Witnessing non-Markovian dynamics through correlations

Dario De Santis\textsuperscript{1}, Markus Johansson\textsuperscript{1}, Bogna Bylicka\textsuperscript{1}, Nadja K. Bernardes\textsuperscript{2,3}, and Antonio Acín\textsuperscript{1,4}

\textsuperscript{1}ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain
\textsuperscript{2} Departamento de Física, Universidade Federal de Minas Gerais, Belo Horizonte, Caixa Postal 702, 30161-970, Brazil
\textsuperscript{3} Departamento de Física, Universidade Federal de Pernambuco, 50670-901 Recife-PE, Brazil
\textsuperscript{4}ICREA–Institució Catalana de Recerca i Estudis Avançats, Lluis Companys 23, 08010 Barcelona, Spain

(Dated: April 1, 2019)

Non-Markovian effects in an open-system dynamics are usually associated to information backflows from the environment to the system. However, the way these backflows manifest and how to detect them is unclear. A natural approach is to study the backflow in terms of the correlations evolving system displays with another unperturbed system during the dynamics. In this work, we study the power of this approach to witness non-Markovian dynamics using different correlation measures. We identify simple dynamics where the failure of completely-positive divisibility is in one-to-one correspondence with a correlation backflow. We then focus on specific correlation measures, such as those based on entanglement and the mutual information, and identify their strengths and limitations. We conclude with a study of a recently introduced correlation measures based on state distinguishability and see how, for these measures, adding an extra auxiliary system enforces the set of detectable non-Markovian dynamics.

I. INTRODUCTION

The dynamics of open quantum systems \cite{1-5} has been investigated extensively in recent years for both fundamental and applicative reasons. In particular the problem of understanding and characterizing memoryless dynamics, the so-called Markovian regime, and dynamics exhibiting memory effects, the non-Markovian regime, have been considered in a wide range of different ways (for extended reviews see \cite{4, 5}). While a unique agreed upon concept of quantum Markovianity does not exist it is frequently identified with the property of \textit{Completely Positive divisibility} (CP-divisibility). An evolution is CP-divisible if between every two points in time it can be described by a CP-map. This idea generalises the semigroup property \cite{6, 7} of classical Markovian processes.

Intuitively, one expects that non-Markovian effects are associated to a backflow of information from the environment to the system. Several approaches have been pursued to put this intuition in rigorous terms. A standard procedure, see for example \cite{9, 12}, consists of considering operational quantities that are monotonically non-increasing under CP maps. An increase of any such quantities implies that the evolution is non-Markovian, although the converse is not true in general. One of the quantities considered in the context of non-Markovian characterization is correlations, a fundamental concept for our understanding of quantum theory and also a resource for many quantum information protocols. The general idea of this approach consists of monitoring the evolution of the correlations between a system subjected to a dynamics and an additional particle that does not take part in the evolution. If at some point a correlation backflow, that is, an increase in the correlations quantified by a given correlation measure, is observed, then the dynamics must be non-Markovian.

The main goal of this work is to understand the power of correlations to witness non-Markovian evolutions. In particular, we study the strengths and weaknesses of several well-known correlation measures for this task. We derive several results that improve our understanding of this question. First, it is shown how for a class of differentiable evolutions termed single parameter, which includes relevant examples such as depolarization, dephasing, and amplitude damping, any continuously differentiable correlation measure increases during non-Markovian dynamics unless it is time independent on the whole image of the preceding evolution. Secondly, we focus on two fundamental quantum correlation measures: entanglement measures and (quantum) mutual information. For the first, we provide a simple argument explaining how it fails to witness non-Markovianity in many situations. For the second, we study its behaviour in different scenarios. We first show that the mutual information witnesses the non-Markovianity of any bijective unitary and non-P-divisible dynamics on a qubit. We then provide several examples of non-Markovian dynamics where no mutual information backflow is observed when using maximally entangled states. For some of these examples, we demonstrate that a backflow in the mutual information does appear when using non-maximally entangled states, in some cases even arbitrarily weakly entangled pure states. This highlight how a high degree of initial correlations is not necessarily beneficial for the detection of non-Markovianity when using the mutual information as a witness. Lastly, based on the results in \cite{28}, we construct examples of non-Markovian dynamics that do not display any backflow of the mutual information. The last part of our work is devoted to a detailed analysis of the correlation measure introduced in Ref. \cite{28} that detects almost all non-Markovian dynamics, more precisely all but a zero-measure set consisting of all those dynamics that are not bijective in finite time intervals. Interestingly, for all these cases, an initial state arbitrarily close to a product state suffices for the non-Markovianity detection. We also discuss how, in order to fully exploit the power of this correlation measure, one has to consider an enlarged scenario involving a second ancilla $A'$ that is not subjected to the evolution either, and look for correlation backflows along the bipartition $SA'$ versus $A$.

The outline of the manuscript is as follows. In Section \textsection1 we define the notation used in the article, we introduce the key concepts of open-system dynamics and present two classes of evolutions that are very useful for the derivation of our results:
II. NON-MARKOVIAN DYNAMICS

In this section we describe the mathematical representation of open system dynamics and the notation used in what follows. In Section II A we focus on the description of the time evolution of open quantum systems and in Section II B we introduce what we call “single parameter” evolutions, a family of differentiable evolutions. Finally, in Section II C we present the random unitary evolutions, an example of dynamics defined by the random application of unitary operations.

