Generalized trace-distance measure connecting quantum and classical non-Markovianity

Steffen Wißmann and Heinz-Peter Breuer

Physikalisches Institut, Universität Freiburg, Hermann-Herder-Straße 3, D-79104 Freiburg, Germany

Bassano Vacchini

Dipartimento di Fisica, Università degli Studi di Milano, Via Celoria 16, I-20133 Milan, Italy
and INFN, Sezione di Milano, Via Celoria 16, I-20133 Milan, Italy

We establish a direct connection of quantum Markovianity of an open system to its classical counterpart by generalizing the criterion based on the information flow. Here the flow is characterized by the time evolution of Helstrom matrices, given by the weighted difference of statistical operators, under the action of the quantum dynamical map. It turns out that the introduced criterion is equivalent to P divisibility of a quantum process, namely, divisibility in terms of positive maps, which provides a direct connection to classical Markovian stochastic processes. Moreover, it is shown that mathematical representations similar to those found for the original trace-distance-based measure hold true for the associated generalized measure for quantum non-Markovianity. That is, we prove orthogonality of optimal states showing a maximal information backflow and establish a local and universal representation of the measure. We illustrate some properties of the generalized criterion by means of examples.

PACS number(s): 03.65.Yz, 03.65.Ta, 03.67.−a, 02.50.−r

I. INTRODUCTION

Even though experimental techniques are developing rapidly, perfect isolation of fragile quantum systems from noisy environments cannot be warranted in general, which necessitates effective descriptions for nonunitary dynamics [1]. A well-known treatment of the dynamics of such systems is provided by a quantum dynamical semigroup represented by a generator of the Gorini-Kossakowski-Sudarshan-Lindblad form [2,3].

However, this description uses several rather drastic approximations that do not apply to open quantum systems in general. The mismatch is typically associated with the neglect of memory effects. Classically, there exists a well-established mathematical theory dealing with stochastic dynamics featuring memory effects, based on the theory of stochastic processes, yet the definition of classical Markovian stochastic processes cannot be straightforwardly carried over to the quantum regime. This led to an intense debate along with different proposals for the characterization and quantification of memory effects in the dynamics of open quantum systems. Among others, approaches to non-Markovianity were based on deviation from semigroup dynamics [4], on divisibility in terms of completely positive maps [5], on dynamics of entanglement [5] and correlations [6], and on the Fisher information [7] (see [8] for a review).

In this work we focus on a straightforward extension of the measure of non-Markovianity introduced in Refs. [9,10] and reviewed in [11], which characterizes memory effects by an exchange of information between the open quantum system and the environment. This interpretation applies also to the generalized measure that uses Helstrom matrices known from quantum estimation theory [12]. A similar approach was already introduced in Ref. [13], focusing, at variance with the present proposal, on divisibility in terms of completely positive maps (CP divisibility). However, CP divisibility yields only a sufficient condition for quantum Markovianity with respect

to the characterization proposed in this article, which turns out to be equivalent to divisibility in terms of positive maps (P divisibility) and has in addition a clear-cut connection to classical Markov stochastic processes. The generalized criterion introduced in Sec. II thus combines a physical interpretation of non-Markovianity in terms of an information flow and a relation to the well-known classical definition.

Moreover, the associated measure has mathematical features and representations similar to those found for the original definition [14,15], which simplify its analytical, numerical, and experimental determination drastically. Indeed, we prove in Sec. III A that, first, optimal initial states for non-Markovian dynamics, experiencing a maximal backflow of information, must be orthogonal and, second, the measure admits a local representation showing locality and universality of quantum memory effects.

An illustration of the essential features of the generalized characterization of non-Markovianity in comparison with the original definition and approaches based on CP divisibility is provided in Sec. III B. The examples show the sensitivity of the present approach to memory effects of dynamics to which the original definition was unsusceptible and illustrate the existing difference between different types of divisibility of a dynamical process. Finally, we summarize our results in Sec. IV.

II. STATE DISCRIMINATION AND P DIVISIBILITY

Henceforth, $\mathcal{H}$ refers to the Hilbert space of the open quantum system and the corresponding set of physical states, i.e., the set of positive trace class operators with unit trace, is denoted by $\mathcal{S}(\mathcal{H})$. Moreover, we assume that the dynamics of the open quantum system is determined by a one-parameter family $\Phi = \{\Phi_t | 0 \leqslant t \leqslant T\}$ of completely positive, trace-preserving linear maps Φ_t, where Φ_0 refers to the identity [1].

042108-1

A. Ensemble discrimination

The previously introduced definition of quantum non-Markovianity [9,10], which was based on the concept of an information flow between system and environment, relied on the trace distance between two quantum states ρ_1 and ρ_2 defined by [16]

$$\mathcal{D}(\rho_1,\rho_2) \equiv \tfrac{1}{2}\|\rho_1 - \rho_2\|_1. \qquad (1)$$

Here $\|A\|_1 = \mathrm{Tr}|A|$, where the modulus of an operator is given by $|A| = \sqrt{A^\dagger A}$, refers to the so-called trace norm on the set of trace class operators [17]. A natural and interesting generalization of this quantity is obtained when one also allows for biased probability distributions. That is, one considers

$$\|p_1\rho_1 - p_2\rho_2\|_1 = \mathrm{Tr}|p_1\rho_1 - p_2\rho_2|, \qquad (2)$$

where $\{p_i\}$ refers to an arbitrary binary probability distribution, i.e., $p_{1,2} \geqslant 0$ and $p_1 + p_2 = 1$. Clearly, choosing an unbiased distribution, i.e., $p_{1,2} = 1/2$, one obtains the original expression (1). This generalization was first proposed in this context by Chruściński *et al.* [13].

The Hermitian operator $\Delta = p_1\rho_1 - p_2\rho_2$ is also known as the Helstrom matrix [12,18] and we find, applying the triangle inequality to Eq. (2),

$$|p_1 - p_2| \leqslant \|\Delta\|_1 \leqslant p_1 + p_2 = 1. \qquad (3)$$

Clearly, the lower bound is obtained for $\rho_1 = \rho_2$, whereas the upper bound is attained if and only if $\rho_1 \perp \rho_2$, meaning that the two states have orthogonal support (which is defined as the subspace spanned by the eigenvectors with nonzero eigenvalues). This is easily shown by means of the characterization of the trace norm

$$\|\Delta\|_1 = 2 \max_\Pi \mathrm{Tr}\{\Pi\Delta\} + p_2 - p_1, \qquad (4)$$

where the maximum is taken over all projection operators Π. Here Eq. (4) is derived by employing the Jordan-Hahn decomposition of the Hermitian Helstrom matrix Δ in terms of two positive and orthogonal operators, i.e., $\Delta = S - Q$, where $S, Q \geqslant 0$ and $S \perp Q$.

The interpretation of the trace distance of two states $\rho_{1,2} \in \mathcal{S}(\mathcal{H})$ as a measure of their distinguishability directly carries over to the trace norm of a Helstrom matrix Δ. Consider a one-shot two-state discrimination problem where Alice prepares one out of two quantum states $\rho_{1,2}$ with corresponding probability $p_{1,2}$, i.e., we have $p_1 + p_2 = 1$. Hence, we allow for the general situation of a biased preparation by Alice. She finally sends the prepared state to Bob, who performs a single (strong) measurement to infer which state he had received [17]. To guess the state from the measurement with possible outcomes Ω, Bob defines two sets of possible results R and $\Omega \backslash R$ and assigns the state to be ρ_1 if the measurement outcome is in R and ρ_2 if a value in $\Omega \backslash R$ is obtained. This strategy results in an effective two-valued positive-operator-valued measure $\{T_R, \mathbb{1} - T_R\}$, where T_R refers to the collection of effects corresponding to outcomes in R. The probability for correct state discrimination via this strategy is then given by

$$P_{\text{success}} = p_1 \mathrm{Tr}\{T_R\rho_1\} + p_2 \mathrm{Tr}\{(\mathbb{1} - T_R)\rho_2\}$$
$$= p_2 + \mathrm{Tr}\{\Delta T_R\}, \qquad (5)$$

which is maximal if T_R is the projection $\Pi_{\{\Delta \geqslant 0\}}$ on the subspace spanned by eigenvectors corresponding to positive eigenvalues of Δ. Employing that $\mathrm{Tr}|X| = \mathrm{Tr}\{X(\Pi_{\{X \geqslant 0\}} - \Pi_{\{X < 0\}})\}$ for any Hermitian operator X, one shows that the maximal success probability for correct discrimination obeys [17]

$$P_{\text{success}}^{\max} = \max_{0 \leqslant T_R \leqslant \mathbb{1}} P_{\text{success}} = \tfrac{1}{2}(1 + \|\Delta\|_1). \qquad (6)$$

Hence, the trace norm of the Helstrom matrix $\Delta = p_1\rho_1 - p_2\rho_2$ is the bias in favor of the correct state identification of the state prepared by Alice, so it may be interpreted as a measure for the distinguishability of the two states ρ_1 and ρ_2 with associated weights p_1 and p_2.

