Aberystwyth University

Information geometry and local asymptotic normality for multi-parameter estimation of quantum Markov dynamics

Guta, Madalin ; Kiukas, Jukka

Published in:
Journal of Mathematical Physics

DOI:
10.1063/1.4982958

Publication date:
2017

Citation for published version (APA):
Guta, M., & Kiukas, J. (2017). Information geometry and local asymptotic normality for multi-parameter estimation of quantum Markov dynamics. *Journal of Mathematical Physics*, *58*(5), Article 052201. https://doi.org/10.1063/1.4982958

General rights
Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400
email: is@aber.ac.uk

Download date: 27. Apr. 2024

Information geometry and local asymptotic normality for multi-parameter estimation of quantum Markov dynamics

Madalin Guta and Jukka Kiukas

Citation: Journal of Mathematical Physics **58**, 052201 (2017); doi: 10.1063/1.4982958
View online: http://dx.doi.org/10.1063/1.4982958
View Table of Contents: http://aip.scitation.org/toc/jmp/58/5
Published by the American Institute of Physics

Articles you may be interested in

New concise upper bounds on quantum violation of general multipartite Bell inequalities
Journal of Mathematical Physics **58**, 052202 (2017); 10.1063/1.4982961

On invariant analysis of space-time fractional nonlinear systems of partial differential equations. II
Journal of Mathematical Physics **58**, 051503 (2017); 10.1063/1.4982804

The least energy sign-changing solution for a nonlocal problem
Journal of Mathematical Physics **58**, 051505 (2017); 10.1063/1.4982960

Borel functional calculus for quaternionic normal operators
Journal of Mathematical Physics **58**, 053501 (2017); 10.1063/1.4982047

Orbital stability of periodic traveling wave solutions for the Kawahara equation
Journal of Mathematical Physics **58**, 051504 (2017); 10.1063/1.4980016

The non-autonomous YdKN equation and generalized symmetries of Boll equations
Journal of Mathematical Physics **58**, 053507 (2017); 10.1063/1.4982747

JOURNAL OF MATHEMATICAL PHYSICS **58**, 052201 (2017)

Information geometry and local asymptotic normality for multi-parameter estimation of quantum Markov dynamics

Madalin Guta[1,a)] and Jukka Kiukas[2,b)]
[1]*School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom*
[2]*Department of Mathematics, Aberystwyth University, Penglais, Aberystwyth, Ceredigion SY23 3BZ, United Kingdom*

(Received 1 August 2016; accepted 22 April 2017; published online 11 May 2017)

This paper deals with the problem of identifying and estimating dynamical parameters of continuous-time Markovian quantum open systems, in the input-output formalism. First, we characterise the space of identifiable parameters for ergodic dynamics, assuming full access to the output state for arbitrarily long times, and show that the equivalence classes of undistinguishable parameters are orbits of a Lie group acting on the space of dynamical parameters. Second, we define an information geometric structure on this space, including a principal bundle given by the action of the group, as well as a compatible connection, and a Riemannian metric based on the quantum Fisher information of the output. We compute the metric explicitly in terms of the Markov covariance of certain "fluctuation operators" and relate it to the horizontal bundle of the connection. Third, we show that the system-output and reduced output state satisfy local asymptotic normality, i.e., they can be approximated by a Gaussian model consisting of coherent states of a multimode continuous variables system constructed from the Markov covariance "data." We illustrate the result by working out the details of the information geometry of a physically relevant two-level system. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4982958]

I. INTRODUCTION

The input-output formalism[12,23] is fundamental to key areas of quantum open system theory such as Markov dynamics, continuous-time measurements and filtering theory,[7,10] quantum networks,[25] and feedback control.[8,39] The formalism serves as a platform, which integrates in a common language the methods from control engineering, classical and quantum stochastic processes, non-equilibrium statistical mechanics, and quantum information. In this paper, we aim to further expand this platform by adopting a system identification[42] perspective. Concretely, we investigate which dynamical parameters of an open system can be estimated from the output state (identifiability problem), how the associated quantum Fisher information (QFI) arises from the structure of the parameter manifold (information geometry), and how the multi-parameter statistical model defined by the output state can be approximated by a quantum Gaussian model (local asymptotic normality (LAN)).

In a typical quantum input-output setup, an open system (e.g., an atom or a cavity mode) is driven by an input consisting of the vacuum or coherent state of the electromagnetic field, the latter being modelled by continuum of Bosonic modes representing the incoming "quantum noise," see Figure 1. The input interacts with the system in a Markovian fashion, with joint unitary evolution $U_{\mathsf{D}}(t)$ determined by the "dynamical parameters" $\mathsf{D} := (H, L^1, \ldots, L^k)$, where H is the system Hamiltonian and L^i is the coupling operator to the ith input mode.

a)Electronic mail: madalin.guta@nottingham.ac.uk
b)Electronic mail: jek20@aber.ac.uk

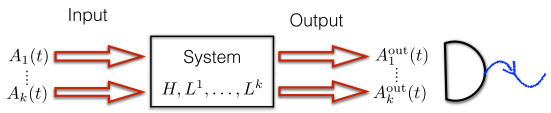

FIG. 1. Continuous-time Markovian dynamics of an open quantum system in the input-output formalism. Input fields $A_i(t)$ interact with the system, so that the joint unitary transformation $U_{\mathsf{D}}(t)$ depends on the dynamical parameter $\mathsf{D} := (H, L^1, \ldots, L^k)$, where H is the system Hamiltonian and L^i are the coupling (jump) operators with the input fields. The output state carry information about D, which can be estimated by measuring the output fields.

The output fields carry information about the dynamical parameter D and can be monitored by means of continuous-time measurements or may be "post-processed," e.g., by using feed-forward or feedback schemes.[53] However, since such schemes often rely on the knowledge of the dynamics, it is important to develop efficient methods for estimating the unknown parameters entering the dynamics. Our goal here is not to propose or analyse specific measurement and estimation schemes (see, e.g., Refs. 14, 20, 21, and 43 for related results), but rather investigate the statistical properties of the output state, which will provide the ultimate limits in estimation precision. We envisage that the structure of the output state uncovered here will be relevant not only for designing efficient measurement schemes (cf. Ref. 28 for optimal estimation of qubit states) but also for applications in quantum metrology[44] and quantum control, including feedback.

In our analysis, we assume that the system is finite dimensional, and the input is stationary (time independent). We also assume that the dynamics is ergodic, i.e., the system has a unique strictly positive stationary state ρ_{ss}^{D}, in which case any initial state converges to ρ_{ss}^{D} and the output becomes stationary in time. From a quantum information perspective, the system-output state $|\Psi_{\mathsf{D}}^{\rm s+o}(t)\rangle$ associated with the time interval $[0, t]$ is a continuous matrix product state,[52] and the output state $\rho_{\mathsf{D}}^{\rm out}(t)$ is a continuous version of a purely generated finitely correlated state.[18] Our results are therefore relevant for the problem of estimating such states, whose discrete version was considered in Ref. 6 from the perspective of quantum tomography of spin chains.

Since we deal with a multi-parameter statistical problem, we adopt a differential geometry viewpoint in the spirit of the theory of information geometry.[2] This allows us to characterise the manifold of identifiable parameters as the quotient of the parameter space with respect to a group of transformations leaving the output state invariant (see Theorem 1), thus extending our previous results for discrete time quantum Markov chains.[31] An analogous differential geometric construction has been presented in Refs. 33 and 34 for parametrisations of discrete matrix product states, and a related approach has been used in studying the manifold of correlation matrices for stationary states of certain specific open quantum systems.[4]

Furthermore, we show that the quantum Fisher information (QFI)[11,36] of the output is closely related to the covariance of certain "fluctuation operators," which we study in detail in Section V. The covariance defines a Riemannian metric on the space of identifiable parameters and provides a complex structure and a positive inner product on the tangent space of identifiable parameters. An alternative approach to computing the quantum Fisher information is described in Ref. 22, see also Refs. 44 and 16.

With the help of this differential geometric structure, we construct an associated algebra of canonical commutation relations (CCR) and a family of coherent states whose QFI is equal to the QFI per time unit of the output state. The latter will play the role of limit Gaussian model below.

Local asymptotic normality (LAN) is a key concept in asymptotic statistics which describes how certain statistical models can be approximated by simpler Gaussian models, with a vanishing error in the limit of large "sample size." This phenomenon occurs, for instance, in the case of models consisting of independent, identically distributed samples,[41] and also for multiple observations from an ergodic Markov process,[37] or hidden Markov process.[9] In quantum statistics, the general theory of convergence of models was discussed in Refs. 24 and 29, and LAN for ensembles of independent finite dimensional systems was established in Ref. 40. For quantum Markov dynamics, LAN for one-dimensional parameter models was discussed in Refs. 27 and 31 for discrete time, and in Ref. 13 for continuous-time. Here we extend the latter to the multi-dimensional model where all

identifiable parameters are assumed to be unknown; this brings forward the information-geometric aspects, which do not play a significant role in a one-parameter setting. Theorem 2 shows that the system-output state and (reduced) output state models converge to the Gaussian model consisting of a family of coherent states of the above mentioned CCR algebra, in the limit of large times.

The present investigation suggests several interesting future lines of research. One direction is to understand the physical significance of the geodesic distance of the Fisher metric and the relation to quantum speed limit[51] and thermodynamic metrics.[50] Another direction is to show that fluctuation operators satisfy the central limit theorem and identify the measurement which achieves the optimal estimation precision. Building on,[32] one can develop a similar theory for the identification of quantum linear input-output systems in the stationary regime, i.e., from the "power spectrum." Moreover, the extension of the current theory to non-ergodic dynamics and the analysis of "metastable"[45] or "near phase transition"[44] systems is important due to its relevance for quantum metrology. Finally, our framework has a number of interesting generalisations connected with other ongoing mathematical work on quantum stochastic evolutions. In particular, when the stationary state manifold is nontrivial (non-ergodic case), one can discuss conserved quantities and adiabatic transport.[1,3,26] From the more technical point of view, our manifold of dynamical parameters actually has a natural Lie group structure;[17] reformulation of our results in this more structured framework could be useful especially for applications to control theory.

In order to increase the accessibility of the paper, we collect the main constructions and results in Sec. II.

II. OVERVIEW OF RESULTS

Section III introduces the input-output formalism of quantum open dynamics, as illustrated in Figure 1. For a given dynamical parameter $\mathsf{D} := (H, L^1, \ldots, L^k)$, the system-output state is given by $|\Psi^{\mathrm{s+o}}_{\mathsf{D}}(t)\rangle = U_{\mathsf{D}}(t)|\varphi \otimes \Omega\rangle$, where $|\varphi\rangle$ is the initial system state, $|\Omega\rangle$ is the input state (taken to be the vacuum), and $U_{\mathsf{D}}(t)$ is the joint unitary evolution given by the quantum stochastic differential equation

$$dU_{\mathsf{D}}(t) = \left(\sum_{i=1}^{k}(iH \otimes \mathbb{1}_{\mathcal{F}}\,dt + L^i \otimes dA_i^*(t) - L^{i*} \otimes dA_i(t)) - \frac{1}{2}\sum_{i=1}^{k} L^{i*}L^i \otimes \mathbb{1}_{\mathcal{F}}\,dt\right) U_{\mathsf{D}}(t),$$

where $dA_i(t)$ and $dA_i^*(t)$ are the time increments of input annihilation and creation operators of k Bosonic input channels, acting on the Fock space $\mathcal{F}$ over $L^2(\mathbb{R}_+) \otimes \mathbb{C}^k$. The reduced system evolution is governed by an ergodic Markov semigroup with Lindblad generator $\mathbb{W}_{\mathsf{D}}$ and unique stationary state ρ^{D}_{ss}. The output state after time t is obtained by tracing out the system, $\rho^{\mathrm{out}}_{\mathsf{D}}(t) = \mathrm{tr}_s(|\Psi^{\mathrm{s+o}}_{\mathsf{D}}(t)\rangle\langle\Psi^{\mathrm{s+o}}_{\mathsf{D}}(t)|)$. For long times, the system converges to the stationary state, and the output becomes stationary in time.

Section IV discusses the identifiability problem in the stationary setting, see Figure 2. We define an equivalence relation between dynamical parameters for which the *stationary output states* are

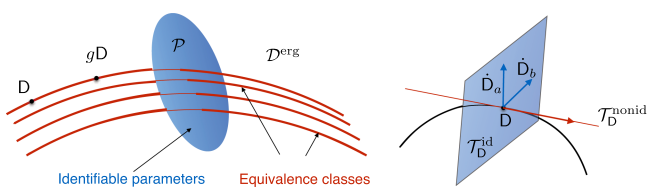

FIG. 2. Left panel: the space of ergodic dynamical parameters $\mathcal{D}^{\mathrm{erg}}$ as principle G-bundle over the base manifold $\mathcal{P}$ of identifiable parameters. Equivalence classes (red lines) of dynamical parameters with identical outputs. Right panel: the tangent space at the point D decomposes as direct sum of the tangent space $\mathcal{T}^{\mathrm{nonid}}_{\mathsf{D}}$ to the orbit of the group action and the space $\mathcal{T}^{\mathrm{id}}_{\mathsf{D}}$ of "identifiable directions" defined by the identity $\mathcal{E}^0_{\mathsf{D}}(\dot{\mathsf{D}}) = 0$. The Markov covariance defines a complex structure and an inner product on $\mathcal{T}^{\mathrm{id}}_{\mathsf{D}}$, such that the QFI rate is $f_{a,b} = 4\mathrm{Re}(\dot{\mathsf{D}}_a, \dot{\mathsf{D}}_a)_{\mathsf{D}}$.

identical for all times t. In Theorem 1 we show that two dynamical parameters D and D' are equivalent *if and only if* they are related by the "gauge transformation" $H' = W^*HW + r\mathbb{1}$ and $L_i' = W^*L_iW$, where W is unitary and r is a real constant. From a differential geometry viewpoint, the space of *identifiable parameters* is the quotient $\mathcal{P} := \mathcal{D}^{erg}/G$, where $\mathcal{D}^{erg}$ is the manifold of dynamical parameters D with ergodic dynamics, and $G = PU(d) \times \mathbb{R}$ is the group of "gauge transformations" whose orbits are the equivalence classes of parameters. In particular, we show that $\mathcal{D}^{erg}$ is a principal G-bundle over the base manifold $\mathcal{P}$. The vertical bundle over $\mathcal{D}^{erg}$ consists of subspaces $\mathcal{T}_{\mathsf{D}}^{\text{nonid}}$ of the tangent space $\mathcal{T}_{\mathsf{D}}$ at D, corresponding to un-identifiable changes of parameters, i.e., infinitesimal changes induced by the action of the group G. Although in general there is no canonical decomposition of the tangent space into "identifiable" and "non-indentifiable" components (i.e., $\mathcal{T}_{\mathsf{D}} = \mathcal{T}_{\mathsf{D}}^{\text{nonid}} \oplus \mathcal{T}_{\mathsf{D}}^{\text{id}}$), such a decomposition can be obtained from a principal connection, in a covariant way. A natural choice of connection is provided by the information geometry, as discussed below. This approach to system identification often appears in the classical setting, and the advantage is that one gains insight into the geometric structure of the parameter manifold, beyond the direct computation of the Fisher information. For instance, the connection can be useful for developing recursive estimation algorithms based on geodesics of the manifold.[35] For the standard theory of connections on principal bundles, see, e.g., Ref. 38.

In Section V, we derive the information geometric structure of the statistical estimation problem at hand. Before discussing the statistical aspects, we describe the basic elements of a theory of "output fluctuations" which is essential for information geometry, but has an interest in its own and deserves to be further investigated. For each $(k+1)$-tuple of system operators $\mathbf{X} = (X^0, X^1, \ldots, X^k) \in M(\mathbb{C}^d)^{k+1}$, we define the associated *fluctuation operator* $\mathbb{F}_{\mathsf{D},t}(\mathbf{X})$ given by the quantum stochastic integral

$$\mathbb{F}_{\mathsf{D},t}(\mathbf{X}) = \frac{1}{\sqrt{t}} \int_0^t \left(i \sum_{i=1}^k j_{\mathsf{D},s}(X^i) dA_i^*(s) + j_{\mathsf{D},s} \circ \mathcal{C}_{\mathsf{D}}(X^0) ds \right), \qquad \mathcal{C}_{\mathsf{D}}(X) := X - \operatorname{tr}[\rho_{ss}^{\mathsf{D}} X]\mathbb{1},$$

where $j_{\mathsf{D},s}(X) := U_{\mathsf{D}}^*(s) X U_{\mathsf{D}}(s)$ is the time-evolved operator X. The covariance of $\mathbb{F}_{\mathsf{D},t}(\mathbf{X})$ converges in the limit of large times, and defines a positive (but degenerate) inner product on $M(\mathbb{C}^d)^{k+1}$ (cf. Proposition 1 for the explicit formula),

$$(\mathbf{X}, \mathbf{Y})_{\mathsf{D}} = \lim_{t\to\infty} \langle \mathbb{F}_{\mathsf{D},t}(\mathbf{X})^* \mathbb{F}_{\mathsf{D},t}(\mathbf{Y}) \rangle.$$

Furthermore, in Propositions 1 and 2, we construct a linear map $R_{\mathsf{D}} : M(\mathbb{C}^d)^{k+1} \to M(\mathbb{C}^d)^{k+1}$ such that R_{D} is a projection onto the subspace of operators of the form $(0, Y^1, \ldots, Y^k) \in M(\mathbb{C}^d)^{k+1}$, and the kernel of R_{D} is the subspace of degenerate vectors of the inner product. With this definition, the inner product takes the following simple form:

$$(\mathbf{X}, \mathbf{Y})_{\mathsf{D}} = \sum_{i=1}^k \operatorname{tr}\left[\rho_{ss}^{\mathsf{D}} R_{\mathsf{D}}(\mathbf{X})^{i*} R_{\mathsf{D}}(\mathbf{Y})^i \right].$$

We denote by $\dot{\mathsf{D}} = (\dot{H}, \dot{L}^1, \ldots, \dot{L}^k)$ an element of the tangent space $\mathcal{T}_{\mathsf{D}}$. The *real* linear map $\mathbf{X}_{\mathsf{D}}$ defined below plays an important role in connecting fluctuation operators with the information geometry,

$$\begin{aligned} \mathbf{X}_{\mathsf{D}} : \mathcal{T}_{\mathsf{D}} &\to M(\mathbb{C}^d)^{k+1} \\ \dot{\mathsf{D}} &\mapsto (\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}}), \dot{L}^1, \ldots, \dot{L}^k), \end{aligned}$$

where $\mathcal{E}_{\mathsf{D}}$ is the map,

$$\begin{aligned} \mathcal{E}_{\mathsf{D}} : \mathcal{T}_{\mathsf{D}} &\to M(\mathbb{C}^d) \\ \mathcal{E}_{\mathsf{D}} : \dot{\mathsf{D}} &\mapsto \dot{H} + \operatorname{Im} \sum_{i=1}^k \dot{L}^{i*} L^i. \end{aligned}$$

Here the second term in $\mathcal{E}_{\mathsf{D}}$ is due to quantum Ito calculus; hence, in some sense it represents the effects of the stochastic output on the information geometry. Using the map $\mathbf{X}_{\mathsf{D}}$, we define a *real* inner product on the tangent space $\mathcal{T}_{\mathsf{D}}$,

$$(\dot{\mathsf{D}}, \dot{\mathsf{D}}') \longmapsto \operatorname{Re}(\mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}}), \mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}}'))_{\mathsf{D}}.$$

Moreover, since $\mathbf{X}_{\mathsf{D}}$ is injective, we can use it to define a projection $P_{\mathsf{D}} = \mathbf{X}_{\mathsf{D}}^{-1} \circ R_{\mathsf{D}} \circ \mathbf{X}_{\mathsf{D}}$ acting on $\mathcal{T}_{\mathsf{D}}$. Its kernel is the vertical space $\mathcal{T}_{\mathsf{D}}^{\text{nonid}}$ whose vectors correspond to infinitesimal "gauge transformations," and are the degenerate vectors of the inner product. The range of P_{D} consists of tangent vectors satisfying the condition $\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}}) = 0$. In particular, since P_{D} is a projection, the tangent space can be decomposed into "identifiable" and "non-identifiable" directions (see right panel of Figure 2),

$$\mathcal{T}_{\mathsf{D}} = \text{ran}P_{\mathsf{D}} \oplus \ker P_{\mathsf{D}} = \mathcal{T}_{\mathsf{D}}^{\text{id}} \oplus \mathcal{T}_{\mathsf{D}}^{\text{nonid}},$$

whose statistical interpretation is discussed below. This split has also an interesting differential geometric interpretation: the above tangent space decomposition and the inner product are covariant with respect to the action of the group G and define a connection on the resulting principal G-bundle, with the associated Lie algebra valued one-form

$$\omega_{\mathsf{D}} : \mathcal{T}_{\mathsf{D}} \to \mathfrak{g}, \quad \omega_{\mathsf{D}}(\dot{\mathsf{D}}) = (-i\mathbb{W}_{\mathsf{D}}^{-1} \circ \mathcal{C}_{\mathsf{D}} \circ \mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}}), \text{tr}[\rho_{ss}^{\mathsf{D}} \mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}})]),$$

explicitly depending on the map $\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}})$ containing the essential quantum Ito correction. Moreover, the strictly positive inner product on $\mathcal{T}_{\mathsf{D}}^{\text{id}}$ induces a strictly positive inner product on the tangent space $\mathcal{T}_{[\mathsf{D}]}$ to the point $[\mathsf{D}]$ in the base space $\mathcal{P} = \mathcal{D}^{erg}/G$ of identifiable parameters. As we will see below, this *Riemannian metric* is closely connected to the quantum Fisher information rate of the output state, so we will refer to it as the *information geometry* of the open quantum system, in analogy to the classical case.[2]

Let us consider now the problem of estimating the dynamical parameter D. Although the key constructions could be introduced in a "coordinate free" way, in order to emphasise the statistical aspects we choose to work with a given (but arbitrary) parametrisation $\theta \mapsto \mathsf{D}_\theta$ of $\mathcal{D}^{erg}$, where θ is an unknown parameter belonging to an open subset of $\mathbb{R}^m$, with $m := \dim(\mathcal{D}^{erg})$. At a given point $\mathsf{D} = \mathsf{D}_\theta \in \mathcal{D}^{erg}$, we define the tangent vectors $\dot{\mathsf{D}}_a := \partial \mathsf{D}/\partial\theta_a = (\dot{H}_a, \dot{L}_a^1, \ldots, \dot{L}_a^k)$ describing infinitesimal changes of the coordinate θ_a, for $a = 1, \ldots, m$; these vectors form a basis of the tangent space $\mathcal{T}_{\mathsf{D}}$.