We consider a quantum system \( S \) with an associated finite dimensional Hilbert space \( \mathcal{H}_S \) of dimension \( d \). The set of bounded linear operators acting on \( \mathcal{H}_S \) is denoted by \( B(\mathcal{H}_S) \) and the subset of states is denoted by \( S(\mathcal{H}_S) \). An ancillary system \( A \) with Hilbert space \( \mathcal{H}_A \) is introduced and the corresponding set of bounded linear operators and set of states are denoted by \( B(\mathcal{H}_A) \) and \( S(\mathcal{H}_A) \), respectively. The set of bounded operators on the combined Hilbert space \( \mathcal{H}_A \otimes \mathcal{H}_S \) is denoted by \( B(\mathcal{H}_A \otimes \mathcal{H}_S) \) and the set of states is denoted by \( S(\mathcal{H}_A \otimes \mathcal{H}_S) \).

The evolution of \( S \) from initial time \( t = 0 \) to a later time \( t \) is described by a dynamical map, i.e., a linear operator \( \Lambda_t : B(\mathcal{H}_S) \rightarrow B(\mathcal{H}_S) \) that is completely positive and trace preserving (CPTP). The dynamics of the system is thus described by the family of maps \( \{\Lambda_t\} \) parametrized by \( t \). An important concept for the study of non-Markovian effects is that of divisibility of the dynamical map \( \Lambda_t \), as well as Positive (P) and Completely Positive (CP)-divisibility in terms of intermediate maps \( V_{s,t} \).

**Definition 1.** A dynamical map \( \Lambda_t \) is called divisible (CP-divisible) if it can be expressed as a sequence of \( \Lambda_t = V_{s,t} \Lambda_{t-s} V_{s,t} \), where \( V_{s,t} \) is a linear trace preserving (P)CP map, for any \( 0 \leq t \leq s \).

CP-divisibility of the family of dynamical maps has been taken by many authors as the definition of Markovian evolution [3]. It is in fact the definition we adopt in this work and, in what follows, whenever it is said that an evolution is non-Markovian, it means that it is not CP-divisible. The idea behind this is that in a CP-divisible process the environment \( E \) of the system can be disregarded and an intermediate map \( V_{s,t} \) can be physically realized just by acting locally on the system and a second environment \( S \) at time \( t \) (see Fig. 1). If \( V_{s,t} \) is not CP it cannot result from a local process involving only \( S \) and \( S' \). Hence, the state of \( S - E \) at time \( t \) holds information that is crucial for the subsequent evolution of \( S \) up to time \( s \). This definition of Markovian evolution is the analogue of the classical Markov process where the evolution at each time interval only depends on the state of the system at the beginning of the interval. In other words both definitions capture the idea that no memory of previous events is needed to describe the process.

![Fig. 1.](image)

**A. The generator of the evolution and the intermediate map**

For a differentiable evolution, any dynamical map \( \Lambda_t \) and any intermediate map \( V_{s,t} \) can be expressed as time ordered exponentials

\[
\Lambda_t(\rho) = T e^{\mathcal{L}_t \rho}, \quad V_{s,t} = T e^{\mathcal{L}_{s-t} \rho},
\]

where \( \mathcal{L}_t \) is the generator of the evolution as first defined by Gorini et al. [6] and Lindblad [7] for dynamical semigroups and later extended to general divisible maps,

\[
\mathcal{L}_t(\rho) \equiv i[H(t),\rho] + \sum_i \gamma_i(t) \left( G_i(t)\rho G_i(t) - \frac{1}{2} \left[ G_i(t) G_i(t), \rho \right] \right).
\]

Here \( \gamma_i(t) \) are real time dependent functions, \( G_i(t) \) are time dependent operators and \( H(t) \) is the Hamiltonian of the system, thus also a Hermitian possibly time-dependent operator. The Hamiltonian term of the generator describes the unitary part of the dynamics generated by \( H(t) \) and the second term describes the dissipative part of the dynamics generated by the
coordinates given by the real numbers $a$. A state $A$ has dimension of $k$. Phasing and amplitude damping in a time independent basis. Let the dimension of $H_5$ be $d_s$ and let $\chi_s$ be $d_s^2 - 1$ traceless Hermitian operators such that $\text{Tr}[\chi_s \chi_s S] = \delta_{s,0}$. Likewise let $d_A$ be the dimension of $H_A$ and let $\chi_A$ be $d_A^2 - 1$ traceless Hermitian operators such that $\text{Tr}[\chi_A \chi_A S] = \delta_{s,0}$. Then one can choose an orthonormal basis $\{e_i\}_{i=0}^d$ for $B(H_A \otimes H_5)$ by constructing the basis vectors $e_i$ as the tensor products $\chi_A \otimes \chi_s$, $\chi_A \otimes \chi_s$, and $\chi_A \otimes \chi_s$ for all $i, j$. We denote $e_0 \equiv 1 \otimes 1 \otimes 1$. A state $\rho \in S(H_A \otimes H_5)$ can be represented in this basis by coordinates given by the real numbers $a_i = \frac{1}{d_A d_s} \text{Tr}(\rho e_i)$, i.e.,

$$\rho = \frac{1}{d_A d_s} 1_A \otimes 1_S + \sum_{i=0}^d a_i e_i. \tag{3}$$

We can now describe the intermediate map $I_A \otimes V_s$ by how it acts on each basis elements $e_i$. For this purpose we define

$$V_s(t, s) \equiv \text{Tr} \left[ e_i I_A \otimes V_s(e_i) \right]. \tag{4}$$

Note that $V_s(t, s)$ is real for all $i, j$ since $V_s$ is hermiticity preserving. Moreover, since the map $V_s$ is trace preserving it follows that $V_{00}(t, s) = 1$ and $V_{0j}(t, s) = 0$ for $j \neq 0$. Let the coordinates $\bar{a} \equiv \{a_i\}$, where $a_0 = 1/d_A d_s$, describe a state at time $t$. This state is mapped by $I_A \otimes V_s(t, s)$ to coordinates $\bar{a}(s) = \{a_i(s) \equiv \sum_j V_i(s, t) a_j\}$. Analogously to Eq. (4), we define the time derivatives of the components $V_i(s, t)$ as

$$\frac{dV_i(s, t)}{ds} \bigg|_{s=t} = \text{Tr} \left[ e_i I_A \otimes \frac{dV_s}{ds}(e_i) \right] \bigg|_{s=t}. \tag{5}$$