Due to a result by Kossakowski [19,20], the trace norm can also be used to witness positivity of a trace-preserving map Λ. That is, a trace-preserving map Λ is positive if and only if it defines a contraction for any Hermitian operator X with respect to the trace norm, i.e.,

$$\|\Lambda X\|_1 \leqslant \|X\|_1 \quad \text{for all } X = X^\dagger. \qquad (7)$$

We thus obtain for two states evolving according to a dynamical map Φ_t,

$$\|p_1\Phi_t(\rho_1) - p_2\Phi_t(\rho_2)\|_1 \leqslant \|p_1\rho_1 - p_2\rho_2\| \qquad (8)$$

for all t where we also used linearity of the map.

Adopting the previous characterization for quantum non-Markovianity [9,10], one then directly derives a generalized criterion that still relies on the concept of an information flow.

Definition 1. A quantum process Φ is defined to be Markovian if $\|p_1\Phi_t(\rho_1) - p_2\Phi_t(\rho_2)\|_1$ is a monotonically decreasing function of $t \geqslant 0$ for all sets $\{p_i, \rho_i\}$ with $p_i \geqslant 0$, $p_1 + p_2 = 1$, and $\rho_i \in \mathcal{S}(\mathcal{H})$.

We stress that the authors of Ref. [13] introduced a different definition for Markovianity, which used the Helstrom matrices of a dilated system making it equivalent to *CP* divisibility of the dynamical process. However, as we shall see in Sec. II B, Markovianity as defined here can be related to the concept of *P* divisibility and provides a clear-cut connection to classical Markov stochastic processes.

We conclude this section by showing that the previously developed explanation to substantiate the interpretation of quantum memory as an information backflow from the environment to the system [21] also applies here. To this end, we define the quantities

$$\mathcal{I}_{\text{int}}(t) = \left\| p_1\rho_S^{(1)}(t) - p_2\rho_S^{(2)}(t) \right\|_1 \qquad (9)$$

and

$$\mathcal{I}_{\text{ext}}(t) = \left\| p_1\rho_{SE}^{(1)}(t) - p_2\rho_{SE}^{(2)}(t) \right\|_1 - \mathcal{I}_{\text{int}}(t), \qquad (10)$$

where $\rho_{SE}^{(i)}(t) = U_t\rho_S^{(i)} \otimes \rho_E U_t^\dagger$ refers to the state of the system and environment at time t subject to a unitary evolution U_t. We thus have $\rho_S^{(i)}(t) = \mathrm{Tr}_E\rho_{SE}^{(i)}(t)$ and, similarly, the state of the environment at time t is given by $\rho_E^{(i)}(t) = \mathrm{Tr}_S\rho_{SE}^{(i)}(t)$.

As discussed in [21], $\mathcal{I}_{\text{int}}(t)$ describes the distinguishability of the open system at time t, while $\mathcal{I}_{\text{ext}}(t)$ can be interpreted as the information on the total system that is not accessible by measurements on the open system only. Due to contractivity of the trace norm under positive and trace-preserving maps

[cf. Eq. (7)], we conclude that both quantities are positive as the partial trace is (completely) positive and preserves the trace. Moreover, the existence of the quantum dynamical map implies that the initial state is factorized, therefore $\mathcal{I}_{\text{ext}}(0) = 0$, so unitary invariance of the trace norm implies

$$\mathcal{I}_{\text{int}}(t) + \mathcal{I}_{\text{ext}}(t) = \mathcal{I}_{\text{int}}(0) = \text{const.} \tag{11}$$

This relation clearly expresses the idea of exchange of information between the open system and the environment as an increase of $\mathcal{I}_{\text{ext}}(t)$ necessarily forces $\mathcal{I}_{\text{int}}(t)$ to decrease. One may also derive upper bounds for the external information that, however, do not allow for a clear interpretation in terms of the formation of correlations between the system and environment and changes in the environmental states as proven for the trace distance in Ref. [22].

B. Distinguishability and divisibility

In this section we introduce the concept of divisibility of a dynamical map, which has been at the basis of several other proposals for quantum non-Markovian dynamics [5,13,23], and elucidate its connection to the distinguishability measure defined before. We start by revisiting the notion of n positivity. A map Λ on $\mathcal{H}$ is said to be n positive if and only if the linear tensor extension of the map given by $(\Lambda \otimes \mathbb{1}_n)(A \otimes B) = \Lambda(A) \otimes B$ for linear operators A on $\mathcal{H}$ and B on $\mathbb{C}^n$ is positive, that is, $(\Lambda \otimes \mathbb{1}_n)C \geqslant 0$ for all positive linear operators C on $\mathcal{H} \otimes \mathbb{C}^n$.

Clearly, 1 positivity is equivalent to the map Λ being positive and the concept of n positivity is hierarchical, which means that n positivity of Λ implies that the map is also k positive for all $1 \leqslant k \leqslant n$ [24]. The converse, however, is not true in general [25,26]. Finally, one speaks of complete positivity if a map is n positive for all $n \in \mathbb{N}$. For a finite-dimensional system with $\dim \mathcal{H} = N$ this property is equivalent to N positivity [27].

If the maps Φ_t comprising a dynamical process are invertible for all $t \geqslant 0$ with inverse Φ_t^{-1}, then one can define a two-parameter family of maps given by

$$\Phi_{t,s} = \Phi_t \Phi_s^{-1} \tag{12}$$

for all $t \geqslant s \geqslant 0$ such that $\Phi_{t,0} = \Phi_t$ and $\Phi_t = \Phi_{t,s}\Phi_s$. We remark that although the existence of a left-inverse requiring injectivity of the maps Φ_t would be sufficient, we assume Φ_t^{-1} to be the left and right inverse. The notion of divisibility deals with the properties concerning n positivity of these maps, i.e.,

Definition 2. A dynamical process Φ is called P divisible (CP divisible) if $\Phi_{t,s}$ is a positive (completely positive) map for all $t \geqslant s \geqslant 0$.

Even though Φ_t is completely positive for all times by definition, this does not imply (complete) positivity of $\Phi_{t,s}$ in general as the inverse of a CP map need not be positive, showing that the above definition is nontrivial. However, the concept of divisibility relies on the existence of the two-parameter family that is limited to processes for which Φ_t^{-1} exists for all times. This property though cannot be taken for granted as, e.g., the damped Jaynes-Cummings model on resonance [1,10,21] or examples on quantum semi-Markov processes [28] show, thus making the concept of divisibility sometimes ill defined.

Typically, however, the inverse maps exist apart from isolated points in time, so one may describe the dynamical process by a time-local master equation for the open system during intermediate intervals. Assuming a sufficiently smooth time dependence, the generator obeys $\mathcal{K}_t = \dot{\Phi}_t \Phi_t^{-1}$ and has the general structure

$$\mathcal{K}_t \rho_S(t) = -i[H_S(t), \rho_S(t)] + \sum_j \gamma_j(t)[A_j(t)\rho_S(t)A_j^\dagger(t)$$
$$- \frac{1}{2}\{A_j^\dagger(t)A_j(t), \rho_S(t)\}], \tag{13}$$

which is similar to the well-known Lindblad form [2,3] apart from the fact that the system Hamiltonian $H_S(t)$, the rates $\gamma_j(t)$, and the Lindblad operators $A_j(t)$ may depend on time as the process might not represent a semigroup.