We consider now the $m \times m$ *quantum Fisher information* (QFI) matrix $F^\theta(t)$ associated with the system-output state $|\Psi_{\mathsf{D}_\theta}^{\text{s+o}}(t)\rangle$. The QFI is proportional to the real part of the covariance matrix of (centred) "generators" $G_{\theta,a}^0(t)$ of infinitesimal changes with respect to parameter component θ_a.[11] We show that the generator $G_{\theta,a}^0(t)$ (normalised by $t^{-1/2}$) can be expressed as a *fluctuation operator* $\mathbb{F}_t(\mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}}_a))$, using the map $\mathbf{X}_{\mathsf{D}}$ defined above. As consequence, the QFI grows linearly in time, and the *QFI rate* per time unit $f^\theta = \lim_{t\to\infty} F^\theta(t)/t$ can be expressed in terms of the Markov covariance as

$$\begin{aligned} f_{a,b}^\theta &= 4\text{Re}(\mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}}_a), \mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}}_b))_{\mathsf{D}} = 4\text{Re}(R_{\mathsf{D}}\mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}}_a), R_{\mathsf{D}}\mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}}_b))_{\mathsf{D}} \\ &= 4\sum_{i=1}^k \text{Re}\,\text{tr}\left[\rho_{ss}^{\mathsf{D}}\left(\dot{L}_a^i - i[L^i, \mathbb{W}_{\mathsf{D}}^{-1} \circ \mathcal{E}_{\mathsf{D}}^0(\dot{\mathsf{D}}_a)]\right)^* \left(\dot{L}_b^i - i[L^i, \mathbb{W}_{\mathsf{D}}^{-1} \circ \mathcal{E}_{\mathsf{D}}^0(\dot{\mathsf{D}}_b)]\right)\right], \end{aligned}$$

where $\mathbb{W}_{\mathsf{D}}$ is the Lindblad operator at D and $\mathcal{E}_{\mathsf{D}}^0 = \mathcal{C}_{\mathsf{D}} \circ \mathcal{E}_{\mathsf{D}}$. In particular, the Fisher information rate associated with directions in vertical bundle $\mathcal{T}_{\mathsf{D}}^{\text{nonid}}$ (gauge transformations) is equal to zero as expected from the invariance of the output state. This follows from the fact that $\mathbf{X}_{\mathsf{D}}$ maps $\mathcal{T}_{\mathsf{D}}^{\text{nonid}}$ into $\ker R_D$.

Above we saw that the *real part* of the Markov covariance $(\cdot,\cdot)_{\mathsf{D}}$ defines a positive definite inner product on the *real* space $\mathcal{T}_{\mathsf{D}}^{\text{id}}$. In fact, $\mathcal{T}_{\mathsf{D}}^{\text{id}}$ can be made into a complex space by introducing the *complex structure*,

$$\begin{aligned} \mathcal{J}_{\mathsf{D}} &: \mathcal{T}_{\mathsf{D}}^{id} \to \mathcal{T}_{\mathsf{D}}^{id} \\ \mathcal{J}_{\mathsf{D}} &: (\dot{H}, \dot{L}^1, \ldots, \dot{L}^k) \mapsto \left(\sum_{i=1}^k \text{Re}\dot{L}^{i*}L^i,\ i\dot{L}^1, \ldots, i\dot{L}^k\right). \end{aligned} \tag{1}$$

With this definition, the map $\mathbf{X}_{\mathsf{D}}$ becomes an isomorphism of complex spaces and $(\cdot,\cdot)_{\mathsf{D}}$ defines a complex inner product on $(\mathcal{T}_{\mathsf{D}}^{\text{id}}, \mathcal{J}_{\mathsf{D}})$. Using the imaginary part σ_D of the inner product, we define

the canonical commutation relations (CCR) algebra $CCR(\mathcal{T}^{\mathrm{id}}_{\mathsf{D}}, \sigma^{\mathsf{D}})$ generated by Weyl operators with commutation relations,

$$W(\dot{\mathsf{D}})W(\dot{\mathsf{D}}') = e^{i\sigma^{\mathsf{D}}(\dot{\mathsf{D}},\dot{\mathsf{D}}')}W(\dot{\mathsf{D}}+\dot{\mathsf{D}}'), \quad W(-\dot{\mathsf{D}}) = W(\dot{\mathsf{D}})^*, \quad \dot{\mathsf{D}}, \dot{\mathsf{D}}' \in \mathcal{T}^{\mathrm{id}}_{\mathsf{D}}.$$

Following a standard construction, we define the Fock representation and the coherent states $|\dot{\mathsf{D}}\rangle := W(\dot{\mathsf{D}})|0\rangle$, where $|0\rangle$ is the vacuum state $\langle 0|W(\dot{\mathsf{D}})|0\rangle = \exp(-(\dot{\mathsf{D}}, \dot{\mathsf{D}})_{\mathsf{D}}/2)$ This model will be interpreted below as limit of the output state model for large times.

Section VI details the above constructions in the case of special one-dimensional models, and for a general multidimensional model for a two-dimensional system.

In Section VII we study the asymptotic statistical structure of the output state. The main result is the *local asymptotic normality* (LAN) Theorem 2 which shows that both the system-output state and the stationary output state can be approximated by coherent states of the CCR algebra $CCR(\mathcal{T}^{\mathrm{id}}_{\mathsf{D}}, \sigma_{\mathsf{D}})$. Below we give a brief description of the result and its interpretation.

Let us consider a parametrisation $u \mapsto [\mathsf{D}]_u$ of (an open subset of) the space of identifiable parameters $\mathcal{P} = \mathcal{D}^{erg}/G$, such that the origin $u = 0$ corresponds to a given parameter of interest $[\mathsf{D}_0] = [\mathsf{D}]_{u=0}$. We define the time-indexed family of "local" statistical models,

$$\tilde{\mathcal{Q}}_t := \left\{ \rho^{\mathrm{out}}_{u/\sqrt{t}}(t) : u \in \mathcal{O} \subset \mathbb{R}^{\dim(\mathcal{P})} \right\}$$

which consist of the output state for unknown parameter values $u/\sqrt{t}$ in a shrinking ball of size scaling as the statistical uncertainty. The local model $\tilde{\mathcal{Q}}_t$ captures the asymptotic properties of the quantum output state for parameters in a neighbourhood of $[\mathsf{D}_0]$, and it can be justified operationally by means of adaptive procedures whereby a "small part" of the output can be used to localise the parameter, while the "remaining part" can be used for estimating the local parameter u (cf. Ref. 30 for a similar argument in the state estimation setup).

To define the system-output statistical model, let us consider a *horizontal section* $s : \mathcal{P} \to \mathcal{D}^{erg}$ of the principal bundle, i.e., the tangent space to $s(\mathcal{P})$ at D is the horizontal space $\mathcal{T}^{id}_{\mathsf{D}}$. Let

$$\mathcal{Q}_t := \left\{ \left|\Psi^{\mathrm{s+o}}_{u/\sqrt{t}}(t)\right\rangle : u \in \mathcal{O} \subset \mathbb{R}^{\dim(\mathcal{P})} \right\},$$

be the quantum statistical model where $\left|\Psi^{\mathrm{s+o}}_{u/\sqrt{t}}(t)\right\rangle$ is the joint system-output state at time t at the dynamical parameter $\mathsf{D}_u = s([D]_u)$. The reasons for using a horizontal section in defining the model are as follows. While the stationary output state depends only on the identifiable parameters in $\mathcal{P}$, the system-output state is also sensitive to the location of the parameter within an orbit. It turns out that the asymptotic properties can be captured most transparently by choosing a horizontal section which sets certain unphysical phase factors to zero and allows us to understand the model directly in terms of the geometric properties of the vector state, as explained below.

Finally, we define the Gaussian model

$$\mathcal{G} := \left\{ \rho_u := |u\rangle\langle u| : u \in \mathcal{O} \subset \mathbb{R}^{\dim(\mathcal{P})} \right\},$$

where $|u\rangle = W(\sum_a u_a \dot{\mathsf{D}}_a)|0\rangle$ is the coherent state of the CCR algebra $CCR(\mathcal{T}^{id}_{\mathsf{D}_0}, \sigma^{\mathsf{D}_0})$. By construction, the QFI of this model is equal to the QFI rate f of the output state at $[\mathsf{D}_0]$, which also is the QFI of the output state with respect to the *local* parameter u, rather than the "true" parameter $u/\sqrt{t}$.

The first version of LAN states that $\mathcal{Q}_t$ converges (weakly) to $\mathcal{G}$ for large times in the sense of convergence of the inner products (uniformly in $u, v \in \mathcal{O}$), as illustrated in Figure 3,

$$\lim_{n\to\infty} \left\langle \Psi^{\mathrm{s+o}}_{u/\sqrt{t}}(t) \middle| \Psi^{\mathrm{s+o}}_{v/\sqrt{t}}(t) \right\rangle = \langle u|v\rangle.$$

Since pure state models are fully characterised by the inner products, the convergence simply means that for large times the geometry of the system-output states is very similar to that of the coherent states. Although intuitive, this notion of convergence is not suitable for mixed states such as that of the output and does not have a direct operational meaning. In the second version of LAN, we show

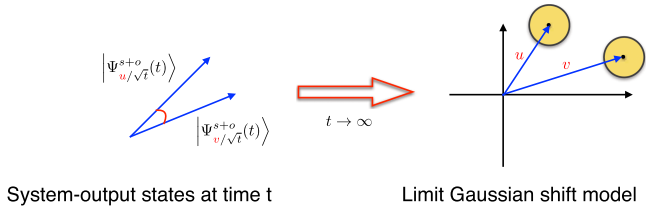

FIG. 3. Local asymptotic normality as weak convergence to the Gaussian limit. The inner products of system-output states with local parameter u, v converge (uniformly) to the inner products of the corresponding coherent state $|u\rangle$ and $|v\rangle$, in the limit of large times.

that the output models $\tilde{\mathcal{Q}}_t$ converge strongly to $\mathcal{G}$ in the sense that there exist channels T_t and S_t such that

$$\lim_{t\to\infty} \sup_{u\in\mathcal{O}} \left\| T_t\left(\rho^{\text{out}}_{u/\sqrt{t}}(t)\right) - \rho_u \right\|_1 = 0,$$
$$\lim_{t\to\infty} \sup_{u\in\mathcal{O}} \left\| S_t\left(\rho_u\right) - \rho^{\text{out}}_{u/\sqrt{t}}(t) \right\|_1 = 0.$$

A concrete consequence of this "convergence to Gaussianity" is that the QFI computed above is asymptotically "achievable" in the sense that the estimation of dynamical parameters reduces to that of estimating a Gaussian displacement family with QFI equal to $f_{a,b}$. Similarly to LAN for ensembles of identical states,[30] the result implies that the optimal measurement is a linear one (i.e., of homodyne and heterodyne type) and the errors are normally distributed. However, since this paper concentrates on the structure of the quantum states, the measurement and estimation procedures are not discussed here.

III. PRELIMINARIES ON QUANTUM MARKOV PROCESSES

We begin by introducing notations and necessary background about the input-output formalism of continuous-time quantum Markov processes.[23] The formalism describes the *joint unitary evolution* of an open quantum system interacting with a Bosonic environment in the Markov regime, cf. Figure 1. From this, one can derive the reduced (master) dynamics of the system, as well as the stochastic Schrödinger equations for quantum trajectories, describing the stochastic evolution of the system conditional on observations produced by a continuous-time measurement on the environment. However, in this paper we will be mainly interested in the *output quantum state*, i.e., the state of the environment after the interaction with the system.

Throughout the paper, we assume the system to be finite-dimensional, with Hilbert space $\mathcal{H} = \mathbb{C}^d$ and associated algebra of observables $\mathcal{A} = M(\mathbb{C}^d)$. As we will detail below, the dynamics is specified by the *system Hamiltonian* H, together with the *quantum jump operators* $L^1, \ldots, L^k$. We denote these collectively by

$$\mathsf{D} = (H, L^1, \ldots, L^k) = (H, \mathbf{L}) \in \mathcal{D} := M_{sa}(\mathbb{C}^d) \times M(\mathbb{C}^d)^k$$

and refer to each $\mathsf{D} \in \mathcal{D}$ as a *dynamical parameter* and to $\mathcal{D}$ as manifold (space) of dynamical parameters.

A. Environment as quantum noise

In the Markov approximation, the interaction can be described as the unitary scattering of incoming vacuum Bosonic fields, caused by the continuous interaction with the system. The environment is modelled by k Bosonic channels whose Hilbert space is the Fock space,

$$\mathcal{F} := \mathcal{F}(\mathfrak{h}_k) = \mathbb{C}|\Omega\rangle \oplus \bigoplus_{l=1}^{\infty} \mathfrak{h}_k^{\otimes_s l},$$

where $\mathfrak{h}_k := L^2(\mathbb{R}_+) \otimes \mathbb{C}^k$ is the one particle space of the k channels and $|\Omega\rangle$ is the vacuum vector. Similarly, we denote by $\mathcal{F}_{(a,b)}$ the Fock space over $L^2((a,b)) \otimes \mathbb{C}^k$. For each time t, the symmetric Fock space decomposes as tensor product $\mathcal{F} = \mathcal{F}_{(0,t)} \otimes \mathcal{F}_{(t,\infty)}$ between the space of excitations up to time t (the past) and after time t (the future). The fundamental environment degrees of freedom are the annihilation and creation operators of the ith channel $A_i(f) := A(|f\rangle \otimes |i\rangle)$ and, respectively, $A_i^*(g) := A_i(g)^*$, which are defined in a standard way[46] for all $|f\rangle, |g\rangle \in L^2(\mathbb{R}_+)$, and satisfy the commutation relations

$$[A_i^*(g), A_j(f)] = \mathrm{Im}\langle f|g\rangle \delta_{i,j} \mathbb{1}.$$

In particular, we will deal with the annihilation and creation *processes* $A_i(t) := A_i(\chi_{[0,t]})$ and $A_i^*(t)$, where $t \in \mathbb{R}_+$ represents time.[23,46] These processes are the quantum analogue of the "classical" Wiener process and can be used to define *quantum stochastic integrals* of the form

$$I(t) = \int_0^t \sum_{i=1}^k \left[M^i(s) dA_i(s) + N^i(s) dA_i^*(s) \right] + P(s) ds,$$

where $M^i(s)$, $N^i(s)$, $P(s)$ are *time-adapted* operator valued integrands, i.e., they are of the form $X(s) \otimes \mathbb{1}_{[s,\infty)}$ with respect to the decomposition $\mathcal{F} = \mathcal{F}_{(0,s)} \otimes \mathcal{F}_{(s,\infty)}$. Quantum stochastic integrals can be formally multiplied, and the product $I_1(t) I_2(t)$ of two such integrals is a stochastic integral whose increment is given by

$$d(I_1(t) I_2(t)) = dI_1(t) \cdot I_2(t) + I_1(t) \cdot dI_2(t) + dI_1(t) \cdot dI_2(t), \tag{2}$$

where the third terms is the *Ito correction* which can computed by using the *quantum Ito rule*,

$$dA_i(t) dA_j^*(t) = \delta_{i,j} dt, \tag{3}$$

while all other products are zero.

B. Interaction as input-output scattering

We now introduce the coupling between system and the Bosonic environment, and the corresponding unitary evolution. Each dynamical parameter $\mathsf{D} = (H, \mathbf{L})$ determines a unique continuous family $U_{\mathsf{D}}(t)(t)$ of unitary operators (cocycles) describing the evolution of the system and environment in the interaction picture with respect to the free evolution of the fields, the latter being given by the second quantisation of the right shift on $L^2(\mathbb{R})$. The unitaries are defined as the solution of the quantum stochastic differential equation (QSDE),

$$dU_{\mathsf{D}}(t) = \left(\sum_{i=1}^k (L^i \otimes dA_i^*(t) - L^{i*} \otimes dA_i(t)) - iH_{\mathrm{eff}} \otimes 1_{\mathcal{F}} dt \right) U_{\mathsf{D}}(t), \tag{4}$$

with initial condition $U_{\mathsf{D}}(0) = \mathbb{1}$. Here, H_{eff} is the effective Hamiltonian $H_{\mathrm{eff}} := H - \frac{i}{2} \sum_{i=1}^k L^{i*} L^i$ which generates a semigroup $S_{\mathsf{D}}(t) = e^{-itH_{\mathrm{eff}}}$ of contractions on the system's space and describes the evolution of the system between consecutive quantum jumps. The imaginary part $\frac{i}{2} \sum_{i=1}^k L^{i*} L^i$ is the Ito correction which insures that $U_{\mathsf{D}}(t)$ is unitary. For simplicity of notation, from now on we will omit the tensor product and simply write $L^i \otimes dA_i^*(t)$ as $L^i dA_i^*(t)$, and similarly for other integrands.

If the system is initialised in state $|\varphi\rangle$ and the input fields are in the vacuum state $|\Omega\rangle$, then the state of the system together with the output after time t is given by

$$|\Psi_{\mathsf{D}}^{\mathrm{s+o}}(t)\rangle = U_{\mathsf{D}}(t) |\varphi \otimes \Omega\rangle = V_{\mathsf{D}}(t) |\varphi\rangle, \tag{5}$$

where $V_{\mathsf{D}}(t) : \mathcal{H} \to \mathcal{H} \otimes \mathcal{F}$ is a family of isometries defined by the second equality. The *output state* is the state of the scattered field modes after the interaction with the system and is obtained by tracing out the system

$$\rho_{\mathsf{D}}^{\mathrm{out}}(t) = \mathrm{tr}_{\mathcal{H}}[|\Psi_{\mathsf{D}}^{\mathrm{s+o}}(t)\rangle\langle\Psi^{\mathrm{s+o}}(t)|].$$

Let us denote by $j_{\mathsf{D},t}(X) := U_{\mathsf{D}}(t)^*(X \otimes \mathbb{1}_{\mathcal{F}})U_{\mathsf{D}}(t)$ the Heisenberg evolved system operator X. Using Equation (4), we find that the operators satisfy the *quantum Langevin equation*,

$$dj_{\mathsf{D},t}(X) = \sum_i \left(j_{\mathsf{D},t}([X, L^i])dA_i^*(t) + j_{\mathsf{D},t}([L^{i*}, X])dA_i(t)\right) + j_{\mathsf{D},t}(\mathbb{W}_{\mathsf{D}}(X))dt, \tag{6}$$

where

$$\mathbb{W}_{\mathsf{D}}(\cdot) = -i(\cdot)H_{\text{eff}} + iH_{\text{eff}}^*(\cdot) + \sum_{i=1}^{k} L^{i*}(\cdot)L^i$$

is called the *Lindblad generator*. Its significance can understood by considering the *reduced Heisenberg evolution* of the system $T_{\mathsf{D},t} : \mathcal{A} \to \mathcal{A}$ defined by taking the expectation over the environment,

$$T_{\mathsf{D},t}(X) := \langle\Omega|j_{\mathsf{D},t}(X)|\Omega\rangle = V_{\mathsf{D}}(t)^*(X \otimes 1_{\mathcal{F}})V_{\mathsf{D}}(t). \tag{7}$$

From (6) we find $dT_{\mathsf{D},t}(X) = d\langle\Omega|j_{\mathsf{D},t}(X)|\Omega\rangle = \mathbb{W}_{\mathsf{D}}(T_{\mathsf{D},t}(X))$, which means that $T_{\mathsf{D},t}$ is a trace preserving completely positive semigroup with generator $\mathbb{W}_{\mathsf{D}}$. The generator is said to be *ergodic* if it has a unique stationary state ρ_{ss}^{D} (i.e., $[\mathbb{W}_{\mathsf{D}}]_*(\rho_{ss}^{\mathsf{D}}) = 0$) which has full rank. In this case,[19]

$$\lim_{t\to\infty} T_{\mathsf{D},t} = \lim_{t\to\infty} \frac{1}{t}\int_0^t T_{\mathsf{D},s}ds = \text{tr}[\rho_{ss}^{\mathsf{D}}(\cdot)]\mathbb{1}. \tag{8}$$

Since we are interested in the long-time asymptotic properties of the output state, we assume that the dynamics has reached stationarity or equivalently that the initial state of the system is ρ_{ss}^{D}. In this case, the output state is time-stationary and is given by

$$\rho_{\mathsf{D}}^{\text{out}}(t) = \text{tr}_{\mathcal{H}}[U_{\mathsf{D}}(t)(\rho_{ss}^{\mathsf{D}} \otimes |\Omega\rangle\langle\Omega|)U_{\mathsf{D}}(t)^*]. \tag{9}$$

Since ρ_{ss}^{D} is a stationary state, the range of $\mathbb{W}_{\mathsf{D}}$ is included in

$$\mathcal{B}_0^{\mathsf{D}} = \{X \mid \text{tr}[\rho_{ss}^{\mathsf{D}}X] = 0\}.$$

By ergodicity, $\mathbb{1}$ is the only fixed point of $e^{t\mathbb{W}_{\mathsf{D}}}$, and hence $\ker(\mathbb{W}_{\mathsf{D}})$ is spanned by $\mathbb{1} \notin \mathcal{B}_0^{\mathsf{D}}$. This implies that the range has dimension $d^2 - 1 = \dim \mathcal{B}_0^{\mathsf{D}}$, i.e., $\mathbb{W}_{\mathsf{D}}$ is surjective onto $\mathcal{B}_0^{\mathsf{D}}$, and the restriction of $\mathbb{W}_{\mathsf{D}}$ onto $\mathcal{B}_0^{\mathsf{D}}$ is injective. Hence $\mathbb{W}_{\mathsf{D}}$ is invertible on $\mathcal{B}_0^{\mathsf{D}}$, and we let $\mathbb{W}_{\mathsf{D}}^{-1} : \mathcal{B}_0^{\mathsf{D}} \to \mathcal{B}_0^{\mathsf{D}}$ denote the inverse. Furthermore, the following limit exists:

$$-\mathbb{W}_{\mathsf{D}}^{-1} = \lim_{t\to\infty}\int_0^t T_{\mathsf{D},s}ds. \tag{10}$$

IV. IDENTIFIABILITY OF CONTINUOUS QUANTUM MARKOV PROCESSES

This section deals with the problem of characterising the equivalence classes of Markov dynamics with identical stationary output states. We restrict ourselves to ergodic Markov processes although similar results are expected to hold more generally. Similar results have been obtained in Ref. 31 for *discrete time* quantum Markov processes.

Let $\mathcal{D}^{\text{erg}}$ denote the open submanifold of $\mathcal{D}$ consisting of dynamical parameters for which the associated Markov process is ergodic; this will be the relevant parameter set for subsequent considerations. Note that $\mathcal{D}^{\text{erg}}$ is indeed an open subset of $\mathcal{D}$ since ergodicity (i.e. non-zero spectral gap and full rank stationary state) is preserved under small perturbations.

Definition 1. Two dynamical parameters $\mathsf{D}, \mathsf{D}' \in \mathcal{D}^{\text{erg}}$ *are* output-equivalent *if the stationary output states* (9) *of the associated continuous-time Markov processes are identical. We denote the set of associated equivalence classes by* $\mathcal{P} := \{[\mathsf{D}] : \mathsf{D} \in \mathcal{D}^{\text{erg}}\}$.

Of course, the same equivalence can be formulated for arbitrary parameters in $\mathcal{D}$. However, it turns out that when restricted to $\mathcal{D}^{\text{erg}}$, the equivalence classes have a simple characterisation in terms of the following transformations:

(PM) Phase conjugation on the Hamiltonian,

$$(H, L^1, \dots, L^k) \mapsto (H + r\mathbb{1}, L^1, \dots, L^k), \quad (r \in \mathbb{R}).$$

(UC) Conjugation by system unitary W,

$$(H, L^1, \dots, L^k) \mapsto (W^*HW, W^*L^1W, \dots, W^*L^kW).$$

Indeed, it is easy to verify that (PM) and (UC) do not change the output of the associated continuous Markov process. The following theorem shows that the converse is also true. The details of the proof can be found in the Appendix.