### B. Single parameter evolutions

A simple family of open-system dynamics is given by those differentiable evolutions $\{\Lambda_t\}$, such that for all $t$ the generator $\mathcal{L}_t$ can be expressed as

$$\mathcal{L}_t(\rho) = i[H(t), \rho] + \gamma(t) \sum_k \left( G_k \rho G_k^\dagger - \frac{1}{2} \{G_k^\dagger G_k, \rho\} \right), \tag{6}$$

where $G_k$ are time-independent operators and $\gamma(t)$ is a continuous function of time. We call those evolutions with this property *single parameter* since $\gamma(t)$ alone describes the time dependence of the dissipative part. Paradigmatic examples of single parameter evolutions are depolarization, as well as dephasing and amplitude damping in a time independent basis.

Using the representation defined in Eq. (5) we can state the single parameter property as

$$\frac{dV_i(s, t)}{ds} \bigg|_{s=t} = g_{ij} \gamma(t) + h_{ij}(t), \tag{7}$$

where $h_{ij}(t)$ is corresponds to the action of $H(t)$ and the time independent parameters $g_{ij}$ to $\sum_k g_k \left( G_k^\dagger G_k - \frac{1}{2} \{G_k^\dagger G_k, \rho\} \right)$.

An important property of single-parameter evolutions is that they can be divided into CP-divisible and not P-divisible time intervals, that is, they never present intermediate maps that are not CP but P.

**Proposition 1.** Let $\{\Lambda_t\}$ be single parameter. Then the intermediate map $V_{t^1, t^2}$ is either CP or not P for any $t'$ and sufficiently small $\epsilon' > 0$.

**Proof.** An intermediate map $V_{s,t'}$ is CP if the Choi matrix $C_{V_s}$ has non-negative eigenvalues [16] [17]. Moreover, $V_{s,t'}$ is CP if $V_{s,t'}(\rho)$ is positive semidefinite for all positive semidefinite Hermitian trace one matrices $\rho \in B(H_s)$. In particular, if $V_{s,t}$ is CP, $V_{s,t'}(\rho)$ is positive semidefinite for any positive semidefinite rank one $\rho$.

The eigenvalues of $V_{s,t'}(\rho)$ for a rank-one pure state $\rho$ and the eigenvalues of $C_{V_s}$, are functions of the parameters $t$ and $V_{s,t'}(s, t)$. Since we assumed that $\gamma(t)$ is continuous it follows that $V_{s,t'}(\rho)$ and $C_{V_s}$ are both continuously differentiable. This in turn implies that the eigenvalues of $V_{s,t'}(\rho)$ and $C_{V_s}$ can be described by continuously differentiable functions [18].

If $\gamma(t) \neq 0$, so that $V_{s,t}$ is not unitary, there always exists at least one eigenvalue $\lambda_{s,t'}$ of the Choi matrix that is zero and has a nonzero time derivative at $s = t$. There also exists at least one rank-one $\rho$ with eigenvalue $\lambda_{s,t'}$ that is zero and its evolution $\lambda(t(s))$, i.e., given by $\rho(s) = 1_A \otimes \rho(s)$, has a nonzero time derivative at $s = t$ (see Appendix A). We start by considering the time derivatives $\frac{d}{dt} \frac{d\lambda(t(s))}{ds} \bigg|_{s=t} = \sum_i \frac{d}{dt} \frac{d\lambda_i(t(s))}{ds} \bigg|_{s=t} = \sum_i \frac{d}{dt} \frac{d\lambda_i(t(s))}{ds} \bigg|_{s=t}$, $h_{ij}(t) = 0$ and $\sum_i \frac{d}{dt} \frac{d\lambda_i(t(s))}{ds} \bigg|_{s=t} = h_{ij}(t) = 0$. If the time derivatives are non-zero they are proportional to $\gamma(t)$. If $V_{t^1, t^2}$ is CP for any sufficiently small $\epsilon'$ it is clear that $\frac{d}{dt} \frac{d\lambda(t)}{ds} \bigg|_{s=t} > 0$ and $\frac{d}{dt} \frac{d\lambda(t)}{ds} \bigg|_{s=t} > 0$. Then if $t'$ exist such that $\text{sign}[\gamma(t')] = -\text{sign}[\gamma(t)]$, there exists an $\epsilon'$ such that $V_{t^1, t^2}$ is neither CP or P since $\frac{d}{dt} \frac{d\lambda(t)}{ds} \bigg|_{s=t} < 0$ and $\frac{d}{dt} \frac{d\lambda(t)}{ds} \bigg|_{s=t} < 0$. From this follows that it is impossible for $V_{t^1, t^2}$ to be P but not CP for sufficiently small $\epsilon'$.

Thus, if the assumptions of Proposition [1] hold we can conclude that the dynamics can be divided into closed intervals where $\gamma(t) \geq 0$ and open time intervals for which $\gamma(t) < 0$.

