Complete parametrizations of diffusive quantum monitorings

A. Chia and H. M. Wiseman

Centre for Quantum Computation and Communication Technology (Australian Research Council) and
Centre for Quantum Dynamics, Griffith University, Brisbane, Queensland 4111, Australia
(Received 24 February 2011; published 26 July 2011)

The master equation for the state of an open quantum system can be unravelled into stochastic trajectories described by a stochastic master equation. Such stochastic differential equations can be interpreted as an update formula for the system state conditioned on results obtained from monitoring the bath. So far only one parametrization (mathematical representation) for arbitrary diffusive unravellings (quantum trajectories arising from monitorings with Gaussian white noise) of a system described by a master equation with L Lindblad terms has been found [H. M. Wiseman and A. C. Doherty, Phys. Rev. Lett. **94**, 070405 (2005)]. This parametrization, which we call the U representation parameterizes diffusive unravellings by $L^2 + 2L$ real numbers, arranged in a matrix U subject to three constraints. In this paper we investigate alternative parametrizations of diffusive measurements. We find, rather surprisingly, the description of diffusive unravellings can be unified by a single equation for a nonsquare complex matrix M if one is willing to allow for some redundancy by lifting the number of real parameters necessary from $L^2 + 2L$ to $3L^2 + L$. We call this parametrization the M representation. Both the M representation and U representation lack a physical picture of what the measurement should look like. We thus propose another parametrization, the B representation that details how the measurement is implemented in terms of beam-splitters, phase shifters, and homodyne detectors. Relations between the different representations are derived.

I. INTRODUCTION

Methods of detection for optical fields such as the heterodyne (and homodyne) [1] enable measurements of the field amplitude (and quadratures) and are useful for detecting nonclassical light [2,3]. The heterodyne (or homodyne) detector is an example of general diffusive measurements [4,5], so called because the statistical fluctuations in the current outputed by the measurement is driven by a Gaussian-white-noise process, a prime example of a diffusion process. This is the process responsible for the broadening of a probability distribution described by the diffusive term appearing in Fokker-Planck equations [6,7]. The heterodyne technique, originally conceived for radio technology more than 100 years ago [8,9], of which the homodyne is a variant, has now become an indispensable part of many quantum information-processing applications. Some of these applications include quantum-state tomography [10–12], teleportation [13–16], and state preparation [17–22]. It also enables tests of quantum mechanics [23–25].

In this paper we are concerned with continuous measurements (also referred to as monitorings) of an open system in a vacuum bosonic environment (Fig. 1). Specifically, we are interested in how one can specify the class of all possible diffusive quantum measurements given how many inputs to the measuring device there are. There is already one solution to this problem [26,27], a parametrization that we shall call the U representation (U-rep). The U-rep parameterizes an arbitrary diffusive measurement by a square, real matrix U, with three constraints imposed on the blocks of U. We will present two more solutions with certain advantages over the U-rep: The M representation (M-rep), a nonsquare complex matrix M, and the B representation (B-rep), a square matrix and two vectors. We will be comparing the different parametrizations in detail later but let us first say why one might consider these alternative parametrizations.

A good reason, and also the more theoretical reason for considering the M-rep, is that aesthetically it is more attractive than the U-rep. As we will show, the M-rep has the ability to unify all diffusive measurements by a single equation whereas three would be required to define U.[1] The aesthetic advantage of the M-rep continues when we come to the quantum theory of multiple-input multiple-output Markovian feedback control based on diffusive measurements [28]. There, the theory is simpler when the equations are expressed in terms of the M-rep rather than the U-rep. Furthermore, the defining equation for M has an intuitive interpretation which makes it easy to remember.

Either the U-rep or the M-rep can be used when we know what the measurement is and would like to parametrize it in order to model the state of the open system conditioned on the results of the monitoring. If, on the other hand, we were given either a U or an M and were asked to describe the measurement with actual optical elements, then this is a much more difficult task for arbitrary measurements. Such a setting where one may be given an M without knowing its physical implementation can in fact arise naturally in quantum feedback control [27,29]: As measurement is an inherent part of the feedback loop, a design of the feedback loop can thus incorporate a design of measurement and it is natural to ask what sort of measurement one should do in order to achieve a particular objective for the control. In the case of optimal control the objective would be to minimize a measurement-dependent cost function. The result of this optimization would

---

[1]However, the M-rep is numerically more cumbersome than the U-rep in the sense that it requires more real numbers to specify a diffusive measurement. In particular, the extra number of parameters that M needs scales as a quadratic in the number of rows of M.
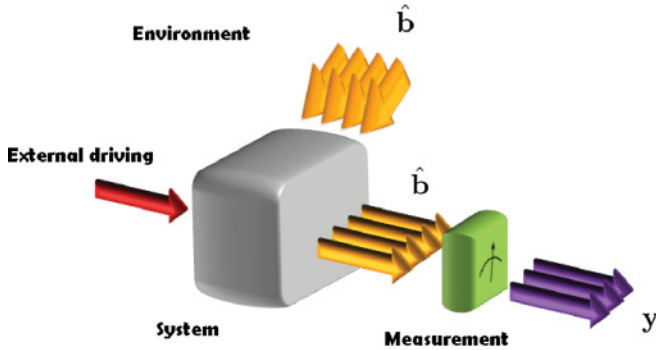
FIG. 1. (Color online) Schematic of some photoemissive system which couples to the environment via some set of operators $\hat{\mathbf{c}}$. The time evolution of the system based on knowledge of the noisy signal y also follows a noisy path, which can be described by a stochastic master equation. The stochastic master equation updates the system state continuously in time as the measurement record grows. If one ignores the measured result y then the system state evolves according to the deterministic master equation. This corresponds to averaging over the stochastic evolution.

be a matrix $\mathsf{M}^\star$, which one would then need to realize in the laboratory. That is, the theorist who now has obtained $\mathsf{M}^\star$ would like to inform his/her experimentalist colleague about how to construct the measurement.

The above considerations motivate us to propose yet another parametrization of diffusive measurements; one which we call the block representation or B-rep for short. Unlike the M- and U-rep, the B-rep is a realization of the measurement in terms of beam-splitters, phase shifters, and homodyne detectors. The B-rep is so called because it parameterizes the diffusive measurement in terms of three distinct blocks, with each block corresponding to a set of parameters (beam-splitter transmission coefficients and phase shifts). Note that being given an M is equivalent to being given a POVM (positive-operator-valued measure[2]). From this point of view the relationship between the M-rep and B-rep is thus one of a POVM and its realization. As we shall see, proving that an arbitrary POVM has a realization defined by the B-rep is much more difficult than translating a given B-rep into an M and hence a POVM. Here we point out that this line of thought had in fact been applied to N-port homodyne detection, a generalization of the standard homodyne measurement but a subset of all diffusive measurements, as early as 1987 [31]. A similar question was raised again in 1994 but for unitary operators instead of a POVM [33]. Our construction of the B-rep in fact relies on this result of Ref. [33].

_____

[2]The outcomes of a measurement can be labeled by a random variable $\mathbf{y}$ (here assumed to be discrete). In a general formulation of quantum measurements the distribution of the particular outcomes $\check{\mathbf{y}}$ (a deterministic variable) of $\mathbf{y}$ is described by a set of "probability operators" [4,30], defined by $\{\hat{E}_{\check{\mathbf{y}}} \mid \forall \, \check{\mathbf{y}} \; \hat{E}_{\check{\mathbf{y}}} \geqslant 0 \,, \; \sum_{\check{\mathbf{y}}} \hat{E}_{\check{\mathbf{y}}} = \hat{1}\}$. This set is known as a POVM and sometimes also as a POM, short for probability-operator measure. We shall find in Sec. III A that $\mathbf{y}$ depends on M [see (36)] so a given M will restrict the set of possible realizations $\check{\mathbf{y}}$. Each choice of M therefore defines a POVM and each value of $\check{\mathbf{y}}$ corresponds to a particular POVM element $\hat{E}_{\check{\mathbf{y}}}$.

Our paper is organized as follows. In Sec. II the unconditional dynamics of open quantum systems is reviewed and the concept of unravellings and its relation to continuous measurements are also briefly sketched. As we frequently make use of vector operators, a comprehensive review of their definitions and algebra is provided in Appendix A. Note that not all of the results in Appendix A will be used in this paper so we will refer to the ones that do appear. However, the reader who is also interested in the theory of multiple-input multiple-output quantum feedback control (for which the M-rep is applied to) [28] may find it worthwhile to embark on a fuller exposition of vector-operator algebra. We also mention that for convenience we will not necessarily reflect the multicomponent nature of vectors or vector operators in our language when they are referred to, such as in "the field $\hat{\mathbf{a}}$" or "the current $\hat{\mathbf{y}}$," as opposed to using plurals as in "the fields $\hat{\mathbf{a}}$" or "the currents $\hat{\mathbf{y}}$." Our first results section begins with Sec. III. The key results here are the definition of the M-rep and its relation with the U-rep. It should be noted that Sec. III begins by postulating the most general diffusive stochastic master equation. The purpose of Sec. II is thus to help the reader gain the intuition required to make this leap. In Sec. IV we define the B-rep and discuss its relation with the M-rep. Here we conjecture a matrix decomposition of M in terms of the matrices in the B-rep triple. We then conclude in Sec. V.

II. DIFFUSIVE CONTINUOUS MEASUREMENTS IN THE SCHRÖDINGER PICTURE

A. Unconditional system dynamics

A general open quantum system with Markovian dynamics [34,35] can be described by a master equation in the Lindblad form [36,37]

$$\hbar \, \dot{\rho} = -i \, [\hat{H}, \rho] + \mathcal{D}[\hat{\mathbf{c}}]\rho \equiv \mathcal{L}\rho, \qquad (1)$$

where $\hat{H}$ is Hermitian and $\hat{\mathbf{c}} \equiv (\hat{c}_1, \hat{c}_2, \ldots, \hat{c}_L)^\top$ (which we will refer to as a vector operator) is a vector of Lindblad operators, all time independent. See Appendix A for a comprehensive review of definitions and conventions that we are adopting for vector operators. We have also defined

$$\mathcal{D}[\hat{\mathbf{c}}] \equiv \sum_{k=1}^{L} \mathcal{D}[\hat{c}_k], \quad \mathcal{D}[\hat{c}]\rho \equiv \hat{c} \, \rho \, \hat{c}^\dagger - \frac{1}{2} \hat{c}^\dagger \hat{c} \rho - \frac{1}{2} \rho \, \hat{c}^\dagger \hat{c}. \tag{2}$$

Note that instead of working in natural units by setting $\hbar = 1$ we will work in units such that a factor of $\hbar$ appears on the left-hand side of (1). This was done in Refs. [4,27] so we will keep their convention for ease of comparison with the results therein. The reason for redefining units this way is to keep track of the correspondence between quantum operators and their classical counterpart (the dynamical variables which they represent) and aspects of quantum control that are not present in the classical theory. Equation (1) can be derived by a unitary operator $\hat{U}(t, t_0)$ acting on the joint Hilbert space of the system and bath [35]. This describes the interaction between the system and its environment, which we assume to be given by the Itô stochastic differential equation

$$\hbar \, d\hat{U}(t, t_0) = \left(-i\hat{H} - \tfrac{1}{2} \hat{\mathbf{c}}^\dagger \hat{\mathbf{c}} + \hat{\mathbf{b}}^\dagger \hat{\mathbf{c}} - \hat{\mathbf{c}}^\dagger \hat{\mathbf{b}} \right) \hat{U}(t, t_0) \, dt, \quad (3)$$

where $d\hat{U}(t,t_0) \equiv \hat{U}(t+dt,t_0) - \hat{U}(t,t_0)$ and

$$\hat{\mathbf{c}}^\dagger \equiv (\hat{c}_1^\dagger, \hat{c}_2^\dagger, \ldots, \hat{c}_L^\dagger). \tag{4}$$

Here $\hat{\mathbf{b}}$ represents a bosonic bath field shown in Fig. 1 and assumed to be in the vacuum state. External driving of the system is included in $\hat{H}$. In this section we are working in the Schrödinger picture so operators are time independent, equal to their initial value.

It is well known that the evolution described by (1) can be decomposed into stochastic paths in Hilbert space, called quantum trajectories (a term coined by Carmichael [38], who also referred to this decomposition as unravelling the master equation) by considering the evolution of ρ conditioned on the results of monitoring [38,39]. We will also refer to the results of monitoring as a current, denoted by a vector $\mathbf{y}(t)$. For Markovian evolution it is sufficient to consider unravellings (the set of possible quantum trajectories the system state may take over time) generated by measurements with either or both of two classes of noise [4,40]: (i) A point (or jump) process [41], also called a Poissonian random variable, or, (ii) a Wiener process also called Gaussian white noise. We will refer to this second case as a diffusive measurement. Physicists usually neglect the technical difference between an "unravelling" and a "measurement," and one often finds the two terms used interchangeably. We will do so as well in this paper for convenience.

B. Conditional system dynamics: Homodyne and heterodyne unravellings

The most familiar forms of diffusive measurements in quantum optics are homodyne and heterodyne detection schemes of a single output field of some photoemissive source. Apart from being characterized by Gaussian white noise (as opposed to point-process noise) these measurements are also determined by physical properties such as the specific arrangement of linear optical elements and the specific parameter values at which they operate. Here we will first show how the master equation (1) is unravelled by the homodyne and heterodyne detection schemes. This will familiarize the reader with the notation used for describing conditional evolution and build some intuition about how the measurement parameters enter into a quantum trajectory equation. For convenience we will set $\hbar = 1$ for the rest of this section.

1. The homodyne stochastic master equation

We begin with the example of a single mode of an intracavity field. For simplicity we will assume the optical cavity to be one-ended with transmission coefficient γ and that the bath is in the vacuum state. In this case the master equation (1) has only one Lindblad term (i.e., $L = 1$), given by $\hat{c} = \sqrt{\gamma}\,\hat{a}$, where $\hat{a}$ is the cavity mode annihilation operator. A homodyne measurement [1] of the field at the leaky port would then give us information about a particular quadrature of $\hat{a}$. Such a continuous measurement would yield a result, for some state $\rho(t)$, given by the Itô formula [4,42] (assuming a measurement efficiency of η)

$$y(t)\,dt = \sqrt{\eta}\,\langle \hat{c} + \hat{c}^\dagger \rangle(t)\,dt + dw(t). \tag{5}$$

We have defined, for an arbitrary operator $\hat{A}$, $\langle \hat{A} \rangle(t) = \mathrm{Tr}[\hat{A}\,\rho(t)]$, reserving angle brackets for quantum (operator) averages. The term $dw(t)$ is a Wiener increment. It is a Gaussian random variable defined by the mean

$$\mathrm{E}[dw(t)] = 0 \tag{6}$$

and variance

$$[dw(t)]^2 = dt, \tag{7}$$

$$dw(t)\,dw(t') = 0 \quad \forall\, t \neq t'. \tag{8}$$

We shall always denote an average taken with respect to a probability distribution by an E as in (6).

From the measurement result $y(t)$ the observer can update her *prior* state from $\rho(t)$ to a *posterior* state $\rho_{y_t}(t+dt)$ according to the stochastic master equation

$$d\rho_{y_t}(t) = \mathcal{L}\rho(t)\,dt + \sqrt{\eta}\,\mathcal{H}[\hat{c}]\,\rho(t)\,dw, \tag{9}$$

where $\mathcal{L}$ is given by (1) and the change in the state is

$$d\rho_{y_t}(t) = \rho_{y_t}(t+dt) - \rho(t). \tag{10}$$

We have also defined [see also (A65) in Appendix A]

$$\mathcal{H}[\hat{A}]\hat{B} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}^\dagger - \mathrm{Tr}(\hat{A}\hat{B} + \hat{B}\hat{A}^\dagger)\hat{B}, \tag{11}$$

for any $\hat{A}$ and $\hat{B}$. Here $\rho_{y_t}(t+dt)$ is the state given a particular realization of $y(t)$. We have also written the time dependence of y as a subscript. For clarity it is best to have this flexibility in writing the time dependence. Averaging (9) over $y(t)$ according to its actual distribution $\wp(\breve{y}_t)$ (a Gaussian with mean $\sqrt{\eta}\,\langle \hat{c} + \hat{c}^\dagger \rangle$ and variance $1/dt$) simply returns the master equation (1). Note that a breve over a random variable denotes its realization.

Equation (9) is a nonlinear stochastic differential equation in the state. This would not have been the case if we did not insist on keeping the state normalized at all times and having the measurement results distributed according to its true statistics described by $\wp(\breve{y}_t)$. An alternative theory is to assign an ostensible distribution $\wp_{\mathrm{ost}}(\breve{y}_t)$ to $y(t)$ and just disregard the norm of the state. When $\wp_{\mathrm{ost}}(\breve{y}_t)$ is a Gaussian with zero mean and variance $1/dt$ the state evolves according to a *linear* stochastic master equation given by

$$d\bar{\rho}_{y_t}(t) = \mathcal{L}\rho(t)\,dt + y(t)dt\,\bar{\mathcal{H}}[\hat{c}]\,\rho(t), \tag{12}$$

where the updated state is now un-normalized

$$d\bar{\rho}_{y_t}(t) = \bar{\rho}_{y_t}(t+dt) - \rho(t), \tag{13}$$

and we have defined a "linear" version of (11)

$$\bar{\mathcal{H}}[\hat{A}]\hat{B} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}^\dagger. \tag{14}$$

The solutions to (12) are called linear quantum trajectories [42]. For linear quantum trajectories to be equivalent to the standard theory [the case when the measurement results are distributed according to their true distribution $\wp(\breve{y}_t)$ and the state follows a nonlinear stochastic master equation] the ostensible distribution must be such that [4]

$$\wp(\breve{y}_t)\,d\breve{y}_t = \mathrm{Tr}\big[\bar{\rho}_{y_t}(t+dt)\big]\,\wp_{\mathrm{ost}}(\breve{y}_t)\,d\breve{y}_t. \tag{15}$$

The form of the linear stochastic master equation is specific to the form of $\wp_{\mathrm{ost}}(\breve{y}_t)$, so choosing a different form for $\wp_{\mathrm{ost}}(\breve{y}_t)$ will result in a different linear stochastic master equation [4].

In the above we have assumed the prior state to be unconditioned [implied by (10)], no matter how it is obtained. In general this need not be the case and certainly will not be once the prior state is propagated for any finite time interval under (12) [or (9)]. Thus for generality and clarity we will introduce a Roman subscript c (for "conditioned") and write the linear stochastic master equation as

$$d\bar{\rho}_{\mathrm{c}}(t) = \mathcal{L}\,\rho(t)\,dt + y(t)\,dt\,\bar{\mathcal{H}}[\hat{c}]\,\rho(t), \qquad (16)$$

where

$$d\bar{\rho}_{\mathrm{c}}(t) = \bar{\rho}_{\mathrm{c}}(t+dt) - \rho(t). \qquad (17)$$

Tacked onto Eqs. (16) and (17) is the system of notation that $\rho(t)$ denotes a state of knowledge that is arbitrarily conditioned on the measurement record during $[0,t)$. Once $\rho(t)$ is determined the precise conditioning of $\bar{\rho}_{\mathrm{c}}(t+dt)$ can then be expressed. We will inform the reader about the dependence of $\rho(t)$ on the measurement record when it is necessary to assume a particular conditioning for $\rho(t)$.

2. The heterodyne stochastic master equation

A heterodyne measurement of efficiency η can be shown to be formally equivalent to two simultaneous measurements of orthogonal quadratures of the signal field by two homodyne detectors each with efficiency $\eta/2$. This latter detection scheme of using two homodyne detectors is also referred to as a dual-homodyne measurement. It is a convenient way of understanding the heterodyne measurement and is how we will think about it. That is, while we use the term "heterodyne" often, we will always realize it using a dual-homodyne scheme. This decomposition of a heterodyne measurement into two homodyne ones will also be of use later in Sec. IV.

