_{\nu,k'l'}\big([L^{(1)}]^{i'j'}_i[L^{(1)}]^{k'l'}_k\Sigma^{ik} + [L^{(2)}]^{i'j'}_{i''j''}$$
$$\times [L^{(2)}]^{k'l'}_{kl}[D]^{i''j''}_{ij}(\Omega^{ik}\Omega^{jl} + \Omega^{il}\Omega^{jk})\big). \quad (69)$$

It should be noted that the tensors appearing in Eq. (69) depend on $\tilde{\Sigma}$, the covariance tensor of the *output* state. In order to finally evaluate the QFI one needs to express $\tilde{\Sigma}$ in terms of the initial quantum state Σ_0 and the channel parameters X. The corresponding relation is well known in the literature [53,56] but we repeat it here for completeness. Arranging the channel parameters X in the *asymptotic* covariance matrix Σ_{ch},

$$\Sigma_{\text{ch}} = \bigoplus_k \begin{pmatrix} \frac{1}{2} + N_k + \text{Re}M_k & \text{Im}M_k \\ \text{Im}M_k & \frac{1}{2} + N_k + \text{Re}M_k \end{pmatrix}, \quad (70)$$

and defining the coupling matrix,

$$\Gamma = \bigoplus_k e^{-\frac{\gamma_k}{2}}\mathbb{1}_2, \quad (71)$$

we have

$$\Sigma = \Gamma(\Sigma_0 - \Sigma_{\text{ch}})\Gamma + \Sigma_{\text{ch}}, \quad (72)$$

where Σ_0 is the CM of the initial state and Σ is the CM of the state going out of the channel.

Equations (69), (70), (71), and (72) provide a systematic way to compute the QFI for any channel $\mathcal{S}(X)$ and any input state ρ_0. Remarkably, this expression is *exact* and *analytical*. However, despite the significant simplification of the problem—namely, from an infinite dimensional operator equation (59) to a finite-dimensional matrix expression, Eq. (69)—explicit analytical expressions are too complex to be of any use, except for very simple channel models.

V. CONCLUSION AND OUTLOOK

Summarizing, for any resource constraint $\phi(\rho_\mathcal{A}) \leqslant \phi^\star$ we have defined a metric tensor on the manifold of Gaussian channels. We have proven that the metric is stable under the addition of ancillas and provided a method for computing numerical approximations, by using convex optimization methods. The resulting distance is an upper bound to the attainable Bures distance between states resulting from the action of the channel onto any given initial Gaussian state fulfilling the constraints, and it allows one to establish a systematic way to measure distances between channels in order to quantify imperfections in quantum information implementations. Moreover, the metric tensor minimizes the volume element assigned at any point of the set of Gaussian channels. This density can be used as an analog of the Jeffrey's (Bures) prior distribution for points in the simplex (density matrices) in Bayesian estimation methods. Our results are a step toward the identification of a useful notion of distance among bosonic channels, which are the basis mathematical framework of quantum communication devices with continuous variables.

Additionally, our approach has provided several results which are of relevance to the field of quantum estimation theory. We have derived closed formulas for computing the QFI of any estimation method that uses Gaussian states and

dissipative Gaussian channels. Moreover, for each one of the channel parameters we have obtained closed expressions for the corresponding optimal observables associated with each one of the parameters (symmetric logarithmic derivatives). Moreover, we have proved that by enlarging sufficiently the number of ancillary modes, the optimal probe states for estimating Gaussian channels are always pure.

These findings should be of immediate use in several practical situations. The most prominent of them is entanglement-enhanced metrology of Gaussian channels. A first relevant application of our results is the evaluation of Gaussian protocols for parameter estimation in dissipative Gaussian channels, such as decay rate, temperature, or degree of squeezing in engineered baths [37]. The stability proof in Sec. III C establishes the maximum number of ancillary modes to be considered for any of these situations. Indeed, such proof is very general and can be extended to several metrology problems, limiting the number of ancillary modes required for optimality and thus restricting the number of setups that need to be explored when designing optimal metrology protocols. In the spirit of the Choi-Jamiolkowski theorem, our arguments show that estimation of single-mode channels cannot be enhanced by multipartite entanglement as compared to bipartite entanglement, at least within the Gaussian framework.

Moreover, the approach taken here may be extended to derive SLDs for other related problems, such as phase estimation under decoherence and phase diffusion, which can be expressed quite naturally in phase space and that would otherwise be difficult to address due to the infinite dimensional character of continuous variable systems.

Our work leaves open several questions and possible extensions. Obviously, obtaining explicit analytic expressions for the channel metric would be the ultimate achievement. However, this seems to be out of reach unless some significant advances are made. In particular we point out three relevant missing points: (a) a clear criterion for the determination of a set of optimal probe states for any given choice of ϕ; (b) a general form of the QFI for all Gaussian channels, not only the dissipative ones; and (c) an analytic expression for the metric in the high resource limit (properly regularized). Concentrating our attention on the regularization schemes, we can point out some interesting questions: What classes of constraints ϕ will always be saturated by the optimizing states? Considering a related problem, preliminary numerical results [37] show that, in the single-mode bosonic lossy channel, at zero-temperature, with constraints of the form $\phi(\rho) = \text{tr}[a^\dagger a \rho]$, an entangled two-mode squeezed state of the probe and ancilla modes with mean photon number $\langle n \rangle$ in the channel mode will perform equally well to any other state with n average photons. If the result of this preliminary analysis is confirmed, it is unlikely to be a coincidence. Finding the reasons behind this surprising match is an interesting problem worth deeper investigation.

Naturally, these are not the only open questions, and in particular, extending our work to non-Gaussian channels and/or non-Gaussian probes would be of utmost importance. There is, in fact, nothing peculiar in our defining scheme that relies crucially on the Gaussianity of the channels and/or the probe states. The metric could thus be defined on more general sets of channels and probe states. It is, however, the

restriction to Gaussian channels and states that allows us to provide computational formulas and prove that our metric satisfies the basic requirements listed in the Introduction, especially the stability property. This does not mean that non-Gaussian extensions will not fulfill the requirements. Indeed, it would be an interesting result if a counterexample to stability could be found in the large set of non-Gaussian states and channels. It would mean that some estimation protocols can be improved by considering multipartite setups, when only one part (mode) undergoes the channel action. This would be an effect exclusively of continuous variable systems which cannot occur in finite dimensions.

Finally, a question remains unanswered regarding the way in which channels are combined. Throughout the text we have assumed that the channel is tested a large number, N, of times, *independently*; that is, a given probe state is prepared, sent through the channel and measured, and the process is repeated N times. The number of trials is immediately regularized and thus does not enter in the discussion. However, in the quantum scenario, regularizing N leaves no room for chaining channels together; so that, for instance, a state may be sent twice through the channel before being measured or entangled states may be used in the simultaneous testing of a single channel and a doubly chained channel. It is evident that a large number of degrees of freedom is not exploited in our approach. However, our framework yields the most natural bounds, which can be obtained in the *iid* ("independent and identically distributed") case, when no correlation is admitted between different samples. A very relevant and interesting open question is whether such sophisticated schemes can enhance the precision of quantum estimation and quantum metrology setups.