In the closed intervals where $\gamma(t) \geq 0$ the dynamics is CP-divisible and in the open intervals where $\gamma(t) < 0$ the dynamics is not P-divisible.

Notice that, if $\Lambda_t$ is CP-divisible in a time interval, we do not necessarily mean that it is Markovian. Indeed, in order for $\Lambda_t$ to be Markovian, the corresponding intermediate map $V_{s,t}$ has to be CP for any $s$ and $t$ such that $0 \leq t \leq s$ (see Definition [1]).
C. Random unitary evolutions

Unital dynamical maps \( \Lambda_t \) on a finite dimensional Hilbert space \( \mathcal{H}_S \) are defined by the property of preserving the identity operator, i.e., \( \Lambda_t(1_S) = 1_S \) for any \( t \geq 0 \). In this section we introduce a useful class of unital evolutions that we repeatedly use in the following: the random unitary dynamics.

We consider a generic finite-dimensional Hilbert space \( \mathcal{H}_S \), where \( \dim(\mathcal{H}_S) = d_S \). The CPTP map \( \Lambda : B(\mathcal{H}_S) \to B(\mathcal{H}_S) \) represents a random unitary channel on \( S \) if its action can be written as

\[
\Lambda(\rho_S) = \sum_k p_k U_k \rho_S U_k^\dagger ,
\]

where \( \{p_k\}_k \) is a probability distribution and \( \{U_k\}_k \) is a set of unitary transformations [24]. If \( d_S = 2 \), i.e., for qubit evolutions, any unital channel is random unitary [36]. However, when \( d_S \geq 3 \), the set of random unitary channels is strictly included in the set of unital channels.

In the following, we consider random unitary dynamical maps \( \Lambda_t : B(\mathcal{H}_S) \to B(\mathcal{H}_S) \) of the form

\[
\Lambda_t(\rho_S) = \sum_{k=0}^{N_S} p_k(t) \sigma_k \rho_S(0) \sigma_k ,
\]

where the unitary operators \( \sigma_k \), for \( k = 1, \ldots, N_S \) and \( N_S = d_S^2 - 1 \), constitute a basis of elements for the Hermitian and traceless operators in \( B(\mathcal{H}_S) \) such that \( \sigma_0 = 1_S \) and \( \text{Tr}[\sigma_i \sigma_j^\dagger] = d_S \delta_{ij} \) for \( i, j = 0, 1, \ldots, d_S^2 - 1 \). Since \( \{p_k(t)\}_k \) is a probability distribution, we have that \( \sum_{k=0}^{N_S} p_k(t) = 1 \) and \( p_k(t) \geq 0 \) for any \( k = 0, \ldots, N_S \). Moreover, from the condition \( \Lambda_0 = 1_S \) it follows that \( p_0(0) = 1 \). For the sake of simplicity, when we refer to a random unitary dynamics we consider dynamical maps of the form given in Eq. (9).

1. Random unitary evolutions for qubits

In the case that the evolution is defined on a qubit, a random unitary dynamics is given by the action of the corresponding dynamical map \( \Lambda_t \) on the Pauli operators \( \sigma_x, \sigma_y, \sigma_z \)

\[
\begin{align*}
\Lambda_t(\sigma_x) &= \exp \left\{ - \int_0^t (\gamma_x(\tau) + \gamma_y(\tau)) d\tau \right\} \sigma_x , \\
\Lambda_t(\sigma_y) &= \exp \left\{ - \int_0^t (\gamma_z(\tau) + \gamma_x(\tau)) d\tau \right\} \sigma_y , \\
\Lambda_t(\sigma_z) &= \exp \left\{ - \int_0^t (\gamma_y(\tau) + \gamma_z(\tau)) d\tau \right\} \sigma_z , \\
\Lambda_t(1) &= 1 ,
\end{align*}
\]

where \( \gamma_i(t) \) are real-valued functions of \( t \). Any random unitary dynamical map is bijective for all \( t \) and the intermediate maps are given by

\[
\begin{align*}
V_{x,t}(\sigma_x) &= \exp \left\{ - \int_0^t (\gamma_x(\tau) + \gamma_y(\tau)) d\tau \right\} \sigma_x , \\
V_{y,t}(\sigma_y) &= \exp \left\{ - \int_0^t (\gamma_z(\tau) + \gamma_x(\tau)) d\tau \right\} \sigma_y , \\
V_{z,t}(\sigma_z) &= \exp \left\{ - \int_0^t (\gamma_y(\tau) + \gamma_z(\tau)) d\tau \right\} \sigma_z , \\
V_{s,t}(1) &= 1 .
\end{align*}
\]

The corresponding generator of the evolution (see Eqs. [1] and [2]) is

\[
\mathcal{L}_t(\rho) = \sum_{k=x,y,z} \gamma_k(t) (\sigma_k \rho \sigma_k - \rho) .
\]

We define

\[
A_{ij}(t) \equiv - 2 \int_0^t (\gamma_x(\tau) + \gamma_z(\tau)) d\tau \geq 0 ,
\]

for \( i \neq j \) and \( i, j \in \{x, y, z\} \). The map defined by Eqs. (10) is CPTP if and only if

\[
B_{ijk}(t) \equiv 1 + A_{ij}(t) - A_{jk}(t) - A_{ki}(t) \geq 0 ,
\]

for any cyclic permutation of \( i, j, k \), i.e., \( (i, j, k) = (x, y, z), (y, z, x), (z, x, y) \) [37]. The dynamics is CP-divisible if and only if every \( \gamma_k(t) \geq 0 \) for all \( k \) and \( t \geq 0 \). Moreover, the dynamics is P-divisible if and only if the conditions

\[
\begin{align*}
gamma_x(t) + \gamma_y(t) &\geq 0 , \\
gamma_z(t) + \gamma_x(t) &\geq 0 , \\
gamma_y(t) + \gamma_z(t) &\geq 0 ,
\end{align*}
\]

are satisfied since the intermediate maps are then contractive in the trace norm [21, 23, 25]. The similarity between Eqs. (10) and (13) allows us to use the following equivalent definition for the random unitary qubit dynamics \( \Lambda_t \)

\[
\Lambda_t(\sigma_i) = \lambda_i(t) \sigma_i , \quad \text{where} \quad \lambda_i(t) = \sqrt{A_{ik}(t)} ,
\]

where \( (i, j, k) = (x, y, z), (y, z, x), (z, x, y) \) and \( \Lambda_0(1) = 1 \).