If we now keep the above example of a leaky cavity and perform a heterodyne (or dual-homodyne) measurement of the output of the cavity then we obtain two currents, y_1 and y_2, which we write as the components of a 2×1 vector $\mathbf{y}$,

$$\mathbf{y}\,dt = \sqrt{\frac{\eta}{2}} \begin{pmatrix} \langle \hat{c} + \hat{c}^\dagger \rangle \\ -i\langle \hat{c} - \hat{c}^\dagger \rangle \end{pmatrix} dt + d\mathbf{w}, \qquad (18)$$

where $d\mathbf{w} = (dw_1, dw_2)^\top$ and dw_1 and dw_2 are independent Wiener increments,

$$dw_1(t)\,dw_2(t') = 0 \quad \forall\, t, t'. \qquad (19)$$

Conditioning the system state on (18) then leads to the heterodyne stochastic master equation

$$d\rho_{\mathrm{c}}(t) = \mathcal{L}\,\rho(t)\,dt$$
$$+ \sqrt{\frac{\eta}{2}}\,\mathcal{H}[\hat{c}]\,\rho(t)\,dw_1 + \sqrt{\frac{\eta}{2}}\,\mathcal{H}[-i\hat{c}]\,\rho(t)\,dw_2. \qquad (20)$$

We can write this more compactly by defining the row vector

$$\mathsf{M} = \sqrt{\frac{\eta}{2}}\,(1, i). \qquad (21)$$

Equation (20) is then equivalent to

$$d\rho_{\mathrm{c}}(t) = \mathcal{L}\,\rho(t)\,dt + d\mathbf{w}^\top \mathcal{H}[\mathsf{M}^\dagger \hat{c}]\,\rho(t). \qquad (22)$$

Note here the use of the vector-operator-valued superoperator (A66), defined for any $\hat{\mathbf{A}}$ and $\hat{B}$ as

$$\mathcal{H}[\hat{\mathbf{A}}]\hat{B} \equiv \hat{\mathbf{A}}\hat{B} + \hat{B}\hat{\mathbf{A}}^\ddagger - \mathrm{Tr}[\hat{\mathbf{A}}\hat{B} + \hat{B}\hat{\mathbf{A}}^\ddagger]\hat{B}, \qquad (23)$$

where $\hat{\mathbf{A}}^\ddagger \equiv (\hat{\mathbf{A}}^\top)^\dagger$. The heterodyne measurement illustrates how $\mathbf{y}$ can be a two-component vector even though $L = 1$ so there is only one output field $\hat{b}_{\mathrm{out}}$.

III. THE M REPRESENTATION

In this section we show how a single complex matrix can capture any such properties of diffusive measurements of an arbitrary number of system outputs in the Schrödinger picture. In Sec. IV a second and physically intuitive method of parameterizing diffusive unravellings is formulated in the Heisenberg picture.

A. Measurement statistics

Based on the foregoing examples of homodyne and heterodyne unravellings in Sec. II B we propose the following form of a general diffusive stochastic master equation

$$\hbar\,d\rho_{\mathrm{c}}(t) = \mathcal{L}\,\rho(t)\,dt + d\mathbf{w}^\top \mathcal{H}[\mathsf{M}^\dagger \hat{\mathbf{c}}]\,\rho(t), \qquad (24)$$

conditioned on the $2L \times 1$ real vector

$$\mathbf{y}(t)\,dt = \mathrm{E}[\mathbf{y}(t)]dt + d\mathbf{w}(t), \qquad (25)$$

which represents the measured current. The length of $\mathbf{y}$ is motivated by the example of heterodyne detection in Sec. II B 2 and the fact that here we are allowing L components in $\hat{\mathbf{b}}_{\mathrm{out}}$. The measurement is a noisy process. The noise in the measurement is given by a vector of independent Wiener increments $d\mathbf{w}$ (also $2L \times 1$ by default). The vector $d\mathbf{w}(t)$ therefore has zero mean

$$\mathrm{E}[d\mathbf{w}(t)] = \mathbf{0} \quad \forall\, t, \qquad (26)$$

and correlation matrices

$$d\mathbf{w}(t)\,d\mathbf{w}^\top(t) = \mathrm{I}_{2L}\,dt, \qquad (27)$$

$$d\mathbf{w}(t)\,d\mathbf{w}^\top(t') = 0 \quad \forall\, t \neq t'. \qquad (28)$$

In (24) we have used (23) and defined M to be an $L \times 2L$ complex matrix that parameterizes the unravelling.[3] The constraints which define M will be determined in the following sections, here we will obtain the form of the measured current consistent with (24).

The precise of form of $\mathrm{E}[\mathbf{y}(t)]$ can in fact be derived from the theory of linear quantum trajectories which allows us to calculate the actual statistics of the measured current by using

---

[3]In the case when some of the columns of M are zero, say $R\ (\leqslant 2L)$, then $d\mathbf{w}$ can be defined to be $R \times 1$ instead of $2L \times 1$ (in fact we do this in Ref. [28]). Defining M to be $L \times 2L$ and $d\mathbf{w}$ to be $2L \times 1$ here just simplifies our theory.

an ostensible distribution in the following way: Here we are only concerned with the average of $\mathbf{y}_t$ and this is

$$\mathrm{E}[\mathbf{y}(t)]dt = \bar{\mathrm{E}}[\mathbf{y}(t)\,\mathrm{Tr}\{\rho(t) + d\bar{\rho}_{\mathrm{c}}(t)\}]\,dt, \qquad (29)$$

where $\rho(t)$ is normalized and $d\bar{\rho}_{\mathrm{c}}(t)$ is given by

$$\hbar\, d\bar{\rho}_{\mathrm{c}}(t) = \mathcal{L}\rho(t)\,dt + \mathbf{y}^{\top}(t)\bar{\mathcal{H}}[\mathsf{M}^{\dagger}\hat{\mathbf{c}}]\,\rho(t)\,dt. \qquad (30)$$

Here we have used (15) to rewrite the average on the left-hand side as an average with respect to the ostensible distribution (denoted by an overbar on E)

$$\wp_{\mathrm{ost}}(\breve{\mathbf{y}}_t) = \left(\frac{dt}{\sqrt{2\pi}}\right)^{L}\exp\left[-\frac{1}{2}\breve{\mathbf{y}}_t^{\top}\breve{\mathbf{y}}_t\,dt\right]. \qquad (31)$$

This is a Gaussian with mean zero and covariance

$$\bar{\mathrm{E}}[\mathbf{y}(t)\,\mathbf{y}^{\top}(t)] = \mathrm{I}_{2L}/dt. \qquad (32)$$

Defining a linear version of (23),

$$\mathcal{H}[\hat{\mathbf{A}}]\hat{B} \equiv \hat{\mathbf{A}}\hat{B} + \hat{B}\hat{\mathbf{A}}^{\ddagger}, \qquad (33)$$

we then obtain, on substituting (30) into (29),

$$\mathrm{E}[\mathbf{y}(t)]dt = \hbar^{-1}\,\bar{\mathrm{E}}[\mathbf{y}(t)dt\,\mathbf{y}^{\top}(t)dt\,\mathrm{Tr}\{\bar{\mathcal{H}}[\mathsf{M}^{\dagger}\hat{\mathbf{c}}]\,\rho(t)\}]$$
$$= \hbar^{-1}\,\mathrm{Tr}\{\mathsf{M}^{\dagger}\hat{\mathbf{c}}\,\rho(t) + \rho(t)\mathsf{M}^{\top}\hat{\mathbf{c}}^{\ddagger}\} \qquad (34)$$
$$= \frac{dt}{\hbar}\,\langle\mathsf{M}^{\dagger}\hat{\mathbf{c}} + \mathsf{M}^{\top}\hat{\mathbf{c}}^{\ddagger}\rangle. \qquad (35)$$

We have used the facts that $\mathcal{L}\rho$ is traceless and that $\bar{\mathrm{E}}[\mathbf{y}] = \mathbf{0}$ in (34). Later (Sec. IV) we will be considering a different representation of diffusive measurements. In order to distinguish between different representations we will use a subscript on the current. Substituting (35) back into (25) we will thus write

$$\hbar\,\mathbf{y}_{\mathsf{M}}\,dt = \langle\mathsf{M}^{\dagger}\hat{\mathbf{c}} + \mathsf{M}^{\top}\hat{\mathbf{c}}^{\ddagger}\rangle\,dt + \hbar\,d\mathbf{w}. \qquad (36)$$

It is clear from (35) that M determines what property of the system gets measured, represented by some Hermitian operator $\hat{f}(\hat{\mathbf{c}})$, and also the measurment statistics (i.e., the statistics of $\mathbf{y}_t$) for a measurement of $\hat{f}(\hat{\mathbf{c}})$. That M defines the diffusive measurement could have been anticipated from (24) since M appears in the $\mathcal{H}$ superoperator which stems from considering measurement. Its appearance in (24) also means that not every M will be a valid parametrization of a diffusive measurement. In Sec. III B we find the necessary and sufficient condition for M to be valid.

The careful reader who is familiar with the quantum theory of indirect measurements [4,43] will know that the conditioning for the state is in general just a label which allows us to distinguish between measurement outcomes, and as such it does not have to be real. Our choice of conditioning the state on a real vector is motivated by its use in control theory in which the prevalent treatments regard $\mathbf{y}$ as a real process [44].

The assumption that $\rho(t)$ is unnormalized can be relaxed. Say the measurement began at time 0 and that $\bar{\rho}_{\mathrm{c}}(t)$ is obtained by evolving $\rho(0)$ via (30) to time t. In this case the correct modification to the above calculation is to divide the trace in (29) by $\mathrm{Tr}\{\bar{\rho}_{\mathrm{c}}(t)\}$. Keeping the same ostensible distribution as (31) we get

$$\mathrm{E}[\mathbf{y}_t | \mathbf{y}_{[0,t)}]dt = \bar{\mathrm{E}}\left[\mathbf{y}_t\,\frac{\mathrm{Tr}\{d\bar{\rho}_{\mathrm{c}}(t)\}}{\mathrm{Tr}\{\bar{\rho}_{\mathrm{c}}(t)\}}\right]dt, \qquad (37)$$

where on the right-hand side the state $\bar{\rho}_{\mathrm{c}}$ is conditioned on the measurement record

$$\mathbf{y}_{[0,t)} \equiv \{\mathbf{y}(s) \mid 0 \leqslant s < t\}. \qquad (38)$$

The end result from calculating (37) is to replace the unconditioned averages in (35) by conditional ones,

$$\langle\hat{\mathbf{c}}\rangle \longrightarrow \langle\hat{\mathbf{c}}\rangle_{\mathrm{c}} = \mathrm{Tr}[\hat{\mathbf{c}}\bar{\rho}_{\mathrm{c}}(t)], \qquad (39)$$

as one would have guessed, where

$$\rho_{\mathrm{c}}(t) = \frac{\bar{\rho}_{\mathrm{c}}(t)}{\mathrm{Tr}[\bar{\rho}_{\mathrm{c}}(t)]}. \qquad (40)$$

B. The set of allowed M

1. Necessary condition for M

If M is to be a valid parametrization of a diffusive quantum measurement it must be such that (24) evolves a valid state at time t to another valid state $\rho(t + dt)$ for all t. Here we remind the reader that a valid state is represented by an Hermitian operator that is normalized and positive. From this it follows that its eigenvalues all lie in the interval [0,1], which we can write as an operator inequality $0 \leqslant \rho \leqslant \hat{1}$. The positivity condition is the only nontrivial criterion because $\mathcal{L}\rho$ and $\mathcal{H}\rho$ are each Hermitian and traceless so (24) will always preserve Hermiticity and normalization for any M. Since these conditions must hold for any measurement process, they are necessary for a given M to be valid and we establish the following implication

$$\mathsf{M}\ \text{is valid}\ \Rightarrow\ \mathsf{M}: 0 \leqslant \rho(t) + d\rho_{\mathrm{c}}(t) \leqslant \hat{1}, \qquad (41)$$

where $\rho(t)$ is understood to be any valid state and $d\rho_{\mathrm{c}}(t)$ is given by (24). It is simple [given $d\rho_{\mathrm{c}}(t)$ is traceless] to see that

$$0 \leqslant \rho(t) + d\rho_{\mathrm{c}}(t) \leqslant \hat{1} \implies \mathrm{Tr}\{[\rho(t) + d\rho_{\mathrm{c}}(t)]^2\} \leqslant 1. \quad (42)$$

Without loss of generality we can assume that $\rho(t)$ is a pure state, because of the convexity of $\mathrm{Tr}[\rho^2]$ in ρ. We will show

$$\mathrm{Tr}\{[\rho(t) + d\rho_{\mathrm{c}}(t)]^2\} \leqslant 1 \iff \mathsf{M}\mathsf{M}^{\dagger}/\hbar \in \mathfrak{H}, \qquad (43)$$

where $\mathfrak{H}$ is

$$\mathfrak{H} = \{\mathsf{H} = \mathrm{diag}(\boldsymbol{\eta}) \,|\, \forall\, k,\, \eta_k \in [0,1]\}, \qquad (44)$$

thereby establishing the right-hand side of the equivalence (43) as a necessary condition for M to be valid.

The equivalence (43) can be proven by first proving the following lemma:

$$\mathrm{Tr}\{[\rho(t) + d\rho_{\mathrm{c}}(t)]^2\} = 1 \iff \mathsf{M}\mathsf{M}^{\dagger}/\hbar = \mathrm{I}_L, \qquad (45)$$

where I_L is the $L \times L$ identity matrix. It can then be shown (see Appendix B) that

$$\mathrm{Tr}\{[\rho(t) + d\rho_{\mathrm{c}}(t)]^2\} = 1 + \frac{2}{\hbar}\,\mathrm{tr}[H(\mathsf{M}\mathsf{M}^{\dagger}/\hbar - \mathrm{I}_L)^{\top}]\,dt, \quad (46)$$

where H is the Hermitian, positive-semidefinite matrix (not to be confused with the Hamiltonian $\hat{H}$, which is an operator)

$$H = \langle(\hat{\mathbf{c}} - \langle\hat{\mathbf{c}}\rangle)^{\ddagger}(\hat{\mathbf{c}} - \langle\hat{\mathbf{c}}\rangle)^{\top}\rangle. \qquad (47)$$

Since $\rho(t)$ is an arbitrary pure state, H can be assumed to be strictly positive (i.e., positive definite). Thus (46) will be 1 if

and only if $\mathsf{M}\mathsf{M}^\dagger/\hbar = \mathsf{I}_L$. This completes the proof of (45), a necessary condition on M in the case of efficient monitoring (i.e., monitoring which preserves the purity of the state).

To prove (43) we must consider the case when the purity at time $t+dt$ drops to below 1. Given that we had a pure state at time t this is possible if and only if we had inefficient monitoring. For a single decay channel we can bundle the sum of all losses[4] into a single parameter η and consider only a fraction η (between 0 and 1) of the system operator $\hat{c}$ to be measured perfectly. This can by modelled by introducing an imaginary beam-splitter with transmission coefficient η in the path of the decay channel followed by an ideal detector [45]. When multiple decay channels are present we simply repeat this setup for each channel. This motivates us to rewrite the master equation (1) as

$$\hbar\,\dot\rho = -i\,[\hat{H},\rho] + \mathcal{D}[\sqrt{\mathsf{I}_L - \mathsf{H}}\,\hat{\mathbf{c}}]\rho + \mathcal{D}[\sqrt{\mathsf{H}}\,\hat{\mathbf{c}}]\rho, \quad (48)$$

where $\mathsf{H} \in \mathfrak{H}$ [recall (44)] and unravel the last term with unit detection efficiency. Such an unravelling is defined by the stochastic master equation

$$\hbar\,d\rho_{\mathrm{c}} = \mathcal{L}\rho\,dt + d\mathbf{w}^\top \mathcal{H}[\mathsf{M}'^\dagger \hat{\mathbf{c}}']\rho, \quad (49)$$

where $\hat{\mathbf{c}}' \equiv \sqrt{\mathsf{H}}\,\hat{\mathbf{c}}$ and (49) is conditioned on the current

$$\mathbf{y}_{\mathsf{M}'}\,dt = \frac{dt}{\hbar}\,\langle\mathsf{M}'^\dagger \hat{\mathbf{c}}' + \mathsf{M}'^\top \hat{\mathbf{c}}'^\ddagger\rangle + d\mathbf{w}. \quad (50)$$

Since (49) and (50) describe efficient monitoring, from the lemma of (45), M' must satisfy

$$\mathsf{M}'\mathsf{M}'^\dagger/\hbar = \mathsf{I}_L. \quad (51)$$

The unravelling defined by (49), (50), and (51) is equivalent to (24), (25), and (35), but the latter make the measurement efficiency H explicit. The set of quantum trajectories generated by (24) must therefore be the same as the set generated by (49). This will be the case if and only if the two equations have the same stochastic term, i.e.,

$$d\mathbf{w}^\top \mathcal{H}[\mathsf{M}^\dagger \hat{\mathbf{c}}]\rho = d\mathbf{w}^\top \mathcal{H}[\mathsf{M}'^\dagger \hat{\mathbf{c}}']\rho. \quad (52)$$

This is true for any ρ and $\hat{\mathbf{c}}$ if and only if $\mathsf{M} = \sqrt{\mathsf{H}}\,\mathsf{M}'$. This gives

$$\mathsf{M}\mathsf{M}^\dagger = \mathsf{H}. \quad (53)$$

We have arrived at (53) by considering inefficient detection. However, there are other properties of the measurement that one would like to capture with M so it would seem that (53) is necessary but not sufficient. As we show next, (53) is in fact, surprisingly, a sufficient constraint on M for it to be a valid parametrization of general diffusive measurements.

2. Sufficient condition for M

To show that (53) is a sufficient condition we will connect M to another parametrization U, which is a different way of representing diffusive measurements and has conditions that have previously been shown to be necessary and sufficient

———————

[4] Note that "loss" here refers to *any* process which leads to a loss of information, not necessarily imperfections of the measuring device.

[26,27]. In this paper we introduce an elegant way to connect M and U by considering a generalized diffusion operator which we denote by $\hat{\mathsf{D}}$. If the system state is in an N-dimensional Hilbert space $\mathbb{H}$, then $\hat{\mathsf{D}}$ is an operator in $\mathbb{H} \otimes \mathbb{H}$, defined by

$$\hat{\mathsf{D}}\,dt \equiv d\rho_{\mathrm{c}} \otimes d\rho_{\mathrm{c}}. \quad (54)$$

This is the mathematical object which characterizes the set of all equivalent representations (i.e., M matrices) of a given unravelling as we now explain.