Theorem 1. *Let* $\mathsf{D}, \mathsf{D}' \in \mathcal{D}^{\mathrm{erg}}$. *Then* D *and* D' *are output-equivalent if and only if they can be obtained from each other via the transformations (UC) and (PM).*

The interpretation of the result is that parameters along the equivalence classes described by the transformations (UC) and (PM) are not identifiable, while the identifiable parameters are "transversal" to these classes, as illustrated in Figure 2. It is now convenient to formulate the equivalence classes in terms of an action of the appropriate Lie group $G := PU(d) \times \mathbb{R}$. On $\mathcal{D}^{\mathrm{erg}}$ we use transformations (PM) and (UC) to set up the action,

$$\begin{aligned} &G \times \mathcal{D}^{\mathrm{erg}} \to \mathcal{D}^{\mathrm{erg}}, \\ &(g, \mathsf{D}) \mapsto g\mathsf{D} := (W^*HW, W^*L^1W, \dots, W^*L^kW) + a(\mathbb{1}, 0, \dots, 0), \\ &g = (W, a) \in PU(d) \times \mathbb{R}, \quad \mathsf{D} = (H, L^1, \dots, L^k) \in \mathcal{D}^{\mathrm{erg}}. \end{aligned} \tag{11}$$

Here $PU(d) = U(d)/U(1)$ is the projective unitary group, equipped with its unique Lie group structure, and the above action is defined as the natural lift of the corresponding one for $U(d)$. The reason to use $PU(d)$ instead of $U(d)$ will become clear from the proof of the lemma below.

The above theorem implies that the equivalence class $[\mathsf{D}] \in \mathcal{P}$ is the orbit of $D \in \mathcal{D}^{\mathrm{erg}}$ under the action of G, such that $\mathcal{P}$ can be identified with the quotient $\mathcal{D}^{\mathrm{erg}}/G$. The following lemma, which relies on the ergodicity assumption, is essential for understanding the structure of the quotient, as we will see below. In order to avoid confusion with the output equivalence, we identify $W \in PU(d)$ with a representative unitary operator without explicit indication.

Lemma 1. The Lie group action $G \times \mathcal{D}^{\mathrm{erg}} \to \mathcal{D}^{\mathrm{erg}}$ *is smooth, proper, and free.*

Proof. The action defined via (11) is clearly smooth with respect to W, and a; hence, its lift to the quotient Lie group $PU(d) \times \mathbb{R}$ is smooth as well. Since the group $PU(d)$ is compact, and the rest is just a translation, it follows from elementary arguments that the smooth map $G \times \mathcal{D}^{\mathrm{erg}} \to \mathcal{D}^{\mathrm{erg}} \times \mathcal{D}^{\mathrm{erg}}$ given by $(g, \mathsf{D}) \mapsto (g\mathsf{D}, \mathsf{D})$ is proper, i.e., preimage of every compact set is compact. This means that the action is proper. In order to show that the action is free, we need to use ergodicity as follows: suppose that $g\ \mathsf{D} = g'\mathsf{D}$ for some D, and $g = (W, a)$, $g' = (W', a')$; then a direct computation similar to the one in the proof of Lemma 2 in the Appendix shows that $\mathbb{W}_{\mathsf{D}}(W^*W') = i(a - a')W^*W'$. Since ergodicity requires

$$\lim_{\tau \to \infty} e^{i\tau(a-a')} W^*W' = \mathrm{tr}[W^*W'\rho_{ss}]\mathbb{I},$$

we must have $a = a'$ and W^*W' a multiple of the identity. But this exactly means that W equals W' as an element of the *projective* unitary group; hence, $g = g'$. This proves that the action is free. □

The fact that the group action preserves the equivalence classes, that is, $g\mathsf{D} \in [\mathsf{D}]$ for all $\mathsf{D} \in \mathcal{D}^{\mathrm{erg}}$, can now be formulated in differential geometric terms. Indeed, using the standard theory of Lie group actions on manifolds, we conclude from the above Lemma that the space of output equivalence classes,

$$\mathcal{P} = \{[\mathsf{D}] \ : \ \mathsf{D} \in \mathcal{D}^{\mathrm{erg}}\} = \mathcal{D}^{\mathrm{erg}}/G,$$

admits a unique smooth structure such that the quotient map $\pi : \mathcal{D}^{\mathrm{erg}} \to \mathcal{P}$,

$$\pi(\mathsf{D}) = [\mathsf{D}], \quad \text{for all } \mathsf{D} \in \mathcal{D}^{\mathrm{erg}},$$

is a submersion, and $\mathcal{D}^{\mathrm{erg}}$ is a principal G-bundle over $\mathcal{P}$.[38] Here the equivalence classes $[\mathsf{D}]$ are considered as *fibres* of the *fiber bundle* over the *base manifold* $\mathcal{P}$, that is, the map π has the local triviality property: each $[D] \in \mathcal{P}$ has an open neighbourhood U such that there exists a diffeomorphism,

$$\phi : \pi^{-1}(U) \to U \times G,$$

which is G-equivariant, i.e., $\phi(g\mathsf{D}) = g\phi(\mathsf{D})$, where G acts on $U \times G$ as $g([\mathsf{D}], g') := ([\mathsf{D}], g'g^{-1})$. The term principal G-bundle refers to the fact that the group action preserves the fibres.

We can now use this differential geometric framework to describe local changes of identifiable parameters via the tangent bundle $\mathcal{T}$of the manifold $\mathcal{D}^{\mathrm{erg}}$. In particular, the *non-identifiable* parameter changes along the equivalence classes correspond to the *vertical bundle* over $\mathcal{D}^{\mathrm{erg}}$ with the fibres,

$$\mathcal{T}^{\mathrm{nonid}}_{\mathsf{D}} := \ker \pi_*|_{\mathsf{D}} \subset \mathcal{T}_{\mathsf{D}}, \quad \mathsf{D} \in \mathcal{D}^{\mathrm{erg}},$$

where $\pi_*|_{\mathsf{D}}$ is the push-forward tangent map of the canonical projection π at point D, and $\mathcal{T}_{\mathsf{D}}$ is the full tangent space at that point.

The group action is reflected in two ways at the level of tangent spaces. On one hand, given any fixed $g \in G$, the push-forward g_* of the map $\mathsf{D} \mapsto g\mathsf{D}$ maps the fibres into each other as

$$g_* \mathcal{T}^{\mathrm{nonid}}_{\mathsf{D}} = \mathcal{T}^{\mathrm{nonid}}_{g\mathsf{D}}.$$

This push-forward is simply obtained by differentiating the parameters in the standard chart,

$$g_*(\dot{H}, \dot{L}^1, \ldots, \dot{L}^k) = (W^* \dot{H} W, W^* \dot{L}^1 W, \ldots, W^* \dot{L}^k W), \quad g = (W, a).$$

On the other hand, for any fixed D, the push-forward of $g \mapsto g\mathsf{D}$ defines a Lie algebra isomorphism,

$$\mathsf{D}_* : \mathfrak{g} \to \mathcal{T}^{\mathrm{nonid}}_{\mathsf{D}}, \tag{12}$$

so that different fibres all have the same dimension, which is that of the Lie algebra $\mathfrak{g}$. We can now explicitly compute this action. First of all, the Lie algebra of G can be conveniently written as

$$\mathfrak{g} = \{(-iK, r) \mid K \in M_{sa}(\mathbb{C}^d)/\mathbb{R}\mathbb{1},\ r \in \mathbb{R}\} = \{(-iK, r) \mid K \in M_{sa}(\mathbb{C}^d),\ \mathrm{tr}[\rho^{ss} K] = 0,\ r \in \mathbb{R}\}, \tag{13}$$

where the choice of the sign as well as the last identification is for later convenience. In particular, the subspace of non-identifiable directions is in one-to-one correspondence with this linear space. From this, we already find the number of non-identifiable directions,

$$\dim \mathcal{T}^{\mathrm{nonid}}_{\mathsf{D}} = d^2 - 1 + 1 = d^2.$$

We stress that this result is crucially based on ergodicity, which ensures that the action is free; this is required for the push-forward D_* to be an isomorphism. Now D_* acts on an element $X = (-iK, r) \in \mathfrak{g}$ as

$$\begin{aligned}
\mathsf{D}_*(X) &= \frac{d}{dt}(\exp(t\,(-iK, r))\mathsf{D})|_{t=0} \\
&= \frac{d}{dt}\left(\left(e^{itK} H e^{-itK}, e^{itK} L^1 e^{-itK}, \ldots, e^{itK} L^k e^{-itK}\right) + t\, r(\mathbb{1}, 0, \ldots, 0)\right)|_{t=0} \\
&= (i[H, K], i[L^1, K], \ldots, i[L^k, K]) + r(\mathbb{1}, 0, \ldots, 0).
\end{aligned} \tag{14}$$

Having now characterised the vertical bundle $\mathcal{T}^{\mathrm{nonid}}$ of non-identifiable directions, an obvious question arises: is there a natural way to choose complementary subspaces for *identifiable directions* in each fibre? This means choosing subspaces $\mathcal{T}^{\mathrm{id}}_{\mathsf{D}}$ such that

$$\mathcal{T}_{\mathsf{D}} = \mathcal{T}^{\mathrm{nonid}}_{\mathsf{D}} \oplus \mathcal{T}^{\mathrm{id}}_{\mathsf{D}}, \quad \mathsf{D} \in \mathcal{D}^{\mathrm{erg}}.$$

If the subspaces are chosen smoothly, i.e., so as to define a fibre bundle $\mathcal{T}^{\mathrm{id}}$ over $\mathcal{D}^{\mathrm{erg}}$, the result is called a *horizontal bundle* $\mathcal{T}^{\mathrm{id}}$, and in case it respects the group action, that is,

$$g_* \mathcal{T}^{\mathrm{id}}_{\mathsf{D}} = \mathcal{T}^{\mathrm{id}}_{g\mathsf{D}}, \tag{15}$$

it defines a *principal connection* on the manifold $\mathcal{D}^{\mathrm{erg}}$.

There is a natural way of defining a principal connection via its associated *connection one-form*; since this approach turns out to be relevant in our situation, we briefly explain the idea in the general

level. As we have shown above, any $\dot{\mathsf{D}} \in \mathcal{T}_{\mathsf{D}}^{\text{nonid}}$ can be generated by the action of the Lie algebra; $\dot{\mathsf{D}} = \mathsf{D}_*(X)$ for some $X \in \mathfrak{g}$. Now suppose that we can associate with *every* tangent vector $\dot{\mathsf{D}} \in \mathcal{T}_{\mathsf{D}}$ an element $\omega_{\mathsf{D}}(\dot{\mathsf{D}}) \in \mathfrak{g}$ which somehow describes the "part" of the parameter that results from the non-identifiable group action. Such a map should define a one-form $\omega_{\mathsf{D}} : \mathcal{T} \to \mathfrak{g}$ satisfying the compatibility condition (sometimes called *nondegeneracy*),

$$\dot{\mathsf{D}} = \mathsf{D}_*(\omega_{\mathsf{D}}(\dot{\mathsf{D}})), \quad \dot{\mathsf{D}} \in \mathcal{T}_{\mathsf{D}}^{\text{nonid}}, \tag{16}$$

and the G-covariance condition

$$g^*\omega = \mathrm{Ad}_{g^{-1}} \circ \omega, \tag{17}$$

where g^* is the pull-back of the action by g on the cotangent bundle, which simply acts as $g^*\omega(\dot{\mathsf{D}}) = \omega_{g\mathsf{D}}(g_*\dot{\mathsf{D}}) = \omega_{g\mathsf{D}}(W^*\dot{\mathsf{D}}W)$ for $g = (W, a)$, and the adjoint action is given by $\mathrm{Ad}_{g^{-1}}(X) = (W^*XW, r+a)$.

Given such a map, we can then define the "back-action" $\mathsf{D}_* \circ \omega_{\mathsf{D}}$ on the tangent space; due to the above compatibility condition, the back-action is a special projection of the tangent space onto the subspace $\mathcal{T}_{\mathsf{D}}^{\text{nonid}}$. Hence, we can use its complementary projection

$$P_{\mathsf{D}} := \mathrm{Id} - \mathsf{D}_* \circ \omega_{\mathsf{D}}$$

to define the above horizontal bundle and the associated principal connection via

$$\mathcal{T}_{\mathsf{D}}^{\text{id}} := \operatorname{ran} P_{\mathsf{D}}.$$

Indeed, condition (15) holds because of (17). The map ω is called the *connection one-form*, and P is the *horizontal projection*.

Any principal connection gives a possible way of extracting the parameter changes relevant for our system identification problem. In Sec. V we show that there is actually a natural connection associated with the *information geometric structure* of the problem, given by the Fisher information of the output state. We will obtain it by explicitly constructing the associated connection one-form, which arises neatly from the quantum Ito calculus.

V. INFORMATION GEOMETRY FOR DYNAMICAL PARAMETER ESTIMATION FROM THE OUTPUT STATE

Our goal is to describe quantitatively the precision with which unknown dynamical parameters can be estimated by making measurements on the output state. As noted above, we will restrict our attention to dynamical parameters D which belong to the open subset $\mathcal{D}^{\text{erg}}$ of $\mathcal{D}$ of *ergodic* Markov dynamics. As we will consider this problem in the limit of large times, the relevant dynamical regime is the stationary one; moreover, the statistical properties of the output state can be understood *locally*, by focusing on a shrinking neighbourhood of the parameter manifold $\mathcal{D}^{\text{erg}}$ whose size is of the order of the statistical uncertainty.[13,31] This will lead to the concept of local asymptotic normality discussed in Section VII. In this section however, we focus on the information geometry of the system identification problem, more precisely on the quantum Fisher information matrix of the output state and its asymptotic behaviour, and its relationship with the covariance of certain quanta stochastic integrals called "fluctuation operators." We will start by introducing the latter in a general setup and then show how the former fits in this theory.

Section V A derives the quantum Fisher information of the system-output state as covariance of certain "generators"; Section V B analyses more general "fluctuation operators" and looks at their Markov covariance; Section V C deals with the information geometry structure and connects the previous constructions; in particular, it provides an explicit expression of the quantum Fisher information; and Section V D constructs an algebra of canonical commutation relations (multimode continuous variables system) and a family of coherent states which will be relevant later on for the local asymptotic result.

A. Quantum Fisher information of a parametric model

We pass now to a statistical setting where the dynamical parameter D is considered to be unknown. The changes in D are encoded in its (partial) derivatives $\dot{\mathsf{D}} = (\dot{H}, \dot{L}^1, \ldots, \dot{L}^k)$, which will be seen as

vectors in the tangent space $\mathcal{T}_{\mathsf{D}}$ to $\mathcal{D}^{\mathrm{erg}}$ at the point D. Since the dynamics is ergodic, the system converges to a unique stationary state ρ_{ss}^{D} for large times, and we will denote by $\langle\cdot\rangle_{ss}$ the expectation with respect to the state $\rho_{ss}^{\mathsf{D}} \otimes |\Omega\rangle\langle\Omega|$. In this subsection we consider a generic statistical model and analyse the quantum Fisher information (QFI) of the output state; we will show that the QFI grows linearly with time and the rate can be expressed in terms of the Markov covariance inner product introduced below. Let

$$\mathbb{R}^m \ni \theta \mapsto \mathsf{D}_\theta.$$

be a smooth family of dynamics parametrised by an unknown parameter $\theta \in \mathbb{R}^m$ which may be thought to encode our prior knowledge about the dynamics. In this subsection, we work with this parametrisation and hence identify D_θ with θ for simplicity. Note that this could be a complete parametrisation of $\mathcal{D}^{erg}$. The directional derivatives of D_θ are defined as

$$\dot{\mathsf{D}}_{\theta,a} := \left(\frac{\partial H}{\partial \theta_a}, \frac{\partial L^1}{\partial \theta_a}, \ldots, \frac{\partial L^k}{\partial \theta_a}\right) = (\dot{H}_{\theta,a}, \dot{L}^1_{\theta,a}, \ldots, \dot{L}^k_{\theta,a}) \in \mathcal{T}_{\mathsf{D}_\theta}.$$

Recall that the QFI of an arbitrary multiparameter (smooth) family of pure states $|\psi_\theta\rangle$ with $\theta \in \mathbb{R}^m$ is the $m \times m$ positive real matrix with elements[11]

$$F^\theta_{a,b} = 4\mathrm{Re}\left(\left\langle \frac{\partial \psi_\theta}{\partial \theta_a} \middle| \frac{\partial \psi_\theta}{\partial \theta_b}\right\rangle - \left\langle \psi_\theta \middle| \frac{\partial \psi_\theta}{\partial \theta_b}\right\rangle \left\langle \frac{\partial \psi_\theta}{\partial \theta_a} \middle| \psi_\theta\right\rangle\right), \quad 1 \le a, b \le m.$$

We apply this formula to the output state $|\Psi^{\mathrm{s+o}}_\theta(t)\rangle := U_\theta(t)|\varphi \otimes \Omega\rangle$ generated with a θ-dependent dynamical parameter D_θ, cf. Equation (9). By differentiating with respect to θ_a, we get

$$U^*_\theta(t)\frac{\partial}{\partial \theta_a}\left|\Psi^{\mathrm{out}}_\theta(t)\right\rangle = U^*_\theta(t)\dot{U}_{\theta,a}(t)|\varphi \otimes \Omega\rangle, \qquad \dot{U}_{\theta,a}(t) := \frac{\partial U_\theta(t)}{\partial \theta_a}. \tag{18}$$

We will now show that the *generator* $-iG_{\theta,a}(t) := U^*_\theta(t)\dot{U}_{\theta,a}(t)$ can be written as a quantum stochastic integral. From (4) we have

$$\begin{aligned}
dU^*_\theta(t) &= U^*_\theta(t)\left(\sum_i(-L^i_\theta dA^*_i(t) + L^{i*}_\theta dA_i(t)) - (-iH_\theta + \frac{1}{2}\sum_i L^{i*}_\theta L^i_\theta)dt\right),\\
d\dot{U}_\theta(t) &= \left(\sum_i(\dot{L}^i_{\theta,a}dA^*_i(t) - \dot{L}^{i*}_{\theta,a}dA_i(t)) - (i\dot{H}_{\theta,a} + \frac{1}{2}\sum_i(\dot{L}^{i*}_{\theta,a}L^i_\theta + L^{i*}_\theta\dot{L}^i_\theta))dt\right)U_\theta(t)\\
&\quad + \left(\sum_i(L^i_\theta dA^*_i(t) - L^{i*}_\theta dA_i(t)) - (iH_\theta + \frac{1}{2}\sum_i L^{i*}_\theta L^i_\theta)dt\right)\dot{U}_\theta(t).
\end{aligned}$$

Therefore, by applying the Ito rule (3), we get

$$dU^*_\theta(t)\cdot d\dot{U}_{\theta,a}(t) = U_\theta(t)^*\sum_i L^{i*}_\theta\left(\dot{L}^i_{\theta,a}U(t) + L^i_\theta\dot{U}_\theta(t)\right)dt,$$

and using (2) we obtain an explicit differential expression for the generator

$$\begin{aligned}
dG_{\theta,a}(t) &= id(U^*_\theta(t)\dot{U}_{\theta,a}(t))\\
&= i\sum_i\left(j_{\theta,t}(\dot{L}^i_{\theta,a})dA^*_i(t) - j_{\theta,t}(\dot{L}^{i*}_{\theta,a})dA_i(t)\right) + j_{\theta,t}\left(\dot{H}_{\theta,a} + \mathrm{Im}\sum_i \dot{L}^{i*}_{\theta,a}L^i_\theta\right)dt\\
&= i\sum_i\left(j_{\theta,t}(\dot{L}^i_{\theta,a})dA^*_i(t) - j_{\theta,t}(\dot{L}^{i*}_{\theta,a})dA_i(t)\right) + j_{\theta,t}\left(\mathcal{E}_\theta(\dot{\mathsf{D}}_{\theta,a})\right)dt,
\end{aligned} \tag{19}$$

where $\mathcal{E}_\theta = \mathcal{E}_{\mathsf{D}_\theta}$ is the *real linear* map $\mathcal{E}_{\mathsf{D}} : \mathcal{T}_{\mathsf{D}} \to M_{sa}(\mathbb{C}^d)$ given by

$$\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}}) := \dot{H} + \mathrm{Im}\sum_{i=1}^k \dot{L}^{i*}L^i. \tag{20}$$

Later on we will see that this map turns out to play a crucial role in the construction of the horizontal bundle for the identifiable parameters, and in the definition of the CCR algebra in Section V D.

The QFI can be written in terms of the covariance matrix of the generators $G_{\theta,b}(t)$,

$$F^{\theta}_{a,b}(t) = 4\mathrm{Re}\left(\langle\varphi\otimes\Omega|G^{*}_{\theta,a}(t)G_{\theta,b}(t)|\varphi\otimes\Omega\rangle - \langle\varphi\otimes\Omega|G^{*}_{\theta,a}(t)|\varphi\otimes\Omega\rangle\langle\varphi\otimes\Omega|G_{\theta,a}(t)|\varphi\otimes\Omega\rangle\right),$$

where the second term stems from the fact that $G_{\theta,b}(t)$ have non-zero mean. The generators are in fact not uniquely defined: since $dA_i(t)$ annihilates the vacuum state, arbitrary annihilation integrals can be added, while terms proportional to the identity produce only unphysical complex phases which do not change the state. We will therefore define a modified (non-selfadjoint) generator which "centres" $G_{\theta,b}(t)$ for large times and lacks annihilations terms so that it is consistent with the definition of "fluctuation operators" introduced in Subsection V B. The modified generator is given by the quantum stochastic integral with differential form

$$dG^{0}_{\theta,a}(t) = i\sum_{i=1}^{k} j_t(\dot{L}^{i}_{\theta,a})dA^{*}_{i}(t) + \left(j_{\theta,t}(\mathcal{E}_{\mathsf{D}_\theta}(\dot{\mathsf{D}}_{\theta,a})) - \mathrm{Tr}\left(\rho^{\mathsf{D}}_{ss}\mathcal{E}_{\mathsf{D}_\theta}(\dot{\mathsf{D}}_{\theta,a})\right)\right)dt. \tag{21}$$

By ergodicity, its rescaled mean converges to zero

$$\lim_{t\to\infty}\frac{1}{t}\langle\varphi\otimes\Omega|G^{0}_{\theta,a}(t)|\varphi\otimes\Omega\rangle = 0.$$

For large times, the QFI matrix elements scale linearly with t and the leading contribution is given by the *quantum Fisher information rate*,

$$f^{\theta}_{a,b} := \lim_{t\to\infty}\frac{F^{\theta}_{a,b}(t)}{t} = \lim_{t\to\infty}\frac{1}{t}4\mathrm{Re}\langle\varphi\otimes\Omega|G^{0*}_{\theta,a}(t)G^{0}_{\theta,b}(t)|\varphi\otimes\Omega\rangle. \tag{22}$$

In Sec. V B we prove the linear scaling and find an explicit expression of the QFI rate.