Any random unitary dynamical map for qubits can be defined either by the time-dependent probability distribution \( \{p_k(t)\}_k \) that appears in Eq. (9) or by the time-dependent set of rates \( \{\gamma_k(t)\}_k \) that defines the generator \( \mathcal{L}_t \) and the dynamical map \( \Lambda_t \) itself (see Eqs. (10) and (12)). The relations that link these two sets are [25]

\[
\begin{align*}
p_0(t) &= \frac{1}{4} \left( 1 + A_{xy}(t) + A_{xz}(t) + A_{yz}(t) \right) , \\
p_x(t) &= \frac{1}{4} \left( 1 - A_{xy}(t) - A_{xz}(t) + A_{yz}(t) \right) , \\
p_y(t) &= \frac{1}{4} \left( 1 - A_{xy}(t) + A_{xz}(t) - A_{yz}(t) \right) , \\
p_z(t) &= \frac{1}{4} \left( 1 + A_{xy}(t) - A_{xz}(t) - A_{yz}(t) \right) .
\end{align*}
\]
2. Example: the quasi-eternal non-Markovian model

We present a class of non-Markovian random unitary evolutions inspired by the ‘eternal’ non-Markovian model in Refs. [26, 27]. We analyze the dynamical maps, defined by Eqs. (10), with parametrized time-dependent rates

$$\{\gamma_x(t), \gamma_y(t), \gamma_z(t)\} = \frac{\alpha}{2} \{1, 1, -\tanh(t - t_0)\} ,$$

where $\alpha > 0$ and $t_0 \geq 0$.

First, we notice that: $\gamma_x(t) = \gamma_y(t) = \alpha/2 > 0$ for any $t$, $\gamma_z(t) \geq 0$ for $t \leq t_0$ but $\gamma_z(t) < 0$ for any $t > t_0$. We point out that, depending on the value of $\alpha > 0$ and $t_0 \geq 0$, the corresponding dynamical map $\Lambda^{(\alpha, t_0)}_t$ generated by Eqs. (10) may be physical, i.e., CPTP, or not. While the non-positivity of the rates implies the non-CP-divisibility of the dynamics, the P-divisibility conditions given by Eqs. (15) are satisfied for any $\alpha > 0$, $t_0$ and $t$. Therefore, in case $\alpha$ and $t_0$ defines a CPTP dynamical map, it is non-Markovian and P-divisible: the intermediate maps $V_{t, t_0}^{(\alpha, t_0)}$ are P (but not CP) for any time interval $(t_1, t_2)$ such that $t_1 > t_2$.

- $t_0 = 0$ (eternal model): The trace-preserving map $\Lambda^{(\alpha, t_0)}_t$ generated by the rates in Eq. (18) is CP for $\alpha \geq 1$ and only P for $0 < \alpha < 1$ [27].

- $t_0 > 0$ (quasi-eternal model): If $\alpha \geq 1$, the map $\Lambda^{(\alpha, t_0)}_t$ is CPTP for any $t \geq 0$ and $t_0 \geq 0$. Instead, if $0 < \alpha < 1$, we need to wisely choose the value of $t_0 \geq 0$ in order to have a $\Lambda^{(\alpha, t_0)}_t$ that is CPTP for every $t \geq 0$. First of all, we calculate the quantities $A^{(\alpha, t_0)}_{ij}(t)$ that completely define the dynamical map $\Lambda^{(\alpha, t_0)}_t$ (see Eq. (16))

$$A^{(\alpha, t_0)}_{xy}(t) = e^{-\alpha t},$$

$$A^{(\alpha, t_0)}_{yz}(t) = A^{(\alpha, t_0)}_{zx}(t) = \left( e^{-\frac{\alpha \cosh(t - t_0)}{\cosh(0)}} \right)^{\alpha} .$$

Now, we derive the ranges of physicality of the parameters $\alpha > 0$ and $t_0 \geq 0$. In order to satisfy the CPTP conditions given in Eq. (14), we notice that $B^{(\alpha, t_0)}_{xyz}(t) = B^{(\alpha, t_0)}_{xyz}(t) = 1 - A^{(\alpha, t_0)}_{xy}(t) = 1 - e^{-2\alpha t} \geq 0$ for any $\alpha > 0$ and $t \geq 0$. It is straightforward to verify that the last condition $B^{(\alpha, t_0)}_{xyz}(t) = 1 + A^{(\alpha, t_0)}_{xy}(t) - 2A^{(\alpha, t_0)}_{zx}(t) \geq 0$ is satisfied for any $t \geq 0$ if and only if $\lim_{t \rightarrow \infty} B^{(\alpha, t_0)}_{xyz}(t) \geq 0$. Therefore, the relation between $\alpha > 0$ and $t_0 \geq 0$ that implies that $\Lambda^{(\alpha, t_0)}_t$ is CPTP for any $t \geq 0$ is $1 - 2e^{2\alpha t} + 1 - \alpha \geq 0$ or, equivalently

$$t_0 \geq T(\alpha) \equiv \frac{1}{2} \log(2^{1/\alpha} - 1) .$$

III. CORRELATION MEASURES AS WITNESSES OF NON-MARKOVIANITY

In this section we introduce the general notion of correlation measures and see how they can be used for the detection of non-Markovian dynamics. We then show that for single-parameter evolutions, any non-Markovian effect leads to an increase in any continuously differentiable correlation measure that is non-zero during the preceding evolution.