Stochastic paths of the quantum state itself are rather abstract but we can make the trajectories more concrete by considering

$$\boldsymbol{\rho}_{\mathrm{c}} = \mathrm{Tr}[\hat{\mathbf{e}}\,\rho_{\mathrm{c}}], \quad (55)$$

where $\hat{\mathbf{e}} = (\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_{N^2})^\top$ is an operator basis for all linear Hermitian operators. Thus $\boldsymbol{\rho}_{\mathrm{c}}$ is a stochastic process in $\mathbb{R}^{N^2}$, satisfying

$$d\boldsymbol{\rho}_{\mathrm{c}} = \boldsymbol{A}\,dt + B\,d\mathbf{w}. \quad (56)$$

Note that $\boldsymbol{A} = \mathrm{Tr}[\hat{\mathbf{e}}\mathcal{L}\rho]$ is a vector while

$$B = \mathrm{Tr}\{\hat{\mathbf{e}}\mathcal{H}[\hat{\mathbf{c}}^\top \mathsf{M}]\rho\} \quad (57)$$

is a matrix. Recall that $\mathcal{L}$ and $\mathcal{H}[\hat{\mathbf{c}}^\top \mathsf{M}]$ are defined by (1) and (23), respectively. The diffusion matrix for $\boldsymbol{\rho}_{\mathrm{c}}$ is given by

$$D\,dt = d\boldsymbol{\rho}_{\mathrm{c}}\,d\boldsymbol{\rho}_{\mathrm{c}}^\top = BB^\top dt. \quad (58)$$

We can see how $\hat{\mathsf{D}}$ arises by rewriting the right-hand side of the first equality in (58),

$$\begin{aligned} D\,dt &= \mathrm{Tr}[\hat{\mathbf{e}}\,d\rho_{\mathrm{c}}]\,\mathrm{Tr}[\hat{\mathbf{e}}^\top d\rho_{\mathrm{c}}]\\ &= \mathrm{Tr}[(\hat{\mathbf{e}}\,d\rho_{\mathrm{c}}) \otimes (\hat{\mathbf{e}}^\top d\rho_{\mathrm{c}})] = \mathrm{Tr}[(\hat{\mathbf{e}} \otimes \hat{\mathbf{e}}^\top)\hat{\mathsf{D}}]. \quad (59)\end{aligned}$$

Thus $\hat{\mathsf{D}}$ is an operator-valued diffusion coefficient whose trace against $\hat{\mathbf{e}} \otimes \hat{\mathbf{e}}^\top$ gives us the diffusion of the more tangible process (55). Turning ρ_{c}, a stochastic process in $\mathbb{H}$, into $\boldsymbol{\rho}_{\mathrm{c}}$, an equivalent stochastic process in $\mathbb{R}^{N^2}$, allows us to grasp the abstract diffusion operator $\hat{\mathsf{D}}$ by using well-known properties of classical stochastic processes; namely that the diffusion of $\boldsymbol{\rho}_{\mathrm{c}}$ is characterized by the matrix D and that for a given D there are many matrices B such that $BB^\top = D$. The equivalence between $\boldsymbol{\rho}_{\mathrm{c}}$ and ρ_{c}, and more specifically between B and M, and also between D and $\hat{\mathsf{D}}$ [as given by (57) and (59)], mean that $\hat{\mathsf{D}}$ characterizes the diffusion of ρ_{c} just as D characterizes the diffusion of $\boldsymbol{\rho}_{\mathrm{c}}$. There will be many matrices M that generate the same $\hat{\mathsf{D}}$ (through the nonlinear term $\mathcal{H}[\hat{\mathbf{c}}^\top \mathsf{M}]\rho$) just as there are more than one B corresponding to a given D.

We mentioned earlier in Sec. II A that an unravelling can be defined as the set of solutions to the stochastic master equation and this set will correspond to some M. In genera, the set of solutions will change everytime M is changed but it may also be possible to have two different choices of M, say M_1 and M_2, which generate the same solution set. In this case we say that M_1 and M_2 are equivalent representations of the same unravelling. Let us substitute (24) into (54) and define

$$\mathsf{T} \equiv \begin{pmatrix}\mathsf{T}_1\\ \mathsf{T}_2\end{pmatrix} \equiv \begin{pmatrix}\mathrm{Re}[\mathsf{M}]\\ \mathrm{Im}[\mathsf{M}]\end{pmatrix}, \quad (60)$$

where T_1 and T_2 are each real and $L \times 2L$ so T is a real $2L \times 2L$ matrix. We have defined the real and imaginary parts of an arbitrary complex matrix A by

$$\text{Re}[A] \equiv \frac{1}{2}\left(A + A^*\right), \quad \text{Im}[A] \equiv \frac{-i}{2}\left(A - A^*\right). \quad (61)$$

We can then write M as

$$\mathsf{M} = \mathsf{T}_1 + i\,\mathsf{T}_2. \quad (62)$$

The vector operator $\mathsf{M}^\dagger \hat{\mathbf{c}}$ can be rewritten as

$$\mathsf{M}^\dagger \hat{\mathbf{c}} = \mathsf{T}^\top \hat{\mathbf{c}}_\&, \quad (63)$$

where $\hat{\mathbf{c}}_\&^\top \equiv (\hat{\mathbf{c}}^\top, -i\,\hat{\mathbf{c}}^\top)$. Using this, (54) is

$$
\begin{aligned}
\hat{\mathsf{D}}\,dt &= \mathcal{H}[\hat{\mathbf{c}}_\&^\top \mathsf{T}\,d\mathbf{w}]\,\rho \otimes \mathcal{H}[d\mathbf{w}^\top \mathsf{T}^\top \hat{\mathbf{c}}_\&]\,\rho \\
&= \mathcal{H}[\hat{\mathbf{c}}_\&^\top]\,\rho\,\mathsf{T}\,d\mathbf{w} \otimes d\mathbf{w}^\top \mathsf{T}^\top \mathcal{H}[\hat{\mathbf{c}}_\&]\,\rho \\
&= \hbar\,\mathcal{H}[\hat{\mathbf{c}}_\&^\top]\,\rho \otimes \mathsf{U}\,\mathcal{H}[\hat{\mathbf{c}}_\&]\,\rho\,dt.
\end{aligned}
\quad (64)
$$

In the last line we have defined a $2L \times 2L$ real matrix U by

$$\mathsf{T}\mathsf{T}^\top = \hbar\,\mathsf{U} \equiv \hbar \begin{pmatrix} \mathsf{U}_{11} & \mathsf{U}_{12} \\ \mathsf{U}_{21} & \mathsf{U}_{22} \end{pmatrix}, \quad (65)$$

where U_{mn} are $L \times L$ blocks of U. Equation (64) tells us that given the set of Lindblad operators $\hat{\mathbf{c}}$, U is what uniquely specifies $\hat{\mathsf{D}}$, and therefore the unravelling. That U defines the unravelling was independently formulated in Ref. [26] for the case of pure-state trajectories and later generalized to include nonunit detection efficiency in Ref. [27]. The necessary and sufficient conditions for a U to be valid [in the same sense as (41)] are

$$\mathsf{U} \geqslant 0, \quad (66)$$

$$\mathsf{U}_{11} + \mathsf{U}_{22} \in \mathfrak{H}, \quad (67)$$

$$\mathsf{U}_{12} = \mathsf{U}_{21}. \quad (68)$$

We will denote the set of valid U matrices as

$$\mathfrak{U} = \left\{ \begin{pmatrix} \mathsf{U}_{11} & \mathsf{U}_{12} \\ \mathsf{U}_{21} & \mathsf{U}_{22} \end{pmatrix} \,\middle|\, \mathsf{U} \geqslant 0, \mathsf{U}_{11} + \mathsf{U}_{22} \in \mathfrak{H}, \mathsf{U}_{12} = \mathsf{U}_{21} \right\}. \quad (69)$$

The question is now whether any U derived from (53) is valid. Since T is a real matrix, the definition (65) ensures (66) for any T. Using (65) conditions (67) and (68) become, respectively,

$$\mathsf{T}_1 \mathsf{T}_1^\top + \mathsf{T}_2 \mathsf{T}_2^\top \in \mathfrak{H}, \quad (70)$$

$$\mathsf{T}_1 \mathsf{T}_2^\top = \mathsf{T}_2 \mathsf{T}_1^\top. \quad (71)$$

This defines the set of valid T matrices

$$\mathfrak{T} = \left\{ \begin{pmatrix} \mathsf{T}_1 \\ \mathsf{T}_2 \end{pmatrix} \,\middle|\, \mathsf{T}_1 \mathsf{T}_1^\top + \mathsf{T}_2 \mathsf{T}_2^\top \in \mathfrak{H}, \mathsf{T}_1 \mathsf{T}_2^\top = \mathsf{T}_2 \mathsf{T}_1^\top \right\}. \quad (72)$$

It is easy to see, on substituting (62) into (53), that (70) and (71) are satisfied and thereby showing the sufficiency (and in fact necessity) of (53) for M to be a valid representation of an arbitrary diffusive unravelling:

$$\mathsf{U}(\mathsf{M}) \in \mathfrak{U} \iff \mathsf{M} \in \mathfrak{M}, \quad (73)$$

where $\mathfrak{M}$ is the set of valid M matrices

$$\mathfrak{M} = \{\mathsf{M} \mid \mathsf{M}\mathsf{M}^\dagger \in \mathfrak{H}\}. \quad (74)$$

Note that even though we have imported a previous result which proves altogether the sufficiency and necessity of (53), our analysis in Sec. III B 1 remains instrumental; since without it there is no reason to consider the set $\mathfrak{M}$. For ease of reference we call the members of $\mathfrak{M}$ "M-reps" and similarly the members of $\mathfrak{U}$ "U-reps."

C. Comparison between $\mathfrak{M}$ and $\mathfrak{U}$

Here we discuss some general features of M-reps and compare them with U-reps. To do so we first recall the essential features of the U-reps from Refs. [26,27]. There, Wiseman and coworkers derived a stochastic master equation of the form

$$d\rho_{\mathrm{c}} = \mathcal{L}\,\rho(t)\,dt + \mathcal{H}[d\mathbf{z}^\dagger \hat{\mathbf{c}}]\,\rho(t) \quad (75)$$

[with $\mathcal{L}$ given by (1)] that conditioned on an $L \times 1$ complex current

$$\mathbf{J}\,dt = \langle \mathsf{H}\hat{\mathbf{c}} + \mathsf{Y}\hat{\mathbf{c}}^\ddagger \rangle\,dt + d\mathbf{z}, \quad (76)$$

where $d\mathbf{z}$ (also $L \times 1$ by default) is a complex white-noise increment defined by $\mathrm{E}[d\mathbf{z}] = \mathbf{0}$ and the correlations

$$d\mathbf{z}\,d\mathbf{z}^\dagger = \hbar\,\mathsf{H}\quad dt, \quad d\mathbf{z}\,d\mathbf{z}^\top = \hbar\,\mathsf{Y}\,dt. \quad (77)$$

The matrix H is a member of $\mathfrak{H}$, while Y is complex and only has to be symmetric. The matrix U, or the "unravelling matrix" as it was originally called, is defined as the correlations of $\left(\text{Re}[d\mathbf{z}^\top], \text{Im}[d\mathbf{z}^\top]\right)/\sqrt{\hbar}$. This can be shown to be

$$\mathsf{U} = \frac{1}{2}\begin{pmatrix} \mathsf{H} + \text{Re}[\mathsf{Y}] & \text{Im}[\mathsf{Y}] \\ \text{Im}[\mathsf{Y}] & \mathsf{H} - \text{Re}[\mathsf{Y}] \end{pmatrix}, \quad (78)$$

which is consistent with (69).

The first noteworthy feature of M-reps is that they capture physically valid measurements by a single equation, (53), whereas U-reps require three, namely (66)–(68). This point makes our theory of M-reps rather elegant. If, however, we compare the number of real parameters required to specify M to that of U, then we find that U-reps are a more efficient parametrization: We know directly from (78) that the number of real parameters in U is $L^2 + 2L$, being the total number of real parameters required to specify H and Y. To find the number of real parameters required to specify M, we note that M is $L \times 2L$, so the total number of real parameters in M is $4L^2$. But (53) means that $L^2 - L$ of these parameters are redundant (since it imposes $L^2 - L$ independent real conditions on the elements of M). Subtracting this redundancy from $4L^2$ leaves $3L^2 + L$. Thus M-reps require $2L^2 - L$ more parameters than U-reps.

The second attractive feature of M-reps is that M can be understood as the generalization of $\sqrt{\mathsf{H}}$, encapsulating more information than just measurement efficiency. This was in fact part of our motivation for developing M-reps and postulating (24), as stated at the start of Sec. III A. So in hindsight, the fact that M turns out to be the matrix square root of H, i.e., constrained by (53), seems "natural" given that we were thinking of generalizing $\sqrt{\eta}$ to $\sqrt{\mathsf{H}}$ and then from $\sqrt{\mathsf{H}}$ to M from the start. The interpretation of U-reps, on the other hand, was derived [summarized in (75)–(78)] as the correlation between the real and imaginary parts of a complex measurement output. This has the advantage that it reflects directly the quantum

trajectory diffusion as was evident in (64). Equations (66)–(68) are simply constraints on the allowed correlations of $d\mathbf{z}$.

Finally, we mentioned before [two paragraphs under (36)] that we prefer to work with a real current, namely $\mathbf{y}_{\rm M}$, for the purpose of feedback control even though quantum measurement theory does not impose such a restriction. However, by allowing the system state to be conditioned on a real vector, $\mathbf{y}_{\rm M}$ now has the capacity to be an observable (i.e., correspond to an actual physical process) rather than a mere dummy variable for distinguishing the different measurement outcomes. Of course, one may also prefer a real current over a complex one simply because it is real. Motivated by exactly these reasons Ref. [27] defines a real current from $\mathbf{J}$ as

$$\mathbf{y}\,dt \equiv \mathsf{T}^+ \begin{pmatrix} \mathrm{Re}[\mathbf{J}] \\ \mathrm{Im}[\mathbf{J}] \end{pmatrix} dt, \qquad (79)$$

where T^+ is the Moore-Penrose inverse of T [46]. Substituting (76) into (79) and using the properties of T^+ it can be shown that (79) simplifies to the form of (36) [recall (62)]. Thus the U-rep achieves the same current as the M-rep but less directly by deriving $\mathbf{y}$ via $\mathbf{J}$.

D. Relations among $\mathfrak{T}$, $\mathfrak{U}$, and $\mathfrak{M}$

The relation between $\mathfrak{M}$ and $\mathfrak{T}$ is straightforward, given by (60) and (62). From these it should be clear that $\mathfrak{M}$ maps one-to-one and onto $\mathfrak{T}$. We said above that M-reps are parameterized by $3L^2 + L$ real numbers, therefore T also has the same number of real parameters. This number is $2L^2 - L$ more than the number of real parameters in the U-rep. We can see where the extra parameters lie by returning to (65) which defines T as the matrix square root of U. We can always write, using the polar decompostion, $\mathsf{T} = \sqrt{\mathsf{TT}^\top/\hbar}\,\mathsf{O} \equiv \sqrt{\mathsf{U}}\,\mathsf{O}$, where $\sqrt{\mathsf{U}}$ is the positive matrix square root of U[5] and O is some $2L \times 2L$ orthogonal matrix which may or may not be unique given T (see following paragraph). The number of parameters in $\sqrt{\mathsf{U}}$ is $L^2 + 2L$, inherited from U, while the number of parameters in O is $2L^2 - L$.[6] Adding the number of parameters in $\sqrt{\mathsf{U}}$ and O gives $3L^2 + L$. The extra parameters thus lie in O, leading us to define the set of pairs (not to be confused with a block matrix)

$$\mathfrak{U}_{\hat{O}} = \{(\sqrt{\mathsf{U}},\mathsf{O})\,|\,\mathsf{U} \in \mathfrak{U}, \mathsf{O}^\top = \mathsf{O}^{-1}\}, \qquad (80)$$

where O is understood to have dimensions $2L \times 2L$.

We can divide the set $\mathfrak{T}$ into two subsets: a subset of invertible Ts and a subset of noninvertible Ts. The subset of Ts that are invertible maps one-to-one and onto $\mathfrak{U}_{\hat{O}}$, since in this case the polar decomposition of T is unique, i.e., given any T, both entries of the pair $(\sqrt{\mathsf{TT}^\top/\hbar},\mathsf{O}) \in \mathfrak{U}_{\hat{O}}$ are uniquely determined by T. The subset of Ts that are not invertible maps

---

[5]The positive square root of a matrix A is defined by a matrix B such that $B^2 = A$, and is itself positive-semidefinite. The positive square root always exists and is uniquely given by A so it is common to denote it by $\sqrt{A}$.

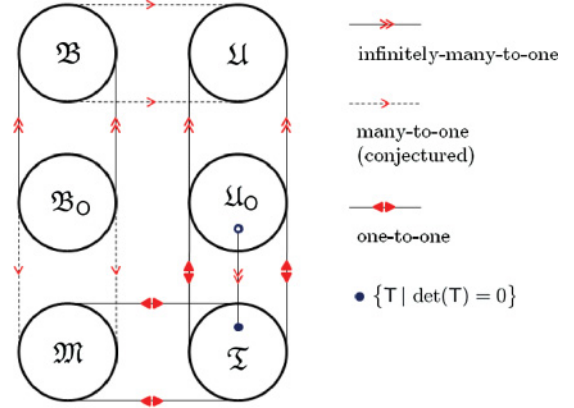[6]An $n \times n$ real orthogonal matrix is parameterized by $n(n-1)/2$ real numbers.

FIG. 2. (Color online) Multivalued mappings among the different parametrizations of diffusive quantum measurements. The lines have been drawn tangent to the circles to denote that one set is mapped onto the other, i.e., each member of one set has a corresponding member in the set to which they are connected to. Note also that we have defined a subset of $\mathfrak{T}$ whose members are noninvertible. This is a set of measure zero and maps to subset of $\mathfrak{U}_{\rm O}$ (shown as an unshaded dot), also a set of measure zero.

infinitely-many-to-one and onto a subset of $\mathfrak{U}_{\rm O}$. In this case the first entry of the pair $(\sqrt{\mathsf{TT}^\top/\hbar},\mathsf{O})$ will still be uniquely determined by T but corresponding to this T there are an infinite number of orthogonal matrices O. We depict these relations in Fig. 2 along with another parametrization $\mathfrak{B}$ to be introduced in Sec. IV. Note the cardinality of the noninvertible subset of $\mathfrak{T}$ is less than the cardinality of $\mathfrak{T}$. That is, the former is a subset of measure zero.

E. The M representation in the Heisenberg picture

We now consider diffusive quantum measurements in the Heisenberg picture where operators evolve and states are static, equal to their initial value. Moving from the Schrödinger picture to the Heisenberg picture requires two conceptual changes which we summarize below. In Sec. IV we will formulate another representation of diffusive measurements in the Heisenberg picture so the discussion here apply to Sec. IV as well.

The bath field $\hat{\mathbf{b}}_{\rm in}$ is now a vector-operator-valued white-noise process. As such, $d\hat{\mathbf{B}}_{\rm in}(t) \equiv \hat{\mathbf{b}}_{\rm in}(t)\,dt$ satisfies Itô rules analogous to (26)–(28) [47,48]; that is,

$$\langle d\hat{\mathbf{B}}_{\rm in}(t)\rangle = \mathbf{0}, \qquad (81)$$

where the angle brackets denote a quantum average and

$$d\hat{\mathbf{B}}_{\rm in}(t)\,d\hat{\mathbf{B}}_{\rm in}^\dagger(t) = \hbar\,\hat{\mathsf{I}}_L\,dt, \qquad (82)$$

with all other second-order moments negligible. Note that

$$\lfloor d\hat{\mathbf{B}}_{\rm in}(t), d\hat{\mathbf{B}}_{\rm in}^\ddagger(t')\rceil = 0 \quad \forall\, t \neq t', \qquad (83)$$

where we have used the matrix-operator bracket (A31), defined for any two vector operators $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ as

$$\lfloor \hat{\mathbf{A}}, \hat{\mathbf{B}}\rceil = \hat{\mathbf{A}}\hat{\mathbf{B}}^\top - (\hat{\mathbf{B}}\hat{\mathbf{A}}^\top)^\top. \qquad (84)$$

Thus all other second-order moments with unequal times also vanish. As a consequence of the singularity of $\hat{\mathbf{b}}_{\rm in}$ it is necessary
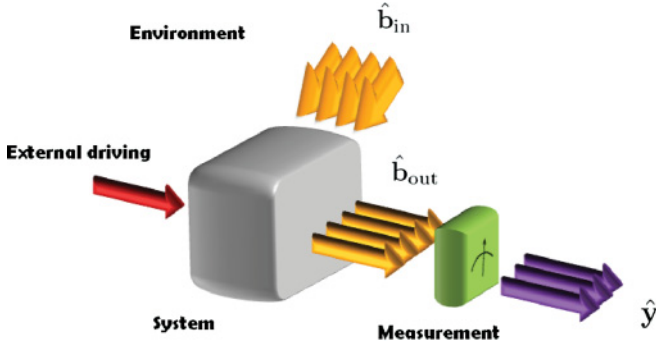
FIG. 3. (Color online) Quantum entanglement and measurement are two processes which have conventionally been described in the Schrödinger picture. When we move to the Heisenberg picture we move to a different mindset, where the system-bath entanglement and the measurement process are all captured by time-dependent operators, given by (86) [49,50] and (87).

to distinguish the bath field before and after its interaction with the system. This gives rise to an output field (Fig. 3)

$$d\hat{\mathbf{B}}_{\text{out}}(t) = \hat{U}^\dagger(t + dt, t) \, d\hat{\mathbf{B}}_{\text{in}}(t) \, \hat{U}(t + dt, t), \quad (85)$$

where $\hat{U}(t + dt, t)$ is defined by (3) but with $\hat{\mathbf{b}}$ replaced by $\hat{\mathbf{b}}_{\text{in}}$. From this one can show that

$$d\hat{\mathbf{B}}_{\text{out}}(t) = \hat{\mathbf{c}}(t) \, dt + d\hat{\mathbf{B}}_{\text{in}}(t). \quad (86)$$

This input-output relation and is what allows us to relate our measurement performed on the bath back to properties of the system [48,51].