The maps of the two-parameter family are then given by

$$\Phi_{t,s} = T_\leftarrow \exp\left[\int_s^t d\tau\, \mathcal{K}_\tau\right], \tag{14}$$

where $T_\leftarrow$ refers to chronological time ordering. We thus have $\mathcal{K}_s = \frac{d}{dt}\Phi_{t,s}|_{t=s}$, so we conclude that Φ is $(C)P$ divisible if and only if $\mathcal{K}_t$ is the generator of a (completely) positive semigroup for all fixed $t \geqslant 0$. The famous Gorini-Kossakowski-Sudarshan-Lindblad theorem [2,3] and a result on generators of positive semigroups by Kossakowski [19] then allow for a characterization of CP and P divisibility in terms of the generator [21].

Theorem 1. The dynamics generated by $\mathcal{K}_t$ (13) (i) is CP divisible if and only if $\gamma_j(t) \geqslant 0$ holds for all j and $t \geqslant 0$ and (ii) is P divisible if and only if for all $n \neq m$

$$\sum_j \gamma_j(t)|\langle m|A_j(t)|n\rangle|^2 \geqslant 0 \tag{15}$$

holds for any orthonormal basis $\{|n\rangle\}$ of $\mathcal{H}$ and all $t \geqslant 0$.

Proof. The first statement is precisely the Gorini-Kossakowski-Sudarshan-Lindblad theorem [2,3] and the second can be derived from Kossakowski's result on generators of positive semigroups [19], which states that the dynamics generated by $\mathcal{L}$ is P divisible if and only if for any set of projections $\Pi = \{\Pi_m\}_{m \in I}$ defining a resolution of the identity, i.e., $\sum_m \Pi_m = \mathbb{1}_\mathcal{H}$, the relations

$$a_{mm}(\Pi) \leqslant 0, \quad m \in I \tag{16}$$

$$a_{mn}(\Pi) \geqslant 0, \quad m \neq n \in I \tag{17}$$

$$\sum_{m \in I} a_{mn}(\Pi) = 0, \quad n \in I \tag{18}$$

for $a_{mn}(\Pi) \equiv \text{Tr}\{\Pi_m(\mathcal{L}\Pi_n)\}$ are satisfied.

It is easily seen that the condition (18) refers to the preservation of trace, which is always met for the generator $\mathcal{K}_t$ (13) by its very structure. Rearranging terms of Eq. (18), one then finds

$$a_{nn}(\Pi) = -\sum_{m \neq n} a_{mn}(\Pi) \tag{19}$$

for all $n \in I$. The constraints thus reduce to the single relation (17), which can additionally be restricted to sets of rank-1 projections Π by virtue of linearity. Evaluating

$a_{mn}(\Pi)$ for $m \neq n$ and rank-1 projections associated with an orthonormal basis $\{|n\rangle\}$ of $\mathcal{H}$, one finds

$$
\begin{aligned}
a_{mn}(\Pi) &= \mathrm{Tr}\{|m\rangle\langle m|\mathcal{K}_t(|n\rangle\langle n|)\} \\
&= \sum_j \gamma_j(t)|\langle m|A_j(t)|n\rangle|^2,
\end{aligned}
\tag{20}
$$

which is condition (15). ∎

We point out that condition (15) only guarantees that the dynamics generated by $\mathcal{K}_t$ is positive. To warrant complete positivity of the dynamics as required for a dynamical map, one has to impose further constraints on the generator, which are unfortunately not completely known. Positivity of the rates $\gamma_j(t)$ for all j and $t \geqslant 0$ represents, for example, a sufficient (but not necessary) condition for this property.

Obviously, the conditions for P and CP divisibility coincide for a master equation with a single decay channel. We may alternatively characterize P divisibility by employing the contraction property (7) and link it to the quantum Markovian behavior defined above (see Definition 1). To this end, we first prove the following lemma.

Lemma 1. For any Hermitian operator $X \neq 0$ there exists a real number $\lambda > 0$ and a Helstrom matrix Δ such that $X = \lambda\Delta$.

Proof. Let $X = X^\dagger \neq 0$ be given. If $X \geqslant 0$, then $\rho_1 = (\mathrm{Tr}X)^{-1}X$ defines a state so that for $\lambda = \mathrm{Tr}X$ the Helstrom matrix characterized by $p_1 = 1$ and $p_2 = 0$ with arbitrary ρ_2 proves the claim, and similarly for $X \leqslant 0$.

Hence, suppose that X is indefinite. Employing the Jordan-Hahn decomposition, we thus find nonzero operators $Y_{1,2} \geqslant 0$, where $Y_1 \perp Y_2$ such that $X = Y_1 - Y_2$ and therefore $\mathrm{Tr}|X| = \mathrm{Tr}Y_1 + \mathrm{Tr}Y_2 > 0$. Clearly, the operators $\rho_i = (\mathrm{Tr}Y_i)^{-1}Y_i$ define states and we have

$$
X = \lambda(p_1\rho_1 - p_2\rho_2)
\tag{21}
$$

for $\lambda = \mathrm{Tr}|X|$ and $p_i = \lambda^{-1}\mathrm{Tr}Y_i$. These quantities are indeed positive and sum to one, thus representing a probability distribution, which concludes the proof. ∎

Since dynamical maps Φ_t as well as the trace norm are homogeneous (with respect to positive numbers) it thus suffices to apply the characterization of positivity (7) to Helstrom matrices, which yields the following theorem.

Theorem 2. If the dynamical maps defining a process Φ are bijective, then Φ is Markovian if and only if it is P divisible.

Proof. We first note that $p_1\Phi_t(\rho_1) - p_2\Phi_t(\rho_2) = \Phi_t(\Delta)$ holds for any probability distribution $\{p_i\}$ and pair of states ρ_i. Hence, it suffices to consider the time evolution of Helstrom matrices when studying quantum Markovianity. Suppose Φ is P divisible. It follows that

$$
\|\Phi_t(\Delta)\|_1 = \|\Phi_{t,s}(\Phi_s(\Delta))\|_1 \leqslant \|\Phi_s(\Delta)\|_1
\tag{22}
$$

for all $t \geqslant s \geqslant 0$ and Helstrom matrices Δ due to positivity of $\Phi_{t,s}$. Hence, $\|\Phi_t(\Delta)\|_1$ is a monotonically decreasing function of time for any Δ showing that the process Φ is Markovian.

For the converse, we first note that the inverse map Φ_t^{-1} exists for all t as the dynamical maps are bijective on the set of Hermitian operators by assumption. Thus, the maps $\Phi_{t,s}$ exist for all $t \geqslant s \geqslant 0$. Now let the process be Markovian, i.e., $\|\Phi_t(\Delta)\|_1 \leqslant \|\Phi_s(\Delta)\|_1$ for all $t \geqslant s \geqslant 0$ and Helstrom matrices Δ. We may rewrite this as

$$
\|\Phi_{t,s}(\Phi_s(\Delta))\|_1 \leqslant \|\Phi_s(\Delta)\|_1
\tag{23}
$$

from which positivity of $\Phi_{t,s}$ follows according to Eq. (7) by employing Lemma 1 and the fact that Φ_t is bijective for all t. Hence, Φ is P divisible. ∎