A. Correlation measures

A correlation measure is a function that quantifies correlations between subsystems. Many different measures have been constructed that quantify correlations in different ways or capture qualitatively different types of correlations. For a function $M$ to be considered an operationally meaningful correlation measure, it has to satisfy the following property:

- $M(\rho) \geq 0$ for any state $\rho$.
- $M(\rho) = 0$ if $\rho$ is a product state.

Now, consider a generic correlation measure $M(\rho)$ that is continuously differentiable on $S(\mathcal{H}_A \otimes \mathcal{H}_S)$. If the intermediate map $\mathcal{I}_A \otimes V_{s,t}$ is differentiable with respect to $t$ the time derivative of $M(\rho(s)) = M(\mathcal{I}_A \otimes V_{s,t}(\rho))$ at time $t$ can be expressed as

$$\sum_i \frac{\partial M(\rho(s))}{\partial a_i(s)} \frac{da_i(s)}{ds} \bigg|_{s=t} = \sum_{ij} \frac{\partial M(\rho)}{\partial a_i} a_j \frac{dV_{ij}(s, t)}{ds} \bigg|_{s=t} ,$$

where $a_i$ ($a_i(s)$) are the coordinates of $\rho$ ($\rho(s)$). We see that $\frac{dM}{dt} |_{s=t}$ depends only on $a_i$, and the time derivatives of the components $V_{ij}(s, t)$ of the intermediate map. As long as $\frac{\partial M(\rho)}{\partial a_i} a_j \neq 0$ for some $i$, $j$ and $\rho$ there exist some intermediate map that will induce either a decrease or increase of $M(\rho)$. However, it is not always the case that a given measure $M$ satisfies $\frac{dM(\rho)}{dt} |_{s=t}$ for a $\rho$ in the image of the dynamical map $\Lambda_t$. In Section IV, we show that entanglement measures and entanglement breaking evolutions [15] can provide examples of this situation.

B. Correlation backflows as a signature of non-Markovianity

The monotonicity of correlation measures under CP maps implies that they are all non-increasing under Markovian, that is, CP-divisible, evolutions. In fact, consider the scenario of Fig. 2 left, where a system $S$, correlated with an additional particle $A$, is subjected to an evolution. If there is an increase in the correlations between these two particles during the evolution between times $t$ and $s$, there cannot be a CP intermediate
map \( V_{a,s} \). This is why correlation measures have been proposed to witness and quantify non-Markovian effects \[^3][^8][^2]\).

From an open-system perspective a decrease in correlations along the bipartition \( S\Lambda \) is used to witness and quantify non-Markovian evolutions \[^8][^12\]. The pure qubit dephasing dynamics is described by the dynamical maps

\[
\Lambda_t(\sigma_x) = e^{-\frac{i}{\hbar} \gamma t} \sigma_x,
\]

\[
\Lambda_t(\sigma_y) = e^{-\frac{i}{\hbar} \gamma t} \sigma_y,
\]

\[
\Lambda_t(\sigma_z) = \sigma_z,
\]

\[
\Lambda_t(\mathbb{1}) = \mathbb{1}.
\]

The dynamical maps are bijective for all times and the intermediate maps \( V_{i,s} \) are therefore given by \( V_{i,s} = \Lambda_t^{-1} \Lambda_s \). Explicitly \( V_{i,s} \) is given by

\[
V_{i,s}(\sigma_x) = e^{-\int_{s}^{t} \gamma(\tau) d\tau} \sigma_x,
\]

\[
V_{i,s}(\sigma_y) = e^{-\int_{s}^{t} \gamma(\tau) d\tau} \sigma_y,
\]

\[
V_{i,s}(\sigma_z) = \sigma_z,
\]

\[
V_{i,s}(\mathbb{1}) = \mathbb{1}.
\]

The generator \( \mathcal{L}_i \) is given by

\[
\mathcal{L}_i(\rho) = \gamma(t) (\sigma_i \rho \sigma_i - \rho),
\]

where \( \gamma(t) \) is the time dependent dephasing rate. The dynamics is CP-divisible if and only if the dephasing rate \( \gamma(t) \geq 0 \). Furthermore, the dephasing dynamics is not P-divisible when \( \gamma(t) < 0 \). This can be seen directly from the two eigenvalues

\[
1 - e^{-\int_{s}^{t} \gamma(\tau) d\tau} \text{ and } 1 + e^{-\int_{s}^{t} \gamma(\tau) d\tau} \text{ of } V_{s,i}(\mathbb{1} + \sigma_i).
\]

We can see from Eq. (25) that for all \( i,j \) such that \( V_{j,s}(s,t) \) has a non-zero time derivative it holds that \( \frac{d V_{j,s}}{d t} \big|_{s=t} = -\gamma(t) \). Thus for any continuously differentiable correlation measure \( M(\rho(t)) \) the dephasing rate \( \gamma(t) \) determines the sign of \( \frac{d M}{d t} \) for \( \gamma(t) \geq 0 \).