We next turn the measured current $\mathbf{y}(t)$ into a vector operator [52–54]. Note that this does not mean we are taking the measured current to be a quantum object. The current is interpreted as a classical quantity but it is represented by a vector operator in the theory. The classicality of the measured current can be formally captured by the "self-nondemolition" property of the vector operator $\hat{\mathbf{y}}(t)$ [55–57],

$$\lfloor \hat{\mathbf{y}}(t), \hat{\mathbf{y}}(t') \rceil = 0 \, , \, \forall \, t, t'. \quad (87)$$

It should also correspond to an observable process so $\hat{\mathbf{y}}(t)$ must be Hermitian. In brief, quantum measurement theory in the Heisenberg picture just assigns an operator $\hat{\mathbf{y}}(t)$ (a quantum stochastic process) which reproduces the correct measurement statistics (i.e., the statistics of $\mathbf{y}(t)$, a classical stochastic process) [4].

If a vector operator $\hat{\mathbf{y}}$ is to represent the measured current $\mathbf{y}$ then it must be such that the statistics of $\mathbf{y}$ are reproduced, namely

$$\langle \hat{\mathbf{y}}(t) \rangle = \mathrm{E}[\mathbf{y}(t)] \quad (88)$$

and

$$[\hat{\mathbf{y}}(t) \, dt][\hat{\mathbf{y}}(t) \, dt]^\top = [\mathbf{y}(t) \, dt][\mathbf{y}(t) \, dt]^\top \hat{\mathbf{l}}. \quad (89)$$

Physically the measurement is never performed directly on the system but rather on the bath which involves detecting the output fields. Therefore $\hat{\mathbf{y}}$ should depend on $\hat{\mathbf{b}}_{\text{out}}$. It thus makes sense to write

$$\hbar \, \hat{\mathbf{y}}_\mathsf{M} \, dt = \mathsf{M}^\dagger d\hat{\mathbf{B}}_{\text{out}} + \mathsf{M}^\top d\hat{\mathbf{B}}_{\text{out}}^\ddagger + \hbar \, d\hat{\boldsymbol{v}}_\mathsf{M}. \quad (90)$$

We have defined the noise increment $d\hat{\boldsymbol{v}}_\mathsf{M}$ to be

$$d\hat{\boldsymbol{v}}_\mathsf{M} = \mathsf{L} \, d\hat{\mathbf{U}} + \mathsf{L}^* \, d\hat{\mathbf{U}}^\ddagger, \quad (91)$$

where L is $2L \times 2L$, to be chosen so that (88) and (89) are satisfied. The vector operator $d\hat{\mathbf{U}}$ is a free vacuum field just like $d\hat{\mathbf{B}}_{\text{in}}$ except that it is $2L \times 1$ so the only nonvanishing moment is

$$d\hat{\mathbf{U}}(t) \, d\hat{\mathbf{U}}^\dagger(t) = \hbar \, \hat{\mathbf{l}}_{2L} \, dt. \quad (92)$$

Note that this means (88) is automatically satisfied. Enforcing (89) we obtain

$$\hbar^2 \, \mathsf{L} \, \mathsf{L}^\dagger = \hbar \, \mathsf{I}_{2L} - \mathsf{M}^\dagger \mathsf{M}. \quad (93)$$

For convenience we will define $\mathsf{Z} = \hbar \, \mathsf{I}_{2L} - \mathsf{M}^\dagger \mathsf{M}$. We can thus choose $\hbar \, \mathsf{L}$ to be the positive square root of the right-hand side of (93) and write

$$\hbar \, \mathsf{L} = \sqrt{\mathsf{Z}}. \quad (94)$$

Note that $d\hat{\boldsymbol{v}}_\mathsf{M}$ is not a Wiener increment (it has correlations given by $\mathsf{Z} \, dt / \hbar$) but we can write $\hat{\mathbf{y}}_\mathsf{M}$ as

$$\hbar \, \hat{\mathbf{y}}_\mathsf{M} \, dt = (\mathsf{M}^\dagger \hat{\mathbf{c}} + \mathsf{M}^\top \hat{\mathbf{c}}^\ddagger) dt + \hbar \, [d\hat{\mathbf{v}}_\text{m}]_\mathsf{M}, \quad (95)$$

where

$$\hbar \, [d\hat{\mathbf{v}}_\text{m}]_\mathsf{M} = \mathsf{M}^\dagger d\hat{\mathbf{B}}_{\text{in}} + \mathsf{M}^\top d\hat{\mathbf{B}}_{\text{in}}^\ddagger + \hbar \, \mathsf{L} \, d\hat{\mathbf{U}} + \hbar \, \mathsf{L}^* \, d\hat{\mathbf{U}}^\ddagger \quad (96)$$

is a Wiener increment (when divided by $\hbar$), i.e.,

$$[d\hat{\mathbf{v}}_\text{m}(t)]_\mathsf{M} \, [d\hat{\mathbf{v}}_\text{m}(t)]_\mathsf{M}^\top = \hat{\mathbf{l}}_{2L} \, dt, \quad (97)$$

$$[d\hat{\mathbf{v}}_\text{m}(t)]_\mathsf{M} \, [d\hat{\mathbf{v}}_\text{m}(t')]_\mathsf{M}^\top = 0 \quad \forall \, t \neq t'. \quad (98)$$

We can show that $\hat{\mathbf{y}}_\mathsf{M}$ satisfies the self-nondemolition property (87) by first substituting either (90), or (95) into (87) and then considering the two cases $t = t'$ and $t \neq t'$ separately. It will be quicker to use (95) and write (87) in terms vacuum increments. Doing so gives

$$\lfloor \hat{\mathbf{y}}(t) \, dt, \hat{\mathbf{y}}(t') \, dt \rceil = \lfloor \, [d\hat{\mathbf{v}}_\text{m}(t)]_\mathsf{M}, [d\hat{\mathbf{v}}_\text{m}(t')]_\mathsf{M} \, \rceil = 0, \quad (99)$$

by virtue of (97) and (98).

Often one would measure, in the laboratory, the autocorrelation function of the photocurrent and its spectrum (which is the Fourier transform of the autocorrelation function at steady state), hence we now consider the autocorrelation function of $\hat{\mathbf{y}}_\mathsf{M}$. This is a two-time average which can be calculated by using the vector-operator quantum regression formula (A55). The derivation of the autocorrelation of $\hat{\mathbf{y}}_\mathsf{M}$ follows the same method as in Ref. [28] which derives the same quantity but with feedback. The only difference here is that these feedback terms would not appear so we will refer the reader to Ref. [28] for details. However, we make the following heuristic argument which leads us to the correct form for the correlation function: First generalizing the correlation function of the homodyne current when $L = 1$ found previously in Ref. [58] to the case of arbitrary L we obtain,

$$\begin{aligned}
&\langle \hat{\mathbf{y}}_\mathsf{M}(t) \, \hat{\mathbf{y}}_\mathsf{M}^\top(t + \tau) \rangle \\
&= (\mathrm{Tr}\{(\sqrt{\mathsf{H}} \, \hat{\mathbf{c}} + \sqrt{\mathsf{H}} \, \hat{\mathbf{c}}^\ddagger) e^{\mathcal{L}\tau} [\hat{\mathbf{c}}^\top \sqrt{\mathsf{H}} \, \rho(t) + \rho(t) \, \hat{\mathbf{c}}^\dagger \sqrt{\mathsf{H}}]\})^\top \\
&\quad + \hbar^2 \, \mathsf{I}_{2L} \, \delta(\tau), \quad (100)
\end{aligned}$$

where $\mathsf{H} \in \mathfrak{H}$ was defined in (44) and $\mathcal{L}$ is given by (1). The action of the superoperator and the trace are defined in (A15) and (A42), respectively. We said in Sec. III C that M is a generalization of $\sqrt{\mathsf{H}}$. So the next logical step would be to replace $\sqrt{\mathsf{H}}$ by M. However, this would not keep the Hermiticity of the terms standing on either side of $\exp(\mathcal{L}\tau)$, so the correct replacements in (100) should be

$$\sqrt{\mathsf{H}}\,\hat{\mathbf{c}} \longrightarrow \mathsf{M}^{\dagger}\hat{\mathbf{c}}, \quad \sqrt{\mathsf{H}}\,\hat{\mathbf{c}}^{\ddagger} \longrightarrow \mathsf{M}^{\top}\hat{\mathbf{c}}^{\ddagger}. \quad (101)$$

We then obtain

$$\begin{aligned}
&\hbar^2 \langle \hat{\mathbf{y}}_{\mathsf{M}}(t)\,\hat{\mathbf{y}}_{\mathsf{M}}^{\top}(t+\tau)\rangle \\
&= (\text{Tr}\{(\mathsf{M}^{\dagger}\hat{\mathbf{c}} + \mathsf{M}^{\top}\hat{\mathbf{c}}^{\ddagger})e^{\mathcal{L}\tau}[\hat{\mathbf{c}}^{\top}\mathsf{M}^{*}\rho(t) + \rho(t)\hat{\mathbf{c}}^{\dagger}\mathsf{M}]\})^{\top} \\
&\quad + \hbar^2 \mathsf{I}_{2L}\,\delta(\tau),
\end{aligned} \quad (102)$$

This is a generalization of the homodyne autocorrelation function found in Ref. [58] to arbitrary diffusive measurements.

IV. THE BLOCK REPRESENTATION

In this section we parametrize diffusive measurements by starting in the Heisenberg picture. The advantage of this is that it allows us to model the measurement by thinking directly about the physical transformations $\hat{\mathbf{b}}_{\text{out}}$ must undergo to produce a suitable current $\hat{\mathbf{y}}$. We show that the measurement block in Fig. 3 can be decomposed in a way shown in Fig. 4.

A. Definition

The blocks of Fig. 4 are as follows.

(i) Perhaps the first property of measurements that comes to one's mind is detection inefficiencies. For a single output this can be modelled by placing a beam-splitter with transmission coefficient η in the path of $\hat{b}_{\text{out}}$. The other input to the beam-splitter is not excited (i.e., no photons) so it is simply a field $\hat{v}$ in the vacuum state. We will choose a phase convention for the beam-splitter such that its output is

$$\hat{s}_{\text{in}} = \sqrt{\eta}\,\hat{b}_{\text{out}} + \sqrt{\bar{\eta}}\,\hat{v}, \quad (103)$$

where $\bar{\eta} \equiv 1 - \eta$. Note that we are retaining only one of the beam-splitter outputs since this is all that is necessary to model detector efficiencies. When multiple outputs from the system are present (103) simply generalizes to

$$\hat{\mathbf{s}}_{\text{in}} = \sqrt{\mathsf{H}}\,\hat{\mathbf{b}}_{\text{out}} + \sqrt{\bar{\mathsf{H}}}\,\hat{\mathbf{v}}, \quad (104)$$

where $\bar{\mathsf{H}} \equiv \mathsf{I}_L - \mathsf{H}$. This is the first block in Fig. 4.

(ii) Next, one is free to mix and phase shift the different components of $\hat{\mathbf{b}}_{\text{out}}$ to obtain an arbitrary linear combination of them. Hence we multiply $\hat{\mathbf{s}}_{\text{in}}$ by an $L \times L$ matrix to get

$$\hat{\mathbf{s}}_{\text{out}} = \mathsf{S}\,\hat{\mathbf{s}}_{\text{in}}. \quad (105)$$

We assume this mixing of the components of $\hat{\mathbf{b}}_{\text{out}}$ conserves total boson number so $\hat{\mathbf{s}}_{\text{in}}^{\dagger}\hat{\mathbf{s}}_{\text{in}} = \hat{\mathbf{s}}_{\text{out}}^{\dagger}\hat{\mathbf{s}}_{\text{out}}$. This constrains S to be a unitary matrix. Note that S can be explicitly constructed from $L(L-1)/2$ beam-splitters and $L(L+1)/2$ phase shifters in a triangular arrangement [33] as shown schematically in Fig. 4. This scheme assumes that each component of $\hat{\mathbf{b}}_{\text{out}}$ has the same polarization, transverse mode structure, and mean frequency. If this were not the case then the appropriate corrections would have to be made prior to entering the array in Fig. 4.

(iii) We then split each component of $\hat{\mathbf{s}}_{\text{out}}$ into quadratures. The process represented by K in Fig. 4 is shown in detail in Fig. 5. In general, the intensities of the quadratures need not be equal so each beam-splitter in Fig. 5 has a different transmission coefficient which we denote by θ_k ($k = 1, 2, \ldots, L$). This gives

$$\hat{\mathbf{d}} = \mathsf{K}\begin{pmatrix} \hat{\mathbf{s}}_{\text{out}} \\ \hat{\mathbf{a}} \end{pmatrix}. \quad (106)$$
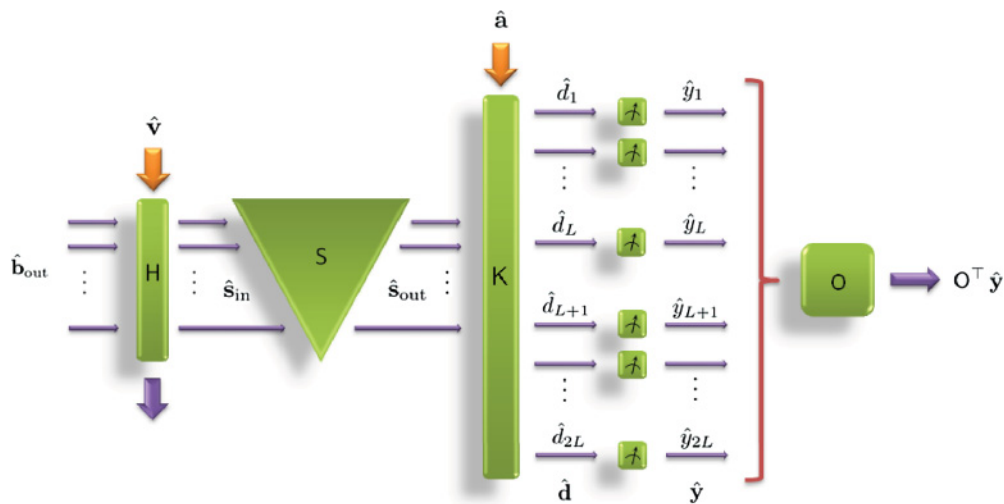

FIG. 4. (Color online) Decomposition of the measurement box in Fig. 3 into blocks. Vacuum inputs are denoted by orange arrows. We are ignoring half of the output modes of H. Note that H and K do not mix the input modes (not counting vacuum inputs), i.e., each output mode of either H or K will not depend on more than one signal mode (mode coming from the left) and one vacuum mode (mode coming from the top). We represent this in the diagram by using highly asymmetric rectangulars. Each output of K then enters a homodyne detector, set to measure the input quadrature as defined in (108) (also see Fig. 5).
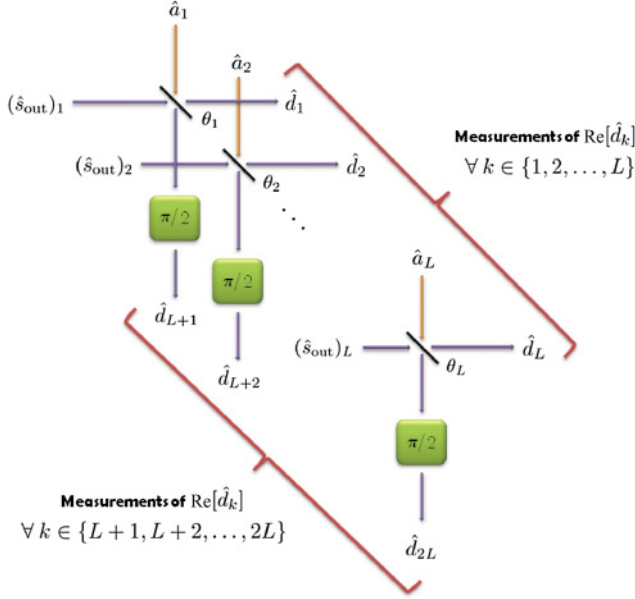
FIG. 5. (Color online) Expanded diagram of K. This prepares the components of $\hat{\mathbf{s}}_{\mathrm{out}}$ into quadratures ready to be measured by units of homodyne detectors.

If we define $\mathsf{Q} = \mathrm{diag}(\boldsymbol{\theta})$ where $\theta_k \in [0,1]$ for every k, and $\bar{\mathsf{Q}} = \mathsf{I}_L - \mathsf{Q}$, then K can be written explicitly as

$$\mathsf{K} = \begin{pmatrix} \sqrt{\mathsf{Q}} & \sqrt{\bar{\mathsf{Q}}} \\ i\sqrt{\bar{\mathsf{Q}}} & -i\sqrt{\mathsf{Q}} \end{pmatrix}. \tag{107}$$

Note that K is also a unitary matrix since $\hat{\mathbf{d}}^\dagger \hat{\mathbf{d}} = \hat{\mathbf{s}}_{\mathrm{out}}^\dagger \hat{\mathbf{s}}_{\mathrm{out}} + \hat{\mathbf{a}}^\dagger \hat{\mathbf{a}}$.

(iv) Individual homodyne detectors are then placed in the path of each component of $\hat{\mathbf{d}}$. The detectors are "tuned" to give a current that is normalized to the local oscillator amplitude that is inside it. Each homodyne detector is set to measure $\mathrm{Re}[\hat{d}_k]$. This final current can therefore be defined to be

$$\hat{\mathbf{y}} = \frac{\hat{\mathbf{d}} + \hat{\mathbf{d}}^\ddagger}{\sqrt{\hbar}}. \tag{108}$$

We see then, that once $\boldsymbol{\eta}$, $\boldsymbol{\theta}$, and S are specified, it seems plausible that steps (i)–(iv) above, by construction, will simulate an arbitrary diffusive measurement. Grouping these parameters into a triple $\mathsf{B} \equiv (\boldsymbol{\eta}, \mathsf{S}, \boldsymbol{\theta})$ we define the block representation (or B-rep for short) of diffusive measurements by

$$\mathfrak{B} = \{(\boldsymbol{\eta}, \mathsf{S}, \boldsymbol{\theta}) \,|\, \forall k, \, \eta_k, \theta_k \in [0,1], \, \mathsf{S}^\dagger = \mathsf{S}^{-1}\}. \tag{109}$$

(v) We shall henceforth label the B-rep current in (108) as $\hat{\mathbf{y}}_{\mathsf{B}}$. Once $\hat{\mathbf{y}}_{\mathsf{B}}$ has been obtained one is free to do some postprocesssing on this current. This will only be so if O is an orthogonal matrix and the postprocessing produces $\mathsf{O}^\top \hat{\mathbf{y}}_{\mathsf{B}}$. We will have more to say about this in Sec. IV C.