C. Connection to classical Markovian stochastic processes

The definition of Markovianity stated above is thus equivalent to P divisibility of a dynamical process Φ if this notion is well defined at all. However, the measure for non-Markovianity (37) may be evaluated for any dynamical process, showing its great benefit. Similarly, one could obtain the equivalence of CP divisibility with Markovian behavior if the dilated process $\Phi \otimes \mathbb{1}_{\mathcal{H}}$ is considered as shown in Ref. [13]. However, P divisible quantum processes offer the relevant feature of a distinct connection to classical Markovian stochastic processes. To show this, we employ the characterization of P divisible processes given in Theorem 1. Let $\rho(t)$ be the solution of the master equation

$$
\frac{d}{dt}\rho(t) = \mathcal{K}_t\rho(t),
\tag{24}
$$

with initial state $\rho(0)$ where $\mathcal{K}_t$ is given by Eq. (13). The time-evolved state admits an instantaneous spectral decomposition

$$
\rho(t) = \sum_m p_m(t)|\phi_m(t)\rangle\langle\phi_m(t)|,
\tag{25}
$$

so $\{|\phi_m(t)\rangle\}$ defines an orthonormal basis on $\mathcal{H}$ and $\{p_m(t)\}$ represents a classical probability distribution for all $t \geqslant 0$. By virtue of the orthonormality, the eigenvalues $p_m(t) = \langle\phi_m(t)|\rho(t)|\phi_m(t)\rangle$ obey the following closed differential equation:

$$
\frac{d}{dt}p_m(t) = \sum_n [W_{mn}(t)p_n(t) - W_{nm}(t)p_m(t)],
\tag{26}
$$

where

$$
W_{mn}(t) = \sum_j \gamma_j(t)|\langle\phi_m(t)|A_j(t)|\phi_n(t)\rangle|^2.
\tag{27}
$$

Obviously, the term in Eq. (26) with $m = n$ drops out. Given any solution of a quantum master equation, one thus obtains a classical jump process. We emphasize that the rates $W_{mn}(t)$ depend in general on the initial eigenbasis $\{|\phi_m(0)\rangle\}$ as well as the initial probability distribution $\{p_m(0)\}$.

According to the theory of classical stochastic processes, any hierarchy of n-point probability distributions $P_n(y_n,t_n;\ldots;y_1,t_1)$ with discrete sample space Ω satisfying the consistency relations

$$
P_n(y_n,t_n;\ldots;y_1,t_1) \geqslant 0,
\tag{28}
$$

$$
\sum_{y_1 \in \Omega} P_1(y_1,t_1) = 1,
\tag{29}
$$

$$
\sum_{y_m \in \Omega} P_n(y_n,t_n;\ldots;y_m,t_m;\ldots;y_1,t_1) = P_{n-1}(y_n,t_n;\ldots;y_1,t_1)
\tag{30}
$$

for all $1 \leqslant m \leqslant n$ and $0 \leqslant t_1 < t_2 < \cdots < t_n$ determines a stochastic process $\mathcal{Y}(t)$ with values in Ω [29–31]. Here $P_n(y_n,t_n;\ldots;y_1,t_1)$ gives the probability to observe the values y_i at times t_i for $i = 1,\ldots,n$ for the stochastic process $\mathcal{Y}(t)$.

For classical Markov processes obeying the relation

$$P_{1|n-1}(y_n,t_n|y_{n-1},t_{n-1};\ldots;y_1,t_1)$$
$$= P_{1|1}(y_n,t_n|y_{n-1},t_{n-1}) \tag{31}$$

for any conditional probability distribution $P_{1|n-1}$, the full hierarchy is completely determined by the one-point probability distribution P_1 and the so-called transition probability $P_{1|1}$. However, these two non-negative functions cannot be chosen arbitrarily but must satisfy

$$P_1(y_2,t_2) = \sum_{y_1 \in \Omega} P_{1|1}(y_2,t_2|y_1,t_1) P_1(y_1,t_1), \tag{32}$$

$$P_{1|1}(y_3,t_3|y_1,t_1) = \sum_{y_2 \in \Omega} P_{1|1}(y_3,t_3|y_2,t_2) P_{1|1}(y_2,t_2|y_1,t_1),$$
$$\tag{33}$$

where Eq. (33) is called the Chapman-Kolmogorov equation. These equations thus characterize uniquely a Markov process.

One may equivalently write this relation for the transition probability in its differential form

$$\frac{d}{dt} P_{1|1}(y,t|x,s)$$
$$= \sum_{z \in \Omega} [\mathcal{W}_{yz}(t) P_{1|1}(z,t|x,s) - \mathcal{W}_{zy}(t) P_{1|1}(y,t|x,s)], \tag{34}$$

where the transition probability per unit time $\mathcal{W}_{yz}(t)$ is non-negative and represents the probability for a transition to y given the classical state was z at time t. It is clear that the (differential) Chapman-Kolmogorov equation characterizes the transition probability $P_{1|1}$ but a similar equation, the so-called Pauli master equation, can be derived for the one-point probability distribution, that is,

$$\frac{d}{dt} P_1(y,t) = \sum_{z \in \Omega} [\mathcal{W}_{yz}(t) P_1(z,t) - \mathcal{W}_{zy}(t) P_1(y,t)]. \tag{35}$$

One thus concludes that Eq. (26) can be interpreted as a Pauli master equation for the one-point probability distribution of a classical Markov process with $\Omega = \{1,\ldots,\dim\mathcal{H}\}$ if and only if

$$W_{mn}(t) = \sum_j \gamma_j(t) |\langle \phi_m(t)|A_j(t)|\phi_n(t)\rangle|^2 \geqslant 0 \tag{36}$$

for all $t \geqslant 0$ and $m \neq n$. Hence, P divisibility of the quantum process is a sufficient condition to warrant positivity of the rates $W_{mn}(t)$ [see Eq. (15)]. This shows that quantum non-Markovianity defined with respect to the generalized trace-distance-based measure not only provides an interpretation in terms of an information backflow but also allows for a connection to classical Markovian stochastic processes. Namely, to each P divisible quantum process, given as the solution of a master equation of the form (13) with Lindblad operators and rates satisfying Eq. (15) for an initial state $\rho(0)$, one associates a classical Markovian process obtained as the solution of the classical master equation (26) with transition rates given by Eq. (27) and initial condition specified by the eigenvalues of $\rho(0)$.

Finally, we note that P divisibility of the quantum process is indeed equivalent to the positivity of the rates $W_{mn}(t)$ of the classical Pauli master equation if the quantum master equation has the property that the eigenbases $\{|\phi_n(t)\rangle\}$ of $\rho(t)$ run over all orthonormal bases when varying the initial state $\rho(0)$. This is satisfied if the maximally mixed state is in the image of the dynamical map Φ_t, which holds true for sufficiently small times due to continuity of the process and as we have $\Phi_0 = \mathbb{1}$. For two-level systems it is shown in Appendix A that this constraint is not only sufficient but also necessary. If any orthonormal basis is encountered, all classical processes derived from the quantum master equation are classically Markovian if and only if the quantum dynamics is P divisible.

III. GENERALIZATION OF THE TRACE-DISTANCE-BASED NON-MARKOVIANITY MEASURE

Having stated the generalized definition for quantum non-Markovianity, which allows for a connection to classical Markov processes, we now consider the expression of the corresponding measure and address its mathematical and physical features. According to the definition of a quantum Markovian process given in Sec. II A, it is natural to consider the following measure for quantum non-Markovianity, quantifying the degree of memory effects with respect to this generalized definition [21]:

$$\mathcal{N}(\Phi) \equiv \max_{\{p_i,\rho_i\}} \int_{\sigma>0} dt\, \sigma(t,p_i,\rho_i) \tag{37}$$

with

$$\sigma(t,p_i,\rho_i) \equiv \frac{d}{dt} \|p_1 \Phi_t(\rho_1) - p_2 \Phi_t(\rho_2)\|_1, \tag{38}$$

where the integration runs over all intervals where the distinguishability $\|p_1 \Phi_t(\rho_1) - p_2 \Phi_t(\rho_2)\|_1$ increases. A process Φ is said to be non-Markovian if $\mathcal{N}(\Phi) > 0$. As discussed in Sec. II A, this measure still admits the interpretation as a quantifier for the information flow from the environment back to the open system. In addition, as discussed in Sec. II B, Markovianity will now be equivalent to P divisibility of a quantum process. Note that this feature was only sufficient in the original definition (see Sec. III B).