B. Fluctuation operators and the Markov covariance form

Our goal is now to formulate the QFI rate (22) in terms of certain quantum fluctuation operators and subsequently compute it using quantum stochastic calculus. These fluctuation operators can be formulated in a slightly more general setting, which is naturally *complex linear* instead of real linear and is also independent of the map $\mathcal{E}_{\mathsf{D}}$ special to our setting. The dynamical parameter D will remain fixed throughout the section.

Recall that for any $X\in M(\mathbb{C}^d)$ we let $j_{\mathsf{D},t}(X)$ denote the Heisenberg evolved system observable defined by the Langevin equation (6). For an arbitrary $(k+1)$-tuple $\mathbf{X} := (X^0, X^1, \ldots, X^k)\in M(\mathbb{C}^d)^{1+k}$, we define the associated *centered fluctuation operator* by the quantum stochastic integral,

$$\mathbb{F}_{\mathsf{D},t}(\mathbf{X}) = \frac{1}{\sqrt{t}}\int_0^t\left(i\sum_{i=1}^{k} j_{\mathsf{D},s}(X^i)dA^{*}_{i}(s) + j_{\mathsf{D},s}\circ\mathcal{C}_{\mathsf{D}}(X^0)ds\right), \tag{23}$$

where the map

$$\mathcal{C}_{\mathsf{D}}(X) := X - \mathrm{tr}[\rho^{\mathsf{D}}_{ss}X]\mathbb{1}$$

"centers" the *stationary mean* of $\mathbb{F}_{\mathsf{D},t}(\mathbf{X})$ to zero,

$$\langle\mathbb{F}_{\mathsf{D},t}(\mathbf{X})\rangle_{ss} = \frac{1}{\sqrt{t}}\int_0^t \mathrm{tr}[\rho^{\mathsf{D}}_{ss}T_{\mathsf{D},s}(X^0 - \mathrm{tr}[\rho^{\mathsf{D}}_{ss}X^0]\mathbb{1})]ds = \frac{1}{\sqrt{t}}\int_0^t \mathrm{tr}[\rho^{\mathsf{D}}_{ss}(X^0 - \mathrm{tr}[\rho^{\mathsf{D}}_{ss}X^0]\mathbb{1})]dt = 0.$$

The proof of the following crucial result is based on quantum Ito calculus and can be found in the Appendix.

Proposition 1 (Markov covariance for fluctuation operators). The following limit exists, is independent of the unit vector $|\varphi\rangle\in\mathcal{H}$, and defines a positive sesquilinear form $(\cdot,\cdot)_{\mathsf{D}}$ on the complex linear space $M(\mathbb{C}^d)^{1+k}$ via

$$(\mathbf{X},\mathbf{Y})_{\mathsf{D}} := \lim_{t\to\infty}\langle\varphi\otimes\Omega|\mathbb{F}_{\mathsf{D},t}(\mathbf{X})^{*}\mathbb{F}_{\mathsf{D},t}(\mathbf{Y})|\varphi\otimes\Omega\rangle = \sum_{i=1}^{k}\mathrm{tr}\left[\rho^{\mathsf{D}}_{ss}R_{\mathsf{D}}(\mathbf{X})^{i*}R_{\mathsf{D}}(\mathbf{Y})^{i}\right],$$

where

$$R_{\mathsf{D}}(\mathbf{X}) = (\mathcal{C}_{\mathsf{D}}(X^0), X^1, \ldots, X^k) - \mathcal{L}_{\mathsf{D}} \circ \mathbb{W}_{\mathsf{D}}^{-1} \circ \mathcal{C}_{\mathsf{D}}(X^0) \qquad \text{and}$$
$$\mathcal{L}_{\mathsf{D}}(X) = \left(\mathbb{W}_{\mathsf{D}}(X), i[L^1, X], \ldots, i[L^k, X]\right).$$

We call $(\cdot,\cdot)_{\mathsf{D}}$ *the Markov covariance inner product.*

From this proposition, it is clear that the map R_{D} plays a central role; in particular, since ρ_{ss}^{D} has full rank, the kernel of the Markov covariance coincides with $\ker R_{\mathsf{D}}$. Also the range of R_{D} turns out to be relevant. These subspaces can be characterised explicitly as follows.

Proposition 2. The operator R_{D} *is a projection, i.e.,* $R_{\mathsf{D}}^2 = R_{\mathsf{D}}$, *with range and kernel*

$$\ker R_{\mathsf{D}} = \left\{(\mathbb{W}_{\mathsf{D}}(K) + r\mathbb{1}, i[L^1, K], \ldots, i[L^k, K]) \,|K \in M(\mathbb{C}^d),\ r \in \mathbb{C}\right\},$$
$$\operatorname{ran} R_{\mathsf{D}} = \left\{\left(0, Y^1, \ldots, Y^k\right) \,\middle|\, Y^1, \ldots, Y^k \in M(\mathbb{C}^d)\right\}.$$

Proof. First of all, $\mathbf{X} \in \ker R_{\mathsf{D}}$ if and only if $X^i = i[L^i, \mathbb{W}^{-1}(X^0 - \operatorname{tr}[\rho_{ss}^{\mathsf{D}} X^0]1)]$ for all $i = 1, \ldots, k$. Since $\operatorname{tr}[\rho_{ss}^{\mathsf{D}} \mathbb{W}_{\mathsf{D}}(K)] = 0$ for any K, the given form of the kernel follows. The range is clear from the definition, and the property $R_{\mathsf{D}}^2 = R_{\mathsf{D}}$ is straightforward to check. □

C. Markov covariance from a principal connection

We now proceed to show how the Markov covariance is naturally associated with a specific horizontal bundle for the principal G-bundle $\mathcal{D}^{\mathrm{erg}}$, and we also define a Riemannian metric on the manifold $\mathcal{D}^{\mathrm{erg}}$. In order to motivate this, we continue the discussion from Subsection V A. Indeed, the modified generator (21) can be expressed as a fluctuation operator

$$G_{\theta,a}^0(t) = \sqrt{t}\, \mathbb{F}_t(\mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}}_{\theta,a})),$$

where we have used the suggestive notation $\mathbf{X}_{\mathsf{D}}$ for the *real linear* isomorphism

$$\mathbf{X}_{\mathsf{D}} : \mathcal{T}_{\mathsf{D}} \to M_{sa}(\mathbb{C}^d) \times M(\mathbb{C}^d)^k \tag{24}$$

$$\dot{\mathsf{D}} = (\dot{H}, \dot{L}^1, \ldots, \dot{L}^k) \mapsto (\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}}), \dot{L}^1, \ldots, \dot{L}^k), \tag{25}$$

where $M(\mathbb{C}^d)^k$ is now considered as a real linear space with dimension $2kd^2$, while $M_{sa}(\mathbb{C}^d)$ is naturally a real linear space. Therefore, using the explicit expression provided in Proposition 1, we obtain the following expression of the QFI rate (22) in the coordinates $\mathsf{D} = \mathsf{D}_\theta$ used in Subsection V A:

$$\begin{aligned} f_{a,b}^{\theta} &= 4\mathrm{Re}\left(\mathbf{X}_{\mathsf{D}}[\dot{\mathsf{D}}_{\theta,a}],\, \mathbf{X}_{\mathsf{D}}[\dot{\mathsf{D}}_{\theta,b}]\right)_{\mathsf{D}} \\ &= 4\sum_{i=1}^{k} \mathrm{Re}\,\mathrm{tr}\left[\rho_{ss}^{\mathsf{D}}\left(\dot{L}_{\theta,a}^i - i[L_\theta^i, \mathbb{W}_{\mathsf{D}}^{-1} \circ \mathcal{E}_{\mathsf{D}}^0(\dot{\mathsf{D}}_{\theta_a})]\right)^* \left(\dot{L}_{\theta,b}^i - i[L_\theta^i, \mathbb{W}_{\mathsf{D}}^{-1} \circ \mathcal{E}_{\mathsf{D}}^0(\dot{\mathsf{D}}_{\theta_b})]\right)\right], \end{aligned} \tag{26}$$

where $\mathcal{E}_{\mathsf{D}}^0 = \mathcal{C}_{\mathsf{D}} \circ \mathcal{E}_{\mathsf{D}}$. The QFI rate inherits the positivity property of the Markov covariance, and also the fact that it may not be positive *definite*. In conclusion, the *real part* of the form

$$(\dot{\mathsf{D}}, \dot{\mathsf{D}}')_{\mathsf{D}} := (\mathbf{X}_{\mathsf{D}}[\dot{\mathsf{D}}], \mathbf{X}_{\mathsf{D}}[\dot{\mathsf{D}}'])_{\mathsf{D}} \tag{27}$$

has a natural interpretation in terms of the output Fisher information. Its explicit form (see Proposition 1) suggests the definition of the following projection on the tangent bundle over $\mathcal{D}^{\mathrm{erg}}$,

$$P : \mathcal{T} \to \mathcal{T}, \quad P_{\mathsf{D}} = \mathbf{X}_{\mathsf{D}}^{-1} \circ R_{\mathsf{D}} \circ \mathbf{X}_{\mathsf{D}}. \tag{28}$$

Indeed, the bilinear form essentially depends on this projection:

$$(\dot{\mathsf{D}}, \dot{\mathsf{D}}')_{\mathsf{D}} = \sum_{i=1}^{k} \operatorname{tr}\left[\rho_{ss}^{\mathsf{D}}[P_{\mathsf{D}}(\dot{\mathsf{D}})^i]^* P_{\mathsf{D}}(\dot{\mathsf{D}}')^i\right].$$

In order to understand the intuitive meaning of P_{D}, we now proceed to make a fundamental observation concerning the relation between the push-forward map D_* defined in (12), and the map $\mathcal{E}_{\mathsf{D}}$ defined in (20),

$$\mathcal{E}_{\mathsf{D}} \circ \mathsf{D}_*(X) = r\mathbb{1} + \mathbb{W}_{\mathsf{D}}(K), \quad X = (-iK, r). \tag{29}$$

This relation has appeared before in a different context.[3] In order to verify it, we recall that $\mathsf{D}_*(-iK, r) = (i[H, K] + r\mathbb{1}, i[L^1, K], \ldots, i[L^k, K])$, so that

$$\begin{aligned}
\mathcal{E}_{\mathsf{D}}(\mathsf{D}_*(-iK, r)) &= i[H, K] + r\mathbb{1} + \frac{1}{2i}\sum_{i=1}^{k}\left((i[L^i, K])^* L^i - L^{i*}(i[L^i, K])\right) \\
&= r\mathbb{1} + i[H, K] - \frac{1}{2}\sum_{i=1}^{k}\left([K, L^{i*}]L^i + L^{i*}[L^i, K]\right) \\
&= r\mathbb{1} + iHK - iKH - \frac{1}{2}\sum_{i=1}^{k}(KL^{i*}L^i - L^{i*}KL^i + L^{i*}L^iK - L^{i*}KL^i) \\
&= r\mathbb{1} - iK\left(H - \frac{i}{2}\sum_i L^{i*}L^i\right) + i\left(H + \frac{i}{2}\sum_i L^{i*}L^i\right)K + \sum_i L^{i*}KL^i \\
&= r\mathbb{1} + \mathbb{W}_{\mathsf{D}}(K).
\end{aligned}$$

Equation (29) is the key to defining the horizontal bundle for the identifiable parameters. Indeed, we get the following crucial result.

Proposition 3 (Principal connection). The map $\omega : \mathcal{T}_{\mathsf{D}} \to \mathfrak{g}$, *defined by*

$$\omega_{\mathsf{D}}(\dot{\mathsf{D}}) = (-i\mathbb{W}_{\mathsf{D}}^{-1} \circ \mathcal{C}_{\mathsf{D}} \circ \mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}}), \mathrm{tr}[\rho_{ss}^{\mathsf{D}}\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}})]),$$

is the one-form of a unique principal connection on $\mathcal{D}^{\mathrm{erg}}$, *having* P_{D} *as its horizontal projection. In particular,* $P_{\mathsf{D}} = \mathrm{Id} - \mathsf{D}_* \circ \omega_{\mathsf{D}}$, *with the vertical subspaces* $\ker P_{\mathsf{D}} = \mathrm{ran}\, \mathsf{D}_* = \mathcal{T}_{\mathsf{D}}^{\mathrm{nonid}}$, *so this connection is compatible with the vertical bundle defining the "non-identifiable directions".*

Proof. We first verify the important relation $P_{\mathsf{D}} = \mathrm{Id} - \mathsf{D}_* \circ \omega_{\mathsf{D}}$. Denote $K = \mathbb{W}_{\mathsf{D}}^{-1} \circ \mathcal{C}_{\mathsf{D}} \circ \mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}})$ and $r = \mathrm{tr}[\rho_{ss}^{\mathsf{D}}\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}})]$, so that $\mathcal{C}_D(\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}})) = \mathbb{W}_{\mathsf{D}}(K)$. On one hand, using the formulas of R_{D} and $\mathcal{L}_{\mathsf{D}}$ in Proposition 1, we get

$$\begin{aligned}
R_{\mathsf{D}}(\mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}})) &= R_{\mathsf{D}}(\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}}), \dot{L}^1, \ldots, \dot{L}^k) = \left(\mathcal{C}_D(\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}})), \dot{L}^1, \ldots, \dot{L}^k\right) - \mathcal{L}_{\mathsf{D}}\left(\mathbb{W}_{\mathsf{D}}^{-1}(\mathcal{C}_{\mathsf{D}}(\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}}))\right) \\
&= (\mathbb{W}_{\mathsf{D}}(K), \dot{L}^1, \ldots, \dot{L}^k) - \mathcal{L}_{\mathsf{D}}(K) = (0, \dot{L}^1 - i[L^1, K], \ldots, \dot{L}^k - i[L^k, K]).
\end{aligned}$$

On the other hand, $\omega_{\mathsf{D}}(\dot{\mathsf{D}}) = (-iK, r)$ by definition, so using formula (14) of the push-forward D_*, we get

$$(\mathrm{Id} - \mathsf{D}_* \circ \omega_{\mathsf{D}})(\dot{\mathsf{D}}) = (\dot{H} - i[H, K] - r\mathbb{1}, \dot{L}^1 - i[L^1, K], \ldots, L^k - i[L^k, K]).$$

The crucial Equation (29) gives $\mathcal{E}_{\mathsf{D}}(\mathsf{D}_*(\omega_{\mathsf{D}}(\dot{\mathsf{D}}))) = \mathbb{W}_{\mathsf{D}}(K) + r\mathbb{1}$, and hence

$$\mathcal{E}_{\mathsf{D}}\left(\dot{\mathsf{D}} - \mathsf{D}_*(\omega_{\mathsf{D}}(\dot{\mathsf{D}}))\right) = \mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}}) - r\mathbb{1} - \mathbb{W}_{\mathsf{D}}(K) = \mathcal{C}_{\mathsf{D}}(\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}})) - \mathbb{W}_{\mathsf{D}}(K) = 0.$$

Consequently,

$$\mathbf{X}_{\mathsf{D}}\left(\dot{\mathsf{D}} - \mathsf{D}_*(\omega_{\mathsf{D}}(\dot{\mathsf{D}}))\right) = (0, \dot{L}^1 - i[L^1, K], \ldots, L^k - i[L^k, K]) = R_{\mathsf{D}}(\mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}})),$$

showing that $\mathbf{X}_{\mathsf{D}} \circ (\mathrm{Id} - \mathsf{D}_* \circ \omega_{\mathsf{D}}) = R_{\mathsf{D}} \circ \mathbf{X}_{\mathsf{D}}$. By the definition of P_{D} in (28), this implies that $P_{\mathsf{D}} = \mathrm{Id} - \mathsf{D}_* \circ \omega_{\mathsf{D}}$.

For a given $X = (-iK, r) \in \mathfrak{g}$ we can clearly find $\dot{\mathsf{D}}$ such that $r = \mathrm{tr}[\rho_{ss}^{\mathsf{D}}\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}})]$ and $\mathcal{C}_D(\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}})) = \mathbb{W}_{\mathsf{D}}(K)$, and hence the range of the one-form ω_{D} fills the whole Lie algebra. Furthermore, we can easily verify the compatibility condition (16), and G-covariance (17) follows from the fact that

$$\mathcal{E}_{g\mathsf{D}}(g_*(\dot{\mathsf{D}})) = W^*\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}})W, \quad \rho_{ss}^{g\mathsf{D}} = W^*\rho_{ss}^{\mathsf{D}}W, \quad g = (W, a) \in G, \tag{30}$$

which is straightforward to check. This completes the proof. □

With this result, we therefore achieve the aim described at the end of Section IV by *defining* the subspace of the identifiable directions to be the horizontal subspace,

$$\mathcal{T}_{\mathsf{D}}^{\mathrm{id}} := \operatorname{ran} P_{\mathsf{D}} = \{\dot{\mathsf{D}} \mid \mathcal{E}(\dot{\mathsf{D}}) = 0\}.$$

The associated split $\mathcal{T}_{\mathsf{D}} = \mathcal{T}_{\mathsf{D}}^{\mathrm{nonid}} \oplus \mathcal{T}_{\mathsf{D}}^{\mathrm{id}}$ now follows immediately from the general theory; in particular, the number of identifiable parameters is

$$\dim \mathcal{T}_{\mathsf{D}}^{\mathrm{id}} = 2kd^2.$$

As a consequence of the G-invariance of the horizontal projection, the form $(\cdot,\cdot)_{\mathsf{D}}$ on $\mathcal{T}_{\mathsf{D}}$ is G-invariant in the sense that

$$(\dot{\mathsf{D}}, \dot{\mathsf{D}}')_{\mathsf{D}} = (g_* \dot{\mathsf{D}}, g_* \dot{\mathsf{D}}')_{g\mathsf{D}}, \qquad \dot{\mathsf{D}}, \dot{\mathsf{D}}' \in \mathcal{T}_{\mathsf{D}}^{id}, g \in G. \tag{31}$$

Hence, this form only depends on the equivalence class, so its real part determines a unique bilinear form on the base manifold $\mathcal{P}$. Moreover, it also only depends on the horizontal projection $P_{\mathsf{D}}(\dot{\mathsf{D}})$ of the tangent vectors; hence it becomes nondegenerate on the horizontal bundle, thereby defining a Riemannian metric on the base manifold $\mathcal{P}$.

We emphasise that the principal connection (together with the stationary state) completely determines the metric and the associated Fisher information. In this way the connection provides geometric insight into how the (in practice rather complicated) expression of the Fisher information arises; for a discussion on a classical analogy, see, e.g., Ref. 35. We demonstrate this in a concrete example in Section VI.

D. Symplectic structure and CCR-algebra for identification

In Proposition 1 we defined the Markov covariance on the complex linear space $M(\mathbb{C}^d)^{k+1}$ and used the *real linear* maps,

$$\mathbf{X}_{\mathsf{D}} : \mathcal{T}_{\mathsf{D}} \to M(\mathbb{C}^d)^{k+1},$$

to induce an associated *real* inner product $(\cdot,\cdot)_{\mathsf{D}}$ on the identifiable part of the tangent space, cf. Equation (27); up to a constant factor, this inner product is the QFI rate. It is then natural to ask if the *imaginary part* of the Markov covariance has any physical interpretation. We will show that the latter can be used to define an algebra of the canonical commutation relations (CCR) over the real space of identifiable parameters $\mathcal{T}_{\mathsf{D}}^{id}$, which will play the role of limit Gaussian model in Sec. VII B.

On the real linear space $\mathcal{T}_{\mathsf{D}}^{\mathrm{id}} = \{\dot{\mathsf{D}} \mid \mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}}) = 0\} = \operatorname{ran} P_{\mathsf{D}}$ we now define a *complex structure* via

$$\begin{aligned} \mathcal{J}_{\mathsf{D}} &: \mathcal{T}_{\mathsf{D}}^{id} \to \mathcal{T}_{\mathsf{D}}^{id} \\ \mathcal{J}_{\mathsf{D}} &: (\dot{H}, \dot{L}^1, \ldots, \dot{L}^k) \mapsto \left(\sum_{i=1}^{k} \mathrm{Re}\dot{L}^{i*} L^i \,,\, i\dot{L}^1, \ldots, i\dot{L}^k \right). \end{aligned} \tag{32}$$

Using the property that $\mathcal{E}_{\mathsf{D}}(\dot{\mathsf{D}}) = 0$ for all vectors $\dot{\mathsf{D}} \in \mathcal{T}_{\mathsf{D}}^{id}$, it is easy to check that $\mathcal{J}_{\mathsf{D}}$ satisfies the defining property of a complex structure on $\mathcal{T}_{\mathsf{D}}^{id}$, i.e., $\mathcal{J}_{\mathsf{D}}^2 = -\mathrm{Id}$. Furthermore, since $P_{\mathsf{D}} = \mathbf{X}_{\mathsf{D}}^{-1} \circ R_{\mathsf{D}} \circ \mathbf{X}_{\mathsf{D}}$, we immediately see from Proposition 2 that

$$\mathbf{X}_{\mathsf{D}}[\mathcal{J}_{\mathsf{D}}(\dot{\mathsf{D}})] = i\mathbf{X}_{\mathsf{D}}[\dot{\mathsf{D}}] = (0, i\dot{L}^1, \ldots, i\dot{L}^k), \qquad \dot{\mathsf{D}} \in \mathcal{T}_{\mathsf{D}}^{id},$$

that is, the map $\mathbf{X}_{\mathsf{D}}$ is compatible with the natural complex structure of $M(\mathbb{C}^d)^{k+1}$. In fact, this is the *only* way of defining a complex structure on $\mathcal{T}_{\mathsf{D}}^{id}$ in such a way that the restriction of $\mathbf{X}_{\mathsf{D}}$ to $\mathcal{T}_{\mathsf{D}}^{id}$ is a *complex* linear map.

When endowed with the complex structure $\mathcal{J}_{\mathsf{D}}$, the space $\mathcal{T}_{\mathsf{D}}^{\mathrm{id}}$ becomes a complex linear space; this is Hilbert space with respect to the inner product induced by the Markov covariance,

$$(\dot{\mathsf{D}}, \dot{\mathsf{D}}')_{\mathsf{D}} := \left(\mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}}) \,,\, \mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}}') \right)_{\mathsf{D}}, \qquad \dot{\mathsf{D}}, \dot{\mathsf{D}}' \in \mathcal{T}_{\mathsf{D}}^{id}. \tag{33}$$

The real part of this form gives the Riemannian metric and QFI rate on the real linear tangent space $\mathcal{T}_{\mathsf{D}}^{\mathrm{id}}$ as discussed above. In addition, the imaginary part can be used to construct a representation of the canonical commutation relations (CCR) over $\mathcal{T}_{\mathsf{D}}^{id}$, together with a distinguished Fock state whose statistical interpretation is discussed in Section VII.