B. The current vector operator

Working out the input-output relation for each block in Fig. 4 we arrive at an equation for $\hat{\mathbf{y}}$ that is a function of

$\mathsf{B} = (\boldsymbol{\eta}, \mathsf{S}, \boldsymbol{\theta})$. Let us denote this dependence by using a subscript on $\hat{\mathbf{y}}$

$$\hbar \, \hat{\mathbf{y}}_{\mathsf{B}} \, dt = \mathsf{J} \, d\hat{\mathbf{B}}_{\mathrm{out}} + \mathsf{J}^* \, d\hat{\mathbf{B}}_{\mathrm{out}}^\ddagger + \hbar \, d\hat{\mathbf{v}}_{\mathsf{B}}, \tag{110}$$

where

$$\mathsf{J} \equiv \sqrt{\hbar} \begin{pmatrix} \sqrt{\mathsf{Q}} \, \mathsf{S}\sqrt{\mathsf{H}} \\ i\sqrt{\bar{\mathsf{Q}}} \, \mathsf{S}\sqrt{\mathsf{H}} \end{pmatrix} \tag{111}$$

and $\hbar \, d\hat{\mathbf{v}}_{\mathsf{B}}$ is given by

$$\hbar \, d\hat{\mathbf{v}}_{\mathsf{B}} = \mathsf{V} \, d\hat{\mathbf{V}} + \mathsf{V}^* \, d\hat{\mathbf{V}}^\ddagger + \mathsf{A} \, d\hat{\mathbf{A}} + \mathsf{A}^* \, d\hat{\mathbf{A}}^\ddagger. \tag{112}$$

The noise increments $d\hat{\mathbf{V}}$ and $d\hat{\mathbf{A}}$ are mutually uncorrelated quantum Wiener increments. The coefficient matrices in (112) are defined by

$$\mathsf{V} = \frac{1}{\sqrt{\hbar}} \begin{pmatrix} \sqrt{\mathsf{Q}} \, \mathsf{S}\sqrt{\bar{\mathsf{H}}} \\ i\sqrt{\bar{\mathsf{Q}}} \, \mathsf{S}\sqrt{\bar{\mathsf{H}}} \end{pmatrix}, \tag{113}$$

$$\mathsf{A} = \frac{1}{\sqrt{\hbar}} \begin{pmatrix} \sqrt{\bar{\mathsf{Q}}} \\ -i\sqrt{\mathsf{Q}} \end{pmatrix}. \tag{114}$$

Like the current in the M-rep, $d\hat{\mathbf{v}}_{\mathsf{B}}$ represents correlated noise but we can write

$$\hbar \, \hat{\mathbf{y}}_{\mathsf{B}} \, dt = \mathsf{J}\hat{\mathbf{c}} + \mathsf{J}^* \hat{\mathbf{c}}^\ddagger + [d\hat{\mathbf{v}}_{\mathrm{m}}]_{\mathsf{B}}, \tag{115}$$

where

$$[d\hat{\mathbf{v}}_{\mathrm{m}}]_{\mathsf{B}} = \mathsf{J} \, d\hat{\mathbf{B}}_{\mathrm{in}} + \mathsf{J}^* \, d\hat{\mathbf{B}}_{\mathrm{in}}^\ddagger$$
$$+ \mathsf{V} \, d\hat{\mathbf{V}} + \mathsf{V}^* \, d\hat{\mathbf{V}}^\ddagger + \mathsf{A} \, d\hat{\mathbf{A}} + \mathsf{A}^* \, d\hat{\mathbf{A}}^\ddagger \tag{116}$$

has a mean of zero

$$\langle [d\hat{\mathbf{v}}_{\mathrm{m}}(t)]_{\mathsf{B}} \rangle = \mathbf{0}, \tag{117}$$

and the only nonvanishing second-order moment

$$[d\hat{\mathbf{v}}_{\mathrm{m}}(t)]_{\mathsf{B}} \, [d\hat{\mathbf{v}}_{\mathrm{m}}(t)]_{\mathsf{B}}^\top = \hat{\mathsf{I}}_{2L} dt. \tag{118}$$

These can be understood physically by considering a vacuum input, i.e., $\hat{\mathbf{b}}_{\mathrm{out}} = \hat{\mathbf{b}}_{\mathrm{in}}$. None of the transformations applied to the vacuum inputs in Fig. 4 alter their statistics, only add one vacuum field to another, which is again vacuum. We can also verify this intuition mathematically by using (111), (114), and (113) to derive (118). All vanishing statistics should be obvious from the properties of the vacuum inputs. One may also verify that (87) is satisfied by $\hat{\mathbf{y}}_{\mathsf{B}}$.

1. Example: dual-homodyne detection

To illustrate how the B-rep works we will consider the familiar example of ideal heterodyne detection from quantum optics in the case of $L = 1$, for which $\mathsf{H} \equiv \eta = 1$, $\mathsf{S} \equiv e^{i\phi}$, and $\mathsf{Q} \equiv \theta = 1/2$. Since the measurement has unit detection efficiency, there is no vacuum field $\hat{v}$. It is simple to see that

$$\mathsf{K} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}, \tag{119}$$

which transforms $\hat{s}_{\mathrm{out}} = e^{i\phi} \hat{b}_{\mathrm{out}}$ and $\hat{a}$ into $\hat{\mathbf{d}} = (\hat{d}_1, \hat{d}_2)^\top$. From $\hat{\mathbf{d}}$ the final output $\hat{\mathbf{y}}_{\mathsf{B}}$ can be seen to be

$$\hat{\mathbf{y}}_{\mathsf{B}} = \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \end{pmatrix} = \frac{2}{\sqrt{2\hbar}} \begin{pmatrix} \hat{X}_{\mathrm{out}}(\phi) + \mathrm{Re}[\hat{a}] \\ \hat{X}_{\mathrm{out}}(\phi + \pi/2) + \mathrm{Im}[\hat{a}] \end{pmatrix}, \tag{120}$$

where we have defined the the quadrature operator

$$\hat{X}_{\text{out}}(\phi) = (\hat{b}_{\text{out}}\, e^{i\phi} + \hat{b}_{\text{out}}^{\dagger}\, e^{-i\phi})/2. \tag{121}$$

C. Relations among $\mathfrak{B}$, $\mathfrak{M}$, and $\mathfrak{U}$

Now that we have the M-rep, B-rep, and also the U-rep found previously in Refs. [26,27], it would be instructive to study how one may convert from one representation to another. We can first derive a map $\mathfrak{B} \to \mathfrak{M}$ by noting that if the B-rep is to realize an arbitrary diffusive measurement then it must be the case that

$$\langle \hat{\mathbf{y}}_{\mathsf{B}} \rangle = \langle \hat{\mathbf{y}}_{\mathsf{M}} \rangle. \tag{122}$$

Note that this is all we need since the second-order moments have already been made equal. From (90), (110), and (111) we obtain

$$\mathsf{M} = \mathsf{J}^{\dagger} = \sqrt{\hbar}\,\mathsf{H}\,\mathsf{S}^{\dagger}(\sqrt{\mathsf{Q}},\, -i\sqrt{\bar{\mathsf{Q}}}\,). \tag{123}$$

One can verify from this that $\mathsf{M}\mathsf{M}^{\dagger} = \hbar\,\mathsf{H}$, as required if M is to be a valid representation of the measurement. Taking the real and imaginary parts of (123) we get,

$$\mathsf{T} = \sqrt{\hbar}\begin{pmatrix} \sqrt{\mathsf{H}}\,\mathrm{Re}[\mathsf{S}^{\top}]\sqrt{\mathsf{Q}} & -\sqrt{\mathsf{H}}\,\mathrm{Im}[\mathsf{S}^{\top}]\sqrt{\bar{\mathsf{Q}}} \\ -\sqrt{\mathsf{H}}\,\mathrm{Im}[\mathsf{S}^{\top}]\sqrt{\mathsf{Q}} & -\sqrt{\bar{\mathsf{Q}}}\,\mathrm{Im}[\mathsf{S}^{\top}]\sqrt{\mathsf{H}} \end{pmatrix}. \tag{124}$$

From T we can construct U using the definition (65). This gives

$$\begin{aligned} \mathsf{U}_{11} = &\;\sqrt{\mathsf{H}}\,\mathrm{Re}[\mathsf{S}^{\top}]\mathsf{Q}\,\mathrm{Re}[\mathsf{S}]\,\sqrt{\mathsf{H}} \\ &+ \sqrt{\mathsf{H}}\,\mathrm{Im}[\mathsf{S}^{\top}]\bar{\mathsf{Q}}\,\mathrm{Im}[\mathsf{S}]\,\sqrt{\mathsf{H}}, \end{aligned} \tag{125}$$

$$\begin{aligned} \mathsf{U}_{12} = &-\sqrt{\mathsf{H}}\,\mathrm{Re}[\mathsf{S}^{\top}]\mathsf{Q}\,\mathrm{Im}[\mathsf{S}]\,\sqrt{\mathsf{H}} \\ &+ \sqrt{\mathsf{H}}\,\mathrm{Im}[\mathsf{S}^{\top}]\bar{\mathsf{Q}}\,\mathrm{Re}[\mathsf{S}]\,\sqrt{\mathsf{H}}, \end{aligned} \tag{126}$$

$$\begin{aligned} \mathsf{U}_{21} = &-\sqrt{\mathsf{H}}\,\mathrm{Im}[\mathsf{S}^{\top}]\mathsf{Q}\,\mathrm{Re}[\mathsf{S}]\,\sqrt{\mathsf{H}} \\ &+ \sqrt{\mathsf{H}}\,\mathrm{Re}[\mathsf{S}^{\top}]\bar{\mathsf{Q}}\,\mathrm{Im}[\mathsf{S}]\,\sqrt{\mathsf{H}}, \end{aligned} \tag{127}$$

$$\begin{aligned} \mathsf{U}_{22} = &\;\sqrt{\mathsf{H}}\,\mathrm{Im}[\mathsf{S}^{\top}]\mathsf{Q}\,\mathrm{Im}[\mathsf{S}]\,\sqrt{\mathsf{H}} \\ &+ \sqrt{\mathsf{H}}\,\mathrm{Re}[\mathsf{S}^{\top}]\bar{\mathsf{Q}}\,\mathrm{Re}[\mathsf{S}]\,\sqrt{\mathsf{H}}. \end{aligned} \tag{128}$$

By construction this U is positive semidefinite. By noting that if S is unitary,

$$\mathrm{Re}[\mathsf{S}^{\top}]\,\mathrm{Re}[\mathsf{S}] + \mathrm{Im}[\mathsf{S}^{\top}]\,\mathrm{Im}[\mathsf{S}] = \mathsf{I}_{L}, \tag{129}$$

$$\mathrm{Re}[\mathsf{S}^{\top}]\,\mathrm{Im}[\mathsf{S}] = \mathrm{Im}[\mathsf{S}^{\top}]\,\mathrm{Re}[\mathsf{S}], \tag{130}$$

we can verify that the remaining conditions for U to be a valid parametrization, (67) and (68), are indeed satisfied by (125)–(128).

Equation (123) says that given a B we can always find an M. But if we are given a valid M can we always find a B-rep? To answer this question we begin by counting the number of parameters required to specify B. The total number of real parameters in B is simply the sum of the parameters in each block. It should be clear that H and K each contribute L real parameters. The number of real parameters in an $L \times L$ unitary matrix is L^2 so the B-rep has $L^2 + 2L$ real parameters, the same as the U-rep. We said earlier that M-reps require $3L^2 + L$ real parameters, which is $2L^2 - L$ more than the

B-rep implying that (123) is one-to-many from $\mathfrak{B}$ to $\mathfrak{M}$. However, observe that for any $2L \times 2L$ orthogonal matrix O

$$\mathsf{M} = \sqrt{\hbar}\,\mathsf{H}\,\mathsf{S}^{\dagger}(\sqrt{\mathsf{Q}},\, -i\sqrt{\bar{\mathsf{Q}}}\,)\mathsf{O} \tag{131}$$

is not only a valid M-rep but its right-hand side has $3L^2 + L$ number of real parameters, exactly the number of parameters required to specify M. It is then very natural to ask whether every valid M-rep can be factorized in the form of (131). There does not seem to be a simple proof in general (for any L) but we do analyze the simplest ($L = 1$) instance of (131) in Appendix C. In summary our analysis shows that for $L = 1$ and for any M we can find a (B,O) such that (131) is satisfied. However, it involves the solutions of transcedental equations. The appearance of these transcedental equations, even in the simplest case, suggests why proving (131) in general may be nontrivial. But based on the parameter-count and the verification of the $L = 1$ case we conjecture that every $\mathsf{M} \in \mathfrak{M}$ has a matrix decomposition in the form of (131). This motivates us to define

$$\mathfrak{B}_{\mathsf{O}} = \{(\mathsf{B},\mathsf{O}) \,|\, \mathsf{B} \in \mathfrak{B}, \mathsf{O}^{\top} = \mathsf{O}^{-1}\}. \tag{132}$$

Equation (131) is then shown in Fig. 2 by the dotted line. This is a many-to-one map from $\mathfrak{B}_{\mathsf{O}}$ to $\mathfrak{M}$ as there will be more than one orthogonal matrix O that will factorize M. This can be seen directly by considering the $L = 1$ case which we have analyzed in Appendix C. From Fig. 6 we can see that given an M (and therefore r and δ) there is a range of values for φ [which in turn determine O via (C1)] at one's disposal for factoring M in the form of (131).

We can connect the orthogonal matrix in (131) and (132) to Fig. 4. Since given an $\mathsf{M} \in \mathfrak{M}$ we can write $\mathsf{M}' = \mathsf{MO}$ which is also in $\mathfrak{M}$, (90) and (122) imply that

$$\langle \hat{\mathbf{y}}_{\mathsf{M}'} \rangle = \langle \hat{\mathbf{y}}_{\mathsf{MO}} \rangle = \mathsf{O}^{\top}\langle \hat{\mathbf{y}}_{\mathsf{M}} \rangle = \mathsf{O}^{\top}\langle \hat{\mathbf{y}}_{\mathsf{B}} \rangle. \tag{133}$$

That is to say, if $\hat{\mathbf{y}}_{\mathsf{B}}$ emulates $\hat{\mathbf{y}}_{\mathsf{M}}$, then $\mathsf{O}^{\top}\hat{\mathbf{y}}_{\mathsf{B}}$ (obtained from postprocessing $\hat{\mathbf{y}}_{\mathsf{B}}$) will emulate $\hat{\mathbf{y}}_{\mathsf{MO}}$.

The completeness of the B-rep also means that every valid U must have a matrix decomposition of the form given in (125)–(128). The map $\mathfrak{B} \to \mathfrak{U}$ is thus onto. Even though B and U are parameterized by the same number of real parameters the map $\mathfrak{B} \to \mathfrak{U}$ is actually many-to-one. To see this we can consider (125)–(128) for $L = 1$. In this case U is 2×2 and $\mathsf{S} \equiv e^{i\phi}$, $\mathsf{H} \equiv \eta$, and $\mathsf{Q} \equiv \theta$ as in Sec. IV B 1. It is sufficient to consider $\eta = 1$, in which case (125)–(128) simplify to

$$\mathsf{U}_{11} = \theta\,\cos^2\phi + (1 - \theta)\,\sin^2\phi, \tag{134}$$

$$\mathsf{U}_{12} = (1 - 2\theta)\cos\phi\,\sin\phi, \tag{135}$$

$$\mathsf{U}_{22} = \theta\,\sin^2\phi + (1 - \theta)\,\cos^2\phi, \tag{136}$$

where U_{12} is identically equal to U_{21} as can be verified directly from (126) and (127) for any L. Remembering that $\cos(\phi + \pi) = -\cos\phi$ and $\sin(\phi + \pi) = -\sin\phi$, we can see that $\mathsf{B}_1 \equiv (1,e^{i\phi},\theta)$ and $\mathsf{B}_2 \equiv (1,e^{i(\phi+\pi)},\theta)$ [recall the order of the B-rep triple from (109)] both map to the same U but yet $\mathsf{B}_1 \neq \mathsf{B}_2$. We can consider another example in which $\mathsf{B}_3 \equiv (1,e^{i(\phi+\pi/2)},1 - \theta)$ and B_1 map to the same U but $\mathsf{B}_1 \neq \mathsf{B}_3$ [this time noting $\cos(\phi + \pi/2) = -\sin\phi$ and $\sin(\phi + \pi/2) = \cos\phi$]. The relation between $\mathfrak{B}$ and $\mathfrak{U}$ is shown in Fig. 2.
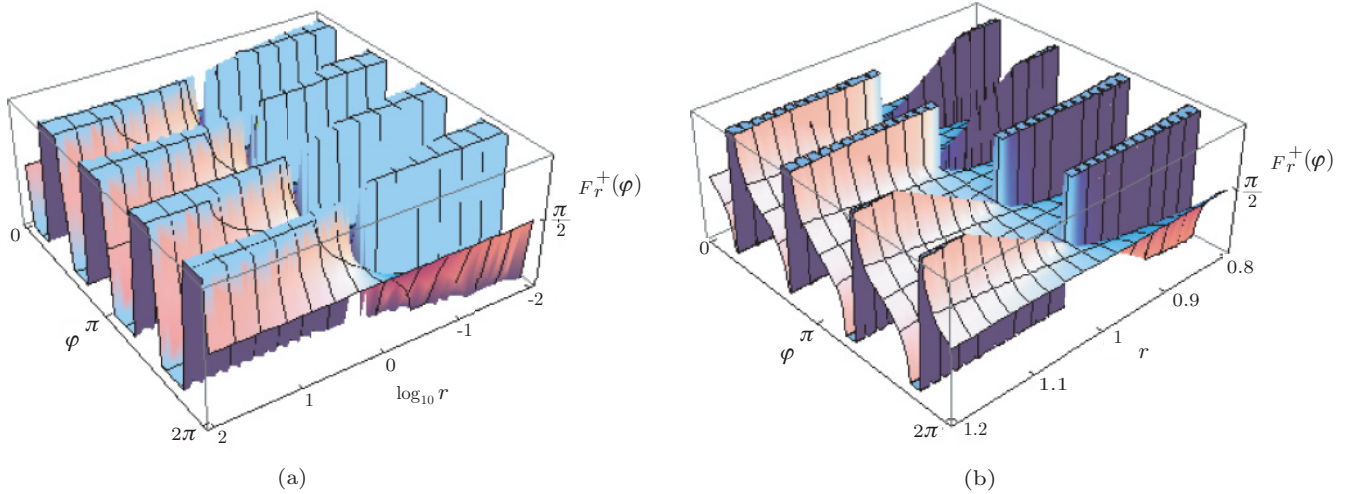
FIG. 6. (Color online) ($F_r^+(\varphi)$ for different values of $r \equiv |m_1/m_2|^2$ and φ. (a) Behavior of $F_r^+(\varphi)$ for large values of r. (b) $F_r^+(\varphi)$ around $r = 1$. $F_r^+(\varphi)$ becomes $\pi/2$ at $r = 1$ [see (C18)] and has a "twist" in crossing $r = 1$. Remember that $|m_1|^2$ and $|m_2|^2$ are constrained by (C6) so given an $|m_2|^2 \neq 0$, $|m_1|^2$ cannot be an arbitrary multiple of $|m_2|^2$ with $|m_2|^2$ staying constant. The limit $r \to \infty$ is thus equivalent to the limits $|m_1|^2 \to 1$ and $|m_2|^2 \to 0$. Similarly, if we are given an $|m_1|^2 \neq 0$, then $|m_2|^2$ cannot be an arbitrary multiple of $|m_1|^2$ with $|m_1|^2$ staying constant so the limit $r \to 0$ is equivalent to the limits $|m_1|^2 \to 0$ and $|m_2|^2 \to 1$.

V. CONCLUDING REMARKS

A. Summary of key results

In this paper we have put forth two new parametrizations of diffusive quantum measurements. The first is the M-rep introduced in Sec. III, defined by (74). This is a characterization of diffusive measurements in the Schrödinger picture, and as a consequence is a rather abstract parametrization. However, it embodies all diffusive measurements by a single equation which relates simply to the efficiency with which each output field is measured $\sqrt{\mathsf{H}}$. The M-rep has been shown to be a complete parametrization; every diffusive quantum measurement can always be associated with an $\mathsf{M} \in \mathfrak{M}$, and every $\mathsf{M} \in \mathfrak{M}$ must correspond to a diffusive measurement.