To begin with, we concentrate on the maximization procedure contained in the quantifier for non-Markovianity (37), which, contrary to the original definition [9], now even requires one to sample also over binary probability distributions. Fortunately, a similar characterization of pairs of states maximizing Eq. (37) can be proven [15] and yet also a local representation is admitted [14], simplifying the sampling significantly. We finally illustrate the present definition of non-Markovianity by means of examples that show the difference from the original characterization and other approaches to non-Markovianity.

A. Expressions of the generalized non-Markovianity measure

A set $\{p_i,\rho_{1,2}\}$, where $\rho_{1,2} \in \mathcal{S}(\mathcal{H})$ and $\{p_i\}$ defines a binary probability distribution, is said to be optimal if the maximum in Eq. (37) is attained for it. Note that the quantum states of an optimal set are necessarily nonequal as the dynamics of the norm of a nonindefinite Helstrom matrix under any dynamical map is trivial due to trace preservation and positivity of the map.

Theorem 3. The states of an optimal set must be orthogonal, i.e.,

$$\mathcal{N}(\Phi) = \max_{\{p_i, \rho_1 \perp \rho_2\}} \int_{\sigma > 0} dt\, \sigma(t, p_i, \rho_i). \tag{39}$$

Proof. Suppose $\{p_i, \rho_{1,2}\}$ is optimal with $\rho_1 \not\perp \rho_2$. Along the lines of the proof of Lemma 1 we obtain a probability distribution $\{q_i\}$ and two orthogonal states $\varrho_{1,2}$ such that

$$p_1 \rho_1 - p_2 \rho_2 = \lambda(q_1 \varrho_1 - q_2 \varrho_2), \tag{40}$$

where $\lambda = \mathrm{Tr}|p_1 \rho_1 - p_2 \rho_2|$. As $\rho_1 \not\perp \rho_2$, we have $0 < \lambda < 1$ according to Eq. (3). Here $\lambda > 0$ holds as states of an optimal set are nonequal by definition. By means of linearity of the dynamical maps Φ_t and homogeneity of the trace norm one finally obtains

$$\|q_1 \varrho_1(t) - q_2 \varrho_2(t)\|_1 = \frac{1}{\lambda} \|p_1 \rho_1(t) - p_2 \rho_2(t)\|_1 \tag{41}$$

for all $t \geq 0$ where $\lambda^{-1} > 1$. This shows that any increase of $\|p_1 \rho_1(t) - p_2 \rho_2(t)\|_1$ is exceeded by the increase of $\|q_1 \varrho_1(t) - q_2 \varrho_2(t)\|_1$. Hence, $\{q_i, \varrho_{1,2}\}$ yields a non-Markovianity strictly larger than the set $\{p_i, \rho_{1,2}\}$ contradicting its optimality. ∎

We define $\mathring{\mathcal{S}}(\mathcal{H})$ to be the interior of the state space, i.e., the set of all quantum states ρ for which there is an $\epsilon > 0$ such that all Hermitian operators X with unit trace satisfying $||\rho - X||_1 \leq \epsilon$ belong to $\mathcal{S}(\mathcal{H})$. Hence, all states in the interior have full rank. Note that $\mathring{\mathcal{S}}(\mathcal{H}) = \emptyset$ if $\dim \mathcal{H} = \infty$, which implies that the local representation of the measure (37) introduced below is not available in infinite dimensions. However, these quantum systems can frequently be accurately described by finite-dimensional Hilbert spaces.

Based on the orthogonality of states of an optimal set, we can also establish a local representation for the generalized measure as provided for the original definition in Ref. [14]. To this end, one first has to prove a characterization of enclosing surfaces, which are defined as follows (see Fig. 1): A set $\partial U(\rho) \subset \mathcal{S}(\mathcal{H})$ not containing $\rho \in \mathring{\mathcal{S}}(\mathcal{H})$ is called an enclosing surface of ρ if and only if for any nonzero, Hermitian, and traceless operator Y there exists a real number $\mu > 0$ such that

$$\rho + 2\mu Y \in \partial U(\rho). \tag{42}$$

Lemma 2. Let $\partial U(\rho)$ be an enclosing surface and X a nonzero, Hermitian, and indefinite operator. Then there exists
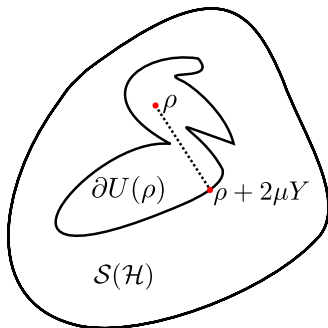

FIG. 1. (Color online) Illustration of an enclosing surface $\partial U(\rho)$ of an inner point ρ of the state space $\mathcal{S}(\mathcal{H})$.

a real number $\lambda > 0$ with $\lambda|\mathrm{Tr}X| < 1$ such that

$$\frac{1}{p_-}[p_+ \rho - \mathrm{sgn}(\mathrm{Tr}X)\lambda X] \in \partial U(\rho), \tag{43}$$

where $p_\pm \equiv \frac{1}{2}(1 \pm \lambda|\mathrm{Tr}X|)$ and

$$\mathrm{sgn}(x) = \begin{cases} -1 & \text{for } x \leq 0 \\ +1 & \text{otherwise.} \end{cases} \tag{44}$$

Proof. Let $\partial U(\rho)$ be an enclosing surface of ρ. Consider a nonzero, Hermitian, indefinite operator X. The operator

$$Y = \mathrm{sgn}(\mathrm{Tr}X)[(\mathrm{Tr}X)\rho - X] \tag{45}$$

defines a nonzero, Hermitian, and traceless operator. By definition, there exists a real number $\mu > 0$ such that

$$\varrho = \rho + 2\mu Y \in \partial U(\rho). \tag{46}$$

We then define a real number λ by means of

$$\mu = \frac{\lambda}{1 - \lambda|\mathrm{Tr}X|} \tag{47}$$

so that $\lambda = \mu/(1 + \mu|\mathrm{Tr}X|) > 0$ holds and hence $\lambda|\mathrm{Tr}X| < 1$. For $p_\pm = \frac{1}{2}(1 \pm \lambda|\mathrm{Tr}X|)$ one finally obtains

$$\begin{aligned} \varrho &= \rho + \frac{2\lambda}{1 - \lambda|\mathrm{Tr}X|}\{\mathrm{sgn}(\mathrm{Tr}X)[(\mathrm{Tr}X)\rho - X]\} \\ &= \rho - \frac{\lambda}{p_-}\mathrm{sgn}(\mathrm{Tr}X)X + \frac{p_+ - p_-}{p_-}\rho \\ &= \frac{1}{p_-}[p_+ \rho - \mathrm{sgn}(\mathrm{Tr}X)\lambda X], \end{aligned} \tag{48}$$

which is Eq. (43). ∎

As proven in Appendix B, Lemma 2 can actually be augmented to show that the characterization (43) of an enclosing surface is equivalent to Eq. (42).

Theorem 4. The generalized measure of quantum non-Markovianity admits a local representation, i.e.,

$$\mathcal{N}(\Phi) = \max_{\{p_i\}, \rho_2 \in \partial U(\rho_1)} \int_{\tilde{\sigma} > 0} dt\, \tilde{\sigma}(t, p_i, \rho_i) \tag{49}$$

with

$$\tilde{\sigma}(t, p_i, \rho_i) \equiv \frac{1}{\|p_1 \rho_1 - p_2 \rho_2\|_1} \frac{d}{dt}\|p_1 \Phi_t(\rho_1) - p_2 \Phi_t(\rho_2)\|_1, \tag{50}$$

where $\rho_1 \in \mathring{\mathcal{S}}(\mathcal{H})$ is any fixed inner point of the state space and $\partial U(\rho_1)$ refers to an arbitrary enclosing surface of ρ_1.

We emphasize that $\tilde{\sigma}(t, p_i, \rho_i)$ is well defined as $0 < \|p_1 \rho_1 - p_2 \rho_2\|_1$ for any state $\rho_2 \in \partial U(\rho_1)$ and probability distribution $\{p_i\}$ as by definition of an enclosing surface we have $\rho_2 \neq \rho_1$ [cf. Eq. (3)].