APPENDIX A: CONVERGENCE OF THE APPROXIMATION ALGORITHM

We show here that the approximation method given in Sec. III D does converge to the true value of the metric $\mathfrak{J}$. Throughout this appendix we make extensive use of the Hausdorff distance $d_H(\cdot,\cdot)$ [57], which defines a distance among subsets of a metric space W. In our case W is the set of positive semidefinite matrices, equipped with the operator norm distance $d(x,y) = \|x - y\|$, where $\|x\|$ is the operator norm $\|x\| = \sup_{|\vec{v}|=1} |x\vec{v}|$. Define the ϵ-ball centered at a point x as the set of points within a distance ϵ of

x, $B_\epsilon(x) = \{x'|d(x,x') < \epsilon\}$. Then, the ϵ-neighborhood of a subset X is the union of all ϵ-balls of X, $B_\epsilon(X) = \cup_{x \in X} B_\epsilon(x)$. Then, the Hausdorff distance is defined as

$$d_H(X,Y) = \inf\{\delta | X \subset B_\delta(Y) \quad \text{and} \quad Y \subset B_\delta(X)\}. \quad (A1)$$

Let us first show some preliminary facts about the randomly generated sets of constraints, C_n.

Theorem. Assuming that the set $C^\star = C(\mathcal{S}|\phi^\star)$ of QFI matrices associated with a given channel $\mathcal{S}$ [within the corresponding probe state restrictions $\phi(\rho) \leqslant \phi^\star$] is bounded, we have the following property: For any $\epsilon > 0$ the probability $\Pr(d_H(C_n, C^\star) \geqslant \epsilon)$ decreases exponentially with n.

Proof: Since the set $C^\star$ is bounded it can be covered by a finite number k of $\epsilon/2$-balls centered at points $\{y_i \in C^\star\}$, $i = 1, \ldots, k$, thus $C^\star \subseteq B_{\epsilon/2}(\{y_i\})$. If

$$B_{\epsilon/2}(\{y_i\}) \subset B_\epsilon(C_n), \quad (A2)$$

then $C^\star \subseteq B_{\epsilon/2}(\{y_i\}) \subset B_\epsilon(C_n)$. Since also $C_n \subset C^\star \subset B_\epsilon(C^\star)$ we have that Eq. (A2) implies $d_H(C_n, C^\star) \leqslant \epsilon$.

We now compute an upper bound to the probability that Eq. (A2) is true. To each random selection $x \in C_n$ we assign the point $y \in \{y_i\}$ which is closest to x. Since $d(x,y) < \epsilon/2$, then $B_{\epsilon/2}(y) \subseteq B_\epsilon(x)$. Let n_i be the number of points of C_n assigned to ball $B_{\epsilon/2}(y_i)$. A sufficient condition for Eq. (A2) to hold is $n_i \neq 0 \forall i$ (all points in $\{y\}$ have been assigned *at least* one point $x \in C_n$). Thus,

$$n_i \neq 0 \ \forall i \implies d_H(C_n, C^\star) \leqslant \epsilon. \quad (A3)$$

Consequently,

$$\Pr(d_H(C_n, C^\star) \leqslant \epsilon) \geqslant \Pr(n_i \neq 0 \ \forall i). \quad (A4)$$

Let p_i be the probability that a random sample x is assigned to point y_i. The probability that after n samplings a number $n_1, \ldots, n_k$ of points has been assigned to each ball is given by the multinomial distribution

$$\Pr(n_1, \ldots, n_k) = \frac{n!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k}.$$

We now bound the complementary probabilities

$$\Pr(d_H(C_n, C^\star) > \epsilon) \leqslant \Pr(\exists i \text{ s.t. } n_i = 0)$$
$$= \sum_{n_2, \ldots, n_k \neq 0} \Pr(0, n_2, \ldots, n_k) + \text{perm.}$$
$$+ \sum_{n_3, \ldots, n_k \neq 0} \Pr(0, 0, \ldots, n_k) + \text{perm.}$$
$$+ \cdots + \Pr(0, 0, \ldots, 0, n) + \text{perm.},$$
$$(A5)$$

where perm. represents all permutations among arguments of the multinomial distribution and it is implicit that summation is over all n_k values such that $\sum_k n_k = n$. We can now upper bound $\Pr(\exists i \text{ s.t. } n_i = 0)$ by completing the sums to include indices equal to zero and by using the multinomial theorem:

$$\Pr(\exists i \text{ s.t. } n_i = 0) \leqslant (p_2 + p_3 + \cdots p_k)^n + \text{perm.}$$
$$+ (p_3 + \cdots + p_k)^n + \text{perm.}$$
$$+ \cdots + p_k^n + \text{perm.}$$

$$= (1 - p_1)^n + \text{perm.}$$
$$+ (1 - p_1 - p_2)^n + \text{perm.}$$
$$+ \cdots$$
$$+ (1 - p_1 - p_2 - \cdots - p_{k-1})^n + \text{perm.}$$
$$(A6)$$

To conclude we can further upper bound this quantity by replacing each term $(\cdot)^n$ by the maximum value $(1 - \min_i p_i)^n$ and counting the number of terms $\sum_{l=1}^{k-1} \binom{k}{l} = 2^k - 2$,

$$\Pr(\exists i \text{ s.t. } n_i = 0) \leqslant (2^k - 2)(1 - \min_i p_i)^n. \quad (A7)$$

Notice that this is exponentially decreasing with n. Thus, we have that for any $\epsilon > 0$ the probability $\Pr(d_H(C_n, C^\star) > \epsilon)$ decreases exponentially with n. This means that by picking enough samples we have $d_H(C_n, C^\star) \leqslant \epsilon$ with arbitrarily high probability. ∎

Let $x \in X$ be a positive semidefinite matrix within a set X. Define $M(x)$ as the set of all upper bounds to x, and define $M(X)$ as the set of upper bounds to all elements in X, that is, $M(X) = \{m | m \geqslant x \ \forall x \in X\}$. We call $M(X)$ the set of *feasible points* to the optimization problem.

Lemma. Given two matrices x and y such that $d(x,y) < \delta$, it holds that $x - \delta\mathbb{1} < y < x + \delta\mathbb{1}$.

Proof. Define $\Delta = y - x$. Then $\|\Delta\| = d(x,y) < \delta$. Moreover $-\|\Delta\|\mathbb{1} \leqslant \Delta \leqslant \|\Delta\|\mathbb{1}$, so that $-\delta\mathbb{1} < \Delta < \delta\mathbb{1}$. Thus,

$$x - \delta\mathbb{1} < y < x + \delta\mathbb{1}. \quad (A8)$$

∎

Theorem. The map $M : X \mapsto M(X)$ is continuous in the Hausdorff distance. Namely, for any $\epsilon > 0$, there exists a $\delta > 0$ such that $d_H(X,Y) < \delta \Rightarrow d_H(M(X), M(X)) < \epsilon$.