Definition 2 (CCR algebra for identifiable parameters). Let $(\mathcal{T}^{id}_{\mathsf{D}}, \mathcal{J}_{\mathsf{D}})$ *and* $(\dot{\mathsf{D}}, \dot{\mathsf{D}}')_{\mathsf{D}}$ *be the complex linear space, and, respectively, inner product defined above. On* $\mathcal{T}^{id}_{\mathsf{D}}$ *we define the symplectic form*

$$\sigma^{\mathsf{D}}(\dot{\mathsf{D}}, \dot{\mathsf{D}}') := \mathrm{Im}(\dot{\mathsf{D}}, \dot{\mathsf{D}}')_{\mathsf{D}} = \sum_{i=1}^{k} \mathrm{Im}\,\mathrm{tr}[\rho^{\mathsf{D}}_{ss} \dot{L}^*_i \dot{L}'_i].$$

We define the CCR algebra $CCR(\mathcal{T}^{id}_{\mathsf{D}}, \sigma^{\mathsf{D}})$ *generated by unitary Weyl operators* $W(\dot{\mathsf{D}})$ *with* $\dot{\mathsf{D}} \in \mathcal{T}^{id}_{\mathsf{D}}$ *satisfying the relations*

$$W(\dot{\mathsf{D}})W(\dot{\mathsf{D}}') = e^{i\sigma^{\mathsf{D}}(\dot{\mathsf{D}},\dot{\mathsf{D}}')} W(\dot{\mathsf{D}} + \dot{\mathsf{D}}'), \quad W(-\dot{\mathsf{D}}) = W(\dot{\mathsf{D}})^*.$$

On $CCR(\mathcal{T}^{id}_{\mathsf{D}}, \sigma^{\mathsf{D}})$ we define the Gaussian state φ determined by the characteristic function

$$\varphi(W(\dot{\mathsf{D}})) = e^{-\frac{1}{8} f^{\mathsf{D}}(\dot{\mathsf{D}},\dot{\mathsf{D}})},$$

where

$$f^{\mathsf{D}}(\dot{\mathsf{D}}, \dot{\mathsf{D}}') := 4\mathrm{Re}(\dot{\mathsf{D}}, \dot{\mathsf{D}}')_{\mathsf{D}} = 4\mathrm{Re} \sum_i \mathrm{tr}[\rho^{\mathsf{D}}_{ss} \dot{L}^*_i \dot{L}'_i]. \tag{34}$$

By a standard construction,[47] the CCR algebra can be represented on the Fock space $\mathcal{F}_{\mathsf{D}}$ over the Hilbert space $(\mathcal{T}^{id}_{\mathsf{D}}, \mathcal{J}_{\mathsf{D}}, (\dot{\mathsf{D}}, \dot{\mathsf{D}}')_{\mathsf{D}})$, in such a way that that $\varphi(W(\dot{\mathsf{D}})) = \langle 0|W(\dot{\mathsf{D}})|0\rangle$, where $|0\rangle \in \mathcal{F}_{\mathsf{D}}$ is the vacuum state, and the Weyl operators $W(\dot{\mathsf{D}})$ can be written in terms of canonical quadrature operators Q_j, P_j satisfying the Heisenberg form of the CCR $[Q_k, P_{k'}] = i\delta_{kk'}\mathbb{1}$.

In order to explicitly find such a representation, we need to choose a *symplectic basis* $\{\dot{D}_1, \ldots, \dot{D}_{2m}\}$ of $\mathcal{T}^{id}_{\mathsf{D}}$ (recall that $m = \dim \mathcal{T}^{id}_{\mathsf{D}} = 2kd^2$) which is also compatible with the complex structure. This means that $f^{\mathsf{D}}(\dot{D}_j, \dot{D}_{j'}) = \delta_{jj'}$ for all j, and $\mathcal{J}_{\mathsf{D}}(\dot{D}_{2j+1}) = \dot{D}_{2j+2}$, $\sigma^{\mathsf{D}}(\dot{D}_{2j+1}, \dot{D}_{2j'+1}) = \sigma^{\mathsf{D}}(\dot{D}_{2j+2}, \dot{D}_{2j'+2}) = 0$, $\sigma^{\mathsf{D}}(\dot{D}_{2j+1}, \dot{D}_{2j'+2}) = \delta_{jj'}$ for $j = 0, \ldots, m-1$. We then define the canonical operators Q_j, P_j as generators of the one-parameter groups $W(u\dot{D}_{2j+1}) = e^{-iuP_j}$ and $W(u\dot{D}_{2j+2}) = e^{iuQ_j}$. In the generic case, the basis will depend smoothly on the coordinates.

A fairly canonical choice for the basis is obtained by first defining the rank-1 matrices

$$E_{j;l} = \frac{1}{\sqrt{p^{\mathsf{D}}_l}} |j\rangle\langle\varphi^{\mathsf{D}}_l|,$$

where p^{D}_l and φ^{D}_l, $l = 1, \ldots, d$, are eigenvalues and eigenvectors of the stationary state, counted according to their multiplicities. We then define

$$\dot{D}_{i;j;l} := (-\mathrm{Im}E^*_{j,l} L^i, 0, \ldots, 0, E_{j;l}, 0, \ldots, 0) \in \mathcal{T}^{id}_{\mathsf{D}}$$

for each $i = 1, \ldots, k$ and $j, l = 1, \ldots, d$, where the nonzero element in the middle is at the ith place. Then $\mathcal{J}_{\mathsf{D}}(\dot{D}_{i;j;l}) = (\mathrm{Re}E^*_{j,l} L^i, 0, \ldots, 0, iE_{j;l}, 0, \ldots, 0)$, and it is easy to check that $\{\dot{D}_{i;j;l}, \mathcal{J}_{\mathsf{D}}(\dot{D}_{i;j;l})\}$ is a symplectic basis compatible with the complex structure, and depends smoothly on the coordinates (since the stationary state does), except possibly at some special points. However, the explicit expressions of these vectors are often rather lengthy and complicated. In the two-parameter example below we will demonstrate the geometry using a more tractable basis.

The main point of the above construction is that an arbitrary local parameter change can be associated with a linear combination of quadratures "generating" it in the effective continuous variable system. We expect that using local asymptotic normality (Section VII), together with a suitable variant of the central limit theorem (which will be the topic of a forthcoming publication), this observation generator can be used to find an optimal output measurement strategies for the system identification problem.

VI. EXAMPLES OF PARAMETRIC MODELS

To illustrate the general theory, we analyse several examples of one-parameter and multi-parameter models.

A. One parameter models

Let $\mathsf{D}_\theta := (H, e^{-i\theta}L)$ be a one-parameter family, where we have chosen $m = 1$ for simplicity. The corresponding one-dimensional tangent vector at $\mathsf{D}_{\theta=0}$ is $\dot{\mathsf{D}} = (0, iL)$. By applying Equation (18) we find that the corresponding generator has differential equation

$$dG_\theta(t) = \sum_{i=1}^{k} \left[j_t(L)dA^*(t) + j_t(L^*)dA(t) + j_t(L^*L)dt \right]. \tag{35}$$

Then $\mathcal{E}^0_{\mathsf{D}}(\dot{\mathsf{D}}) = (-L^*L + \langle L^*L\rangle_{ss}\mathbb{1})$, and $\mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}}) = (\mathcal{E}^0_{\mathsf{D}}(\dot{\mathsf{D}}), iL)$. The QFI rate is

$$f^\theta = 4\mathrm{tr}\left[\rho^{\mathsf{D}}_{ss}\left(L + [L, \mathbb{W}^{-1}(L^*L - \langle L^*L\rangle_{ss}\mathbb{1})]\right)^2\right].$$

Physically, this transformation can be implemented by placing a phase-shifter in each output channel, which gives each photon a phase shift $e^{i\theta}$.[44] This phase parameter is identifiable, and it is easy to see that

$$|\Psi^{\mathrm{s+o}}_\theta(t)\rangle = \exp(-i\theta N(t))|\Psi^{\mathrm{s+o}}(t)\rangle,$$

where $N(t)$ is the counting process associated with the number of photons up to time t in the Bosonic environment. Equivalently, this can be written as $U^*(t)|\Psi^{\mathrm{s+o}}_\theta(t)\rangle = \exp(-i\theta N^{\mathrm{out}}(t))|\phi \otimes \Omega\rangle$, where $N^{\mathrm{out}}(t) := U(t)^*N(t)U(t)$ is the output number of photons operator, whose differential form is

$$N^{\mathrm{out}}(t) = dN(t) + j_t(L)dA^*(t) + j_t(L^*)dA(t) + j_t(L^*L)dt. \tag{36}$$

By comparing (35) and (36), we see that the two generators are not identical. However, the difference is the term $dN(t)$ which annihilates the vacuum state, so the resulting *action* of the generators *is* identical. This illustrates that in general the generator is not unique but one can add terms which annihilate the vacuum, such as annihilation or number operator terms.

The second example we consider is that of the coupling constant, where $L_\theta = \theta L$, with unknown parameter $\theta \in \mathbb{R}$. The tangent vector is $\dot{\mathsf{D}} := (0, L)$ and $\mathcal{E}^0_{\mathsf{D}}(T) = 0$. Therefore, $\mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}}) = (0, L)$ and the QFI rate is

$$f = 4\mathrm{tr}\left[\rho^{\mathsf{D}}_{ss}L^*L\right],$$

which is simply the photon emission rate in the stationary regime.

In the third example, we consider the model where the hamiltonian is known up to a multiplicative constant $H_\theta = \theta H$. The tangent vector is $\dot{\mathsf{D}} := (H, 0)$ and $\mathcal{E}^0_{\mathsf{D}}(T) = H - \langle H\rangle_{ss}$. Therefore, $\mathbf{X}_{\mathsf{D}}(\dot{\mathsf{D}}) = (H - \langle H\rangle_{ss}, 0)$ and the QFI rate is

$$4\mathrm{tr}\left(\rho^{\mathsf{D}}_{ss}[L, \mathbb{W}^{-1}(H - \langle H\rangle_{ss})]^*[L, \mathbb{W}^{-1}(H - \langle H\rangle_{ss})]\right).$$

B. Simplest multiparameter setting

The geometric aspects are naturally trivial in a one-parameter model. In order to illustrate the full use of the theory developed above, we now consider the simplest nontrivial setting with $d = 2$ and $k = 1$, that is, $\mathcal{D}^{\mathrm{erg}}$ is the open subset of $\{(H, L) \mid H \in M_s(\mathbb{C}^2),\ L \in M(\mathbb{C}^2)\}$ consisting of ergodic dynamical parameters. The dimension of this manifold is 12, and the number of identifiable parameters is 8. Hence, full treatment of this simplest setting is still rather tedious, and we settle for looking at points on a physically relevant submanifold, extended suitably so as to allow for the full description of the relevant geometry. The model is the following 7-dimensional submanifold:

$$H_{\Delta,\Omega,\mathbf{v}} = \frac{1}{2}\begin{pmatrix} \Delta & \Omega + v_1 - iv_2 \\ \Omega + v_1 + iv_2 & -\Delta + v_0 \end{pmatrix}, \quad L_{\alpha,\theta,\mathbf{v}} = \alpha e^{i\theta}\begin{pmatrix} (iv_1 - v_2)/\alpha^2 & 1 + iv_0/\alpha^2 \\ 0 & (-iv_1 + v_2)/\alpha^2 \end{pmatrix}.$$

Here the three parameters v are auxiliary, and the rest have physical meaning at $\mathbf{v} = 0$. In fact, we are looking at the off-resonant laser-driven two-level system with Rabi frequency Ω and detuning Δ, in contact with a zero-temperature heat bath, with emission rate α^2 and emitted photons monitored in the environment. In addition, we include the above discussed phase shift θ to the emitted photons. The auxiliary parameters are chosen such that their tangent vectors lie in the identifiable subspace

at $\mathbf{v} = 0$; their span is needed in order to describe the horizontal projections of the physical tangent vectors, as we will see below.

The quantum Fisher information associated with the three parameters (Δ, Ω, α) of this model has been compared with particular measurement strategies;[22] we emphasise geometric aspects not discussed there and have also included the phase parameter θ. The main idea is to demonstrate how the rather complicated expressions of the Fisher information arise from considerably simpler geometric ingredients as a result of straightforward linear algebra. This provides insight into the structure of the physical system from the operational identification point of of view and may eventually be useful in developing global estimation strategies in analogy to classical cases (see, e.g., Ref. 35).

Accordingly, we let $\mathcal{D}_{\text{ext}}$ denote the whole (extended) 7-dimensional manifold and $\mathcal{D}_{\text{phys}} = \{\mathsf{D} \in \mathcal{D}_{\text{ext}} \mid \mathbf{v} = 0\}$ the physical submanifold. The dynamical parameters $\mathsf{D} \in \mathcal{D}_{\text{phys}}$ are ergodic except at special points; the unique stationary state is

$$\rho_{ss} = \frac{\Omega}{\gamma} \begin{pmatrix} \gamma/\Omega - \Omega & \xi \\ \bar{\xi} & \Omega \end{pmatrix},$$

where $\gamma = \alpha^4 + 4\Delta^2 + 2\Omega^2$ and $\xi = 2\Delta + i\alpha^2$. In the following we only consider points $\mathsf{D} \in \mathcal{D}_{\text{phys}}$.

1. The Lie algebra and unidentifiable directions

We begin the description of the geometry by finding the unidentifiable part of the tangent space of the physical manifold (Fig. 4). We let E_{ij} denote the natural basis matrices of $M(\mathbb{C}^2)$, so that, e.g., $\sigma_z = E_{00} - E_{11}$. Note that matrices such as E_{ij} and iE_{ij} are linearly independent since we look at the *real linear* version of $M(\mathbb{C}^2)$. We parametrise the Lie algebra by

$$\mathfrak{g} = \{X[\mathbf{w}, r] \mid \mathbf{w} \in \mathbb{R}^3, r \in \mathbb{R}\},$$

where $X[\mathbf{w}, r] = (-iK_{\mathbf{w}}, r)$, with $K_{\mathbf{w}} = f(\mathbf{w})\mathbb{1} + \mathbf{w} \cdot \sigma$, and the "gauge" function $f(\mathbf{w})$ is irrelevant to the action of Lie algebra on the parameter manifold. As discussed above, this gauge can be fixed so that $K_{\mathbf{w}}$ has zero mean; this is convenient since the back-action given by the connection one-form will automatically have this gauge. The zero mean gauge for the Lie algebra is

$$f(\mathbf{w}) = -\gamma^{-1}\left(4\Delta(w_1\Omega + \Delta w_3) - 2\alpha^2 w_2 \Omega + \alpha^4 w_3\right).$$

In order to describe the action of the Lie algebra on the tangent space, we compute explicitly the image of the push-forward D_*, corresponding to the unidentifiable part of the tangent space. Omitting the

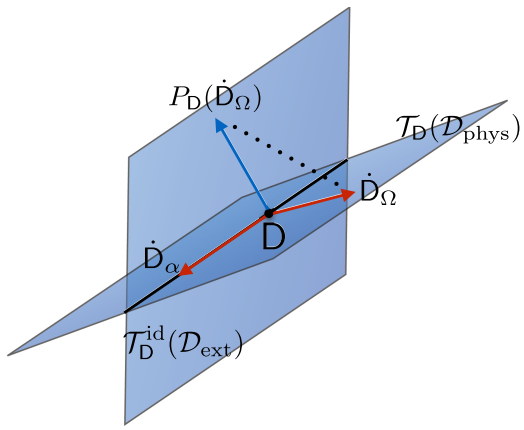

FIG. 4. Sketch of the information geometry of two parameters in the driven two-level system. The tangent space of the physical manifold contains the directions of the decoupling parameter α and the driving frequency Ω (in red). Only the α-direction lies in the identifiable subspace which supports the Fisher information metric; Ω-direction needs to be projected there via the horizontal projection P_{D} of the principal connection.

subscript for simplicity, the tangent vectors induced by basic rotations and the unidentifiable phase are

$$\dot{\mathsf{D}}_x^{nonid} := \mathsf{D}_*(X[(1,0,0),0]) = (-\Delta\sigma_y, i\alpha e^{i\theta}\sigma_z), \quad \dot{\mathsf{D}}_y^{nonid} := \mathsf{D}_*(X[(0,1,0),0]) = (\Delta\sigma_x - \Omega\sigma_z, -\alpha e^{i\theta}\sigma_z),$$
$$\dot{\mathsf{D}}_z^{nonid} := \mathsf{D}_*(X[(0,0,1),0]) = (\Omega\sigma_y, -2\alpha i e^{i\theta}E_{01}), \quad \dot{\mathsf{D}}_{phase}^{nonid} := \mathsf{D}_*(X[(0,0,0),1]) = (\mathbb{1}, 0).$$

2. Symplectic structure of the identifiable subspace

The 8-dimensional identifiable subspace supports the principal connection and Markov covariance. It is characterised by the condition $\mathcal{E}(\dot{\mathsf{D}}) = 0$. Using this condition, one easily finds the following basis for this subspace:

$$\dot{D}_1 = (0, e^{i\theta}E_{01}/\alpha), \qquad \dot{D}_3 = (\sigma_x/2, ie^{i\theta}\sigma_z/\alpha), \qquad \dot{D}_5 = (0, E_{10}), \qquad \dot{D}_7 = (0, E_{11}),$$
$$\dot{D}_2 = (E_{11}, ie^{i\theta}E_{01}/\alpha), \qquad \dot{D}_4 = (\sigma_y/2, -e^{i\theta}\sigma_z/\alpha), \qquad \dot{D}_6 = (0, iE_{10}), \qquad \dot{D}_8 = (0, iE_{11}).$$

This basis is compatible with the complex structure in the sense that $\mathcal{J}_{\mathsf{D}}(\dot{D}_{2k+1}) = \dot{D}_{2k+2}$ for $k = 0, \ldots, 3$. We observe that the first pair on the left depends on the parameters, the second also exhibits nontrivial dependencies between the $\dot{H}$ and $\dot{L}$ matrices, and the pairs on the right are trivial. In fact, $\dot{D}_2, \dot{D}_3, \dot{D}_4$ are exactly the tangent vectors of the auxiliary parameters v, and it turns out that the vectors $\dot{D}_5, \ldots, \dot{D}_8$ are irrelevant for the physical model. The Markov covariance can be directly computed on this part of of the identifiable subspace; the corresponding matrix M is given by

$$M_{ij} := (\dot{D}_i, \dot{D}_j)_{\mathsf{D}} = \mathrm{tr}[\rho_{ss}^{\mathsf{D}} \dot{L}_i^* \dot{L}_j],$$

where j indexes the basis vectors. However, a better choice is to replace the first four vectors by

$$\dot{D}_1^{\mathrm{sym}} = \frac{\alpha\sqrt{\gamma}}{\Omega}\left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \frac{e^{i\theta}}{\alpha}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right), \quad \dot{D}_3^{\mathrm{sym}} = \frac{\alpha\sqrt{\gamma}}{\sqrt{2}\Omega^2}\left(\begin{pmatrix} 0 & -i\Omega/2 \\ i\Omega/2 & \alpha^2 \end{pmatrix}, \frac{e^{i\theta}}{\alpha}\begin{pmatrix} -\Omega & \xi \\ 0 & \Omega \end{pmatrix}\right),$$
$$\dot{D}_2^{\mathrm{sym}} = \frac{\alpha\sqrt{\gamma}}{\Omega}\left(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \frac{e^{i\theta}}{\alpha}\begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix}\right), \quad \dot{D}_4^{\mathrm{sym}} = \frac{\alpha\sqrt{\gamma}}{\sqrt{2}\Omega^2}\left(\begin{pmatrix} 0 & -\Omega/2 \\ \Omega/2 & 2\Delta \end{pmatrix}, \frac{e^{i\theta}}{\alpha}\begin{pmatrix} -i\Omega & i\xi \\ 0 & i\Omega \end{pmatrix}\right).$$

Now $\mathcal{J}_{\mathsf{D}}(\dot{D}_{2k+1}^{\mathrm{sym}}) = \dot{D}_{2k+2}^{\mathrm{sym}}$ for $k = 0, 1$ still holds, but in addition the Markov covariance (restricted to the relevant subspace) is in the *standard form*

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + i\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

This means that we have a *symplectic basis* which is also compatible with the complex structure. In particular, the two vectors $\dot{D}_1^{\mathrm{sym}}$ and $\dot{D}_3^{\mathrm{sym}}$, regarded as complex vectors (in the sense of their second component matrices), while *orthonormal* with respect to the Markov covariance, and $\dot{D}_2^{\mathrm{sym}} = \mathcal{J}_{\mathsf{D}}(\dot{D}_1^{\mathrm{sym}})$ and $\dot{D}_4^{\mathrm{sym}} = \mathcal{J}_{\mathsf{D}}(\dot{D}_3^{\mathrm{sym}})$ is obtained from them by multiplying the second component matrices by i. Observe that the relationship between the first component matrices is *not* simply a multiplication by i; the induced complex structure $\mathcal{J}_{\mathsf{D}}$ is nontrivial on this part.

In the CCR-algebra, the pairs $(\dot{D}_1^{\mathrm{sym}}, \dot{D}_2^{\mathrm{sym}})$ and $(\dot{D}_3^{\mathrm{sym}}, \dot{D}_4^{\mathrm{sym}})$ correspond to pairs of quadratures (Q_1, P_1) and (Q_2, P_2) satisfying the canonical commutation relations $[Q_k, P_l] = i\delta_{kl}\mathbb{1}$. In fact, the correspondence is given by the Weyl operators $W(u\dot{D}_1^{\mathrm{sym}}) = e^{-iuP_1}$, $W(u\dot{D}_2^{\mathrm{sym}}) = e^{iuQ_1}$, $W(u\dot{D}_3^{\mathrm{sym}}) = e^{-iuP_2}$, and $W(u\dot{D}_4^{\mathrm{sym}}) = e^{iuQ_2}$, $u \in \mathbb{R}$. This means, for instance, that parameter changes along the tangent vector $\dot{D}_1^{\mathrm{sym}}$ are generated by the "momentum operator" P_1 in the effective continuous variable system. A generating operator for any tangent vector can be found by writing the tangent vector in this basis.