Proof. Let $\partial U(\rho_1)$ be an enclosing surface of ρ_1. To prove that the corresponding local representation yields a value smaller than or equal to the original definition (39) one follows the lines of the proof of Theorem 3. Let $\rho_2 \in \partial U(\rho_1)$ and a probability distribution $\{p_i\}$ be given. Now, following the proof of Lemma 1 there exist two states $\varrho_1 \perp \varrho_2$ and a probability distribution $\{q_i\}$ such that

$$p_1 \rho_1 - p_2 \rho_2 = \lambda(q_1 \varrho_1 - q_2 \varrho_2), \tag{51}$$

where $\lambda = \|p_1\rho_1 - p_2\rho_2\|_1 > 0$. We then obtain

$$\tilde{\sigma}(t, p_i, \rho_i) = \sigma(t, q_i, \varrho_i) \tag{52}$$

for all $t \geqslant 0$ due to linearity of Φ_t and homogeneity of the trace norm and the derivative. We conclude that the right-hand side of Eq. (49) is smaller than or equal to $\mathcal{N}(\Phi)$ as defined in Eq. (39).

Conversely, let $\varrho_1 \perp \varrho_2$ be two orthogonal states and denote by $\{q_i\}$ a probability distribution. Then $\Delta = q_1\varrho_1 - q_2\varrho_2$ defines a nonzero, Hermitian, indefinite operator. Thus, according to Lemma 2, there exists a real number $\lambda > 0$ with $\lambda|\mathrm{Tr}\Delta| < 1$ such that

$$p_+\rho_1 - p_-\rho_2 = c\lambda\Delta, \tag{53}$$

where $c = \mathrm{sgn}(\mathrm{Tr}\Delta)$ and $p_\pm = \frac{1}{2}(1 \pm \lambda|\mathrm{Tr}\Delta|)$ for some quantum state $\rho_2 \in \partial U(\rho_1)$ of the enclosing surface. As $|c| = 1$ it follows that $\|p_+\rho_1 - p_-\rho_2\|_1 = \lambda > 0$. Linearity of the dynamical map and homogeneity of the trace norm and the derivative then yields

$$\tilde{\sigma}(t, p_\pm, \rho_i) = \sigma(t, q_i, \varrho_i), \tag{54}$$

showing that the original definition (39) leads to a value that is smaller than or equal to the right-hand side of Eq. (49). This thus concludes the proof, that is, the maximization over an enclosing surface with an information flux rescaled by the initial distinguishability reproduces the dynamics of the trace norm of orthogonal states. ∎

A careful reformulation of an enclosing surface thus allows us to establish an equivalent local and universal representation of the generalized trace-distance-based measure as for the original definition [14]. This shows that also for the present characterization, non-Markovianity is a universal feature appearing everywhere in state space. Similarly, it suffices if Eq. (42) holds for exactly one $\mu > 0$ as λ is uniquely determined by this parameter given a nonzero, Hermitian, and indefinite operator (cf. proof of Lemma 2). In addition, also here, no assumption on the enclosing surface concerning, for example, the shape or the smoothness is needed, implying a great benefit for the analytical, numerical, and experimental determination of the generalized measure.

B. Examples

As already stated before, the fundamental difference between the trace-distance-based measure for non-Markovianity and its generalization is that in the trace-distance approach P divisibility is only a sufficient criterion for Markovianity. This fact becomes particularly apparent when uniform translations of states are encountered in the dynamics that do not describe positive maps but leave the trace distance unchanged [32]. However, by choosing unequal weights p_i, the Helstrom matrices are no longer invariant under such operations, thus making it possible to detect the effect of such maps.

To illustrate this property, we consider a two-level open quantum system, i.e., $\mathcal{H} = \mathbb{C}^2$, undergoing a dynamics that, as depicted in Fig. 2, can be described in the Bloch sphere as an isotropic contraction followed by a translation along the z axis. The dynamics is analytically described by a time-local

master equation (24) where the generator $\mathcal{K}_t$ obeys

$$\mathcal{K}_t\rho(t) = \sum_{j=1}^3 \frac{\gamma(t)}{4}[\sigma_j\rho(t)\sigma_j - \rho(t)] \tag{55}$$

for $0 \leqslant t \leqslant t_1$ with $t_1 < T$ and

$$\mathcal{K}_t\rho(t) = \frac{-b(t)}{2}\left[\sigma_-\rho(t)\sigma_+ - \frac{1}{2}\{\sigma_+\sigma_-, \rho(t)\}\right]$$
$$+ \frac{b(t)}{2}\left[\sigma_+\rho(t)\sigma_- - \frac{1}{2}\{\sigma_-\sigma_+, \rho(t)\}\right] \tag{56}$$

if $t_1 < t \leqslant T$ holds. Here σ_j refers to the Pauli matrices and $\sigma_{+(-)}$ describes the usual raising (lowering) operator with respect to the eigenstates $|\pm\rangle$ of σ_3. Moreover, the rates satisfy $\gamma(t), b(t) \geqslant 0$ for all t in their respective domains, so the second phase of the process is neither CP nor P divisible, which can be easily shown using Theorem 1. This dynamical process exhibits the essential feature of the generalized amplitude damping channel studied in Ref. [32].

By employing the Bloch vector representation for two-level systems, the master equation is equivalently described by a differential equation for the Bloch vector

$$\frac{d}{dt}\vec{v}(t) = \begin{cases} A(t)\vec{v}(t), & 0 \leqslant t \leqslant t_1 < T \\ \vec{b}(t), & t_1 < t \leqslant T, \end{cases} \tag{57}$$

where $A(t) = \mathrm{diag}[-\gamma(t), -\gamma(t), -\gamma(t)]$ and $\vec{b}(t) = [0, 0, b(t)]^T$. Hence, the process's first phase corresponds to an isotropic contraction of the Bloch sphere $B_1 = \{\vec{v} \mid |\vec{v}| \leqslant 1\}$ to the smaller sphere $B_r = \{\vec{v} \mid |\vec{v}| \leqslant r\}$ with radius

$$r = \exp\left[-\int_0^{t_1} dt\, \gamma(t)\right] \in (0, 1), \tag{58}$$

which is clearly CP divisible and therefore Markovian. On the other hand, the second phase describes a uniform translation of the Bloch sphere along the z axis, i.e., $\vec{v}(t_1) \mapsto \vec{v}(t_1) + (0, 0, a)^T$ with

$$a = \int_{t_1}^T dt\, b(t) > 0, \tag{59}$$

where we thus have to require $a \leqslant 1 - r$ in order to maintain positivity of the dynamical map (see Fig. 2). It is easily shown that this condition is also necessary and sufficient for complete positivity of the process.

The trace norm of the Helstrom matrix $\Delta = p_1\rho_1 - p_2\rho_2$ at time t for two quantum states $\rho_{1,2} = \frac{1}{2}(\mathbb{1}_2 + \vec{v}_{1,2} \cdot \vec{\sigma})$, evolving
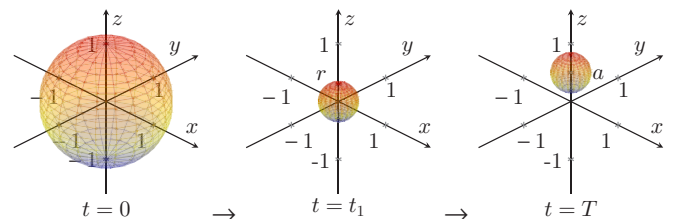

FIG. 2. (Color online) Bloch sphere representation of the action of the dynamical map Φ_t corresponding to the time-local generator $\mathcal{K}_t$ defined by Eqs. (55) and (56) for $t = 0$, t_1, and T.

according to the described dynamical map, is given by

$$\|\Delta(t)\|_1 = \tfrac{1}{2}\{|p_1 - p_2 + |\vec{w}(t)|| + |p_1 - p_2 - |\vec{w}(t)||\}, \tag{60}$$

where $\vec{w}(t) = p_1 \vec{v}_1(t) - p_2 \vec{v}_2(t)$. Clearly, if $p_1 = p_2 = 1/2$, then $\|\Delta(t)\|_1 = \tfrac{1}{2}|\vec{v}_1(t) - \vec{v}_2(t)|$, showing that the unbiased case is unable to detect the non-Markovianity of the process resulting from the uniform translation.