Proof. For any $\epsilon > 0$, take $\delta < \epsilon$. We prove constructively by showing that if $d_H(X,Y) < \delta$ it follows that $M(X) \subset B_\epsilon(M(Y))$ and $M(Y) \subset B_\epsilon(M(X))$, which in turn implies that $d_H(M(X), M(Y)) < \epsilon$.

We have to show that $a \in B_\epsilon(M(Y)) \ \forall a \in M(X)$. Take any element $a \in M(X)$ and construct $b = a + \delta\mathbb{1}$. It must follow that (i) $b \in M(Y)$ and (ii) $d(a,b) < \epsilon$. (i) Since $d_H(X,Y) < \delta$, then for all $y \in Y$ we have an $x \in X$ such that $d(x,y) < \delta$ and, by the lemma, $x + \delta\mathbb{1} > y$. Thus, $b = a + \delta\mathbb{1} > x + \delta\mathbb{1} > y$ for all $y \in Y$. Thus $b \in M(Y)$. (ii) $d(a,b) = \delta < \epsilon$. This proves that $M(X) \subset B_\epsilon(M(Y))$. The exact same reasoning with X and Y interchanged shows the converse; thus $d_H(M(X), M(Y)) < \epsilon$. ∎

The last two theorems combined show that, given any $\epsilon > 0$, one can always generate a set of feasible points $M(C_n)$ which has Hausdorff distance from $M(C^\star)$ smaller than ϵ with arbitrarily high probability, that is, $\Pr(d_H(M(C_n), M(C^\star)) < \epsilon) \geqslant 1 - k\exp(-n)$. Define the function

$$f(X) = \arg\inf_{x \in X} \det x. \quad (A9)$$

It only remains to prove that $\mathfrak{J}_n = f(M(C_n))$ approaches $\mathfrak{J}$ as $M(C_n)$ approaches $M(C^\star)$. Let us remark that, by construction, $M(C^\star) \subset M(C_n)$. As said in the text, we assume that the solution to the real problem $\mathfrak{J} = f(M(C^\star))$ is nondegenerate. This means that $\det(\mathfrak{J} + \Delta) > \det\mathfrak{J}$ whenever $\mathfrak{J} + \Delta \in M(C^\star)$ and $\|\Delta\| \neq 0$.

Theorem. Assuming that the miminum of $\{\det y | y \in Y\}$ is nondegenerate, the map $\{f(X_n)\}$ converges to $f(Y)$ in the Hausdorff distance whenever $\{X_n | X_n \supset Y \forall n\}$ converges to Y. Namely, for any $\epsilon > 0$, there is a $\delta > 0$ such that if $Y \subset X$ and $d_H(X, Y) < \delta$ then $\|f(X) - f(Y)\| < \epsilon$.

Proof. Let $x^\star = f(X)$ and $y^\star = f(Y)$. Consider an open ball of radius $d < \epsilon/2$ around $y^\star$, and let Γ be the smallest gap between determinants of points in Y outside $B_d(y^\star)$ and $\det y^\star$, that is,

$$\Gamma = \inf_{\substack{y \notin B_d(y^\star) \\ y \in Y}} \det y - \det y^\star. \qquad (A10)$$

By the nondegeneracy assumption, $\Gamma > 0$. Consider the function

$$h(\delta) = n(\|y^\star\| + D + \delta)^{n-1}\delta, \qquad (A11)$$

where D is the maximum distance between points in the boundary of Y. Clearly h is monotonically increasing and $h(0) = 0$. Let

$$\delta < \min\left(h^{-1}(\Gamma), \epsilon/2\right). \qquad (A12)$$

By assumption $X \subset B_\delta(Y)$ and $Y \subset X$. Let $y \in Y$ be the closest point to $x^\star$. It is easy to see that y must lie in the boundary of Y, since if it did not there would be a ball $B_\tau(y) \subset Y$ and $y' = y - \tau\|x^\star - y\|\mathbb{1} \in B_\tau(y) \subset Y$ would have $\|y' - x^\star\| = (1 - \tau)\|y - x^\star\|$, which contradicts the assumption that y is the closest point to $x^\star$ in Y.

We know that $\|x^\star - y\| < \delta$. Let $y = y^\star + \Delta$. By the triangle inequality,

$$\|x^\star - y^\star\| \leqslant \|\Delta\| + \delta. \qquad (A13)$$

Also, since $Y \subset X$ we have $\det x^\star \leqslant \det y^\star$; thus

$$\det y - \det y^\star \leqslant \det y - \det x^\star. \qquad (A14)$$

Moreover (see [48]),

$$\det y - \det x^* \leqslant n \max(\|y\|, \|x^\star\|)^{n-1} \|x^\star - y\|. \qquad (A15)$$

We can further simplify this by noticing that $\|x^\star\| = \|y + (x^\star - y)\| \leqslant \|y\| + \|x^\star - y\| \leqslant \|y\| + \delta$; thus $\max(\|y\|, \|x^\star\|) \leqslant \|y\| + \delta \leqslant \|y^\star\| + \|\Delta\| + \delta$, so we have

$$\det y - \det x^* \leqslant n(\|y^\star\| + \|\Delta\| + \delta)^{n-1}\delta. \qquad (A16)$$

Notice that $\|\Delta\| = \|y - y^\star\|$ is the distance between two points in the boundary of Y and thus $\|\Delta\| \leqslant D$, so that

$$\det y - \det x^* \leqslant h(\delta). \qquad (A17)$$

On the other hand, let us show by contradiction that $\|\Delta\| < d$. Suppose $y \notin B_d(y^\star)$. By Eq. (A17), the monotonicity of h, and the definition of δ, we have

$$\det y - \det x^\star \leqslant h(\delta) < \Gamma, \qquad (A18)$$

which means that $\det x^\star > \det y - \Gamma$. But $y \notin B_d(y^\star)$ means $\det y - \det y^\star \geqslant \Gamma$, which yields $\det x^\star > \det y^\star$. This contradiction shows that $\|\Delta\| \leqslant d$, which in turn shows that $\|\Delta\| < \epsilon/2$.

Finally, recovering Eq. (A13) we obtain

$$\|x^\star - y^\star\| < \epsilon. \qquad (A19)$$

■

Taking $\mathfrak{J}_n = f(M(C_n))$ and $\mathfrak{J} = f(M(C^\star))$ and applying this theorem to $f(M(C_n))$ and $f(M(C^\star))$, where $C_n \subset C^\star$ and $M(C^\star) \subseteq M(C)$ with

$$\Pr(d_H(M(C_n), M(C^\star)) > \delta) \overset{n \to \infty}{\longrightarrow} 0 \qquad (A20)$$

will ensure that

$$\Pr(\|\mathfrak{J}_n - \mathfrak{J}\| > \epsilon) \overset{n \to \infty}{\longrightarrow} 0 \qquad (A21)$$

Since the probability to generate a C_n within any given Hausdorff distance to $C^\star$ approaches 1 exponentially, we see that any desired precision can be attained with arbitrarily high certainty by sampling enough matrices from $C^\star$.