3. The connection on the physical manifold

In order to investigate the geometry of the physical manifold, we first find the tangent space $\mathcal{T}(\mathcal{D}_{\mathrm{phys}}) \subset \mathcal{T}$, consisting of meaningful directions in the model. It is the span of the following tangent

vectors:

$$\dot{D}_\Omega = (\tfrac{1}{2}\sigma_x, 0), \qquad \dot{D}_\Delta = (\tfrac{1}{2}\sigma_z, 0),$$
$$\dot{D}_\alpha = (0, e^{i\theta}E_{01}), \qquad \dot{D}_\theta = \alpha(0, ie^{i\theta}E_{01}).$$

These vectors span a 4-dimensional subspace of $\mathcal{T}$; note that the dependence on the manifold point only comes with the phase parameter θ. We can now determine the connection one-form on the physical manifold; this is a straightforward computation involving the inversion of the generator $\mathbb{W}$ on the zero-mean subspace. The result is

$$\omega = \omega(\dot{D}_\Delta)d\Delta + \omega(\dot{D}_\Omega)d\Omega + \omega(\dot{D}_\alpha)d\alpha + \omega(\dot{D}_\theta)d\theta,$$

where the components are given by

$$\begin{aligned}
\omega(\dot{D}_\Delta) &= \gamma^{-1}X[-\alpha^{-2}(4\Delta\Omega, 2\Omega\alpha^2, |\xi|^2), |\xi|^2/2],\\
\omega(\dot{D}_\Omega) &= 2\gamma^{-1}X[-\alpha^{-2}(\alpha^4 + 2\Omega^2, -2\Delta\alpha^2, 2\Delta\Omega), \Delta\Omega],\\
\omega(\dot{D}_\alpha) &= 0,\\
\omega(\dot{D}_\theta) &= -\gamma^{-1}X[(4\Delta\Omega, 2\alpha^2\Omega, |\xi|^2), \alpha^2\Omega^2].
\end{aligned}$$

Similarly, we could determine the connection on the extended manifold; however, the result is considerably more complicated and is not very illuminating. Using the above components together with the push-forward D_*, we obtain the horizontal projection of the physical tangent space on the identifiable subspace; a direct computation shows that it coincides with the above four-dimensional subspace, with components in the above chosen symplectic basis given by

$$\begin{aligned}
P(\dot{D}_\Delta) &= \frac{\Omega}{\alpha\gamma^{\frac{3}{2}}}\begin{pmatrix} -4\alpha^2\Delta \\ -2\alpha^4 \end{pmatrix} \oplus \begin{pmatrix} \sqrt{2}\alpha^2\Omega \\ -2\sqrt{2}\Delta\Omega \end{pmatrix},\\
P(\dot{D}_\Omega) &= \frac{1}{\alpha\gamma^{\frac{3}{2}}}\begin{pmatrix} -\alpha^2\left(\gamma - 8\Delta^2\right) \\ 4\alpha^4\Delta \end{pmatrix} \oplus \begin{pmatrix} -2\sqrt{2}\alpha^2\Delta\Omega \\ -\sqrt{2}\Omega\left(\alpha^4 + 2\Omega^2\right) \end{pmatrix},\\
P(\dot{D}_\alpha) &= \frac{\Omega}{\sqrt{\gamma}}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \oplus \begin{pmatrix} 0 \\ 0 \end{pmatrix},\\
P(\dot{D}_\theta) &= \frac{\alpha\Omega}{\gamma^{\frac{3}{2}}}\begin{pmatrix} -4\alpha^2\Delta \\ \gamma - 2\alpha^4 \end{pmatrix} \oplus \begin{pmatrix} \sqrt{2}\alpha^2\Omega \\ -2\sqrt{2}\Delta\Omega \end{pmatrix}.
\end{aligned}$$

4. The Fisher information of the physical parameters

We can now easily find the QFI rate for any tangent vector written in the symplectic basis, by simply computing its norm. This gives

$$f_\Delta = 4\frac{2\Omega^2|\xi|^2\left(2\alpha^4 + \Omega^2\right)}{\alpha^2\gamma^3}, f_\Omega = 4\frac{\alpha^{12} + \alpha^8\left(8\Delta^2 + 6\Omega^2\right) + 4\alpha^4\left(4\Delta^4 - 2\Delta^2\Omega^2 + 3\Omega^4\right) + 8\Omega^6}{\alpha^2\gamma^3},$$
$$f_\alpha = 4\frac{\Omega^2}{\gamma}, \qquad f_\theta = 4\frac{\alpha^2\Omega^2\left(-2\Omega^2\left(\alpha^4 - 12\Delta^2\right) + |\xi|^4 + 4\Omega^4\right)}{\gamma^3}.$$

Note that the reason why these expressions are rather complicated is partially due to the fact that the physical directions do not lie in the identifiable subspace, but need to be projected there.

5. The canonical coordinates of the physical parameters

This can be read off from the components of the above column vectors; for instance, at $\Delta = 0$ (resonance) we have

$$W(P(\dot{D}_\Delta)) = \exp -i\frac{\alpha\Omega}{\gamma^{\frac{3}{2}}}\left(2\alpha^2 Q_1 + \sqrt{2}\Omega P_2\right), W(P(\dot{D}_\Omega)) = \exp\frac{1}{\alpha\gamma^{\frac{3}{2}}}\left(\alpha^2\gamma P_1 - \sqrt{2}\Omega\left(\alpha^4 + 2\Omega^2\right)Q_2\right),$$
$$W(P(\dot{D}_\alpha)) = \exp -i\frac{\Omega}{\sqrt{\gamma}}P_1, \qquad W(P(\dot{D}_\theta)) = \exp i\frac{\alpha\Omega}{\gamma^{\frac{3}{2}}}\left((\gamma - 2\alpha^4)Q_1 - \sqrt{2}\alpha^2\Omega P_2\right),$$

that is, change in each physical parameter is generated by a linear combination of at most two canonical quadratures.

Using the connection form, one could further investigate the global structure of the information geometry, in terms of the curvature, geodesics and parallel transport. This would be relevant for some of the future lines of research mentioned in the introduction, but beyond the scope of this paper. We only note that for instance the curvature two-form can easily be determined by straightforward although somewhat tedious computer algebra; this shows in particular that the connection is not flat, i.e., the horizontal bundle is not integrable.

VII. LOCAL ASYMPTOTIC NORMALITY IN THE MULTIPARAMETER SETTING

In Subsection V A we showed that the quantum Fisher information of the output state increases linearly in time as $F^\theta(t) \approx t f^\theta$, and we identified the QFI rate f^θ as the real part of the Markov covariance matrix of tangent vectors corresponding to changes in the parameter θ, cf. Equations (22) and (26). In this section we extend the statistical analysis by proving that the *output state is asymptotically Gaussian* in the limit of large times, in a sense which will be defined precisely below. In effect this means that the output states for parameters in a local neighbourhood of a given dynamical parameter D_0 can be approximated by a limit model which consists of a family of pure Gaussian states of the CCR algebra $CCR(\mathcal{T}^{id}_{\mathsf{D}_0}, \sigma^{\mathsf{D}_0})$ defined above, with mean determined by local changes in the unknown parameter, and covariance given by the QFI rate. Before stating the asymptotic normality result, we briefly review the general statistical concepts involved in its formulation. For more details about the general theory of quantum statistical models we refer to Refs. 29 and 31.

A. Convergence of quantum statistical models

A *quantum statistical model* over the parameter space $\Theta \subset \mathbb{R}^k$ is a family $\mathcal{Q} := \{\rho^\theta : \theta \in \Theta\}$ of quantum states on a fixed Hilbert space $\mathcal{H}$, which are indexed by an unknown parameter $\theta \in \Theta$. We are interested in characterising the asymptotic behaviour of an *ordered set* of statistical models, in particular the convergence to a limit model. Such problems arise in quantum state estimation where the statistical models consist of ensembles of identically prepared systems, and the order parameter is the size of the ensemble,[40] or in the estimation of dynamical parameters (system identification) where time plays the role of "sample size." The latter case is the topic of this paper.

We start by noting that the space of statistical models is equipped with a natural notion of equivalence. Two models $\mathcal{Q}_1 := \{\rho_1^\theta : \theta \in \Theta\}$ and $\mathcal{Q}_2 := \{\rho_2^\theta : \theta \in \Theta\}$ (possibly on different Hilbert spaces) are statistically equivalent if there exist quantum channels T, S between the appropriate state spaces such that

$$T(\rho_1^\theta) = \rho_2^\theta, \qquad S(\rho_2^\theta) = \rho_1^\theta \tag{37}$$

for all $\theta \in \Theta$. A consequence of the equivalence is that the probability distribution of any measurement M on $\mathcal{Q}_1$ can be reproduced by a measurement on $\mathcal{Q}_2$ obtained by applying S followed by M and vice versa. Therefore the two models have exactly the same optimal risks (figures of merit) for any statistical problem concerning the parameter θ. In the special case, when $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are pure state models, it can be shown[15] that the models are equivalent if and only if there exist representative vectors (i.e. $\rho_1^\theta = |\psi_1^\theta\rangle\langle\psi_1^\theta|$, $\rho_2^\theta = |\psi_2^\theta\rangle\langle\psi_2^\theta|$) such that the overlaps of all pairs of vectors in the two models coincide

$$\langle\psi_1^\theta|\psi_1^{\theta'}\rangle = \langle\psi_2^\theta|\psi_2^{\theta'}\rangle, \qquad \theta, \theta' \in \Theta.$$

This shows that the intrinsic statistical properties of the model are encoded in the overlaps, up to an ambiguity in choosing the phases.

More generally, the theory of *quantum sufficiency*[48] deals with the situation when the models are related by the channel transformation (37) only in one direction, so that one of the models is more informative that the other. However, such a relationship is still rather restrictive; in asymptotic statistics one is often interested in approximating a given model by a "simpler" one which is "close" to it in a statistical sense. The above discussion suggests two ways of formalising this idea. The first one is to define a notion of distance between models,[29] inspired by the classical theory developed by Le Cam.[41]

Definition 3. Let $\mathcal{Q}_1$ and $\mathcal{Q}_2$ be two quantum statistical models over Θ, defined as above. The deficiencies of one model with respect to the other are defined as

$$\delta(\mathcal{Q}_1, \mathcal{Q}_2) = \inf_T \sup_{\theta\in\Theta} \|T(\rho_1^\theta) - \rho_2^\theta\|_1, \qquad \delta(\mathcal{Q}_2, \mathcal{Q}_1) = \inf_S \sup_{\theta\in\Theta} \|S(\rho_2^\theta) - \rho_1^\theta\|_1,$$

where the infima are taken over all quantum channels between the appropriated spaces, and the distance is given by the trace-norm $\|\tau\|_1 := \mathrm{Tr}(|\tau|)$. The Le Cam distance between the models $\mathcal{Q}_1$ and $\mathcal{Q}_2$ is defined as $\Delta(\mathcal{Q}_1, \mathcal{Q}_2) = \max(\delta(\mathcal{Q}_1, \mathcal{Q}_2), \delta(\mathcal{Q}_2, \mathcal{Q}_1))$.

A set of model $\mathcal{Q}_t := \{\rho_t^\theta : \theta \in \Theta\}$ indexed by t in $\mathbb{N}$ or $\mathbb{R}$ converges strongly (or in the sense of Le Cam) to a limit model $\mathcal{Q} := \{\rho^\theta : \theta \in \Theta\}$ if $\Delta(\mathcal{Q}, \mathcal{Q}_t) \to 0$ as $t \to \infty$.

It can be shown that two models are equivalent if and only if the Le Cam distance between them is zero. More generally, the Le Cam distance provides an upper bound to the the difference between optimal risks of statistical decision problems with bounded loss functions.[29] Furthermore, the convergence to a simpler limit model can be used to identify asymptotically optimal measurement procedures for a given statistical decision problem, e.g., state estimation. This can be done by mapping the state ρ_t^θ through the channel T_t onto the space of the limit model, followed by applying the optimal measurement for the limit model. An instance of this the phenomenon is local asymptotic normality for state estimation[40] which we illustrate below in the simplified setup of pure states. For this we formulate the second notion of convergence of models, based on the fidelity of the state vectors.

*Definition 4 (weak convergence of pure states statistical models). Let $\mathcal{Q}_t := \{\rho_t^\theta : \theta \in \Theta\}$ be a set of pure states quantum statistical models on Hilbert spaces $\mathcal{H}_t$ over parameter space $\Theta \subset \mathbb{R}^k$, where the index t is chosen from $\mathbb{N}$ or $\mathbb{R}$. The family $\mathcal{Q}_t$ is said to converge weakly to a model $\mathcal{Q} := \{\rho^\theta : \theta \in \Theta\}$ on a Hilbert space $\mathcal{H}$, if there exists a choice of representative vectors (*i.e. $\rho_t^\theta = |\psi_t^\theta\rangle\langle\psi_t^\theta|$, $\rho_t^\theta = |\psi_t^\theta\rangle\langle\psi_t^\theta|$*) such that*

$$\lim_{t\to\infty} \langle\psi_t^\theta|\psi_t^{\theta'}\rangle = \langle\psi^\theta|\psi^{\theta'}\rangle, \qquad \theta, \theta' \in \Theta.$$

Given that each statistical model is completely determined by the overlaps of pairs of vectors with different parameters, the definition captures the intuitive idea that two models are "close" to each other if they have similar overlaps. As a simple multidimensional example we consider the weak convergence of ensembles of identically prepared qubits to coherent states of a one mode continuous variables system, which is closely related to the theory of coherent spin states.[49] Let

$$|\psi_n^u\rangle = \left[\exp\left(\frac{i}{\sqrt{2n}}\left(u_y\sigma_x - u_x\sigma_y\right)\right)|0\rangle\right]^{\otimes n}, \qquad u = (u_x, u_y) \in \mathbb{R}^2$$

be a 2-dimensional family of i.i.d. qubit states obtained by rotating the basis vector $|0\rangle$ with generators given by the Pauli matrices σ_x, σ_y. Since the ensemble has size n, the statistical uncertainty in estimating rotation parameters is of the order of $n^{-1/2}$. It is then meaningful to restrict the attention to a shrinking region in the parameter space and write the rotation parameters as $u/\sqrt{n}$.[30] Due to the rescaling, the QFI of the "local parameter" u is a constant 2×2 matrix $f = 2\mathbb{1}_2$ which plays a similar role to the QFI rate per unit of time defined in Equations (22) and (26). We will now show that the sequence of local models $\mathcal{Q}_n = \{|\psi_n^u\rangle : u \in \mathbb{R}^2\}$ converges weakly to the quantum Gaussian model $\mathcal{Q} = \{|u\rangle : u \in \mathbb{R}^2\}$, where $|u\rangle$ denotes the coherent state of a one mode continuous variables system with mean values for the canonical variables given by $\langle Q\rangle = u_x, \langle P\rangle = u_y$. Indeed, since $\langle 0|\sigma_x|0\rangle = \langle 0|\sigma_y|0\rangle = 0$, by expanding in powers of $n^{-1/2}$ we obtain

$$\begin{aligned}\lim_{n\to\infty} \langle\psi_n^u|\psi_n^v\rangle &= \lim_{n\to\infty}\left(1 - \frac{1}{4n}\langle 0|(u_y\sigma_x - u_x\sigma_y)^2|0\rangle + o(n^{-1})\right)^n \\ &= \exp(\|u - v\|^2/4) = \langle u|v\rangle, \qquad u, v \in \mathbb{R}^2.\end{aligned}$$

In particular, the limit model has QFI equal to $f = 2\mathbb{1}_2$ which is the inverse of the covariance of the vacuum state. Furthermore, one can show that the convergence holds also in the stronger sense of Le Cam, so that optimal estimation procedures for the limit Gaussian model can be "pulled back" to asymptotically optimal measurements for the n qubits ensemble. When the figure of merit (or risk) is

the mean square error $\mathbb{E}(\|\hat{u}-u\|^2)$, the optimal measurement for estimating u in the limit model is the heterodyne measurement; this can be seen as a noisy joint measurement of the canonical variables Q and P and it outcome $\hat{u}$ is an unbiased estimator of u which has Gaussian distribution $N(u, \mathbb{1})$. The variance of $\hat{u}$ can be written as $V = f^{-1} + \frac{1}{2}\mathbb{1}$ where the first term comes from the quantum covariance while the second is the minimum amount of "noise" required for the simultaneous estimation of the means of the non-commuting observables Q and P. Moreover, the estimator is normally distributed, which allows one to devise confidence regions for large n. By a central limit argument, one can show that Q and P are the are the appropriately rescaled limits of the total spin observables L_x and L_y so that the optimal measurement is essentially a joint measurement of collective spin observables.

As we will see below, the key features of the i.i.d. qubit model are also present in the more complicated Markovian output setup, which we now proceed to consider.

B. Multiparameter LAN for quantum Markov processes

We start by considering a completely general model in which all identifiable parameters are unknown, and show how this model can be approximated locally by a Gaussian model on the CCR algebra of Definition 2. This result can then be applied to the situation where some prior information is available and we deal with a lower dimensional model.

1. Estimation of identifiable parameters

We will consider that the physical dynamics is governed by an unknown dynamical parameter D; however, since the latter cannot be completely identified from the stationary output state, we will focus on the estimation of *all identifiable parameters* given by the equivalence classes $[\mathsf{D}] \in \mathcal{P}$. Similarly to the i.i.d. setup described in Section VII A, we will be interested in the properties of the quantum output statistical model in the limit of large times. It is then meaningful to consider parameters $[\mathsf{D}]$ lying in a shrinking neighbourhood of a fixed point $[\mathsf{D}_0]$ in $\mathcal{P}$, whose size is of the order of the statistical uncertainty $t^{-1/2}$. We will formulate two convergence results: the first one concerns the weak convergence of the system-output state, while the second deals with the strong convergence of the output state. Since the latter depends only on the equivalence class $[\mathsf{D}]$, the strong convergence can be formulated solely in terms of the parameter space $\mathcal{P} = \mathcal{D}^{\mathrm{erg}}/G$. On the other hand, since the system-output state is *not* invariant over equivalence classes, the weak convergence depends on the specific choice of dynamical parameters for each equivalence class. Geometrically, this choice is determined by a *section* of the principal bundle, i.e., a smooth map $s : \mathcal{P} \to \mathcal{D}^{\mathrm{erg}}$ such that $\pi \circ s([\mathsf{D}]) = [\mathsf{D}]$ for $[\mathsf{D}]$ in a local neighbourhood of $[\mathsf{D}_0]$. We will assume that s is "horizontal" in the sense that the tangent space to $s(\mathcal{P})$ at D_0 is the horizontal space $\mathcal{T}^{id}_{\mathsf{D}_0}$. The intuition here is that the changes along equivalence classes of dynamical parameters are not observable in the output state, while those along tangent vectors in $\mathcal{T}^{id}_{\mathsf{D}_0}$ describe all the identifiable parameters. Although the theory can be developed in a coordinate-free way, for concreteness we consider a local coordinates chart in a neighbourhood of $[\mathsf{D}_0]$ defined by

$$\mathcal{C} : \mathcal{P} \to \mathcal{O} \subset \mathbb{R}^{\delta^{id}},$$

where $\mathcal{O}$ is a open ball centred at the origin, and $\mathcal{C}([\mathsf{D}_0]) = 0$. For simplicity we denote the parameter with coordinate u by $[\mathsf{D}]_u$ and the corresponding "lifted" dynamical parameter by $\mathsf{D}_u := s([\mathsf{D}]_u)$. The tangent vectors

$$[\dot{D}]_a := \left.\frac{\partial [\mathsf{D}]_u}{\partial u_a}\right|_{u=0}, \qquad \dot{D}_a := \left.\frac{\partial \mathsf{D}_u}{\partial u_a}\right|_{u=0} \qquad a = 1, \ldots, \delta^{id}$$

form a basis of the space $\mathcal{T}_{[\mathsf{D}_0]}$, and respectively $\mathcal{T}^{id}_{\mathsf{D}_0}$. With these notations we define two local statistical models corresponding to the system-output state and respectively the output state at time t.

Definition 5. Let $s, \mathcal{C}, [\mathsf{D}]_u \in \mathcal{P}, \mathsf{D}_u \in \mathcal{D}^{\mathrm{erg}}$ be define as above with coordinate $u \in \mathcal{O} \subset \mathbb{R}^{\delta^{id}}$ in a neighbourhood of the origin. The quantum statistical models of system-output state and respectively

the stationary output state at time t are defined by

$$\mathcal{Q}_t := \left\{ \left| \Psi^{\mathrm{s+o}}_{u/\sqrt{t}}(t) \right\rangle : u \in \mathcal{O} \subset \mathbb{R}^{\delta^{id}} \right\}, \quad \tilde{\mathcal{Q}}_t := \left\{ \rho^{\mathrm{out}}_{u/\sqrt{t}}(t) : u \in \mathcal{O} \subset \mathbb{R}^{\delta^{id}} \right\}.$$

with dynamics generated by $\mathsf{D}_{u/\sqrt{t}}$. *Furthermore, we define the (pure states) Gaussian model*

$$\mathcal{G} := \left\{ \rho_u := |u\rangle\langle u| : u \in \mathcal{O} \subset \mathbb{R}^{\delta^{id}} \right\}$$

where $|u\rangle = W(\sum_a u_a \dot{\mathsf{D}}_a)|\Omega\rangle$ *is the coherent state of the CCR algebra* $CCR(\mathcal{T}^{id}_{\mathsf{D}_0}, \sigma^{\mathsf{D}_0})$, cf. *Definition 2*.

The overlaps of the coherent states $|u\rangle$ can be computed from Definition 2 and are given by

$$\langle u|u'\rangle = \exp\left(-\frac{1}{8}(u-u')^T f^{\mathsf{D}_0}(u-u') + iu^T \sigma^{\mathsf{D}_0} u' \right). \tag{38}$$

From this one can deduce that the Gaussian model $\mathcal{G}$ has quantum Fisher information f^{D_0}, equal to the QFI rate of the system-output model $\mathcal{Q}_t$. The following theorem explains this connection by showing that the system-output and respectively output local models converge the to the Gaussian limit model. From the practical viewpoint, this means that the linear QFI scaling with rate f^{D_0} is *asymptotically achievable*, and moreover, the optimal measurement has asymptotically Gaussian distribution, cf. Ref. 40 for a detailed discussion of the interpretation of local asymptotic normality.

Theorem 2 (local asymptotic normality). *Let* $\mathcal{Q}_t, \tilde{\mathcal{Q}}_t, \mathcal{G}$ *be the system-output, output, and Gaussian models introduced in Definition 5. The following statements hold.*

(1) The pure state models $\mathcal{Q}_t$ *converge weakly to the Gaussian model* $\mathcal{G}$. *More precisely, there exists a particular choice of the (unphysical) phase angle* $\phi(u)$ *of the coherent state* $|u\rangle$ *such that*

$$\lim_{t\to\infty} \left\langle \Psi^{\mathrm{s+o}}_{u/\sqrt{t}}(t) \middle| \Psi^{\mathrm{s+o}}_{u'/\sqrt{t}}(t) \right\rangle = e^{i\phi(u')-i\phi(u)} \langle u|u'\rangle, \qquad u, u' \in \mathcal{O} \subset \mathbb{R}^{\delta^{id}}. \tag{39}$$

(2) The mixed state models $\tilde{\mathcal{Q}}_t$ *converge strongly to the the Gaussian model* $\mathcal{G}$, *i.e.,* $\Delta(\tilde{\mathcal{Q}}_t, \mathcal{G}) \to 0$. *More precisely, there exist quantum channels* T_t, S_t *such that*

$$\lim_{t\to\infty} \sup_{u\in\mathcal{O}} \left\| T_t\left(\rho^{\mathrm{out}}_{u/\sqrt{t}}(t)\right) - \rho_u \right\|_1 = 0,$$
$$\lim_{t\to\infty} \sup_{u\in\mathcal{O}} \left\| S_t\left(\rho_u\right) - \rho^{\mathrm{out}}_{u/\sqrt{t}}(t) \right\|_1 = 0.$$

In the reminder of this section we give the main idea of the proof and discuss the physical interpretation. The technical details can be found in the Appendix. Recall that the system-output state is given by $|\Psi^{\mathrm{s+o}}_{\mathsf{D}}(t)\rangle = U_{\mathsf{D}}(t)|\varphi\rangle \otimes |\Omega\rangle$ where $U_{\mathsf{D}}(t)$ is the unitary defined by the QSDE (4). By using Ito calculus it can be shown[13] that the overlaps of system-output states for different dynamical parameters can be expressed in terms of a contractive (non-CP) semigroup

$$\left\langle \Psi^{\mathrm{s+o}}_{\mathsf{D}}(t) \middle| \Psi^{\mathrm{s+o}}_{\mathsf{D}'}(t) \right\rangle = \langle \varphi | \, e^{t\mathbb{W}_{\mathsf{D},\mathsf{D}'}}(\mathbb{1}) \, |\varphi\rangle \, ,$$

where $W_{\mathsf{D},\mathsf{D}'}$ is the "off-diagonal" semigroup generator

$$\mathbb{W}_{\mathsf{D},\mathsf{D}'}(X) = i(HX - XH') + \sum_i \left[L_i^* X L_i' - \frac{1}{2}\left(L_i^* L_i X + X L_i'^* L_i' \right) \right], \quad \mathsf{D} = (H, \mathbf{L}),\ \mathsf{D}' = (H', \mathbf{L}')$$

which coincides with the usual Markov generator $\mathbb{W}_{\mathsf{D}}$ for $\mathsf{D} = \mathsf{D}'$. When choosing $\mathsf{D} = \mathsf{D}_{u/\sqrt{t}}$ and $\mathsf{D}' = \mathsf{D}_{u'/\sqrt{t}}$ the generator can be expanded as

$$\mathbb{W}_{\mathsf{D},\mathsf{D}'} = \mathbb{W}_{\mathsf{D}_0} + \frac{1}{\sqrt{t}}\mathbb{L}_1[u,u'] + \frac{1}{2t}\mathbb{L}_2[u,u'] + O(t^{3/2}).$$

Using a version of the Trotter-Kato second order perturbation Theorem for semigroups (cf. Theorem 2.2 in Ref. 13) one can show that (39) holds with an explicit choice of the phase angle $\phi(u)$ as a

quadratic form in u. The details of the calculations can be found in the Appendix. Note that since the phase $e^{i\phi(u)}$ is unphysical, it could have been incorporated in the definition of the coherent state $|u\rangle$, or in that of the system-output state $|\Psi^{s+o}_{u/\sqrt{t}}(t)\rangle$.