Without restriction we may assume $p_1 \geqslant p_2$ so that

$$\|\Delta(t)\|_1 = \begin{cases} p_1 - p_2 & \text{for } p_1 - p_2 > |\vec{w}(t)| \\ |\vec{w}(t)| & \text{for } p_1 - p_2 \leqslant |\vec{w}(t)| \end{cases} \tag{61}$$

and, due to Theorem 3, we may restrict ourselves to orthogonal states corresponding to antipodal unit vectors. That is, $\vec{v}_1 = -\vec{v}_2$ with $|\vec{v}_1| = 1$, which implies $\vec{w}(t) = \vec{v}_1(t)$.

If the probability distribution is such that $p_1 - p_2 > |\vec{v}_1(t)|$ holds for all $t_1 < t \leqslant T$, then $\|\Delta(t)\|_1$ is a monotonically decreasing function of t. However, if $p_1 - p_2 < |\vec{v}_1(T)|$, then

$$\int_{\sigma > 0} dt\, \sigma(t) = |\vec{v}_1(T)| - (p_1 - p_2) > 0, \tag{62}$$

indicating non-Markovianity. The change of the trace norm thus increases with decreasing difference $p_1 - p_2$, which is bounded by $p_1 - p_2 = r = |\vec{v}_1(t_1)|$. Calculating explicitly the change of the trace norm for $p_1 - p_2 \leqslant r$ one finds

$$\|\Delta(T)\|_1 - \|\Delta(t_1)\|_1 = |\vec{v}_1(t_1) + (p_1 - p_2)\vec{a}| - |\vec{v}_1(t_1)|, \tag{63}$$

which shows that the trace norm attains its maximal value

$$\|\Delta(T)\|_1 - \|\Delta(t_1)\|_1 = r|\vec{a}| \tag{64}$$

if $\vec{v}_1$ is parallel to $\vec{a}$, i.e., $\vec{v}_1 = c\,\vec{a}$ for some $c > 0$, and $p_1 - p_2 = r$. Obviously, this is satisfied by the probability distribution $\{p_i\} = \{p_\pm = \tfrac{1}{2}(1 \pm r)\}$. Unlike the original definition, the generalized trace-distance-based measure is thus able to capture non-Markovian dynamics arising from uniform translations contained in the dynamical process.

The dynamics generated by $\mathcal{K}_t$ for $0 \leqslant t \leqslant t_1$ [cf. Eq. (55)] for any $t_1 > 0$ provides also an example for a process that is not CP but P divisible. Choosing $\gamma_1(t) = \gamma_2(t) = 1$ and $\gamma_3(t) = -\tanh(t)$ as proposed in Ref. [33], the dynamical process is not CP divisible according to Theorem 1. In particular, not even one interval exists for which CP divisibility is restored as $\gamma_3(t) < 0$ for all $t > 0$. A further example of this property was constructed in Ref. [28]. However, the dynamics is always P divisible since

$$\sum_{i=1}^{2} |\langle m|\sigma_i|n\rangle|^2 - \tanh(t)|\langle m|\sigma_3|n\rangle|^2 \geqslant 0 \tag{65}$$

is valid for all $0 \leqslant t \leqslant t_1$ and $m \neq n$, which is condition (15). Hence, the trace-distance-based measure and its generalization are equal to zero in this case, while measures for non-Markovianity relying on CP divisibility [5,13,33] are of course nonvanishing. This random unitary evolution [34,35] illustrates the persisting and significant difference between the two major approaches for the characterization of quantum non-Markovianity.

IV. CONCLUSION AND OUTLOOK

We have introduced a generalization of the criterion of quantum non-Markovianity based on the flow of information. This characterization relies on the trace norm of Helstrom matrices, which can also be interpreted as a measure for the distinguishability. By virtue of this property, the generalized measure still admits an interpretation as quantifier of an information backflow from the environment to the open quantum system.

It was shown that the generalized criterion is equivalent to P divisibility of the dynamical process, which has an explicit connection to classical Markovian stochastic processes. That is, any rate equation obtained from a quantum master equation of a P divisible process can be interpreted as the Pauli master equation of a classical Markov process. However, the presented approach is more general since it can also be applied even when the notion of divisibility is ill defined and can be experimentally tested.

The experimental determination is substantially simplified by the derived mathematical representations of the generalized measure, which are similar to those for the original definition. First, we have demonstrated that optimal initial states for non-Markovian dynamics must be orthogonal and, based on this result, we could finally establish a local representation for the measure. Hence, orthogonal states, corresponding to a maximal information content, exhibit maximal memory effects that can be revealed locally and anywhere in the quantum state space as provided by the local representation.

An essential feature of the generalized approach to non-Markovianity in comparison with the original definition is its sensitivity to memory effects arising from uniform translations of states. To illustrate this, we constructed a dynamical process for a two-level system that comprises a uniform translation of the Bloch sphere. However, there exist dynamical processes that are not CP divisible but P divisible, as we have shown in our second example, which manifests the existing difference of the generalized definition with other approaches to non-Markovianity.

We believe that the definition of non-Markovianity given in this paper is of great relevance for the study of memory effects in the field of complex quantum systems and quantum information due to the experimental accessibility and its clear-cut interpretation and connection to classical Markov processes.

APPENDIX A: ORTHONORMAL BASES AND THE MAXIMALLY MIXED STATE

We prove the statement made in Sec. II C about the relation between the maximally mixed state and the eigenbases of

the set of time-evolved states $\mathrm{Im}\Phi_t = \{\Phi_t(\rho)|\rho \in \mathcal{S}(\mathcal{H})\}$ for two-level systems, i.e., $\mathcal{H} = \mathbb{C}^2$. Note that $\mathrm{Im}\Phi_t$ defines a nonempty, convex, and compact set for any finite-dimensional Hilbert space due to linearity of Φ_t along with compactness and convexity of the state space.

Lemma 3. Any orthonormal basis $\{|\Psi_i\rangle\}$ of $\mathcal{H}$ represents the eigenbasis of a quantum state $\rho(t) \in \mathrm{Im}\Phi_t$ if and only if $\frac{1}{2}\mathbb{1}_2 \in \mathrm{Im}\Phi_t$.

The "if" statement is clear from the fact that any orthonormal basis defines a resolution of identity. Hence, if $\frac{1}{2}\mathbb{1}_2 \in \mathrm{Im}\Phi_t$ then any basis $\{|\Psi_i\rangle\}$ defines at least the eigenbasis of the maximally mixed state, proving the claim. It is clear that the same reasoning actually applies to any finite-dimensional Hilbert space.

To show the reverse we employ the Hahn-Banach separation theorem for normed vector spaces. Suppose that $\frac{1}{2}\mathbb{1}_2 \notin \mathrm{Im}\Phi_t$. Then there exists a real-valued, linear, and continuous functional φ_A separating the disjoint nonempty, convex, and compact sets $\mathrm{Im}\Phi_t$ and $\{\frac{1}{2}\mathbb{1}_2\}$. That is, we have

$$\varphi_A\left(\tfrac{1}{2}\mathbb{1}_2\right) < \inf\{\varphi_A(\rho(t))|\rho(t) \in \mathrm{Im}\Phi_t\}, \tag{A1}$$

where $\varphi_A(X) = \mathrm{Tr}\{AX\}$ for some Hermitian operator A due to the Riesz representation theorem. By means of the transformation $A \mapsto A - \mathrm{Tr}A$ one may also assume that the operator A is traceless. The set $\{X|X = X^\dagger, \mathrm{Tr}\{X\} = 1, \varphi_A(X) = 0\}$ thus describes a hyperplane that intersects the maximally mixed state and separates it from $\mathrm{Im}\Phi_t$.