APPENDIX B: DERIVATIVES WITH RESPECT TO THE CHANNEL PARAMETERS

In order to compute the derivatives of $\rho(X) = \mathcal{S}(X)\rho_0$, note that

$$\partial_\mu \rho(X) = \partial_\mu \mathcal{S}(X)\rho_0. \qquad (B1)$$

Thus we only need to compute the derivative of the superoperator $\mathcal{S}(X)$. Using the relation [58]

$$\frac{\partial}{\partial X^\mu} \exp \mathcal{G}(X) = \int_0^1 e^{u\mathcal{G}(X)} \partial_\mu \mathcal{G}(X) e^{(1-u)\mathcal{G}(X)} du, \qquad (B2)$$

the channel derivatives can be written as

$$\partial_\mu \mathcal{S}(X) = \int_0^1 e^{u\mathcal{G}(x_\kappa)} \partial_\mu \mathcal{G}(x_\kappa) e^{-u\mathcal{G}(x_\kappa)} du \, \mathcal{S}(X), \qquad (B3)$$

where κ stands for the mode to which the parameter X^μ corresponds. This expression allows for the handy relation

$$\partial_\mu \mathcal{S}(X) = \mathcal{D}_\mu \mathcal{S}(X), \qquad (B4)$$

with

$$\mathcal{D}_\mu = \int_0^1 e^{u\mathcal{G}(x_\kappa)} \partial_\mu \mathcal{G}(x_\kappa) e^{-u\mathcal{G}(x_\kappa)} du. \qquad (B5)$$

Using the Hadamard lemma,

$$\exp(uA)B\exp(-uA) = \sum_{m=0}^{\infty} \frac{u^m}{m!} [A, B]_m, \qquad (B6)$$

with $[A, B]_0 \equiv B$ and $[A, B]_m \equiv [A, [A, B]_{m-1}]$, and the commutation relations

$$[L[a_k^\dagger], L[a_{k'}]] = 2\delta_{kk'}(L[a_k] + L[a_k^\dagger]), \qquad (B7a)$$

$$[D[a_k^\dagger], D[a_{k'}]] = 0, \qquad (B7b)$$

$$[D[a_k], L[a_{k'}]] = 2\delta_{kk'}D[a_k], \qquad (B7c)$$

$$[L[a_k^\dagger], D[a_{k'}]] = 2\delta_{kk'}D[a_k], \qquad (B7d)$$

$$[D[a_k^\dagger], L[a_{k'}]] = 2\delta_{kk'}D[a_k^\dagger], \qquad (B7e)$$

$$[L[a_k^\dagger], D[a_{k'}^\dagger]] = 2\delta_{kk'}D[a_k^\dagger], \qquad (B7f)$$

one can show by induction that

$$[\mathcal{G}(X), L[a_k]]_m = \gamma_k^m L[a_k] + 2\gamma_k^{m-1}\mathcal{G}(x_k), \qquad (B8a)$$

$$[\mathcal{G}(X), L[a_k^\dagger]]_m = \gamma_k^m L[a_k^\dagger] - 2\gamma_k^{m-1}\mathcal{G}(x_k), \qquad (B8b)$$

$$[\mathcal{G}(X), D[a_k]]_m = \gamma_k^m D[a_k], \qquad (B8c)$$

$$[\mathcal{G}(X), D[a_k^\dagger]]_m = \gamma_k^m D[a_k^\dagger], \qquad (B8d)$$

and the $\mathcal{D}$ superoperators become

$$\mathcal{D}_{\gamma_k} = \frac{1}{\gamma_k}\mathcal{G}(x_k), \tag{B9a}$$

$$\mathcal{D}_{N_k} = (e^{\gamma_k} - 1)\frac{L[a_k] + L[a_k^\dagger]}{2}, \tag{B9b}$$

$$\mathcal{D}_{\mathrm{Re}M_k} = (e^{\gamma_k} - 1)\frac{D[a_k] + D[a_k^\dagger]}{2}, \tag{B9c}$$

$$\mathcal{D}_{\mathrm{Im}M_k} = (e^{\gamma_k} - 1)\frac{D[a_k] - D[a_k^\dagger]}{2i}. \tag{B9d}$$

Finally, notice that the $\mathcal{D}$ superoperators allow for evaluating derivatives of the state ρ in a simple form,

$$\partial_\mu \rho = \partial_\mu S(X)\rho_0 = \mathcal{D}_\mu S(X)\rho_0 = \mathcal{D}_\mu \rho, \tag{B10}$$

where all have the structure

$$\mathcal{D}_\mu \rho = \alpha_{\mu,ij}[\chi^i \rho \chi^j - (\chi^j \chi^i) \circ \rho]. \tag{B11}$$

APPRNDIX C: COMPUTING THE SLD

We proceed here to show how Eqs. (64), (67), and (68) are derived. We remark that the relations obtained apply to all Gaussian states in any number of modes. We use the notations introduced in Sec. IV.

Lemma. $[\rho_{\mathrm{th}}, \chi^i] = [\exp M - \mathbb{1}]^i_j \chi^j \rho_{\mathrm{th}}$, where $M = \bigoplus_k \mathrm{diag}(\beta_k, -\beta_k)$.

Proof. We begin by showing

$$\rho_{\mathrm{th}}\chi^i \rho_{\mathrm{th}}^{-1} = [\exp M]^i_j \chi^j. \tag{C1}$$

Define $A = -\bigoplus_k \beta_k a_k^\dagger a_k$ and $B_m = [A, B_{m-1}]$; $B_0 = \chi^i$. One can show by induction that $B_m = [M^m]^i_j \chi^j$. Using the Hadamard lemma,

$$\rho_{\mathrm{th}}\chi^i \rho_{\mathrm{th}}^{-1} = \sum_m \frac{1}{m!}B_m = [\exp M]^i_j \chi^j, \tag{C2}$$

thus, finally,

$$[\rho_{\mathrm{th}}, \chi^i] = (\rho_{\mathrm{th}}\chi^i \rho_{\mathrm{th}}^{-1} - \chi^i)\rho_{\mathrm{th}}$$
$$= [\exp M - \mathbb{1}]^i_j \chi^j \rho_{\mathrm{th}}. \tag{C3}$$

∎

Lemma. Let u be the symplectic representation of S as defined in Eq. (43a) and s as defined in Eq. (46), where S is defined from ρ in Eq. (39). Let Σ be the covariance matrix of ρ [Eq. (42)] and $E = \exp M$. Then

$$uEu^{-1} = H^\dagger f(\tilde\Sigma)H,$$

where

$$f(x) = \frac{x - i/2}{x + i/2}.$$

Proof. Observe that

$$f(\tilde\Sigma) = \frac{\tilde\Sigma^2 - \mathbb{1}/4 - i\tilde\Sigma}{\tilde\Sigma^2 + \mathbb{1}/4}$$
$$= s\frac{\tilde\Sigma_{\mathrm{th}}^2 - \mathbb{1}/4 - i\tilde\Sigma_{\mathrm{th}}}{\tilde\Sigma_{\mathrm{th}}^2 + \mathbb{1}/4}s^{-1}$$

$$= s\bigoplus_k \begin{pmatrix} \cosh\beta_k & i\sinh\beta_k \\ -i\sinh\beta_k & \cosh\beta_k \end{pmatrix}s^{-1}$$
$$= sHEH^\dagger s^{-1}$$
$$= HuEu^{-1}H^\dagger. \tag{C4}$$