We then turned to the Heisenberg picture in Sec. IV and proposed a realization for arbitrary diffusive measurements. This gave another parametrization which we have called the B-rep, defined by (109). By construction, every $\mathsf{B} \in \mathfrak{B}$ must correspond to a diffusive measurement. In particular we derived (131) and used it to investigate the existence of B-reps for arbitrary diffusive measurements. An analysis of (131) for $L = 1$ reveals that it requires solutions to transcedental equations, which do indeed exist, for any M. The occurrence of transcedental equations even in this simplest case suggests that proving (131) for any L may be nontrivial or even formidable. The verification of (131) and the fact that this equation (131) has the same number of parameters on each side prompts us to propose it as a general matrix decomposition for every $\mathsf{M} \in \mathfrak{M}$ and thereby conjecturing the completeness of the B-rep.

Finally, we have explained the relations between the different parametrizations (including the U-rep) and Fig. 2 provides a summary of these relations.

B. Future prospects

Quantum measurements are an integral part of many quantum computing architectures. The monitoring of qubits (or qunats [59]) and their unwanted environmental coupling provides a natural setting to apply quantum-trajectory theory. Indeed, the applicability of quantum trajectories to quantum information was recognized in 1995 by Schack and colleagues [60]. At the same time a second group (Pellizzari *et al.*) had already *applied* quantum trajectories to study decoherence in the cavity QED architecture of quantum computing [61]. Subsequently quantum-trajectory theory was adopted to study the effects of decoherence on the quantum information processing of discrete-level systems to some generality in Ref. [62] and, more specifically, in quantum error correction [63–66], gate performance [67], and teleportation (of both discrete and continuous variables) [68–70].

As we scale up our quantum processors one would expect quantum trajectories to remain useful. A scale-up of existing protocols would inevitably mean a scale-up of the necessary measurement from being a few-port device to many-port ones. In fact, the error syndrome which conditions the corrective displacement of just one qunat in the nine-qunat error correction scheme of Braunstein [71,72] already requires eight homodyne detectors. If one wanted to simulate the conditioned evolution based on this eight-homodyne output using (24), then it is necessary to parametrize this measurement as an M, in this case of dimensions of 8×16. If all we knew was $\mathsf{MM}^\dagger/\hbar \in \mathfrak{H}$, then it is not so easy to see what M should be, at least not without some thought. Parameterizing this measurement into a B-rep, on the other hand, is much easier. One could then use (131) to obtain the appropriate M. This illustrates that (131) can be useful when given a $(\mathsf{B}, \mathsf{O}) \in \mathfrak{B}_\mathsf{O}$. Given that simulating (pure-state) quantum trajectories requires less time and computer memory than a direct integration of the

corresponding master equation [4] one can expect applications of quantum trajectories to quantum information to continue in the foreseeable future.

APPENDIX A: VECTOR-OPERATOR ALGEBRA

1. Basic operations

In the same manner that we write a set of equations compactly using vectors and matrices in classical theories, we would also like to write a set of operator equations compactly. In order to do this we have to introduce a vector whose elements are operators which we call vector operators. This Appendix will make clear particular definitions that we adopt for vector operators and their algebra.

We begin with the definitions of the vector operator, its transpose, and its Hermitian conjugate. For an arbitrary vector operator $\hat{\mathbf{A}}$ consisting of n components,

$$\hat{\mathbf{A}} \equiv \begin{pmatrix} \hat{A}_1 \\ \hat{A}_2 \\ \vdots \\ \hat{A}_n \end{pmatrix} \equiv (\hat{A}_1, \hat{A}_2, \ldots, \hat{A}_n)^\top, \tag{A1}$$

we define its Hermitian conjugate as

$$\hat{\mathbf{A}}^\dagger \equiv (\hat{A}_1^\dagger, \hat{A}_2^\dagger, \ldots, \hat{A}_n^\dagger). \tag{A2}$$

Note the Hermitian conjugate of $\hat{\mathbf{A}}$ is defined by transposing $\hat{\mathbf{A}}$ (as usually defined) and then taking the Hermitian conjugate of each element of $\hat{\mathbf{A}}$. We will say that a vector operator is Hermitian if and only if $\hat{\mathbf{A}}^\top = \hat{\mathbf{A}}^\dagger$. Matrix operators also appear in this article and we will define the Hermitian conjugate of an $m \times n$ matrix operator $\hat{\mathsf{A}}$ as follows:

$$\hat{\mathsf{A}} \equiv \begin{pmatrix} \hat{\mathsf{A}}_{11} & \hat{\mathsf{A}}_{12} & \cdots & \hat{\mathsf{A}}_{1n} \\ \hat{\mathsf{A}}_{21} & \hat{\mathsf{A}}_{22} & & \\ \vdots & & \ddots & \\ \hat{\mathsf{A}}_{m1} & & & \hat{\mathsf{A}}_{mn} \end{pmatrix}, \tag{A3}$$

$$\hat{\mathsf{A}}^\dagger \equiv \begin{pmatrix} \hat{\mathsf{A}}_{11}^\dagger & \hat{\mathsf{A}}_{21}^\dagger & \cdots & \hat{\mathsf{A}}_{m1}^\dagger \\ \hat{\mathsf{A}}_{12}^\dagger & \hat{\mathsf{A}}_{22}^\dagger & & \\ \vdots & & \ddots & \\ \hat{\mathsf{A}}_{1n}^\dagger & & & \hat{\mathsf{A}}_{mn}^\dagger \end{pmatrix}. \tag{A4}$$

We will denote, for this Appendix only, matrix operators (which we refer to as matrix operators), by san serif letters with a hat on top. We retain the familiar properties of the Hermitian conjugate for complex matrices:

$$(\hat{\mathsf{A}}^\top)^\dagger = (\hat{\mathsf{A}}^\dagger)^\top, \tag{A5}$$

$$(\hat{\mathsf{A}}\hat{\mathsf{B}})^\dagger = \hat{\mathsf{B}}^\dagger \hat{\mathsf{A}}^\dagger, \tag{A6}$$

where $\hat{\mathsf{A}}$ and $\hat{\mathsf{B}}$ are respectively $m \times n$ and $n \times k$ matrix operators. Equation (A5) crops up often so for convenience we define

$$\hat{\mathsf{A}}^\ddagger = (\hat{\mathsf{A}}^\top)^\dagger . \tag{A7}$$

It will also be useful for us to define the real and imaginary parts of an arbitrary scalar operator $\hat{c}$ to be the Hermitian operators

$$\mathrm{Re}[\hat{c}] \equiv \frac{1}{2}(\hat{c} + \hat{c}^\dagger), \quad \mathrm{Im}[\hat{c}] \equiv \frac{1}{2i}(\hat{c} - \hat{c}^\dagger). \tag{A8}$$

Since we also deal with complex matrices, we have also defined the real and imaginary parts of a complex matrix C to be

$$\mathrm{Re}[C] \equiv \frac{1}{2}(C + C^*), \quad \mathrm{Im}[C] \equiv \frac{1}{2i}(C - C^*), \tag{A9}$$

where C^* denotes the complex conjugate of C. We generalize the definitions of the real and imaginary parts of a scalar operator to a matrix operator:

$$\mathrm{Re}[\hat{\mathsf{A}}] \equiv \begin{pmatrix} \mathrm{Re}[\hat{\mathsf{A}}_{11}] & \mathrm{Re}[\hat{\mathsf{A}}_{12}] & \cdots & \mathrm{Re}[\hat{\mathsf{A}}_{1n}] \\ \mathrm{Re}[\hat{\mathsf{A}}_{21}] & \mathrm{Re}[\hat{\mathsf{A}}_{22}] & & \\ \vdots & & \ddots & \\ \mathrm{Re}[\hat{\mathsf{A}}_{m1}] & & & \mathrm{Re}[\hat{\mathsf{A}}_{mn}] \end{pmatrix} \tag{A10}$$

$$= \frac{1}{2}(\hat{\mathsf{A}} + \hat{\mathsf{A}}^\ddagger), \tag{A11}$$

$$\mathrm{Im}[\hat{\mathsf{A}}] \equiv \begin{pmatrix} \mathrm{Im}[\hat{\mathsf{A}}_{11}] & \mathrm{Im}[\hat{\mathsf{A}}_{12}] & \cdots & \mathrm{Im}[\hat{\mathsf{A}}_{1n}] \\ \mathrm{Im}[\hat{\mathsf{A}}_{21}] & \mathrm{Im}[\hat{\mathsf{A}}_{22}] & & \\ \vdots & & \ddots & \\ \mathrm{Im}[\hat{\mathsf{A}}_{m1}] & & & \mathrm{Im}[\hat{\mathsf{A}}_{mn}] \end{pmatrix} \tag{A12}$$

$$= \frac{1}{2i}(\hat{\mathsf{A}} - \hat{\mathsf{A}}^\ddagger). \tag{A13}$$

Note the necessity of taking both the transpose and Hermitian conjugate in (A10) and (A12) in order to retain an $m \times n$ matrix operator.

Multiplication of $\hat{\mathsf{A}}$ by an arbitrary scalar operator $\hat{c}$ is defined component-wise by

$$\hat{c}\,\hat{\mathsf{A}} \equiv \begin{pmatrix} \hat{c}\,\hat{\mathsf{A}}_{11} & \hat{c}\,\hat{\mathsf{A}}_{12} & \cdots & \hat{c}\,\hat{\mathsf{A}}_{1n} \\ \hat{c}\,\hat{\mathsf{A}}_{21} & \hat{c}\,\hat{\mathsf{A}}_{22} & & \\ \vdots & & \ddots & \\ \hat{c}\,\hat{\mathsf{A}}_{m1} & & & \hat{c}\,\hat{\mathsf{A}}_{mn} \end{pmatrix}, \tag{A14}$$

and similarly for the action of an arbitrary superoperator $\mathcal{N}$,

$$\mathcal{N}\,\hat{\mathsf{A}} \equiv \begin{pmatrix} \mathcal{N}\,\hat{\mathsf{A}}_{11} & \mathcal{N}\,\hat{\mathsf{A}}_{12} & \cdots & \mathcal{N}\,\hat{\mathsf{A}}_{1n} \\ \mathcal{N}\,\hat{\mathsf{A}}_{21} & \mathcal{N}\,\hat{\mathsf{A}}_{22} & & \\ \vdots & & \ddots & \\ \mathcal{N}\,\hat{\mathsf{A}}_{m1} & & & \mathcal{N}\,\hat{\mathsf{A}}_{mn} \end{pmatrix}. \tag{A15}$$

It is straightforward to see (and will be useful to remember) from (A14) that

$$\hat{c}\,\hat{\mathsf{A}}^\top = (\hat{c}\,\hat{\mathsf{A}})^\top, \tag{A16}$$

$$(C\hat{\mathsf{A}})^\top = \hat{\mathsf{A}}^\top C^\top, \tag{A17}$$

for any $k \times m$ matrix C.

Two special cases of (A6) that we are interested in are the inner and outer products between vector operators for which (A6) reads

$$(\hat{\mathbf{A}}^\top \hat{\mathbf{B}})^\dagger = \hat{\mathbf{B}}^\dagger \hat{\mathbf{A}}^\ddagger, \tag{A18}$$

$$(\hat{\mathbf{A}} \hat{\mathbf{B}}^\top)^\dagger = \hat{\mathbf{B}}^\ddagger \hat{\mathbf{A}}^\dagger. \tag{A19}$$

The following identities will also be useful for manipulating vector operator products

$$(\hat{\mathbf{A}}^\top \hat{\mathbf{B}})\hat{\mathbf{C}} = (\hat{\mathbf{A}}^\top \hat{\mathbf{B}}\hat{\mathbf{C}}^\top)^\top, \tag{A20}$$

$$(\hat{\mathbf{A}} \hat{\mathbf{B}}^\top)^\top \hat{\mathbf{C}} = [\hat{\mathbf{A}}^\top (\hat{\mathbf{B}}\hat{\mathbf{C}}^\top)^\top]^\top. \tag{A21}$$

The tensor product for vector operators are defined analogous to standard multiplication rules. If $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ have different dimensions then we define

$$\hat{\mathbf{A}} \otimes \hat{\mathbf{B}}^\top \equiv \begin{pmatrix} \hat{A}_1 \otimes \hat{B}_1 & \hat{A}_1 \otimes \hat{B}_2 & \cdots & \hat{A}_1 \otimes \hat{B}_n \\ \hat{A}_2 \otimes \hat{B}_1 & \hat{A}_2 \otimes \hat{B}_2 & & \\ \vdots & & \ddots & \\ \hat{A}_m \otimes \hat{B}_1 & & & \hat{A}_m \otimes \hat{B}_n \end{pmatrix}. \tag{A22}$$

When $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ have the same dimension then

$$\hat{\mathbf{A}}^\top \otimes \hat{\mathbf{B}} \equiv \sum_k \hat{A}_k \otimes \hat{B}_k. \tag{A23}$$

We will assign the following symbols for identity operations with respect to multiplication: (a) $\hat{1}$ for a scalar-operator identity, (b) I_N for an $N \times N$ identity matrix, and (c) $\hat{\mathbf{I}}_N$ for a matrix-operator identity (defined by $\hat{\mathbf{I}}_N = \mathrm{I}_N \hat{1}$).

While it is common to find special symbols for identity operations for multiplication in the literature, it is also of common practice to *not* denote the identity with respect to addition by any distinguished symbol except for 0. We will follow this convention as well except for the zero vector which we will write as **0**.

2. Commutativity

a. Scalar-operator brackets

Unlike normal vectors the order of the vector operators in an inner product may not be changed in general,

$$\hat{\mathbf{B}}^\top \hat{\mathbf{A}} \neq \hat{\mathbf{A}}^\top \hat{\mathbf{B}}, \tag{A24}$$

and this motivates us to define the scalar-operator bracket

$$\lceil \hat{\mathbf{A}}, \hat{\mathbf{B}} \rfloor = \hat{\mathbf{A}}^\top \hat{\mathbf{B}} - \hat{\mathbf{B}}^\top \hat{\mathbf{A}}, \tag{A25}$$

so called because it maps two vector operators to a scalar operator. We then have the following sufficient (but not necessary) condition for reordering the inner product,

$$\forall k, [\hat{A}_k, \hat{B}_k] = 0 \Rightarrow \lceil \hat{\mathbf{A}}, \hat{\mathbf{B}} \rfloor = 0. \tag{A26}$$

That $[\hat{A}_k, \hat{B}_k] = 0$ is not necessary for $\lceil \hat{\mathbf{A}}, \hat{\mathbf{B}} \rfloor = 0$ can be seen in the following example,

$$\hat{\mathbf{A}} = (\hat{q}, \hat{p})^\top, \quad \hat{\mathbf{B}} = (\hat{p}, \hat{q})^\top, \quad [\hat{q}, \hat{p}] = i\hbar \hat{1}. \tag{A27}$$

Here $\hat{A}_1$ does not commute with $\hat{B}_1$, and $\hat{A}_2$ does not commute with $\hat{B}_2$ but we still have $\hat{\mathbf{A}}^\top \hat{\mathbf{B}} = \hat{\mathbf{B}}^\top \hat{\mathbf{A}}$. Note that we have not overloaded the usual notation for commutator brackets in

(A25) because while the scalar-operator bracket satisfies the following (which is easy to check)

$$\lceil \hat{\mathbf{A}}, \hat{\mathbf{B}} \rfloor = -\lceil \hat{\mathbf{B}}, \hat{\mathbf{A}} \rfloor, \tag{A28}$$

$$\lceil \hat{\mathbf{A}} + \hat{\mathbf{B}}, \hat{\mathbf{C}} \rfloor = \lceil \hat{\mathbf{A}}, \hat{\mathbf{C}} \rfloor + \lceil \hat{\mathbf{B}}, \hat{\mathbf{C}} \rfloor, \tag{A29}$$

it cannot satisfy a property analogous to the Jacobi identity which the standard commutator does.

The inner product is a scalar operator so by definition $\hat{\mathbf{A}}^\top \hat{\mathbf{B}} = (\hat{\mathbf{A}}^\top \hat{\mathbf{B}})^\top$. Thus (A26) may also be seen as a sufficient condition for $(\hat{\mathbf{A}}^\top \hat{\mathbf{B}})^\top = \hat{\mathbf{B}}^\top \hat{\mathbf{A}}$.

b. Matrix-operator brackets

For the vector-operator outer product we do have a necessary and sufficient condition for writing the transpose of $\hat{\mathbf{A}}\hat{\mathbf{B}}^\top$ as a product of transposes:

$$\forall j,k, [\hat{A}_j, \hat{B}_k] = 0 \iff (\hat{\mathbf{A}} \hat{\mathbf{B}}^\top)^\top = \hat{\mathbf{B}} \hat{\mathbf{A}}^\top. \tag{A30}$$

Therefore we define the matrix-operator bracket

$$\lfloor \hat{\mathbf{A}}, \hat{\mathbf{B}} \rceil \equiv \hat{\mathbf{A}}\hat{\mathbf{B}}^\top - (\hat{\mathbf{B}}\hat{\mathbf{A}}^\top)^\top, \tag{A31}$$

which maps vector operators to matrix operators and refer to $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ as commuting vector operators if and only if $\lfloor \hat{\mathbf{A}}, \hat{\mathbf{B}} \rceil = 0$. It is easy to see that this satisfies

$$\lfloor \hat{\mathbf{A}}, \hat{\mathbf{B}} \rceil = -\lfloor \hat{\mathbf{B}}, \hat{\mathbf{A}} \rceil^\top, \tag{A32}$$

$$\lfloor \hat{\mathbf{A}} + \hat{\mathbf{B}}, \hat{\mathbf{C}} \rceil = \lfloor \hat{\mathbf{A}}, \hat{\mathbf{C}} \rceil + \lfloor \hat{\mathbf{B}}, \hat{\mathbf{C}} \rceil. \tag{A33}$$

The symbol for the matrix-operator bracket in (A31) has the ceiling ($\rceil$) on $\hat{\mathbf{B}}$ to remind us of $\hat{\mathbf{A}}\hat{\mathbf{B}}^\top$, which is a matrix operator. For the scalar-operator bracket of (A25) the ceiling is placed on $\hat{\mathbf{A}}$ to remind us of $\hat{\mathbf{A}}^\top \hat{\mathbf{B}}$. Note the commutativity of vector operators thus defined means that an arbitrary vector operator will not necessarily commute with itself as that is a question about its components.