2. Example: Generalized amplitude damping evolution

Generalized amplitude damping evolution in a fixed basis is a second example that is single-parameter and satisfies the
conditions in Propositions 1 and 2 except where it is non-differentiable. The dynamics of generalized amplitude damping on a qubit is described by the dynamical maps
\[
\Lambda_0(\sigma_x) = G(t)\sigma_x,
\Lambda_0(\sigma_y) = G(t)\sigma_y,
\Lambda_0(\sigma_z) = G^2(t)\sigma_z,
\Lambda(t)(\mathbb{1}) = \mathbb{1} + (2p-1)(1-G^2(t))\sigma_z,
\]
where \(0 \leq G(t) \leq 1\) and \(0 \leq p \leq 1\). For \(G(t) > 0\) the dynamical maps are bijective and the intermediate maps are given by
\[
V_{s,t}(\sigma_x) = \frac{G(s)}{G(t)}\sigma_x,
V_{s,t}(\sigma_y) = \frac{G(s)}{G(t)}\sigma_y,
V_{s,t}(\sigma_z) = \left(\frac{G(s)}{G(t)}\right)^2\sigma_z,
V_{s,t}(\mathbb{1}) = \mathbb{1} + (2p-1)\left(1 - \left(\frac{G(s)}{G(t)}\right)^2\right)\sigma_z.
\]

For \(s\) and \(t\) such that \(G(t) = 0\) the intermediate map only exists if \(G(s) = 0\) and can be defined as the identity map. If \(G(t) = 0\) and \(G(s) \neq 0\) the intermediate map does not exist since the evolution is many-to-one. For \(t\) where \(\mathcal{L}_t\) is well defined, it is given by
\[
\mathcal{L}_t(\rho) = p\gamma(t)(\sigma_+\rho\sigma_- - 1/2(\sigma_+\sigma_-)\rho)
+ (1-p)\gamma(t)(\sigma_-\rho\sigma_+ - 1/2(\sigma_-\sigma_+)\rho),
\]
where \(\sigma_\pm = 1/2(\sigma_x \pm i\sigma_y)\) and \(\gamma(t)\) is given by
\[
\gamma(t) = -2\left.\frac{d}{ds}\frac{G(s)}{G(t)}\right|_{s=t} = -2\frac{d}{dt}\frac{G(t)}{G(t)},
\]
whenever \(G(t) > 0\) and differentiable. The dynamics is CP-divisible in a generic time interval \([t_1, t_2]\) when \(\gamma(t) \geq 0\) for any \(t \in [t_1, t_2]\). Furthermore, the amplitude damping dynamics is not P-divisible in \([t_1, t_2]\) when \(\gamma(t) < 0\) for any \(t \in [t_1, t_2]\).

We can easily see that when \(G(t) > 0\) and differentiable all non-zero time derivatives of the components of the intermediate map are proportional to \(\gamma(t)\). Thus \(\gamma(t)\) determines the sign of \(\frac{d}{dt}M(\tilde{a}, t)\) for any continuously differentiable correlation measure \(M(\tilde{a}, t)\).

IV. ENTANGLEMENT MEASURES

After presenting the general definition of correlation measures, in the next sections we focus on different examples of these measures and study their behavior under non-Markovian effects. We start with entanglement measures, denoted by \(M_E\), which are originally constructed to capture only non-classical correlations. This is why they do not only not increase under local operations, but even when these operations are assisted by classical communication (LOCC). This implies that \(M_E(\rho) = 0\) if \(\rho\) is a separable state. Ref. [8] introduced the idea of using entanglement measures to witness non-Markovianity.

It is relatively straightforward to see that entanglement measures cannot detect all possible examples of non-Markovian dynamics, see also [28]. In fact, consider the situation in which a dynamics \(\Lambda_{EB}\) becomes and remains entanglement breaking (EB) after a given time \(t_{EB}\), see Fig. 3. Any entanglement measure remains equal to zero for \(t > t_{EB}\) and will therefore be unable to detect any non-Markovian effect taking place for \(t < t_{EB}\). This is for instance the case for dynamics that are \(P\)-divisible. At \(t = t_{EB}\) all the states in the image are separable, and remain separable after it since local positive maps do not create any entanglement when acting on separable states.

For the sake of clarity, in what follows we provide an example of a non-Markovian qubit dynamics \(\Lambda_{EB}\) which is CP-divisible in the time interval \([0, t_0]\), while it is only \(P\)-divisible for \(t > t_0\). Showing that \(t_{EB}(\Lambda_{EB}) < t_0\) we prove that this non-Markovian dynamics does not display any entanglement backflow.

1. Example: the quasi-eternal non-Markovian model

We consider a bipartite system, where \(A\) and \(S\) are qubits. A necessary and sufficient condition for a bipartite qubit state \(\rho_{SA} \in \mathcal{S}(\mathcal{H}_S \otimes \mathcal{H}_A)\) to be entangled is that its negativity, i.e., the entanglement measure
\[
N(\rho_{SA}) \equiv \left|\left|\Gamma_A^{\mathcal{U}}\rho_{SA}\right|\right|_1 - 1
\]

is positive, where \(\Gamma_A\) represents the operation of partial transposition on the system \(A\).

We consider the dynamical map \(\Lambda^{(2/5,2)}\) that belongs to the class of quasi-eternal non-Markovian (\(P\)-divisible) evolutions.
Therefore, since \( t \) is the condition of physicality given by Eq. (21), it is satisfied. Since \( \gamma_1(t) = \gamma_2(t) = 1/5 \) and \( \gamma_3(t) < 0 \) if and only if \( t < t_0 \), the corresponding intermediate map \( V_{s;2}^{(2/5,2)} \) of this evolution is \( P \)-but not \( CP \)-divisible for any \( s > t > t_0 \), while it is \( CP \) when \( t_0 \leq s \leq t \).