∎

Lemma. Given $\rho > 0$ with CM Σ, define $F = H^\dagger f(\tilde\Sigma)H$. The following relation holds:

$$e^{-v\rho}\tilde\chi^i e^{v\rho} = \sum_m \frac{(-v)^m}{m!}[(F - \mathbb{1})^m]^i_{i'}\tilde\chi^{i'}\rho^m.$$

Proof. Again, using the Hadamard lemma,

$$e^{-v\rho_{\mathrm{th}}}\chi^i e^{v\rho_{\mathrm{th}}} = \sum_m \frac{(-v)^m}{m!}C_m, \tag{C5}$$

where $C_0 = \chi^i$ and $C_m = [\rho_{\mathrm{th}}, C_{m-1}]$. One can show by induction that $C_m = [(E - \mathbb{1})^m]^i_j \chi^j \rho_{\mathrm{th}}^m$. Now observe that

$$e^{-v\rho}\tilde\chi^i e^{v\rho} = DSe^{-v\rho_{\mathrm{th}}}S^\dagger \chi^i Se^{v\rho_{\mathrm{th}}}S^\dagger D^\dagger$$
$$= s^i_{i'}DSe^{-v\rho_{\mathrm{th}}}\chi^{i'}e^{v\rho_{\mathrm{th}}}S^\dagger D^\dagger$$
$$= \sum_m \frac{(-v)^m}{m!}s^i_{i'}[(E - \mathbb{1})^m]^{i'}_{i''}DS\chi^{i''}\rho_{\mathrm{th}}^m S^\dagger D^\dagger$$
$$= \sum_m \frac{(-v)^m}{m!}u^i_{i'}[(E - \mathbb{1})^m]^{i'}_{i''}[u^{-1}]^{i''}_{i'''}\tilde\chi^{i'''}\rho^m$$
$$= \sum_m \frac{(-v)^m}{m!}[(F - \mathbb{1})^m]^i_{i'}\tilde\chi^{i'}\rho^m. \tag{C6}$$

∎

Lemma. Given $\rho > 0$ with CM Σ and $\tilde\chi^i = D\chi^i D^\dagger$, the following holds:

$$\int_0^\infty e^{-v\rho}\tilde\chi^i e^{-v\rho}dv = [(\mathbb{1} + F)^{-1}]^i_{i'}\tilde\chi^{i'}\rho^{-1}$$
$$= [(\mathbb{1} + F^{-1})^{-1}]^i_{i'}\rho^{-1}\tilde\chi^{i'}. \tag{C7}$$

Proof.

$$\int_0^\infty e^{-v\rho}\tilde\chi^i e^{-v\rho}dv$$
$$= \sum_m [(F - \mathbb{1})^m]^i_{i'}\tilde\chi^{i'}\rho^m \int_0^\infty dv \frac{(-v)^m}{m!}e^{-2v\rho}$$
$$= \sum_m \frac{1}{2}\left[\left(\frac{\mathbb{1} - F}{2}\right)^m\right]^i_{i'}\tilde\chi^{i'}\rho^{-1}$$
$$= [(\mathbb{1} + F)^{-1}]^i_{i'}\tilde\chi^{i'}\rho^{-1}. \tag{C8}$$

The other identity is proven analogously. ∎

Theorem. Given $\rho > 0$, the following relation holds:

$$\int_0^\infty e^{-v\rho}\tilde\chi^i \rho \tilde\chi^j e^{-v\rho}dv = [(F \otimes \mathbb{1} + \mathbb{1} \otimes F^{-1})^{-1}]^{ij}_{kl}\tilde\chi^k \tilde\chi^l,$$

where the inverse of a tensor T^{ij}_{kl} is T^{-1} such that

$$T^{ij}_{kl}[T^{-1}]^{kl}_{rs} = \delta^i_r \delta^j_s.$$

Proof. Making extensive use of the previous lemma,

$$\int_0^\infty e^{-v\rho} \tilde{\chi}^i \rho \tilde{\chi}^j e^{-v\rho} dv$$

$$= \sum_m \frac{1}{m!} [(F - \mathbb{1})^m]_k^i \tilde{\chi}^k$$

$$\times \int_0^\infty (-v)^m \rho^{m+1} e^{-v\rho} \tilde{\chi}^j e^{-v\rho} dv$$

$$= \sum_{mn} \frac{(-1)^m}{m!n!} [(F - \mathbb{1})^m]_k^i [(F - \mathbb{1})^n]_l^j \tilde{\chi}^k \rho^{1+m+n}$$

$$\times \int_0^\infty v^{m+n} e^{-2v\rho} \rho^{-n} \tilde{\chi}^l \rho^n du$$

$$= \sum_{mn} \frac{(-1)^m (m+n)!}{2^{1+m+n} m!n!} [(F - \mathbb{1})^m]_k^i [(F - \mathbb{1})^n]_l^j$$

$$\times \tilde{\chi}^k \rho^{-n} \tilde{\chi}^l \rho^n$$

$$= \sum_{mn} \frac{(-1)^m}{2^{1+m+n}} \binom{m+n}{n} [(F - \mathbb{1})^m]_k^i [(F - \mathbb{1})^n]_l^j$$

$$\times [F^{-n}]_s^l \tilde{\chi}^k \tilde{\chi}^s$$

$$= \sum_{mn} \frac{(-1)^m}{2^{1+m+n}} \binom{m+n}{n} [(F - \mathbb{1})^m]_k^i [(\mathbb{1} - F^{-1})^n]_l^j$$

$$\times \tilde{\chi}^k \tilde{\chi}^l. \tag{C9}$$

Next, make the change of variables $m + n = q$ and replace $\sum_{m,n=0}^\infty$ by $\sum_{q=0}^\infty \sum_{n=0}^q$,

$$\int_0^\infty e^{-v\rho} \tilde{\chi}^i \rho \tilde{\chi}^j e^{-v\rho} dv$$

$$= \sum_{q=0}^\infty \frac{(-1)^q}{2^{q+1}} \sum_{n=0}^q \binom{q}{n} [(F - \mathbb{1})^{q-n} \otimes (F^{-1} - \mathbb{1})^n]_{kl}^{ij}$$

$$= [(F \otimes \mathbb{1} + \mathbb{1} \otimes F^{-1})^{-1}]_{kl}^{ij}. \tag{C10}$$

This proves the theorem. ∎

Finally we are in position to obtain Eq. (64). Observe that

$$\int_0^\infty e^{-v\rho} \chi^i \rho \chi^j e^{-v\rho} dv$$

$$= \int_0^\infty e^{-v\rho} (\tilde{\chi} + \langle \chi \rangle)^i \rho (\tilde{\chi} + \langle \chi \rangle)^j e^{-v\rho} dv. \tag{C11}$$