From the above one may verify the following formulas which help simplify expressions containing matrix-operator brackets: For any matrix C (with the appropriate dimensions)

$$\lfloor C\hat{\mathbf{A}}, \hat{\mathbf{B}} \rceil = C\lfloor \hat{\mathbf{A}}, \hat{\mathbf{B}} \rceil, \tag{A34}$$

$$\lfloor \hat{\mathbf{A}}, C\hat{\mathbf{B}} \rceil = \lfloor \hat{\mathbf{A}}, \hat{\mathbf{B}} \rceil C^\top. \tag{A35}$$

For any scalar operator $\hat{c}$,

$$\lfloor \hat{c}, \hat{\mathbf{A}} \rceil = [\hat{c}, \hat{\mathbf{A}}^\top], \tag{A36}$$

$$\lfloor \hat{\mathbf{A}}, \hat{c} \rceil = [\hat{\mathbf{A}}, \hat{c}]. \tag{A37}$$

Note that it is the standard commutator which appears on the right-hand side of (A36) and (A37). We also have

$$\lfloor \hat{\mathbf{A}}\hat{c}, \hat{\mathbf{B}} \rceil = \hat{\mathbf{A}}[\hat{c}, \hat{\mathbf{B}}^\top] + \lfloor \hat{\mathbf{A}}, \hat{\mathbf{B}} \rceil \hat{c}, \tag{A38}$$

$$\lfloor \hat{c}\hat{\mathbf{A}}, \hat{\mathbf{B}} \rceil = \hat{c}\lfloor \hat{\mathbf{A}}, \hat{\mathbf{B}} \rceil + ([\hat{c}, \hat{\mathbf{B}}]\hat{\mathbf{A}}^\top)^\top, \tag{A39}$$

$$\lfloor \hat{\mathbf{A}}, \hat{\mathbf{B}}\hat{c} \rceil = (\hat{\mathbf{B}}[\hat{\mathbf{A}}^\top, \hat{c}])^\top + \lfloor \hat{\mathbf{A}}, \hat{\mathbf{B}} \rceil \hat{c}, \tag{A40}$$

$$\lfloor \hat{\mathbf{A}}, \hat{c}\hat{\mathbf{B}} \rceil = \hat{c}\lfloor \hat{\mathbf{A}}, \hat{\mathbf{B}} \rceil + [\hat{\mathbf{A}}, \hat{c}]\hat{\mathbf{B}}^\top. \tag{A41}$$

3. Trace and average

Since both matrices and operators appear we will denote an operator trace by $\mathrm{Tr}[\hat{c}]$ and the trace of a matrix C by $\mathrm{tr}[C]$. This distinction is especially important for operators

with a continuous eigenvalue spectrum where the operator-matrix analogy breaks down. The trace of an $m \times n$ matrix operator is a matrix in $\mathbb{C}^{m \times n}$

$$\mathrm{Tr}[\hat{\mathbf{A}}] \equiv \begin{pmatrix} \mathrm{Tr}[\hat{\mathbf{A}}_{11}] & \mathrm{Tr}[\hat{\mathbf{A}}_{12}] & \cdots & \mathrm{Tr}[\hat{\mathbf{A}}_{1n}] \\ \mathrm{Tr}[\hat{\mathbf{A}}_{21}] & \mathrm{Tr}[\hat{\mathbf{A}}_{22}] & & \\ \vdots & & \ddots & \\ \mathrm{Tr}[\hat{\mathbf{A}}_{m1}] & & & \mathrm{Tr}[\hat{\mathbf{A}}_{mn}] \end{pmatrix}. \quad \text{(A42)}$$

From this it should be simple to see that

$$\mathrm{Tr}[\hat{\mathbf{A}}^\top] = (\mathrm{Tr}[\hat{\mathbf{A}}])^\top, \quad \text{(A43)}$$

$$(\mathrm{Tr}[\hat{\mathbf{A}}])^* = \mathrm{Tr}[\hat{\mathbf{A}}^\ddagger]. \quad \text{(A44)}$$

Particular cases of (A43) and (A44) which we are interested in are

$$(\mathrm{Tr}[\hat{\mathbf{A}}])^\dagger = \mathrm{Tr}[\hat{\mathbf{A}}^\dagger], \quad \text{(A45)}$$

$$(\mathrm{Tr}[\hat{\mathbf{A}}\hat{\mathbf{B}}^\top])^* = \mathrm{Tr}[\hat{\mathbf{A}}^\ddagger\hat{\mathbf{B}}^\dagger]. \quad \text{(A46)}$$

If $\hat{\mathbf{A}}$ is any $m \times n$ matrix operator and $\hat{\mathbf{B}}$ is $n \times k$, then we have the "cyclic" property

$$\mathrm{Tr}[(\hat{\mathbf{A}}\hat{\mathbf{B}})^\top] = \mathrm{Tr}[\hat{\mathbf{B}}^\top\hat{\mathbf{A}}^\top]. \quad \text{(A47)}$$

Special cases of the "cyclic" property (A47) are

$$\mathrm{Tr}[\hat{c}\,\hat{\mathbf{A}}] = \mathrm{Tr}[\hat{\mathbf{A}}\,\hat{c}], \quad \text{(A48)}$$

$$\mathrm{Tr}[(\hat{\mathbf{A}}\hat{\mathbf{B}}^\top)^\top] = \mathrm{Tr}[\hat{\mathbf{B}}\hat{\mathbf{A}}^\top]. \quad \text{(A49)}$$

On some occasions we will want to permute a product of three vector operators inside a trace and the following identity may be verified,

$$\mathrm{Tr}[\hat{\mathbf{A}}\,\hat{\mathbf{B}}^\top\hat{\mathbf{C}}] = \mathrm{Tr}[\hat{\mathbf{B}}^\top\hat{\mathbf{C}}\,\hat{\mathbf{A}}] = \mathrm{Tr}[(\hat{\mathbf{C}}\hat{\mathbf{A}}^\top)^\top\hat{\mathbf{B}}]. \quad \text{(A50)}$$

An important tool from quantum statistics that we will use is the quantum regression formula, also known as the quantum regression theorem, formulated by Lax [73,74]. A good discussion of this can be found in Ref. [37]. In essence this result expresses the average of a product of two operators, each at a different time, as a trace in the Schrödinger picture. To write down the vector-operator counterpart of this result we first recall the regression formulas for scalar operators: Given any two operators $\hat{A}$ and $\hat{B}$ in the system Hilbert space, and the solution to the Markovian master equation $\dot{\rho} = \mathcal{L}\rho$,

$$\langle \hat{A}(t)\,\hat{B}(t+\tau)\rangle = \mathrm{Tr}\{\hat{B}(0)\,e^{\mathcal{L}\tau}[\rho(t)\,\hat{A}(0)]\}, \quad \text{(A51)}$$

$$\langle \hat{A}(t+\tau)\,\hat{B}(t)\rangle = \mathrm{Tr}\{\hat{A}(0)\,e^{\mathcal{L}\tau}[\hat{B}(0)\,\rho(t)]\}, \quad \text{(A52)}$$

for any $\tau > 0$. Let us define the quantum average of a matrix operator $\hat{\mathbf{A}}$ at time t as

$$\langle \hat{\mathbf{A}}\rangle = \mathrm{Tr}[\rho(t)\hat{\mathbf{A}}]. \quad \text{(A53)}$$

A corollary which follows from this definition and (A45) is

$$\langle \hat{c}\,\hat{\mathbf{A}}\rangle^\dagger = \langle \hat{\mathbf{A}}^\dagger\,\hat{c}^\dagger\rangle, \quad \text{(A54)}$$

obtained by letting $\hat{\mathbf{A}} \to \rho\hat{c}\hat{\mathbf{A}}$ in (A45). We can now extend (A51) and (A52) to

$$\langle \hat{\mathbf{A}}(t)\,\hat{\mathbf{B}}^\top(t+\tau)\rangle = (\mathrm{Tr}\{\hat{\mathbf{B}}(0)\,e^{\mathcal{L}\tau}[\rho(t)\,\hat{\mathbf{A}}^\top(0)]\})^\top, \quad \text{(A55)}$$

$$\langle \hat{\mathbf{A}}(t+\tau)\,\hat{\mathbf{B}}^\top(t)\rangle = \mathrm{Tr}\{\hat{\mathbf{A}}(0)\,e^{\mathcal{L}\tau}[\hat{\mathbf{B}}^\top(0)\,\rho(t)]\}. \quad \text{(A56)}$$

4. Superoperators of special interest

When considering continuous measurements terms of certain forms arise frequently. Here we will define superoperators whose forms allow us to write these frequently occurring terms compactly. For an arbitrary scalar operator $\hat{s}$ we define

$$\mathcal{J}[\hat{A}]\hat{B} \equiv \hat{A}\,\hat{B}\,\hat{A}^\dagger. \quad \text{(A57)}$$

Note that $\hat{A}$ here should be treated as a parameter and $\hat{B}$ the input for $\mathcal{J}[\hat{A}]$, the output of the superoperator is another scalar operator given by the right-hand side of (A57). The definition (A57) thus implies that $\mathcal{J}[\hat{A}^\dagger]\hat{B} = \hat{A}^\dagger\,\hat{B}\,\hat{A}$.

Often we encounter a sum of terms in the form of (A57) so we define

$$\mathcal{J}[\hat{\mathbf{A}}]\hat{B} \equiv \sum_k \mathcal{J}[\hat{A}_k]\hat{B} = \sum_k \hat{A}_k\,\hat{B}\,\hat{A}_k^\dagger = \hat{\mathbf{A}}^\top\hat{B}\,\hat{\mathbf{A}}^\ddagger. \quad \text{(A58)}$$

From (A15) and (A58) the action of $\mathcal{J}[\hat{\mathbf{A}}]$ on a vector operator produces another vector operator, given by

$$\mathcal{J}[\hat{\mathbf{A}}]\hat{\mathbf{B}} = \sum_k \hat{A}_k\,\hat{\mathbf{B}}\,\hat{A}_k^\dagger = (\hat{\mathbf{A}}\hat{\mathbf{B}}^\top)^\top\hat{\mathbf{A}}^\ddagger. \quad \text{(A59)}$$

This also gives $\mathcal{J}[\hat{\mathbf{A}}^\ddagger]\hat{\mathbf{B}} = (\hat{\mathbf{A}}^\ddagger\hat{\mathbf{B}}^\top)^\top\hat{\mathbf{A}}$. Here we note that $\mathcal{J}[\hat{\mathbf{A}}^\ddagger]$ is formally equivalent to its superoperator adjoint.[7]

From (A57) we can define what is sometimes referred to as the "dissipator" [35]

$$\mathcal{D}[\hat{A}]\hat{B} \equiv \mathcal{J}[\hat{A}]\hat{B} - \tfrac{1}{2}\{\hat{A}^\dagger\hat{A}, \hat{B}\}, \quad \text{(A60)}$$

where $\{\hat{A}, \hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}$. When $\hat{A} = \hat{A}^\dagger$ one may prefer to write the dissipator as a nested commutator,

$$\mathcal{D}[\hat{A}]\hat{B} = -\tfrac{1}{2}\,[\hat{A},[\hat{A},\hat{B}]]. \quad \text{(A61)}$$

As in (A58), we shorthand a sum of dissipators by

$$\mathcal{D}[\hat{\mathbf{A}}]\hat{B} \equiv \sum_k \mathcal{D}[\hat{A}_k]\hat{B}. \quad \text{(A62)}$$

Equations (A15) and (A62) then permit us to write

$$\mathcal{D}[\hat{\mathbf{A}}]\hat{\mathbf{B}} = \mathcal{J}[\hat{\mathbf{A}}]\hat{\mathbf{B}} - \tfrac{1}{2}\{\hat{\mathbf{A}}^\dagger\hat{\mathbf{A}}, \hat{\mathbf{B}}\}$$
$$= (\hat{\mathbf{A}}\hat{\mathbf{B}}^\top)^\top\hat{\mathbf{A}}^\ddagger - \tfrac{1}{2}\{\hat{\mathbf{A}}^\dagger\hat{\mathbf{A}}, \hat{\mathbf{B}}\}. \quad \text{(A63)}$$

In the case when $\hat{\mathbf{A}} = \hat{\mathbf{A}}^\ddagger$ the vector-operator generalization of (A61) is

$$\mathcal{D}[\hat{\mathbf{A}}]\hat{\mathbf{B}} = -\tfrac{1}{2}\{(\hat{\mathbf{A}}^\top\lfloor\hat{\mathbf{A}},\hat{\mathbf{B}}\rceil)^\top + \lfloor\hat{\mathbf{B}},\hat{\mathbf{A}}\rceil\hat{\mathbf{A}}\}. \quad \text{(A64)}$$

The appearance of a dissipator $\mathcal{D}[\hat{A}]$ in the master equation is associated with the coupling of the system to the environment via $\hat{A}$. If the system is also under continuous observation, then measurement back-action terms arise and they can be written concisely by defining

$$\mathcal{H}[\hat{A}]\hat{B} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}^\dagger - \mathrm{Tr}(\hat{A}\hat{B} + \hat{B}\hat{A}^\dagger)\hat{B}. \quad \text{(A65)}$$

---

[7] The standard definition of the adjoint of a superoperator $\mathcal{N}$ [35] is another superoperator $\mathcal{N}^\dagger$, such that for any scalar operator $\hat{A}$, and any state ρ, $\mathrm{Tr}[\hat{A}(\mathcal{N}\rho)] = \mathrm{Tr}[(\mathcal{N}^\dagger\hat{A})\rho]$. We can generalize $\hat{A}$ in this definition to a vector operator $\hat{\mathbf{A}}$ so $\mathcal{N}^\dagger$ is such that $\mathrm{Tr}[\hat{\mathbf{A}}(\mathcal{N}\rho)] = \mathrm{Tr}[(\mathcal{N}^\dagger\hat{\mathbf{A}})\rho]$. From this it follows that $(\mathcal{J}[\hat{\mathbf{A}}])^\dagger = \mathcal{J}[\hat{\mathbf{A}}^\ddagger]$.

We will invariably be considering $\mathcal{H}[\mathbf{C}^\top\hat{\mathbf{A}}]$ with $\mathbf{C} = \mathbf{C}^*$. In this case one may prefer to generalize (A65) to

$$\mathcal{H}[\hat{A}]\hat{B} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}^\ddagger - \mathrm{Tr}[\hat{A}\hat{B} + \hat{B}\hat{A}^\ddagger]\hat{B} \qquad (A66)$$

and write

$$\mathcal{H}[\mathbf{C}^\top\hat{\mathbf{A}}] = \mathbf{C}^\top\mathcal{H}[\hat{\mathbf{A}}]. \qquad (A67)$$

Note that (A65) and (A66) each contain a term which is nonlinear in $\hat{B}$. One may find it useful to also define the linear versions of (A65) and (A66),

$$\bar{\mathcal{H}}[\hat{A}]\hat{B} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}^\dagger, \qquad (A68)$$

$$\bar{\mathcal{H}}[\hat{A}]\hat{B} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}^\ddagger. \qquad (A69)$$

APPENDIX B: DERIVATION Of EQ. (46)

We first write (suppressing the time dependence)

$$\mathrm{Tr}[(\rho + d\rho_{\mathrm{c}})^2] = 1 + \frac{2}{\hbar}\mathrm{Tr}[\rho\,(\hbar\,d\rho_{\mathrm{c}})] + \frac{1}{\hbar^2}\mathrm{Tr}[(\hbar\,d\rho_{\mathrm{c}})^2], \quad (B1)$$

where the change in the state is given by

$$\hbar\,d\rho_{\mathrm{c}} = \mathcal{L}\rho\,dt + \mathcal{H}[d\mathbf{w}^\top\mathsf{M}^\dagger\hat{\mathbf{c}}]\rho. \qquad (B2)$$

Recall that $\mathcal{L}$ is defined by (1). We will assume that $\rho(t)$ is unconditioned and such that $\mathrm{Tr}[\rho(t)] = 1$ and $\rho^2(t) = \rho(t)$.

We first work out the second term in (B1). Taking the trace of (B2) against ρ,

$$\mathrm{Tr}\{\rho\,(\hbar\,d\rho_{\mathrm{c}})\} = \mathrm{Tr}\{\rho\,\mathcal{L}\rho\}dt + \mathrm{Tr}\{\rho\mathcal{H}[d\mathbf{w}^\top\mathsf{M}^\dagger\hat{\mathbf{c}}]\rho\}. \quad (B3)$$

From (A65) we find, for any $\hat{A}$,

$$\mathrm{Tr}\{\rho\,\mathcal{H}[\hat{A}]\rho\} = 0. \qquad (B4)$$

From (1) we get

$$\mathrm{Tr}\{\rho\mathcal{L}\rho\} = \langle\hat{\mathbf{c}}^\dagger\rangle\langle\hat{\mathbf{c}}\rangle - \langle\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}\rangle. \qquad (B5)$$

Therefore (B3) is

$$\mathrm{Tr}\{\rho\,(\hbar\,d\rho_{\mathrm{c}})\} = \mathrm{Tr}\{\rho\mathcal{L}\rho\}dt = (\langle\hat{\mathbf{c}}^\dagger\rangle\langle\hat{\mathbf{c}}\rangle - \langle\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}\rangle)\,dt. \quad (B6)$$

To order dt the third term in (B1) is proportional to

$$\mathrm{Tr}[(\hbar\,d\rho_{\mathrm{c}})^2] = \mathrm{Tr}[(\mathcal{H}[d\mathbf{w}^\top\mathsf{M}^\dagger\hat{\mathbf{c}}]\rho)^2]. \qquad (B7)$$

For any $\hat{A}$,

$$\begin{aligned}
(\mathcal{H}[\hat{A}]\rho)^2 &= (\hat{A}\rho)(\hat{A}\rho) + (\hat{A}\rho)(\rho\hat{A}^\dagger) - (\hat{A}\rho)\mathrm{Tr}[\hat{A}\rho + \rho\hat{A}^\dagger]\rho \\
&\quad + (\rho\hat{A}^\dagger)(\hat{A}\rho) + (\rho\hat{A}^\dagger)(\rho\hat{A}^\dagger) - (\rho\hat{A}^\dagger)\mathrm{Tr}[\hat{A}\rho + \rho\hat{A}^\dagger]\rho \\
&\quad + \mathrm{Tr}[\hat{A}\rho + \rho\hat{A}^\dagger]\mathrm{Tr}[\hat{A}\rho + \rho\hat{A}^\dagger]\rho^2. \qquad (B8)
\end{aligned}$$

Taking the trace and using $\rho^2 = \rho$,

$$\mathrm{Tr}\{(\mathcal{H}[\hat{A}]\rho)^2\} = 2(\langle\hat{A}^\dagger\hat{A}\rangle - \langle\hat{A}\rangle\langle\hat{A}\rangle). \qquad (B9)$$

This gives, for $\hat{A} = d\mathbf{w}^\top\mathsf{M}^\dagger\hat{\mathbf{c}}$,

$$\mathrm{Tr}\{(\mathcal{H}[d\mathbf{w}^\top\mathsf{M}^\dagger\hat{\mathbf{c}}]\rho)^2\} = 2(\langle\hat{\mathbf{c}}^\dagger\mathsf{M}\mathsf{M}^\dagger\hat{\mathbf{c}}\rangle - \langle\hat{\mathbf{c}}\rangle\mathsf{M}\mathsf{M}^\dagger\langle\hat{\mathbf{c}}\rangle)\,dt. \tag{B10}$$

Thus substituting (B6) and (B10) into (B1),

$$\begin{aligned}
&\mathrm{Tr}[(\rho + d\rho_{\mathrm{c}})^2] \\
&= 1 + \frac{2}{\hbar}\mathrm{tr}[\langle(\hat{\mathbf{c}} - \langle\hat{\mathbf{c}}\rangle)^\ddagger(\hat{\mathbf{c}} - \langle\hat{\mathbf{c}}\rangle)^\top\rangle(\mathrm{I}_L - \mathsf{M}\mathsf{M}^\dagger/\hbar)^\top]dt \\
&= 1 + \mathrm{tr}[H(\mathrm{I}_L - \mathsf{M}\mathsf{M}^\dagger/\hbar)^\top]\,dt. \qquad (B11)
\end{aligned}$$

The matrix H was defined in (47).