Employing the Bloch representation, this hyperplane defines an ordinary plane in $\mathbb{R}^3$ that contains the origin and therefore intersects the surface of the Bloch sphere in a unit circle. As pure, orthogonal states on $\mathbb{C}^2$ correspond to antipodal points on the surface of the Bloch sphere, we thus conclude that the (hyper)plane contains an infinite set of orthonormal bases that do not represent eigenbases of quantum states $\rho(t) \in \mathrm{Im}\Phi_t$ (cf. Fig. 3). ∎

APPENDIX B: CHARACTERIZATIONS OF ENCLOSING SURFACES

In this section we show that the relation for enclosing surfaces in terms of nonzero, Hermitian, and indefinite operators derived in Lemma 2 is equivalent to the original definition, i.e., the following lemma holds.

Lemma 4. A set $\partial U(\rho)$ defines an enclosing surface of ρ if and only if for any nonzero, Hermitian, indefinite operator X there exists a real number $\lambda > 0$ with $\lambda|\mathrm{Tr}X| < 1$ such that

$$\frac{1}{p_-}[p_+\rho - \mathrm{sgn}(\mathrm{Tr}X)\lambda X] \in \partial U(\rho), \tag{B1}$$

where $p_\pm \equiv \frac{1}{2}(1 \pm \lambda|\mathrm{Tr}X|)$ and

$$\mathrm{sgn}(x) = \begin{cases} -1 & \text{for } x \leqslant 0 \\ +1 & \text{otherwise.} \end{cases} \tag{B2}$$
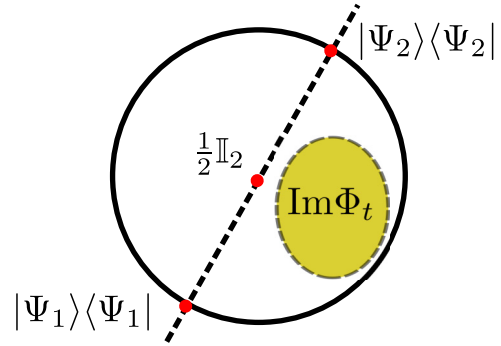

FIG. 3. (Color online) Two-dimensional cut of a hyperplane in the Bloch representation separating the two disjoint nonempty, convex, and compact sets $\{\frac{1}{2}\mathbb{1}_2\}$ and $\mathrm{Im}\Phi_t$. As the plane intersects the maximally mixed state its intersection with the state space comprises an infinite set of pure and orthogonal states defining orthonormal bases of $\mathbb{C}^2$.

Proof. That any enclosing surface obeys a characterization in terms of nonzero, Hermitian, and indefinite operators has already been proven in Lemma 2. Conversely, suppose the states in $\partial U(\rho)$ are characterized by Eq. (B1). Hence, for a nonzero, indefinite, and Hermitian operator X there exists a real number $\lambda > 0$ with $\lambda|\mathrm{Tr}X| < 1$ such that

$$\varrho = \frac{1}{p_-}[p_+\rho - \mathrm{sgn}(\mathrm{Tr}X)\lambda X] \in \partial U(\rho), \tag{B3}$$

where $p_\pm = \frac{1}{2}(1 \pm \lambda|\mathrm{Tr}X|)$. Now consider the map

$$\Theta_\rho(X) \equiv \mathrm{sgn}(\mathrm{Tr}X)[(\mathrm{Tr}X)\rho - X] \tag{B4}$$

defined on the set of nonzero, Hermitian, and indefinite operators. The operator $Y = \Theta_\rho(X)$ represents a traceless Hermitian operator and $Y = 0$ if and only if $(\mathrm{Tr}X)\rho = X$, contradicting that X is indefinite. Thus, we have $Y \neq 0$ and one finds

$$\begin{aligned} \varrho &= \frac{1}{p_-}\{p_+\rho - \lambda[|\mathrm{Tr}X|\rho - Y]\} \\ &= \frac{p_+ - \lambda|\mathrm{Tr}X|}{p_-}\rho + \frac{\lambda}{p_-}Y \\ &= \rho + 2\mu Y, \end{aligned} \tag{B5}$$

where $\mu \equiv \lambda/2p_- > 0$. It remains to show that Θ_ρ defines a surjection on the set of nonzero, Hermitian, and traceless operators. This is obviously true as any traceless, nonzero, Hermitian operator Y is necessarily indefinite and we have $\Theta_\rho(Y) = Y$. Hence, the set $\partial U(\rho)$ is indeed an enclosing surface. ∎

[1] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2007).

[2] G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).

[3] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, J. Math. Phys. **17**, 821 (1976).

[4] M. M. Wolf, J. Eisert, T. S. Cubitt, and J. I. Cirac, Phys. Rev. Lett. **101**, 150402 (2008).

[5] Á. Rivas, S. F. Huelga, and M. B. Plenio, Phys. Rev. Lett. **105**, 050403 (2010).

[6] S. Luo, S. Fu, and H. Song, Phys. Rev. A **86**, 044101 (2012).

[7] X.-M. Lu, X. Wang, and C. P. Sun, Phys. Rev. A **82**, 042103 (2010).

[8] Á. Rivas, S. F. Huelga, and M. B. Plenio, Rep. Prog. Phys. **77**, 094001 (2014).

[9] H.-P. Breuer, E.-M. Laine, and J. Piilo, Phys. Rev. Lett. **103**, 210401 (2009).

[10] E.-M. Laine, J. Piilo, and H.-P. Breuer, Phys. Rev. A **81**, 062115 (2010).

[11] H.-P. Breuer, J. Phys. B **45**, 154001 (2012).

[12] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[13] D. Chruściński, A. Kossakowski, and Á. Rivas, Phys. Rev. A **83**, 052128 (2011).

[14] B.-H. Liu, S. Wißmann, X.-M. Hu, C. Zhang, Y.-F. Huang, C.-F. Li, G.-C. Guo, A. Karlsson, J. Piilo, and H.-P. Breuer, Sci. Rep. **4**, 6327 (2014).

[15] S. Wißmann, A. Karlsson, E.-M. Laine, J. Piilo, and H.-P. Breuer, Phys. Rev. A **86**, 062108 (2012).

[16] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[17] M. Hayashi, *Quantum Information* (Springer, Berlin, 2006).

[18] A. S. Holevo, Trans. Moscow Math. Soc. **26**, 133 (1972).

[19] A. Kossakowski, Bull. Acad. Pol. Sci., Série Math. Astr. Phys. **20**, 1021 (1972).

[20] A. Kossakowski, Rep. Math. Phys. **3**, 247 (1972).

[21] H.-P. Breuer, E.-M. Laine, J. Piilo, and B. Vacchini, arXiv:1505.01385.

[22] E. M. Laine, J. Piilo, and H. P. Breuer, Europhys. Lett. **92**, 60010 (2010).

[23] D. Chruściński and S. Maniscalco, Phys. Rev. Lett. **112**, 120404 (2014).

[24] B. Blackadar, *Operator Algebras* (Springer, Berlin, 2005).

[25] M. D. Choi, Can. J. Math. **24**, 520 (1972).

[26] W. F. Stinespring, Proc. Amer. Math. Soc. **6**, 211 (1955).

[27] M. D. Choi, Linear Algebra Appl. **10**, 285 (1975).

[28] B. Vacchini, A. Smirne, E.-M. Laine, J. Piilo, and H.-P. Breuer, New J. Phys. **13**, 093004 (2011).

[29] A. N. Kolmogorov, *Foundations of the Theory of Probability* (Chelsea, New York, 1950).

[30] C. W. Gardiner, *Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences*, Springer Series in Synergetics Vol. 13 (Springer, Berlin, 1985).

[31] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (North-Holland, Amsterdam, 2007).

[32] J. Liu, X.-M. Lu, and X. Wang, Phys. Rev. A **87**, 042103 (2013).

[33] M. J. W. Hall, J. D. Cresser, L. Li, and E. Andersson, Phys. Rev. A **89**, 042120 (2014).

[34] B. Vacchini, J. Phys. B **45**, 154007 (2012).

[35] D. Chruściński and F. Wudarski, Phys. Lett. A **377**, 1425 (2013).