Collecting the results and using

$$[H \otimes H]_{k'l'}^{kl} \tilde{\chi}^{k'} \tilde{\chi}^{l'} = \tilde{R}^k \tilde{R}^l$$

$$= \tilde{R}^k \circ \tilde{R}^l + \frac{i}{2} \Omega^{kl}, \tag{C12}$$

$$[H \otimes H]_{k'l'}^{kl} \tilde{\chi}^{k'} \langle \chi^{l'} \rangle = \tilde{R}^k \langle R^l \rangle, \tag{C13}$$

$$[H \otimes H]_{k'l'}^{kl} \langle \chi^{k'} \rangle \langle \chi^{l'} \rangle = \langle R^k \rangle \langle R^l \rangle, \tag{C14}$$

we finally obtain

$$\int_0^\infty e^{-u\rho} \chi^i \rho \chi^j e^{-u\rho} du$$

$$= [H^\dagger \otimes H^\dagger]_{i'j'}^{ij} \Big([(f(\tilde{\Sigma}) \otimes \mathbb{1} + \mathbb{1} \otimes f(\tilde{\Sigma})^{-1})^{-1}]_{kl}^{i'j'}$$

$$\times \left(\tilde{R}^k \circ \tilde{R}^l + \frac{i}{2} \Omega^{kl} \right) + [(\mathbb{1} + f(\tilde{\Sigma}))^{-1}]_{i''}^{i'} \tilde{R}^{i''} \langle R^{j'} \rangle$$

$$+ [(\mathbb{1} + f(\tilde{\Sigma})^{-1})^{-1}]_{j''}^{j'} \langle R^{i'} \rangle \tilde{R}^{j''} + \frac{1}{2} \langle R^{i'} \rangle \langle R^{j'} \rangle \Big), \tag{C15}$$

which is Eq. (64).

APPENDIX D: SOME SIMPLIFYING IDENTITIES

In this section we derive some relations that will be useful for the simplification and evaluation of the QFI. Consider the tensor $[f(\tilde{\Sigma}) \otimes \mathbb{1} + \mathbb{1} \otimes f(-\tilde{\Sigma})]^{-1}$. Notice that if $[X, Y] = 0$ we can write

$$[f(X) + f(Y)]^{-1} = \frac{1}{2} \left(\frac{XY - \mathbb{1}/4}{XY + \mathbb{1}/4} + \frac{i}{2} \frac{(X + Y)}{XY + \mathbb{1}/4} \right). \tag{D1}$$

Given that $f(X) \otimes \mathbb{1} = f(X \otimes \mathbb{1})$, $\mathbb{1} \otimes f(Y) = f(\mathbb{1} \otimes Y)$, and $[\tilde{\Sigma} \otimes \mathbb{1}, \mathbb{1} \otimes \tilde{\Sigma}] = 0$, we can write

$$[f(\tilde{\Sigma}) \otimes \mathbb{1} + \mathbb{1} \otimes f(-\tilde{\Sigma})]^{-1}$$

$$= \frac{1}{2} D^{-1} \left(\tilde{\Sigma} \otimes \tilde{\Sigma} + \frac{1}{4} \mathbb{1} \otimes \mathbb{1} - \frac{i}{2} (\tilde{\Sigma} \otimes \mathbb{1} - \mathbb{1} \otimes \tilde{\Sigma}) \right), \tag{D2}$$

where we have defined

$$D = \tilde{\Sigma} \otimes \tilde{\Sigma} - \tfrac{1}{4} \mathbb{1} \otimes \mathbb{1}; \tag{D3}$$

thus

$$L^{(0)} = i D^{-1} \left[\tilde{\Sigma} \otimes \tilde{\Sigma} - \frac{i}{4} (\tilde{\Sigma} \otimes \mathbb{1} - \mathbb{1} \otimes \tilde{\Sigma}) \right], \tag{D4}$$

$$L^{(2)} = \frac{1}{2} D^{-1} [\mathbb{1} \otimes \mathbb{1} - i(\tilde{\Sigma} \otimes \mathbb{1} - \mathbb{1} \otimes \tilde{\Sigma})]. \tag{D5}$$

At this point, the following relations are useful,

$$[\tilde{\Sigma} \otimes \tilde{\Sigma}]_{kl}^{ij} \Omega^{kl} = [\Sigma \Omega \Omega (\Sigma \Omega)^T]^{ij}$$
$$= [\Sigma \Omega \Sigma]^{ij}, \tag{D6a}$$

$$[\tilde{\Sigma} \otimes \mathbb{1} - \mathbb{1} \otimes \tilde{\Sigma}]_{kl}^{ij} \Omega^{kl} = [\Sigma \Omega \Omega - \Omega (\Sigma \Omega)^T]^{ij}$$
$$= -2\Sigma^{ij}, \tag{D6b}$$

$$[\tilde{\Sigma} \otimes \tilde{\Sigma}]_{kl}^{ij} \Sigma^{kl} = [\Sigma \Omega \Sigma (\Sigma \Omega)^T]^{ij}$$
$$= -[\Sigma \Omega \Sigma \Omega \Sigma]^{ij}, \tag{D6c}$$

$$[\tilde{\Sigma} \otimes \mathbb{1} - \mathbb{1} \otimes \tilde{\Sigma}]_{kl}^{ij} \Sigma^{kl} = [\Sigma \Omega \Sigma - \Sigma (\Sigma \Omega)^T]^{ij}$$
$$= 2[\Sigma \Omega \Sigma]^{ij}. \tag{D6d}$$

Thus finally we get

$$[L^{(0)}]_{kl}^{ij} \Omega^{kl} = i \left[D^{-1} \left(\Sigma \Omega \Sigma + \frac{i}{2} \Sigma \right) \right]^{ij}, \tag{D7}$$

$$[L^{(2)}]_{kl}^{ij} \Sigma^{kl} = -i \left[D^{-1} \left(\Sigma \Omega \Sigma + \frac{i}{2} \Sigma \right) \right]^{ij}; \tag{D8}$$

thus

$$[L^{(0)}]_{kl}^{ij} \Omega^{kl} = -[L^{(2)}]_{kl}^{ij} \Sigma^{kl}. \tag{D9}$$

This identity, together with the fact that $\mathrm{tr}[\rho \tilde{R}^i] = 0$, guarantees the consistency check $\mathrm{tr}[\rho \Lambda_\mu] = 0$ [Eq. (7)]. It also

turns out to be very handy in the simplification of the QFI in Appendix E.