APPENDIX C: EQ. (131) FOR $L = 1$

We can show that for $L = 1$ it is always possible to find a decomposition in the form of (131) for an arbitrary M. The most general 2×2 orthogonal matrix can be parameterized by a single variable φ,

$$\mathsf{O} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \mp\sin\varphi & \pm\cos\varphi \end{pmatrix}, \qquad (C1)$$

where the sign in the second row corresponds to $\det(\mathsf{O}) = \pm 1$. This gives, from (131),

$$\mathsf{M}^\top = \sqrt{\hbar\,\eta}\,e^{i\phi}\begin{pmatrix} \sqrt{\theta}\,\cos\varphi \pm i\sqrt{\bar{\theta}}\,\sin\varphi \\ \sqrt{\theta}\,\sin\varphi \mp i\sqrt{\bar{\theta}}\,\cos\varphi \end{pmatrix}. \qquad (C2)$$

If we define

$$\mathsf{M}^\top = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} |m_1|\,e^{i\alpha_1} \\ |m_2|\,e^{i\alpha_2} \end{pmatrix}, \qquad (C3)$$

then our problem is to find $(\eta, \phi, \theta, \varphi)$ such that

$$|m_1|\,e^{i\alpha_1} = \sqrt{\hbar\,\eta}\,e^{i\phi}(\sqrt{\theta}\,\cos\varphi \pm i\sqrt{\bar{\theta}}\,\sin\varphi) \qquad (C4)$$

$$|m_2|\,e^{i\alpha_2} = \sqrt{\hbar\,\eta}\,e^{i\phi}(\sqrt{\theta}\,\sin\varphi \mp i\sqrt{\bar{\theta}}\,\cos\varphi), \qquad (C5)$$

given m_1, m_2, which determine η by

$$|m_1|^2 + |m_2|^2 = \hbar\,\eta. \qquad (C6)$$

Equations (C4) and (C5) are true if and only if the modulus and argument (phase) of each side are equal, i.e.,

$$|m_1|^2 = \hbar\,\eta\,[\theta\cos^2\varphi + \bar{\theta}\sin^2\varphi], \qquad (C7)$$

$$|m_2|^2 = \hbar\,\eta\,[\theta\sin^2\varphi + \bar{\theta}\cos^2\varphi], \qquad (C8)$$

and

$$\alpha_1 = \phi + \arg(\sqrt{\theta}\,\cos\varphi \pm i\sqrt{\bar{\theta}}\,\sin\varphi), \qquad (C9)$$

$$\alpha_2 = \phi + \arg(\sqrt{\theta}\,\sin\varphi \mp i\sqrt{\bar{\theta}}\,\cos\varphi). \qquad (C10)$$

Rearranging (C7) we find

$$\theta = \frac{|m_1|^2 - \hbar\,\eta\sin^2\varphi}{\hbar\,\eta\cos(2\varphi)}. \qquad (C11)$$

If we now substitute (C6) into (C11) and define $r = (|m_1|/|m_2|)^2$, then we get

$$\theta_r(\varphi) \equiv \frac{r\cos^2\varphi - \sin^2\varphi}{(r+1)(\cos^2\varphi - \sin^2\varphi)}. \qquad (C12)$$

Note that φ must also ensure $0 \leqslant \theta \leqslant 1$. For small r we find, from (C12),

$$\theta_0(\varphi) = \frac{-\sin^2\varphi}{\cos^2\varphi - \sin^2\varphi}, \qquad (C13)$$

while for r large $\theta_r(\varphi)$ approaches

$$\theta_\infty(\varphi) \equiv \frac{\cos^2 \varphi}{\cos^2 \varphi - \sin^2 \varphi}. \quad (C14)$$

At these extremes the only values of φ that enforce $\theta \in [0,1]$ are integer multiples of $\pi/2$, for which θ is either 0 or 1.

If we now attempt to eliminate θ in (C8) by substituting in (C12) we simply arrive at (C6), which is independent of φ. Thus (C7) and (C8) are solved simultaneously with any φ and (C12). The value of φ is determined from solving (C9) and (C10) simultaneously. First eliminating ϕ we see that φ must satisfy

$$\delta \equiv \alpha_1 - \alpha_2 = F_r^\pm(\varphi), \quad (C15)$$

where we have defined the argument of m_1/m_2 as

$$F_r^\pm(\varphi) \equiv \arg\left[\frac{\sqrt{\theta_r(\varphi)}\,\cos\varphi \pm i\sqrt{1-\theta_r(\varphi)}\,\sin\varphi}{\sqrt{\theta_r(\varphi)}\,\sin\varphi \mp i\sqrt{1-\theta_r(\varphi)}\,\cos\varphi}\right]. \quad (C16)$$

We have written $\theta_r(\varphi)$ in (C16) to signify that here we are substituting in (C12). Recall the signs here correspond to $\det(\mathsf{O}) = \pm 1$. Equation (C15) is transcendental. We plot

$F_r^+(\varphi)$ in Fig. 7 as a function of r and φ. We find that for $\det(\mathsf{O}) = 1$

$$\mathrm{Im}\left[\frac{\sqrt{\theta}\,\cos\varphi + i\sqrt{1-\theta}\,\sin\varphi}{\sqrt{\theta}\,\sin\varphi - i\sqrt{1-\theta}\,\cos\varphi}\right]$$
$$= \frac{\sqrt{\theta(1-\theta)}}{\theta\sin^2\varphi + (1-\theta)\cos^2\varphi} \geqslant 0, \quad (C17)$$

which means the range of F_r^+ should be between 0 and π as seen in Fig. 7. A noteworthy feature occurs at $r = 1$. From (C12) we get $\theta_1 = 1/2$ which gives

$$F_1^+ = \arg\left[\frac{\cos\varphi + i\sin\varphi}{\sin\varphi - i\cos\varphi}\right] = \arg[i] = \pi/2. \quad (C18)$$

We also plot $F_r^+(\varphi)$ (solid blue line) together with $\theta_r(\varphi)$ (purple dashed line) for selected values of r as a function of φ in Fig. 6.

Observe that $F_r^+(\varphi)$ always has a support which corresponds to $\theta_r \in [0,1]$. We illustrate this region in Fig. 6(b) with a light shade. It can be seen in Fig. 6 that this region persists for the various values of r that we have chosen. This illustrates that there is always a value of φ which solves (C15) for $\delta \in [0,\pi]$, *and* for which $\theta \in [0,1]$. Note that more than one value of φ will solve (C15) for a given δ which means that M can be
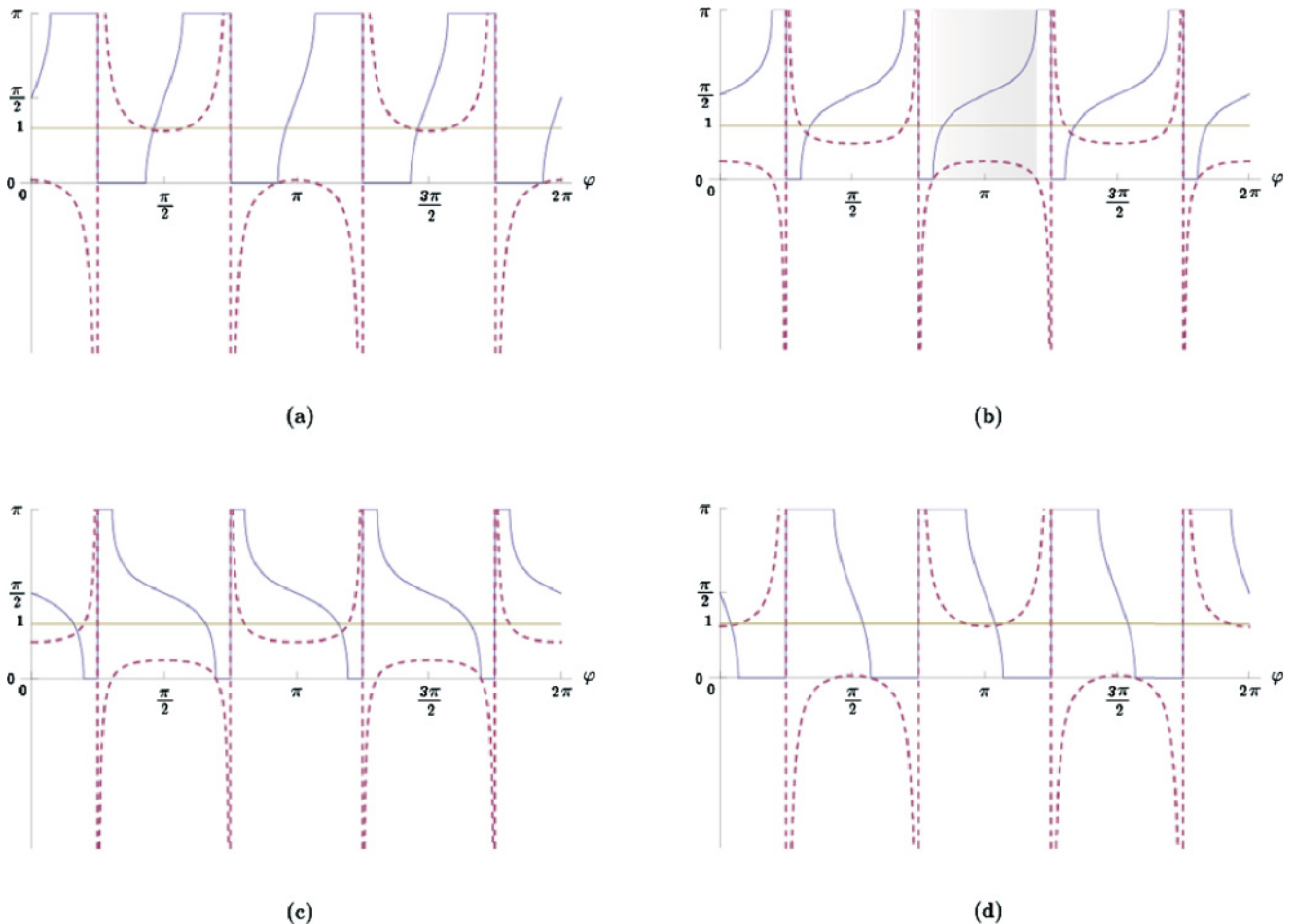

FIG. 7. (Color online) Plots of $F_r^+(\varphi)$ (solid blue line) and $\theta_r(\varphi)$ (dashed purple line) for selected values of r: (a) $r = 0.05$, (b) $r = 0.5$ (c) $r = 2$, and (d) $r = 20$.

factorized by more than one O [recall (C1) and (C2)]. Having obtained a value of φ we can then substitute it back into either (C9) or (C10) to obtain ϕ.

In the above we have concentrated on the case $\delta \in [0,\pi]$. Here we show how M can be factorized by (131) if we were given $\delta \in [-\pi,0]$. The analysis up to (C15) would remain the same but in this case we can first let $\alpha_2' \equiv \alpha_2 \pm \pi$ and solve

$$\alpha_1 - \alpha_2' = F_r^+(\varphi') \qquad (C19)$$

for φ'. The factorization of $\mathsf{M}^\top$ is then given by

$$\begin{pmatrix} |m_1|e^{i\alpha_1} \\ |m_2|e^{i\alpha_2} \end{pmatrix} = \begin{pmatrix} |m_1|e^{i\alpha_1} \\ |m_2|e^{i(\alpha_2' \mp \pi)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} |m_1|e^{i\alpha_1} \\ |m_2|e^{i\alpha_2'} \end{pmatrix}, \qquad (C20)$$

where $(\, |m_1|e^{i\alpha_1},\ |m_2|e^{i\alpha_2'}\,)^\top$ is factorized by

$$\begin{pmatrix} |m_1|e^{i\alpha_1} \\ |m_2|e^{i\alpha_2'} \end{pmatrix} = \sqrt{\hbar\,\eta}\ e^{i\phi}\begin{pmatrix} \cos\varphi' & -\sin\varphi' \\ \sin\varphi' & \cos\varphi' \end{pmatrix}\begin{pmatrix} \sqrt{\theta} \\ -i\sqrt{\bar{\theta}} \end{pmatrix}. \qquad (C21)$$

Substituting this into (C20) and taking the transpose we see that M has the form of (131) with the orthogonal matrix O given by

$$\mathsf{O} = \begin{pmatrix} \cos\varphi' & \sin\varphi' \\ -\sin\varphi' & \cos\varphi' \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad (C22)$$

where φ' is defined by (C19). Note that (C22) now has a determinant of -1. We see then for $\delta \in [-\pi,0]$ M is still factorized by (131) but with an O such that $\det(\mathsf{O}) = -1$.

[1] M. O. Scully and M. S. Zubairy, *Quantum Optics* (Cambridge University Press, New York, 1997).

[2] H. J. Carmichael, H. M. Castro-Beltran, G. T. Foster, and L. A. Orozco, Phys. Rev. Lett. **85**, 1855 (2000).

[3] H. J. Carmichael and H. Nha, J. Opt. B **6**, S645 (2004).

[4] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, New York, 2010).

[5] A. Barchielli and M. Gregoratti, *Quantum Trajectories and Measurements in Continuous Time: The Diffusive Case* (Springer-Verlag, Berlin, 2009).

[6] D. T. Gillespie, *Markov Processes: An Introduction For Physical Scientists* (Acamdemic Press, 1992).

[7] K. Jacobs, *Stochastic Processes for Physicists: Understanding Noisy Systems* (Canbridge University Press, New York, 2010).

[8] J. E. Brittain, Proc. IEEE **92**, 1866 (2004).

[9] I. Brodsky, Global Telecommunications Conference 2008, IEEE.

[10] A. I. Lvovsky and M. G. Raymer, Rev. Mod. Phys. **80**, 299 (2009).

[11] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, Phys. Rev. Lett. **70**, 1244 (1993).

[12] A. Ourjoumtsev, R. Tualle-Brouri, and P. Grangier, Phys. Rev. Lett. **96**, 213601 (2006).

[13] L. Vaidman, Phys. Rev. A **49**, 1473 (1994).

[14] S. L. Braunstein and H. J. Kimble, Phys. Rev. Lett. **80**, 869 (1998).

[15] A. Furusawa, J. L. Sørensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, and E. S. Polzik, Science **282**, 706 (1998).

[16] T. C. Zhang, K. W. Goh, C. W. Chou, P. Lodahl, and H. J. Kimble, Phys. Rev. A **67**, 033802 (2003).

[17] R. Ruskov and A. N. Korotkov, Phys. Rev. B **67**, 241305(R) (2003).

[18] H. Jeong, M. S. Kim, T. C. Ralph, and B. S. Ham, Phys. Rev. A **70**, 061801(R) (2004).

[19] A. Ourjoumtsev, H. Jeong, R. Tualle-Brouri, and P. Grangier, Nature **448**, 784 (2007).

[20] D. A. Rodrigues, C. E. A. Jarvis, B. Györffy, T. P. Spiller, and J. F. Annett, J. Phys. Condens. Matter **20**, 075211 (2008).

[21] M. Ohlinger and J. Eisert.

[22] Z. Liu, L. Kuang, K. Hu, L. Xu, S. Wei, L. Guo, and X.-Q. Li, Phys. Rev. A **82**, 032335 (2010).

[23] A. Kuzmich, I. A. Walmsley, and L. Mandel, Phys. Rev. Lett. **85**, 1349 (2000).

[24] H. Nha and H. J. Carmichael, Phys. Rev. Lett. **93**, 020401 (2004).

[25] R. García-Patrón, J. Fiurášek, N. J. Cerf, J. Wenger, R. Tualle-Brouri, and Ph. Grangier, Phys. Rev. Lett. **93**, 130409 (2004).

[26] H. M. Wiseman and L. Diósi, Chem. Phys. **268**, 91 (2001).

[27] H. M. Wiseman and A. C. Doherty, Phys. Rev. Lett. **94**, 070405 (2005).

[28] A. Chia and H. M. Wiseman, Phys. Rev. A **84**, 012120 (2011).

[29] S. Mancini and H. M. Wiseman, Phys. Rev. A **75**, 012330 (2007).

[30] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, New York, 2010).

[31] N. G. Walker, J. Mod. Opt. **34**, 15 (1987). See also Ref. [32] and the references therein for applications of N-port homodyning.

[32] A. Zucchetti, W. Vogel, and D.-G. Welsch, Phys. Rev. A **54**, 856 (1996).

[33] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Phys. Rev. Lett. **73**, 58 (1994).

[34] R. Alicki and K. Lendi, *Quantum Dynamical Semigroups and Applications* (Springer, Berlin, 2007).

[35] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, UK, 2002).

[36] G. Lindblad, Commun. Math. Phys. **48**, 199 (1976).

[37] H. J. Carmichael, *Statistical Methods in Quantum Optics 1* (Springer-Verlag, Berlin, 1999).

[38] H. J. Carmichael, *An Open Systems Approach to Quantum Optics* (Springer-Verlag, Berlin, 1993).

[39] H. J. Carmichael, *Statistical Methods in Quantum Optics 2* (Springer-Verlag, Berlin, 2008).

[40] C. W. Gardiner, *Handbook of Stochastic Methods*, 3rd ed. (Springer-Verlag, Berlin, 2004).

[41] M. B. Plenio and P. L. Knight, Rev. Mod. Phys. **70**, 101 (1998).

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

[43] V. B. Braginsky and F. Ya. Khalili, *Quantum Measurement* (Cambridge University Press, New York, 1992).

[44] W. S. Levine (ed.), *The Control Handbook* (CRC Press, Boca Raton, FL, 1996).

[45] U. Leonhardt and H. Paul, Prog. Quantum Electron. **19**, 89 (1995).

[46] H. Anton and R. C. Busby, *Contemporary Linear Algebra* (Wiley, New York, 2003).

[47] K. R. Parthasarathy, *An Introduction to Quantum Stochastic Calculus* (Birkhäuser, Berlin, 1992).

[48] C. W. Gardiner and M. J. Collett, Phys. Rev. A **31**, 3761 (1985).

[49] M. A. Rubin, Found. Phys. Lett. **14**, 301 (2001).

[50] B. d'Espagnat, *Veiled Reality: An Analysis of Present-Day Quantum Mechanical Concepts*, (Addison-Wesley, Reading, MA, 1995).

[51] M. J. Collett and C. W. Gardiner, Phys. Rev. A **30**, 1386 (1984).

[52] C. Noh, A. Chia, H. Nha, M. J. Collett, and H. J. Carmichael, Phys. Rev. Lett. **102**, 230501 (2009).

[53] T. C. Ralph, Opt. Lett. **24**, 348 (1998).

[54] J. H. Shapiro, G. Saplakoglu, S. T. Ho, P. Kumar, B. E. A. Saleh, and M. C. Teich, J. Opt. Soc. Am. B **4**, 1604 (1987).

[55] V. P. Belavkin, Phys. Lett. A **140**, 355 (1989).

[56] V. P. Belavkin and P. Staszewski, Phys. Lett. A **140**, 359 (1989).

[57] V. P. Belavkin, O. Hirota, and R. L. Hudson (Eds.), *Quantum Communications and Measurement, Proceedings of an International Workshop on Quantum Communications and Measurement, July 11–16, 1994, Nottingham, England* (Plenum Press, New York, 1995).

[58] H. M. Wiseman and G. J. Milburn, Phys. Rev. A **47**, 642 (1993).

[59] P. Kok and B. W. Lovett, *An Introduction to Optical Quantum Information Processing* (Cambridge University Press, New York, 2010).

[60] R. Schack, T. A. Brun, and I. C. Percival, J. Phys. A **28**, 5401 (1995).

[61] T. Pellizzari, S. A. Gardiner, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **75**, 3788 (1995).

[62] G. G. Carlo, G. Benenti, G. Casati, and C. Mejía-Monasterio, Phys. Rev. A **69**, 062317 (2004).

[63] G. G. Carlo, G. Benenti, and G. Casati, Phys. Rev. Lett. **91**, 257903 (2003).

[64] C. Ahn, H. M. Wiseman, and G. J. Milburn, Phys. Rev. A **67**, 052310 (2003).

[65] R. van Handel and H. Mabuchi, e-print arXiv:quant-ph/0511221v1 (2005).

[66] H. Mabuchi, New J. Phys. **11**, 105044 (2009).

[67] H. Goto and K. Ichimura, Phys. Rev. A **72**, 054301 (2005).

[68] A. Barenco, T. A. Brun, R. Schack, and T. P. Spiller, Phys. Rev. A **56**, 1177 (1997).

[69] R. Wagner Jr. and J. P. Clemens, J. Opt. Soc. Am. B **26**, 541 (2009).

[70] H. J. Carmichael, Int. J. Quantum. Inform. **3**, 11 (2005).

[71] S. L. Braunstein, Nature **394**, 47 (1998).

[72] T. Aoki, G. Takahashi, T. Kajiya, J. Yoshikawa, S. L. Braunstein, P. van Loock, and A. Furusawa, Nat. Phys. **5**, 541 (2009).

[73] M. Lax, Phys. Rev. **129**, 2342 (1963).

[74] M. Lax, Phys. Rev. **157**, 213 (1967).