The second part of the Theorem can be proven by following the lines of an analogous discrete-time result, cf. Theorem 7 in Ref. 31. The main ideas are as follows. Let $\rho^{\mathsf{D}}_{ss} = \sum_m \Lambda_m |e_m\rangle\langle e_m|$ be the spectral decomposition of the stationary state for some dynamical parameter D. The stationary output state is given by

$$\rho^{\text{out}}_{\mathsf{D}}(t) = \sum_{m,m'} \Lambda_m |\psi_{mm'}(t)\rangle\langle\psi_{mm'}(t)|,$$

where, up to normalisation, $|\psi_{mm'}(t)\rangle$ are the conditional output states obtained by initialising the system in state $|e_m\rangle$ and projecting on state $|e_{m'}\rangle$ at time t, cf. proof of Lemma 3 in the Appendix. For large times, the overlaps of the different pure components $|\psi_{mm'}(t)\rangle$ vanish exponential fast; more generally, if $\mathsf{D} = \mathsf{D}_{u/\sqrt{t}}$ and $\mathsf{D}' = \mathsf{D}_{u'/\sqrt{t}}$ are two dynamical parameters in the local neighbourhood of D_0 (i.e. $u, u' \in \mathcal{O}$) then all the overlaps of components with *different* indices decay exponentially uniformly in u, u'. This can be shown by expressing the overlaps in terms of the deformed generator $\mathbb{W}_{\mathsf{D},\mathsf{D}'}$,

$$\langle\psi_{mm'}(t)|\psi_{nn'}(t)\rangle = \left\langle e_{m'}|e^{t\mathbb{W}_{\mathsf{D},\mathsf{D}'}}(|e_n\rangle\langle e_m|)|e_{n'}\right\rangle$$

and following the steps of the proof of Theorem 3 in Ref. 31, in particular the argument following Equation (35). This implies that the components can be distinguished with vanishing error probability, without the knowledge of the local parameter u. Each pure component satisfies the weak version of the local asymptotic normality, which can be upgraded to the strong version as in Theorem 7 of Ref. 31, which in turn employs a general result described in Lemma 5 of Ref. 31. Combining this with the fact that the pure components can be distinguished allows to construct the channels T_t, S_t as in Ref. 31.

2. Estimation for specific model of dynamical parameters

In Subsection VII B 1 we considered the problem of estimating all identifiable parameters, and showed how this becomes a quantum Gaussian estimation problem. Here, we show how this general result can be used for estimating an unknown parameter of the dynamics. Suppose that the the dynamical parameter D is known to depend on $\theta \in \mathbb{R}^m$ as described in Section V A, so that $\mathsf{D} = \mathsf{D}_\theta$. Let θ_0 be a fixed but arbitrary parameter value and let

$$\dot{\mathsf{D}}_a := \left(\left.\frac{\partial H}{\partial \theta_a}\right|_{\theta_0}, \left.\frac{\partial L^1}{\partial \theta_a}\right|_{\theta_0}, \ldots, \left.\frac{\partial L^k}{\partial \theta_a}\right|_{\theta_0} \right) = (\dot{H}_a, \dot{L}^1_a, \ldots, \dot{L}^k_a) \in \mathcal{T}_{\mathsf{D}_{\theta_0}}$$

be the tangent vectors associated with the different directions in the parameter space $\mathbb{R}^m$.

The stationary output's QFI rate matrix f at a given point θ_0 can be computed using the explicit formula (26), and we assume that θ is identifiable so that f is a strictly positive matrix. This condition can be verified as follows. Let $\tilde{\mathcal{T}}_{\mathsf{D}_{\theta_0}}$ be the tangent space at D_{θ_0} to the parametrised submanifold consisting of the points D_θ, and recall that $\mathcal{T}^{\text{nonid}}_{\mathsf{D}_{\theta_0}}$ is the space of directions corresponding to unidentifiable parameters. If the intersection $\tilde{\mathcal{T}}_{\mathsf{D}_{\theta_0}} \cap \mathcal{T}^{\text{nonid}}_{\mathsf{D}_{\theta_0}}$ is trivial then the Fisher information matrix f has no zero eigenvalue and therefore θ is identifiable locally around θ_0. We consider a local parametrisation around θ_0 given by $\theta = \theta_0 + h/\sqrt{t}$, with local parameter $h \in \mathbb{R}^m$. Since the stationary state depends only on the equivalence class $[\mathsf{D}]$, the statistical model can be projected onto the base space $\mathcal{P}$ giving rise to a local model $[D]_{h/\sqrt{t}}$, with $h \in \mathcal{O}' \subset \mathbb{R}^m$, which can be seen as sub-model of the 'full' model considered in Subsection VII B 1. In particular, the asymptotic normality Theorem 2 applies directly to the sub-model. However, in general it may happen that the "full" Gaussian limit model may be "too large," and one can use a restricted model defined as follows. Recall that $\mathcal{T}^{id}_{\mathsf{D}_{\theta_0}}$ is a Hilbert space with inner product (33), which defines the CCR algebra $CCR(\mathcal{T}^{id}_{\mathsf{D}_{\theta_0}}, \sigma^{\mathsf{D}_{\theta_0}})$ and the Gaussian state $|0\rangle$. Let $P(\dot{\mathsf{D}}_a)$ be the projection of the tangent vector $\dot{\mathsf{D}}_a$ onto $\mathcal{T}^{id}_{\mathsf{D}_{\theta_0}}$, and define $\mathcal{T}'$ to be the (complex) subspace spanned by these projections, with $a = 1, \ldots, m$. The subspace defines a CCR subalgebra

$CCR(\mathcal{T}', \sigma^{\mathsf{D}_{\theta_0}})$, and the restriction of the Fock state $|0\rangle$ to this subalgebra is also a Fock state which we denote by the same symbol.

As a concrete example, consider the driven two-level model of Section VI B, where the total identifiable subspace is 8-dimensional, while the subspace spanned by the projections of the physical tangent vectors is four-dimensional, associated with the CCR-algebra of four canonical quadratures.

We now obtain the following asymptotic normality result for the model D_θ in the neighbourhood of θ_0.

Corollary 1. Let

$$\mathcal{Q}'_t := \left\{\rho^{\text{out}}_{h/\sqrt{t}}(t) : h \in \mathcal{O}' \subset \mathbb{R}^m\right\}, \qquad \mathcal{G}' := \left\{\rho'_h := |h\rangle\langle h| : h \in \mathcal{O}' \subset \mathbb{R}^m\right\}$$

denote the local quantum statistical model of the output state associated with the dynamical parameter $\mathsf{D}_{\theta_0+h/\sqrt{t}}$*, and, respectively, the Gaussian model associated with the algebra* $CCR(\mathcal{T}', \sigma^{\mathsf{D}_{\theta_0}})$*, with* $|h\rangle := W(h)|0\rangle$*. Then* $\mathcal{Q}'_t$ *converge strongly to the the Gaussian model* $\mathcal{G}'$, i.e., $\Delta(\mathcal{Q}'_t, \mathcal{G}') \to 0$*. More precisely, there exist quantum channels* T'_t, S'_t *such that*

$$\lim_{t\to\infty} \sup_{h\in\mathcal{O}'} \left\| T'_t\left(\rho^{\text{out}}_{h/\sqrt{t}}(t)\right) - \rho'_h \right\|_1 = 0,$$
$$\lim_{t\to\infty} \sup_{h\in\mathcal{O}'} \left\| S'_t\left(\rho'_h\right) - \rho^{\text{out}}_{h/\sqrt{t}}(t) \right\|_1 = 0.$$

ACKNOWLEDGMENTS

Dedicated to Hans Maassen on the occasion of his 66th birthday. This work was supported by the EPSRC Project No. EP/J009776/1. J.K. also acknowledges additional support from the EPSRC Project No. EP/M01634X/1. We thank K. Macieszczak, J. P. Garrahan, and I. Lesanovsky for useful discussions during the preparation of this work.

APPENDIX: PROOFS

A. Proof theorem 1

The proof of Theorem 1 is inspired by a related argument from Ref. 5 and requires some auxiliary lemmas.

Lemma 2. Let $\mathsf{D}_l := (H_l, L^1_l, \ldots, L^k_l)$, $l = 1, 2$ *be two dynamical parameters with system spaces* $\mathcal{H}_l$ *and assume that both dynamics are ergodic. We define the maps*

$$\mathbb{W}_{ll'} : \mathcal{B}(\mathcal{H}_{l'}, \mathcal{H}_l) \to \mathcal{B}(\mathcal{H}_{l'}, \mathcal{H}_l), \qquad \mathbb{W}_{ll'}(X) = -iXH_{l',\text{eff}} + iH^*_{l,\text{eff}}X + \sum_{i=1}^k L^{i*}_l X L^i_{l'}$$

for $l, l' = 1, 2$*. Then the following conditions are equivalent:*

(i) $\mathbb{W}_{12}$ *has a purely imaginary eigenvalue;*
(ii) $\mathbb{W}_{21}$ *has an purely imaginary eigenvalue;*
(iii) there exists a unitary operator $U : \mathcal{H}_2 \to \mathcal{H}_1$*, and* $r \in \mathbb{R}$*, such that* $L^i_2 = U^* L^i_1 U$ *for all* $i = 1, \ldots, k$*, and* $H_2 = U^* H_1 U - r\mathbb{1}$*.*

If any of these conditions hold, then $\mathbb{W}_{12}(U) = irU$ *and* $\mathbb{W}_{21}(U^*) = -irU^*$*.*

Proof. Conditions (i) and (ii) are clearly equivalent: $\mathbb{W}_{ll'}(F) = irF$ with some $r \in \mathbb{R}$ and $F \in \mathcal{B}(\mathcal{H}_{l'}, \mathcal{H}_l)$, then $\mathbb{W}_{l'l}(F^*) = -irF^*$. Assuming (iii) we have

$$\mathbb{W}_{12}(U) = -iUH_{2,\text{eff}} + iH^*_{1,\text{eff}}U + \sum_{i=1}^k L^{i*}_1 U L^i_2 = U\mathbb{W}_{22}(\mathbb{1}) + irU = irU,$$

i.e., (i) holds, with ir the corresponding eigenvalue. Thus, the only nontrivial implication is (i) $\Longrightarrow$ (iii).

Let us define the families of isometries $V_l(t):\mathcal{H}\to\mathcal{H}_l\otimes\mathcal{F}$ such that $V_l(t)|\varphi\rangle = U_l(t)|\varphi\rangle\otimes\Omega\rangle$, with $U_l(t)$ the unitary generated by the dynamical parameter D_l, cf. Equation (4). Then $T_{ll',t}(X) := V_l(t)^*(X\otimes 1_{\mathcal{F}})V_{l'}(t) = e^{it\mathbb{W}_{ll'}}(X)$. Assume now (i), and let r and F be such that $\mathbb{W}_{12}(F) = irF$. Then $T_{12,t}(F) = e^{itr}F$, and since $\mathbb{W}_{21}(F^*) = -irF^*$, we also have $T_{21,t}(F^*) = e^{-irt}F^*$. For each t we have $V_1(t)V_1(t)^* \le 1_{\mathcal{H}_1\otimes\mathcal{F}}$, so

$$\begin{aligned} T_{22,t}(F^*F) &= V_2^*(t)(F^*\otimes \mathbb{1}_{\mathcal{F}})(F\otimes \mathbb{1}_{\mathcal{F}})V_2(t) \\ &\ge V_2^*(t)(F^*\otimes \mathbb{1}_{\mathcal{F}})V_1(t)V_1(t)^*(F\otimes 1_{\mathcal{F}})V_2(t) = T_{21,t}(F^*)T_{12,t}(F) = F^*F. \end{aligned}$$

Let P be the projection onto the eigenspace of F^*F corresponding to its largest eigenvalue $\|F^*F\|$. Now $\lim_{t\to\infty} T_{22,t}(X) = \mathrm{tr}[\rho_{ss,2}X]\mathbb{1}_{\mathcal{H}_2}$ by ergodicity, so

$$\mathrm{tr}[\rho_{ss,2}F^*F] = \lim_{t\to\infty}\mathrm{tr}[P]^{-1}\mathrm{tr}[PT_{22,t}(F^*F)] \ge \mathrm{tr}[P]^{-1}\mathrm{tr}[PF^*F] = \|F^*F\|.$$

This implies that $\mathrm{tr}[\rho_{ss,2}F^*F] = \|F^*F\|$, i.e., $\rho_{ss,2}$ is supported in the projection P. But $\rho_{ss,2}$ has full rank in $\mathcal{H}_2$, so $P = \mathbb{1}_{\mathcal{H}_2}$, and, consequently, $F^*F = \|F^*F\|\mathbb{1}_{\mathcal{H}_2}$. By proceeding in exactly the same way starting from $T_{11,t}$, we show that $FF^* = \|FF^*\|\mathbb{1}_{\mathcal{H}_1}$. Denote $\alpha := \|FF^*\| = \|F^*F\|$, and $U := \alpha^{-\frac{1}{2}}F$. Then $U:\mathcal{H}_2\to\mathcal{H}_1$ is a unitary operator between the two Hilbert spaces and in particular, $\dim\mathcal{H}_1 = \dim\mathcal{H}_2$. Moreover, we now have

$$ir\mathbb{1}_{\mathcal{H}_2} = U^*\mathbb{W}_{12}(U) = -i(H_2 - U^*H_1U) - \frac{1}{2}\sum_i(L_2^{i*}L_2^i + U^*L_1^{i*}L_1^iU) + \sum_i(U^*L_1^iU)^*L_2^i, \tag{A1}$$

and taking the real part of the trace of this equation gives

$$\mathrm{Re}\sum_i \mathrm{tr}[(U^*L_1^iU)^*L_{2,i}] = \frac{1}{2}\sum_i \mathrm{tr}[L_2^{i*}L_2^i + U^*L_1^{i*}L_1^iU] \tag{A2}$$

which implies that the (generally valid) inequalities

$$\begin{aligned} \mathrm{Re}\sum_i \mathrm{tr}[(U^*L_1^iU)^*L_2^i] &\le \left|\sum_i \mathrm{tr}[(U^*L_1^iU)^*L_2^i]\right| \le \sqrt{\sum_i \mathrm{tr}[L_2^{i*}L_2^i]\sum_i \mathrm{tr}[(U^*L_1^iU)^*(U^*L_1^iU)]} \\ &\le \frac{1}{2}\sum_i \mathrm{tr}[L_2^{i*}L_2^i + U^*L_1^{i*}L_1^iU] \end{aligned}$$

are in fact equalities. In particular,

$$\mathrm{Im}\sum_i \mathrm{tr}[(U^*L_1^iU)^*L_{2,i}] = 0. \tag{A3}$$

Moreover, since the second inequality is Cauchy-Schwartz for the scalar product $\sum_i \mathrm{tr}[A_i^*B_i]$ of k-tuples $(A_1,\ldots,A_k)$ of Hilbert-Schmidt operators, it follows that there exists a scalar $c\in\mathbb{C}$ such that $U^*L_1^iU = cL_2^i$ for all i. Putting this into (A3) we see that $c\in\mathbb{R}$, and from (A2) it follows that $c = 1$. Finally, from (A1) we then get $r = -H_2 + U^*H_1U$, which proves (iii). □

For reader's convenience we formulate the following simple lemma using the notations of the input-output setting, but the statement holds in a general context.

Lemma 3. Let $\mathcal{F}$ and $\mathcal{H}_l$, $l = 1, 2$, be Hilbert spaces. For each $t\ge 0$ let

$$V_l(t):\mathcal{H}_l\to\mathcal{H}_l\otimes\mathcal{F}, \quad l = 1,2$$

be an isometry, and define the maps

$$T_{ll',t}(X) := V_l(t)^*(X\otimes \mathbb{1}_{\mathcal{F}})V_{l'}(t), \quad X\in\mathcal{B}(\mathcal{H}_{l'},\mathcal{H}_l).$$

Suppose that $\lim_{t\to\infty} T_{ll',t}(\cdot) = \delta_{ll'}\mathrm{tr}[\rho_{ss,l}(\cdot)]\mathbb{1}_{\mathcal{H}_l}$ *holds for some states* $\rho_{ss,l}$ *and define* $\rho_l^{\mathrm{out}}(t) := \mathrm{tr}_{\mathcal{H}_l}[V_l(t)\rho_{ss,l}V_l(t)^*]$. *Then* $\lim_{t\to\infty}\mathrm{tr}[\rho_1^{\mathrm{out}}(t)^2]$ *and* $\lim_{t\to\infty}\mathrm{tr}[\rho_2^{\mathrm{out}}(t)^2]$ *exist and are strictly positive, while*

$$\lim_{t\to\infty}\mathrm{tr}[\rho_1^{\mathrm{out}}(t)\rho_2^{\mathrm{out}}(t)] = 0. \tag{A4}$$

Proof. We write $\rho_{ss,l} = \sum_m \Lambda_{l,m}|e_{l,m}\rangle\langle e_{l,m}|$ where $\Lambda_{l,m} \geq 0$, and $\{e_{l,m}\}_m$ is an orthonormal basis of $\mathcal{H}_l$. Then

$$\rho_l^{\mathrm{out}}(t) = \sum_m \Lambda_{l,m}\mathrm{tr}_{\mathcal{H}_l}[|V_l(t)e_{l,m}\rangle\langle V_l(t)e_{l,m}|] = \sum_{m,m'}\Lambda_{l,m}|\psi_{l,mm'}(t)\rangle\langle\psi_{l,mm'}(t)|,$$

where $\psi_{l,mm'}(t) \in \mathcal{F}$ is the unique vector satisfying $\langle\chi|\psi_{l,mm'}(t)\rangle = \langle e_{l,m}\otimes\chi|V_l(t)e_{l,m'}\rangle$ for all $\chi\in\mathcal{F}$. Now $\langle\psi_{l,mn}(t)|\psi_{l',m'n'}(t)\rangle = \langle e_{l,n}|T_{ll',t}(|e_{l,m}\rangle\langle e_{l'm'}|)|e_{l',n'}\rangle$ so we can write

$$\begin{aligned}\mathrm{tr}[\rho_l^{\mathrm{out}}(t)\rho_{l'}^{\mathrm{out}}(t)] &= \sum_{n,m}\sum_{n',m'}\Lambda_{l,n}\Lambda_{l',n'}|\langle\psi_{l,mn}(t)|\psi_{l',mn}(t)\rangle|^2\\ &= \sum_{n,m}\sum_{n',m'}\Lambda_{l,n}\Lambda_{l',n'}|\langle e_{l,n}|T_{ll',t}(|e_{l,m}\rangle\langle e_{l'm'}|)|e_{l',n'}\rangle|^2 \to \delta_{ll'}\sum_{n,n'}\Lambda_{l,n}^2\Lambda_{l,n'}^2.\end{aligned}$$

□

We can now proceed with the proof of Theorem 1. The the 'if' part is straightforward. Assume now that parameter sets $(H^l, \{L_i^l\}_{i=1}^k)$, $l = 1, 2$ are equivalent and define $T_{ll'}$ as in Lemma 2. We consider the direct sum isometry

$$V_{\mathrm{tot}}(t) := V_1(t)\oplus V_2(t) : \mathcal{H}_1\oplus\mathcal{H}_2 \to \mathcal{H}_1\otimes\mathcal{F}\oplus\mathcal{H}_2\otimes\mathcal{F} = (\mathcal{H}_1\oplus\mathcal{H}_2)\otimes\mathcal{F}.$$

We identify the elements $X\in\mathcal{B}(\mathcal{H}_1\oplus\mathcal{H}_2)$ in the usual way with block matrices

$$X = \begin{pmatrix} X_{11} & X_{12}\\ X_{21} & X_{22}\end{pmatrix},$$

where $X_{ll'}\in\mathcal{B}(\mathcal{H}_{l'},\mathcal{H}_l)$, the set of linear operators $\mathcal{H}_{l'}\to\mathcal{H}_l$. This identifies $\mathcal{B}(\mathcal{H}_l,\mathcal{H}_{l'})$ as a subspace $\mathcal{B}(\mathcal{H}_1\oplus\mathcal{H}_2)$, and each of these four subspaces is invariant under the channels T_t associated with $V_{\mathrm{tot}}(t)$. Explicitly, we have

$$T_t(X) = \begin{pmatrix} T_{11,t}(X_{11}) & T_{12,t}(X_{12})\\ T_{21,t}(X_{21}) & T_{22,t}(X_{22})\end{pmatrix}. \tag{A5}$$

In particular, any eigenvalue of $T_{ll',t}$ is also an eigenvalue of T_t, because the subspaces are invariant. Since each T_t is completely positive and unital by construction, all eigenvalues of $T_{ll',t}$ have modulus at most one, hence the eigenvalues of $\mathbb{W}_{12}$ have real part ≤ 0. If all of these are strictly negative, then we have $\lim_{t\to\infty}T_{12}(t) = \lim_{t\to\infty}e^{t\mathbb{W}_{12}} = 0$, which according to Lemma 3 contradicts the assumption that the output states are equal. Hence $\mathbb{W}_{12}$ must have a purely imaginary eigenvalue, so Lemma 2 concludes the proof. □

B. Proof of proposition 1

We fix a dynamical parameter D. Clearly we may assume $\mathcal{C}_{\mathsf{D}}(X^0) = X^0$ (that is, $X^0\in\mathcal{B}_0$) without loss of generality. The key ingredient is the following lemma which we prove first. In order to avoid cluttering the notation, we do not indicate D-dependence explicitly.