PPENDIX E: COMPUTING THE QFI

The QFI can be computed from Eqs. (6), (67), and (68). Combining Eqs. (6) and (67) we obtain all possible cross terms among $L^{(0)}$, $L^{(1)}$, and $L^{(2)}$. Classifying terms according to the number of $\tilde{R}$ operators they contain, we find the following. Of zero order only one term is obtained (E1a). Terms in first order contain products of $L^{(0)}$ and $L^{(1)}$. These terms vanish when the expectation is taken ($\langle \tilde{R} \rangle = 0$). Second-order terms contain cross products between $L^{(0)}$ and $L^{(2)}$ and "second powers" of $L^{(1)}$, lines (E1b), (E1c), and (E1d). Third-order terms contain expressions of the form $\mathrm{tr}[\rho \tilde{R}^i (\tilde{R}^j \circ \tilde{R}^k)]$, which vanish identically (we argue this claim at the end of the appendix). There is only one fourth-order term (E1e), which comes from double contribution of $L^{(2)}$.

$$\mathcal{J}_{\mu\nu}(X) = \tilde{\alpha}_{\mu,ij}\tilde{\alpha}_{\nu,kl}\big([L^{(0)}]_{i'j'}^{ij}\Omega^{i'j'}[L^{(0)}]_{k'l'}^{kl}\Omega^{k'l'} \tag{E1a}$$

$$+[L^{(1)}]_{i'}^{ij}[L^{(1)}]_{k'}^{kl}\mathrm{tr}[\rho\tilde{R}^{i'}\circ\tilde{R}^{k'}] \tag{E1b}$$

$$+[L^{(0)}]_{i'j'}^{ij}\Omega^{i'j'}[L^{(2)}]_{k'l'}^{kl}\Sigma^{k'l'} \tag{E1c}$$

$$+[L^{(2)}]_{i'j'}^{ij}\Sigma^{i'j'}[L^{(0)}]_{k'l'}^{kl}\Omega^{k'l'} \tag{E1d}$$

$$+[L^{(2)}]_{i'j'}^{ij}[L^{(2)}]_{k'l'}^{kl}\mathrm{tr}[\rho(R^{i'}\circ R^{j'})\circ(R^{k'}\circ R^{l'})]\big). \tag{E1e}$$

Using Eq. (D9) this is reduced to

$$\mathcal{J}_{\mu\nu}(X) = \tilde{\alpha}_{\mu,ij}\tilde{\alpha}_{\nu,kl}\big([L^{(1)}]_{i'}^{ij}[L^{(1)}]_{k'}^{kl}\Sigma^{i'k'}$$
$$+[L^{(2)}]_{i'j'}^{ij}\Sigma^{i'j'}[L^{(0)}]_{k'l'}^{kl}\Omega^{k'l'}+[L^{(2)}]_{i'j'}^{ij}[L^{(2)}]_{k'l'}^{kl}$$
$$\times\mathrm{tr}[\rho(R^{i'}\circ R^{j'})\circ(R^{k'}\circ R^{l'})]\big). \tag{E2}$$

At this point a manageable expression for $T^{ijkl} = \mathrm{tr}[\rho(R^i \circ R^j)\circ(R^k \circ R^l)]$ is an imperative. This tensor is symmetric under index permutations (12), (34) and (13), (24), where we have used standard cycle notation. These permutations generate a subgroup $\mathfrak{S}$ of the full symmetric group S_4. It is convenient to establish the relation between the tensor T^{ijkl} and the fully symmetric one

$$T_W^{ijkl} = \frac{1}{4!}\sum_{\sigma\in S_4}T^{\sigma(ijkl)}, \tag{E3}$$

which can be easily computed with standard phase space methods [50],

$$T_W^{ijkl} = \mathrm{tr}[\rho W(R^i R^j R^k R^l)], \tag{E4}$$

where $W(X)$ is the Weyl-ordered product of X. With Σ being the covariance matrix of ρ, T_W^{ijkl} is just the fourth moment of the Gaussian distribution with covariance matrix Σ,

namely,

$$T_W^{ijkl} = \Sigma^{ij}\Sigma^{kl} + \Sigma^{ik}\Sigma^{jl} + \Sigma^{il}\Sigma^{jk}. \tag{E5}$$

In order to exploit the preexisting symmetry of T we split the sum over $\sigma \in S_4$ over right cosets $S_4/\mathfrak{S}$, of which we pick $\{e,(23),(24)\}$ as representatives, namely, $\mathfrak{S} \cup (23)\mathfrak{S} \cup (24)\mathfrak{S} = S_4$. We can rewrite T_W as

$$T_W^{ijkl} = \frac{1}{4!}\sum_{a\in\{e,(23),(24)\}}\sum_{\sigma\in\mathfrak{S}}T^{a\sigma(ijkl)}. \tag{E6}$$

Since T is invariant under the action of $\mathfrak{S}$ we have

$$T_W^{ijkl} = \frac{|\mathfrak{S}|}{4!}\sum_{a\in\{e,(23),(24)\}}T^{a(ijkl)}$$
$$= \frac{1}{3}(T^{ijkl} + T^{ikjl} + T^{ilkj}). \tag{E7}$$

Therefore

$$T^{ijkl} = T_W^{ijkl} + \tfrac{1}{3}(T^{ijkl} - T^{ikjl}) + \tfrac{1}{3}(T^{ijkl} - T^{ilkj})$$
$$= T_W^{ijkl} + \tfrac{2}{3}\big(T_{[2,3]}^{ijkl} + T_{[2,4]}^{ijkl}\big), \tag{E8}$$

where $T_{[2,3]}^{ijkl}$ is the antisymmetrized tensor in the second and third indices,

$$T_{[2,3]}^{ijkl} = \tfrac{1}{2}(T^{ijkl} - T^{ikjl})$$
$$= \tfrac{1}{8}(\Omega^{ij}\Omega^{kl} - \Omega^{ik}\Omega^{jl}) - \tfrac{1}{4}\Omega^{il}\Omega^{jk}, \tag{E9}$$

and $T_{[2,4]}^{ijkl} = \tfrac{1}{3}(T^{ijkl} - T^{ilkj}) = \tfrac{1}{3}(T^{ijlk} - T^{iljk}) = T_{[2,3]}^{ijlk}$.
With this we obtain

$$T^{ijkl} = T_W^{ijkl} - \tfrac{1}{4}(\Omega^{ik}\Omega^{jl} + \Omega^{il}\Omega^{jk}). \tag{E10}$$

Thus, finally,

$$T^{ijkl} = \Sigma^{ij}\Sigma^{kl} + \Sigma^{ik}\Sigma^{jl} + \Sigma^{il}\Sigma^{jk} - \tfrac{1}{4}(\Omega^{ik}\Omega^{jl} + \Omega^{il}\Omega^{jk}). \tag{E11}$$

It will be convenient to write this result as

$$T^{ijkl} = \Sigma^{ij}\Sigma^{kl} + [D]_{i'j'}^{ij}(\Omega^{i'k}\Omega^{j'l} + \Omega^{i'l}\Omega^{j'k}), \tag{E12}$$

where D is defined in Eq. (D3).

A similar approach shows that $\tilde{R}^i \circ (\tilde{R}^j \circ \tilde{R}^k) = W(\tilde{R}^i\tilde{R}^j\tilde{R}^k)+$ terms containing only first powers of $\tilde{R}$. It is easy to see that expectation values of all these terms vanish identically.

Using the expression for T^{ijkl} we get

$$\mathcal{J}_{\mu\nu}(X) = \tilde{\alpha}_{\mu,i'j'}\tilde{\alpha}_{\nu,k'l'}\big([L^{(1)}]_i^{i'j'}[L^{(1)}]_k^{k'l'}\Sigma^{ik} + [L^{(2)}]_{i''j''}^{i'j'}$$
$$\times[L^{(2)}]_{kl}^{k'l'}[D]_{ij}^{i''j''}(\Omega^{ik}\Omega^{jl} + \Omega^{il}\Omega^{jk})\big), \tag{E13}$$

where, by virtue of Eq. (D9), the second term in Eq. (E2) has been canceled against the first contribution coming from Eq. (E12).

---

[1] G. M. Palma, K. Suominen, and A. K. Ekert, Proc. R. Soc. London, Ser. A **452**, 567 (1996); D. A. Lidar, I. L. Chuang, and K. B. Whaley, Phys. Rev. Lett. **81**, 2594 (1998).

[2] C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993); D. Bouwmeester, J. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, Nature (London) **390**, 575 (1997).

[3] M. Fleischhauer and M. D. Lukin, Phys. Rev. A **65**, 022314 (2002).

[4] J. I. Cirac and P. Zoller, Phys. Rev. A **50**, R2799 (1994).

[5] F. Dell'Anno, S. De Siena, and F. Illuminati, Phys. Rep. **428**, 53 (2006).

[6] D. Aharonov, A. Kitaev, and N. Nisan, in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing* (Assoc. Comput. Mach., Dallas, TX, 1998), pp. 20–30.

[7] C. M. Caves and K. Wodkiewicz, Open Syst. Inf. Dyn. **11**, 309 (2004).

[8] A. Gilchrist, N. K. Langford, and M. A. Nielsen, Phys. Rev. A **71**, 062310 (2005).

[9] J. Watrous, Theory Comput. **5**, 217 (2009).

[10] S. Amari and H. Nagaoka, *Methods of Information Geometry* (Am. Math, Soc., Providence, 2001).

[11] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, UK, 2006).

[12] D. Petz, *Quantum Information Theory and Quantum Statistics* (Springer, Berlin/New York, 2007).

[13] R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).

[14] D. J. C. Bures, Trans. Am. Math. Soc. **135**, 199 (1969).

[15] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, San Diego, 1976).

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

[17] M. Hübner, Phys. Lett. A **163**, 239 (1992).

[18] A. Uhlmann, Rep. Math. Phys. **33**, 253 (1993).

[19] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[20] I. Bengtsson, e-print arXive:quant-ph/0509017 (2005).

[21] P. B. Slater, J. Phys. A: Math. Gen. **29**, L601 (1996).

[22] H. Jeffreys, Proc. R. Soc. London, Ser. A **186**, 453 (1946).

[23] J. M. Bernardo and A. F. M. Smith, *Bayesian Theory* (Wiley, New York, 2000).

[24] A. Jamiolkowski, Rep. Math. Phys. **3**, 275 (1972); M. Choi, Lin. Alg. Appl. **10**, 285 (1975).

[25] A. Kitaev, Russ. Math. Surv. **52**, 1191 (1997).

[26] X. Wang, T. Hiroshima, A. Tomita, and M. Hayashi, Phys. Rep. **448**, 1 (2007).

[27] G. Adesso and F. Illuminati, J. Phys. A: Math. Theor. **40**, 7821 (2007).

[28] Consider, for example, the well-known problem of phase estimation [29–31], in which the phase sensitivity increases proportionally to the average energy of the probe state.

[29] C. M. Caves, Phys. Rev. D **23**, 1693 (1981).

[30] A. Monras, Phys. Rev. A **73**, 033821 (2006).

[31] B. C. Sanders and G. J. Milburn, Phys. Rev. Lett. **75**, 2944 (1995).

[32] G. Lindblad, J. Phys. A: Math. Gen. **33**, 5059 (2000).

[33] W. K. Wootters, Phys. Rev. D **23**, 357 (1981).

[34] R. D. Gill, IMS Lecture notes - Monograph Series **36**, 255 (2001).

[35] M. G. A. Paris, Int. J. Quantum. Inf. **7**, 125 (2009).

[36] G. Fubini, Atti del Regio Istituto Veneto **63**, 502 (1904); E. Study, Math. Ann. **60**, 321 (1905); A. Fujiwara and H. Nagaoka, Phys. Lett. A **201**, 119 (1995); M. Hayashi, J. Phys. A **31**, 4633 (1998).

[37] A. Monras and F. Illuminati (2010) (unpublished).

[38] G. Adesso, F. Dell'Anno, S. De Siena, F. Illuminati, and L. A. M. Souza, Phys. Rev. A **79**, 040305(R) (2009).

[39] A. S. Holevo and R. F. Werner, Phys. Rev. A **63**, 032312 (2001).

[40] G. Giedke, J. Eisert, J. I. Cirac, and M. B. Plenio, Quant. Inf. Comp. **3**, 211 (2003).

[41] A. Botero and B. Reznik, Phys. Rev. A **67**, 052311 (2003).

[42] A. Serafini, G. Adesso, and F. Illuminati, Phys. Rev. A **71**, 032349 (2005).

[43] A. Fujiwara, Phys. Rev. A **65**, 012316 (2001); M. A. Ballester, *ibid.* **69**, 022303 (2004); A. Fujiwara and H. Imai, J. Phys. A **36**, 8093 (2003).

[44] D. Petz and C. Sudár, J. Math. Phys. **37**, 2662 (1996).

[45] A. Fujiwara, Phys. Rev. A **63**, 042304 (2001).

[46] R. Reemtsen and J. J. Rückmann, *Semi Infinite Programming* (Springer, Berlin/New York, 1998).

[47] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, UK, 2004).

[48] R. Bhatia, *Matrix Analysis* (Springer, Berlin/New York, 1996).

[49] M. O. Scully and M. S. Zubairy, *Quantum Optics* (Cambridge University Press, Cambridge, UK, 1997).

[50] S. M. Barnett and P. M. Radmore, *Methods in Theoretical Quantum Optics* (Oxford University Press USA, New York, 2003).

[51] R. R. Puri, *Mathematical Methods of Quantum Optics* (Springer, Berlin/New York, 2001).

[52] J. Williamson, Am. J. Math. **58**, 141 (1936).

[53] A. Serafini, M. G. A. Paris, F. Illuminati, and S. De Siena, J. Opt. B: Quantum Semiclass. Opt. **7**, R19 (2005).

[54] E. Bagan, M. Baig, R. Muñoz-Tapia, and A. Rodriguez, Phys. Rev. A **69**, 010304(R) (2004).

[55] A. Monras and M. G. A. Paris, Phys. Rev. Lett. **98**, 160401 (2007).

[56] A. Serafini, J. Eisert, and M. M. Wolf, Phys. Rev. A **71**, 012320 (2005).

[57] J. Munkres, *Topology* (Prentice Hall, New York, 2000).

[58] R. M. Wilcox, J. Math. Phys. **8**, 962 (1967).