Lemma 4. For any tuple of operators $\mathbf{X} := (X^0, X^1, \ldots, X^k)\in\mathcal{B}_0\otimes M(\mathbb{C}^d)^k$, *and all* $s\geq 0$, *the following equality between maps on* $M(\mathbb{C}^d)$ *holds*

$$\sqrt{s}\,\langle\mathbb{F}_s(\mathbf{X})\Omega\,\lfloor j_s(\cdot)\Omega\rangle = \int_0^s T_t\circ\Phi_{\mathbf{X}}\circ T_{s-t}(\cdot)\,dt, \tag{A6}$$

where $\Phi_{\mathbf{X}} : M(\mathbb{C}^d)\to M(\mathbb{C}^d)$ *is the map given by*

$$\Phi_{\mathbf{X}}(Y) = X^{0*}Y + \sum_{i=1}^{k} X^{i*}[Y, L^i].$$

Proof. We let F_s denote the left hand side of (A6). By applying the quantum Ito formula (3) to the product of two adapted processes inside the conditional expectation, and eliminating the terms involving the annihilation processes acting on the vacuum, we get

$$\begin{aligned} dF_s(B) = \langle\Omega| \left(\sum_{i=1}^{k} j_s(X^{i*}) dA_i(s) + j_s(X^{0*}) ds \right) dj_s(B) \, |\Omega\rangle + \langle\Omega| j_s(X^{0*}) j_s(B) |\Omega\rangle ds \\ + \langle\Omega| \left(\sum_{i=1}^{k} j_s(X^{i*}) dA_i(s) dj_s(B) \right) |\Omega\rangle \end{aligned}$$

for all system operators B. Using now the Langevin equation (6), together with the Ito multiplication rules, and again eliminating the contributions from the annihilation processes, we get

$$\begin{aligned} F_s(B) &= \int_0^s F_t(\mathbb{W}(B))\, dt + \langle\Omega| \int_0^s \left(j_t(X^{0*}) j_t(B) + \sum_{i=1}^{k} j_t(X^{i*}) j_t([B, L^i]) \right) |\Omega\rangle dt \\ &= \int_0^s (F_t \circ \mathbb{W}(B) + T_t \circ \Phi_{\mathbf{X}}(B))\, dt, \end{aligned}$$

where we have also used (7). Hence F_s satisfies the (ordinary) differential equation

$$\frac{dF_t}{dt} = F_t \circ \mathbb{W} + T_t \circ \Phi_{\mathbf{X}},$$

with initial condition $F_0 = 0$. We can easily solve this equation: without the inhomogeneous part $T_t \circ \Phi_{\mathbf{X}}$, the solution would be simply $T_t = e^{t\mathbb{W}}$; hence the actual solution is obtained by concatenating $T_t \circ \Phi_{\mathbf{X}}$ with T_{s-t} and integrating. This gives the claimed result. □

Note that the covariance of the fluctuation operators is sesquilinear with respect to the operator coefficients $\mathbf{X}$ and $\mathbf{Y}$. We show that the limit exists by computing it explicitly using the Ito calculus. The differential of the product $\mathbb{F}_s(\mathbf{X})^* \mathbb{F}_s(\mathbf{Y})$ is given by the quantum Ito formula (3),

$$d(\mathbb{F}_s(\mathbf{X})^* \mathbb{F}_s(\mathbf{Y})) = \mathbb{F}_s(\mathbf{X})^* \cdot d\mathbb{F}_s(\mathbf{Y}) + d\mathbb{F}_s(\mathbf{X})^* \cdot \mathbb{F}_s(\mathbf{Y}) + d\mathbb{F}_s(\mathbf{X})^* \cdot d\mathbb{F}_s(\mathbf{Y}). \tag{A7}$$

For the last term, the Ito rule gives $d\mathbb{F}_s(\mathbf{X})^* \cdot d\mathbb{F}_s(\mathbf{Y}) = \frac{1}{t} \sum_{i=1}^{k} j_s(X^{i*} Y^i) ds$, and hence by using (7) and (8), we get

$$\int_0^t \langle \varphi \otimes \Omega | d\mathbb{F}_s(\mathbf{X})^* \cdot d\mathbb{F}_s(\mathbf{Y}) | \varphi \otimes \Omega \rangle = \frac{1}{t} \int_0^t \left\langle \varphi \left| T_s \left(\sum_{i=1}^{k} X^{i*} Y^i \right) \right| \varphi \right\rangle ds \xrightarrow{t\to\infty} \mathrm{tr}[\rho_{ss} \sum_{i=1}^{k} X^{i*} Y^i].$$

The expectation of the first term in (A7) can be computed by applying Lemma 4 with $\Phi_{\mathbf{X}} := iX^{0*}(\cdot) + \sum_i X^{i*}[(\cdot), L_i]$; we get

$$\begin{aligned} \langle \varphi \otimes \Omega | \int_0^t \mathbb{F}_s(\mathbf{X})^* \cdot d\mathbb{F}_s(\mathbf{Y}) | \varphi \otimes \Omega \rangle &= \frac{1}{\sqrt{t}} \langle \varphi \otimes \Omega | \int_0^t \mathbb{F}_s(\mathbf{X})^* j_s(-iY^0)\, ds | \varphi \otimes \Omega \rangle \\ &= \frac{1}{t} \langle \varphi | \int_0^t ds \int_0^s dr\, T_r \circ \Phi \circ T_{s-r}(-iY^0) \, |\varphi\rangle \\ &= \int_0^t ds\, \langle \varphi | \frac{1}{t} \left(\int_0^{t-s} dr\, T_r \right) \circ \Phi \circ T_s(-iY^0) \, |\varphi\rangle \\ &\xrightarrow{t\to\infty} -\mathrm{tr}[\rho_{ss} \Phi \circ \mathbb{W}^{-1}(-iY^0)], \end{aligned}$$

where we have also used the limit relations (8) and (10). The second term in (A7) is obtained by taking the adjoint of the first term with the roles of $\mathbf{X}$ and $\mathbf{Y}$ interchanged; this gives

$$
\begin{aligned}
(\mathbf{X},\mathbf{Y})_{\mathsf{D}} &= \mathrm{tr}\left[\rho_{ss}^{\mathsf{D}}\left(\sum_i X^{i*}Y^i + \Phi_{\mathbf{X}}\circ\mathbb{W}^{-1}(iY^0) + (\Phi_{\mathbf{Y}}\circ\mathbb{W}^{-1}(iY^0))^*\right)\right]\\
&= \mathrm{tr}\Big[\rho_{ss}^{\mathsf{D}}\Big(\sum_{i=1}^k X^{i*}Y^i - X^{0*}\mathbb{W}^{-1}(Y^0) - \mathbb{W}^{-1}(X^{0*})Y^0 - i\sum_{i=1}^k X^{i*}[L^i,\mathbb{W}^{-1}(Y^0)]\\
&\quad + i\sum_{i=1}^k[\mathbb{W}^{-1}(X^{0*}),L^{i*}]Y^i\Big)\Big].
\end{aligned}
$$

We then apply the identity

$$
\mathbb{W}(X^*Y) - X^*\mathbb{W}(Y) - \mathbb{W}(X^*)Y = \sum_i [L^i,X]^*[L^i,Y],
$$

which holds for arbitrary matrices X, Y, to the case where $X=\mathbb{W}^{-1}(X^0)$ and $Y=\mathbb{W}^{-1}(Y^0)$. Using the fact that $\mathrm{tr}[\rho_{ss}^{\mathsf{D}}\mathbb{W}(X^*Y)]=0$, we obtain the following formula for the inner product

$$
(\mathbf{X},\mathbf{Y})_{\mathsf{D}} = \sum_{i=1}^k \mathrm{tr}\left[\rho_{ss}^{\mathsf{D}}\left(X^i - i[L^i,\mathbb{W}^{-1}(X^0)]\right)^*\left(X^i - i[L^i,\mathbb{W}^{-1}(X^0)]\right)\right] \tag{A8}
$$

This proves the proposition.

C. Proof of theorem 2

Since $\dot{\mathsf{D}}_a \in \mathcal{T}_{\mathsf{D}_0}^{id}$ we have $\mathcal{E}_{\mathsf{D}_0}(\dot{\mathsf{D}}_a)=0$ and using this we find that the first order term is given by

$$
\mathbb{L}_1[u,u'](X) := \sum_{a,i} u_a\dot{L}_a^{i*}[X,L^i] - \sum_{a,i} u'_a[X,L^{i*}]\dot{L}_a^i. \tag{A9}
$$

The second order term is given by

$$
\begin{aligned}
\mathbb{L}_2[u,u'](X) &= \sum_{a,a'} u_a u_{a'}\left(\left(i\ddot{H}_{aa'} - \frac{1}{2}\sum_i(\ddot{L}_{aa'}^{i*}L^i + L^{i*}\ddot{L}_{aa'}^i + 2\dot{L}_a^{i*}\dot{L}_{a'}^i)\right)X + \sum_i \ddot{L}_{aa'}^{i*}XL^i\right)\\
&\quad + \sum_{a,a'} u'_a u'_{a'}\left(X\left(-i\ddot{H}_{aa'} - \frac{1}{2}\sum_i(\ddot{L}_{aa'}^{i*}L^i + L^{i*}\ddot{L}_{aa'}^i + 2\dot{L}_a^{i*}\dot{L}_{a'}^i)\right) + \sum_i L^{i*}X\ddot{L}_{aa'}^i\right)\\
&\quad + 2\sum_{aa'} u_a u'_{a'}\sum_i \dot{L}_a^{i*}X\dot{L}_{a'}^i.
\end{aligned}
$$

Using a version of Trotter-Kato theorem (cf. Theorem 2.2 in Ref. 13), we obtain the limit

$$
\lim_{t\to\infty} e^{t\mathbb{W}_{\mathsf{D}_{u/\sqrt{t}},\mathsf{D}_{u'/\sqrt{t}}}}(\mathbb{1}) = e^{f(u,u')}1,
$$

where

$$
f(u,u') = \mathrm{tr}\left[\rho_{ss}^{\mathsf{D}_0}\left(\frac{1}{2}\mathbb{L}_2[u,u'](\mathbb{1}) - \mathbb{L}_1[u,u']\circ\mathbb{W}_{\mathsf{D}_0}^{-1}\circ\mathbb{L}_1[u,u'](\mathbb{1})\right)\right].
$$

Now, from Equation (A9) we find $\mathbb{L}_1[u,u'](\mathbb{1})=0$ and

$$\begin{aligned}
\mathbb{L}_2[u,u'](\mathbb{1}) &= \sum_{aa'} u_a u_{a'} \left(i\ddot{H}_{aa'} + \frac{1}{2}\sum_i (\ddot{L}^{i*}_{aa'} L^i - L^{i*}\ddot{L}^i_{aa'} - 2\dot{L}^{i*}_a \dot{L}^i_{a'}) \right) \\
&+ \sum_{aa'} u'_a u'_{a'} \left(-i\ddot{H}_{aa'} + \frac{1}{2}\sum_i (-\ddot{L}^{i*}_{aa'} L^i + L^{i*}\ddot{L}^i_{aa'} - 2\dot{L}^{i*}_a \dot{L}^i_{a'}) \right) \\
&- \sum_{aa'} (u_a - u'_a)(u_{a'} - u'_{a'}) \sum_i \dot{L}^{i*}_a \dot{L}^i_{a'} + 2i\mathrm{Im} \sum_{aa'} u_a u'_{a'} \sum_i \dot{L}^{i*}_a \dot{L}^i_{a'} \\
&+ \sum_{aa'} (u_a u_{a'} + u'_a u'_{a'}) \sum_i \dot{L}^{i*}_a \dot{L}^i_{a'} \\
&= \sum_{aa'} \Big[-(u_a - u'_a)(u_{a'} - u'_{a'}) \dot{L}^{i*}_a \dot{L}^i_{a'} + 2i u_a u'_{a'} \mathrm{Im} \dot{L}^{i*}_a \dot{L}^i_{a'} \\
&+ i(u_a u_{a'} - u'_a u'_{a'}) \Big(\ddot{H}_{aa'} + \mathrm{Im} \sum_i \ddot{L}^{i*}_{aa'} L^i \Big) \Big].
\end{aligned}$$

Therefore we have

$$f(u,u') = \frac{1}{2}\mathrm{tr}\left(\rho^{\mathsf{D}_0}_{ss} \mathbb{L}_2[u,u'](\mathbb{1})\right) = -\frac{1}{8}(u-u')^T f^{\mathsf{D}_0}(u-u') + iu^T \sigma^{\mathsf{D}_0} u' + i(u^T S u - u'^T S u').$$

Above, f^{D_0} is the quantum Fisher information matrix at D_0 whose entries have the simple form due to the fact that the tangent vectors $\dot{\mathsf{D}}_a$ belong to the space $\mathcal{T}^{id}_{\mathsf{D}_0}$,

$$f^{\mathsf{D}_0}_{aa'} = 4\mathrm{Re}(\dot{\mathsf{D}}_a, \dot{\mathsf{D}}_{a'})_{\mathsf{D}_0} = 4\mathrm{Re} \sum_i \mathrm{tr}(\rho^{\mathsf{D}_0}_{ss} \dot{L}^{i*}_a \dot{L}^i_{a'}).$$

Moreover, σ^{D_0} is the symplectic matrix at D_0 (see Definition 2)

$$\sigma^{\mathsf{D}_0}_{aa'} = \mathrm{Im}(\dot{\mathsf{D}}_a, \dot{\mathsf{D}}_{a'})_{\mathsf{D}_0} = \mathrm{Im} \sum_i \mathrm{tr}(\rho^{\mathsf{D}_0}_{ss}\ \dot{L}^{i*}_a \dot{L}^i_{a'})$$

and S is the real symmetric matrix

$$S_{aa'} = \frac{1}{2}\mathrm{tr}\left[\rho^{\mathsf{D}_0}_{ss}\left(\ddot{H}_{aa'} + \mathrm{Im}\sum_i \ddot{L}^{i*}_{aa'} L^i\right)\right].$$

In conclusion, the overlaps of the system-output states have the following limit:

$$\begin{aligned}
\lim_{t\to\infty} \left\langle \Psi^{\mathrm{s+o}}_{u/\sqrt{t}}(t) \middle| \Psi^{\mathrm{s+o}}_{u'/\sqrt{t}}(t) \right\rangle &= \exp\left(-\frac{1}{8}(u-u')^T f^{\mathsf{D}_0}(u-u') + iu^T \sigma^{\mathsf{D}_0} u' + i(u^T S u - u'^T S u') \right) \\
&= e^{i\phi(u) - i\phi(u')} \langle u | u' \rangle,
\end{aligned}$$

where $\phi(i) := u^T S u$ is a phase angle and $|u\rangle = W(u)|\Omega\rangle$ is the coherent state on the CCR algebra $CCR(\mathcal{T}^{id}_{\mathsf{D}_0}, \sigma^{\mathsf{D}_0})$ introduced in Definition 2, so that the overlaps of two coherent states are given by Equation (38).

[1] V. V. Albert, B. Bradlyn, M. Fraas, and L. Jiang, "Geometry and response of Lindbladians," Phys. Rev. X **6**, 041031 (2016).
[2] S. Amari and H. Nagaoka, *Methods of Information Geometry*, Volume 191 of Translations of Mathematical Monographs (American Mathematical Society, 2000).
[3] J. E. Avron, M. Fraas, and G. M. Graf, "Adiabatic response for Lindblad dynamics," J. Stat. Phys. **148**, 800–823 (2012).
[4] L. Banchi, P. Giorda, and P. Zanardi, "Quantum information-geometry of dissipative quantum phase transitions," Phys. Rev. A **89**, 022102 (2014).
[5] B. Baumgartner and H. Narnhofer, "Analysis of quantum semigroups with GKS–Lindblad generators: II. General," J. Phys. A: Math. Theor. **41**, 395303 (2008).
[6] T. Baumgratz, D. Gross, M. Cramer, and M. B. Plenio, "Scalable reconstruction of density matrices," Phys. Rev. Lett. **111**, 020401 (2013).
[7] V. P. Belavkin, "Quantum stochastic calculus and quantum nonlinear filtering," J. Multivar. Anal. **42**, 171–201 (1992).
[8] V. P. Belavkin, "Measurement, filtering and control in quantum open dynamical systems," Rep. Math. Phys. **43**, A405–A425 (1999).
[9] P. J. Bickel, Y. Ritov, and T. Ryden, "Inference in hidden Markov models I: Local asymptotic normality in the stationary case," Bernoulli **2**, 199–228 (1996).

[10] L. Bouten, R. van Handel, and M. James, "An introduction to quantum filtering," SIAM J. Control Optim. **46**, 2199–2241 (2007).
[11] S. L. Braunstein and C. M. Caves, "Statistical distance and the geometry of quantum states," Phys. Rev. Lett. **72**, 3439–3443 (1994).
[12] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, 2002).
[13] C. Catana, M. Guţă, and L. Bouten, "Fisher informations and local asymptotic normality for continuous-time quantum Markov processes," J. Phys. A: Math. Theor. **48**, 365301 (2015).
[14] C. Catana, M. Guţă, and T. Kypraios, "Maximum likelihood versus likelihood-free quantum system identification in the atom maser," J. Phys. A: Math. Theor. **47**, 415302 (2013); e-print arXiv:1311.4091.
[15] A. Chefles, R. Josza, and A. Winter, "On the existence of physical transformations between sets of quantum states," Int. J. Quantum Inf. **02**, 11 (2004).
[16] M. Cozzini, R. Ionicioiu, and P. Zanardi, "Quantum fidelity and quantum phase transitions in matrix product states," Phys. Rev. B **76**, 104420 (2007).
[17] D. G. Evans, J. E. Gough, and M. R. James, "Non-abelian Weyl commutation relations and the series product of quantum stochastic evolutions," Philos. Trans. R. Soc., A **370**, 5437–5451 (2012).
[18] M. Fannes, B. Nachtergaele, and R. F. Werner, "Finitely correlated states on quantum spin chains," Commun. Math. Phys. **144**, 443 (1992).
[19] A. Frigerio, "Stationary states of quantum dynamical semigroups," Commun. Math. Phys. **63**, 269–276 (1978).
[20] J. Gambetta and H. M. Wiseman, "State and dynamical parameter estimation for open quantum systems," Phys. Rev. A **64**, 042105 (2001).
[21] S. Gammelmark and K. Molmer, "Bayesian parameter inference from continuously monitored quantum systems," Phys. Rev. A **87**, 032115 (2013).
[22] S. Gammelmark and K. Mølmer, "Fisher information and the quantum Cramér-Rao sensitivity limit of continuous measurements," Phys. Rev. Lett. **112**, 170401 (2014).
[23] C. Gardiner and P. Zoller, *Quantum Noise* (Springer, 2004).
[24] R. D. Gill and M. Guţă, "On asymptotic quantum statistical inference," Inst. Math. Stat. Collect. **9**, 105–127 (2012).
[25] J. Gough and M. R. James, "The series product and its application to quantum feedforward and feedback networks," IEEE Trans. Autom. Control **54**, 2530–2544 (2009).
[26] J. E. Gough, T. S. Ratiu, and O. G. Smolyanov, "Noethers theorem for dissipative quantum dynamical semi-groups," J. Math. Phys. **56**, 022108 (2015).
[27] M. Guţă, "Quantum Fisher information and asymptotic normality in system identification for quantum Markov chains," Phys. Rev. A **83**, 062624 (2011).
[28] M. Guţă, B. Janssens, and J. Kahn, "Optimal estimation of qubit states with continuous time measurements," Commun. Math. Phys. **277**, 127–160 (2008).
[29] M. Guţă and A. Jençová, "Local asymptotic normality in quantum statistics," Commun. Math. Phys. **276**, 341–379 (2007).
[30] M. Guţă and J. Kahn, "Local asymptotic normality for qubit states," Phys. Rev. A **73**, 052108 (2006).
[31] M. Guţă and J. Kiukas, "Equivalence classes and local asymptotic normality in system identification for quantum Markov chains," Commun. Math. Phys. **335**, 1397–1428 (2015).
[32] M. Guţă and N. Yamamoto, "System identification for passive linear quantum systems," IEEE Trans. Autom. Control **61**, 921 (2016); e-print arXiv:1303.3771v2.
[33] J. Haegeman, M. Marien, T. J. Osborne, and F. Verstraete, "Geometry of matrix product states: Metric, parallel transport and curvature," J. Math. Phys. **55**, 021902 (2014).
[34] J. Haegeman, T. J. Osborne, and F. Verstraete, "Post-matrix product state methods: To tangent space and beyond," Phys. Rev. B **88**, 075133 (2013).
[35] B. Hanzon, "On a coordinate free prediction error algorithm for system identification," in *Modelling, Identification and Robust Control*, edited by C. I. Byrnes and A. Lindquist (Elsevier Science Publishers B.V., 1986), p. 595.
[36] A. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland Publishing Company, Amsterdam, New York, Oxford, 1982).
[37] R. Höpfner, J. Jacod, and L. Ladelli, "Local asymptotic normality and mixed normality for Markov statistical models," Probab. Theor. Relat. Fields **86**, 105–129 (1990).
[38] C. J. Isham, *Modern Differential Geometry for Physicists*, 2nd ed. (World Scientific, 2003).
[39] M. R. James, H. I. Nurdin, and I. R. Petersen, "H-infinity control of linear quantum stochastic systems," IEEE Trans. Autom. Control **53**, 1787–1803 (2008).
[40] J. Kahn and M. Guţă, "Local asymptotic normality for finite dimensional quantum systems," Commun. Math. Phys. **289**, 597–652 (2009).
[41] L. Le Cam, *Asymptotic Methods in Statistical Decision Theory* (Springer Verlag, New York, 1986).
[42] L. Ljung, *System Identification Theory for the User* (Prentice Hall, 1999).
[43] H. Mabuchi, "Dynamical identification of open quantum systems," Quantum Semiclassical Opt. **8**, 1103–1108 (1996).
[44] K. Macieszczak, M. Guţă, I. Lesanovsky, and J. P. Garrahan, "Dynamical phase transitions as a resource for quantum enhanced metrology," Phys. Rev. A **93**, 022103 (2016); e-print arXiv:1411.3914.
[45] K. Macieszczak, M. Guţă, I. Lesanovsky, and J. P. Garrahan, "Towards a theory of metastability in open quantum dynamics," e-print arXiv:1512.05801.
[46] K. R. Parthasarathy, *An Introduction to Quantum Stochastic Calculus* (Birkhäuser, Basel, 1992).
[47] D. Petz, *An Invitation to the Algebra of Canonical Commutation Relations* (Leuven University Press, Leuven, 1990).
[48] D. Petz and A. Jencova, "Sufficiency in quantum statistical inference," Commun. Math. Phys. **263**, 259–276 (2006).
[49] J. M. Radcliffe, "Some properties of coherent spin states," J. Phys. A: Gen. Phys. **4**, 313–323 (1971).

[50] D. A. Sivak and G. E. Crooks, "Thermodynamic metrics and optimal paths," Phys. Rev. Lett. **108**, 190602 (2012).
[51] M. M. Taddei, B. M. Escher, L. Davidovich, and R. L. de Matos Filho, "Quantum speed limit for physical processes," Phys. Rev. Lett. **110**, 050402 (2013).
[52] F. Verstraete and J. I. Cirac, "Continuous matrix product states for quantum fields," Phys. Rev. Lett. **104**, 190405 (2010).
[53] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, 2009).