This temporal evolution becomes entanglement breaking. Indeed, let \( \phi_A(t) = I_A \otimes \Lambda(t)^{(2/5)}_{s,A} \) be the temporal evolution of the maximally entangled state \( \phi_{s,A}^+ = |\phi_+\rangle\langle\phi_+|_{s,A} \), where \( |\phi_+\rangle_{s,A} = (|00\rangle_{s,A} + |11\rangle_{s,A})/\sqrt{2} \). We obtain the separability of \( \phi_A(t) \), namely \( N(\phi_{s,A}^+)(t) = 0 \), for any \( t \geq t_{EB}(\Lambda(t)^{(2/5)}_{s,A}) \) \( \equiv 1.47 \). Therefore, since \( t_{EB}(\Lambda(t)^{(2/5)}_{s,A}) \leq t_0 < t \), we conclude that it is not possible to observe a non-monotonic evolution of \( N(\rho_{s,A}(t)) \), and more generally of any entanglement measure, for any initial state \( \rho_{s,A}(0) \). Indeed, the intermediate maps of this evolution are either \( P \) or \( CP \) for \( t < t_{0s,A}(2/5) \leq t_{EB}(\Lambda(t)^{(2/5)}_{s,A}) \).

V. MUTUAL INFORMATION

A commonly used correlation measure that satisfies the conditions of Section II A is the (quantum) mutual information \( I(\rho_{AB}) \) defined as

\[
I(\rho_{AB}) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}),
\]

where \( S \) is the von Neumann entropy, \( \rho_{AB} \) is the state of the system, and \( \rho_A \) (\( \rho_B \)) is the reduced states of \( A \) (\( B \)). The mutual information is a continuous function on the set of states and is analytic on the interior of the set of states, i.e., it is infinitely differentiable and equals its Taylor series in a neighbourhood of any point. Differently from entanglement measures (see Section IV), \( I \) considers both classical and quantum correlations of bipartite systems. The use of the mutual information to detect non-Markovian effects was proposed in Ref. [12].

In the following, we present many different scenarios where the mutual information witnesses non-Markovian effects, but conclude by identifying a situation where it fails in this task. In Section IV A we show that whenever a qubit unital non-Markovian dynamics is not even \( P \)-divisible, an increase in the mutual information appears. In Section IV B we consider non-Markovian random unitary dynamics that are \( P \)-divisible to prove that maximally entangled states are not optimal to detect correlation backflow in terms of the mutual information. In Section IV C we move to a non-unital dynamics, in the form of a generalized amplitude damping channel, and provide a class of initial states that efficiently witness the non-Markovian nature of this evolution. Finally, in Sections IV D and IV E we introduce the tools needed to show the existence of non-Markovian evolutions for which the mutual information does not provide backflows for any initial state.

A. Non-Markovian non-\( P \)-divisible unitit qubit dynamics

A number of commonly studied dynamics including dephasing and random unitary dynamics are unitit. For bijective unital dynamical maps acting on a qubit an increase of the mutual information can be observed for any non-\( P \) intermediate map \( V_{s,t} \).

**Theorem 1.** Let \( \Lambda_t \) be a unital bijective qubit evolution. Furthermore assume that the intermediate map \( V_{s,t} \) is analytic and non-\( P \) at a given time \( t \). Then there exist states \( \rho_{AS} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_S) \) in the image of \( \Lambda_t \) for which \( I(\Lambda(t) \otimes V_{s,t}(\rho_{AS})) > I(\rho_{AS}) \).

**Proof.** Assume that \( V_{s,t} \) is not \( P \). Then there exists a pure state \( |\phi\rangle \) on \( \mathcal{H}_S \) such that \( V_{s,t}(|\phi\rangle\rangle) \) has a negative eigenvalue. Let the eigenvalues of \( V_{s,t}(|\phi\rangle\rangle) \) be \( 1+\epsilon(s) \) and \( -\epsilon(s) \), where \( \epsilon(t) = 0 \) and \( \epsilon(s) \geq 0 \). Consider the state

\[
\rho_{AS} = \frac{1}{2} |0\rangle\langle 0|_A \otimes (p|\phi\rangle\langle \phi|_S + (1-p)\mathbb{1}_S) + \frac{1}{2} |1\rangle\langle 1|_A \otimes (p|\phi^+\rangle\langle \phi^+|_S + (1-p)\mathbb{1}_S),
\]

where \( |\phi^+\rangle \) is the pure state orthogonal to \( |\phi\rangle \), and \( |0\rangle\langle 0|_A \) and \( |1\rangle\langle 1|_A \) are orthogonal states in \( \mathcal{H}_A \). Note that since the evolution is unital, the eigenvalues of \( V_{s,t}(|\phi^+\rangle\langle \phi^+|_S) \) are \( 1+\epsilon(s) \) and \( -\epsilon(s) \). Note also that the reduced density matrices of both the system and the ancilla are maximally mixed. Therefore, the reduced states are unchanged by a unital map \( V_{s,t} \).

The mutual information as a function of time is

\[
I(\Lambda(t) \otimes V_{s,t}(\rho_{AS})) = \left( \frac{1}{2} + p \epsilon(s) \right) \ln \left( \frac{1}{2} + p \epsilon(s) \right) + \left( 1 - \frac{1}{2} - p \epsilon(s) \right) \ln \left( \frac{1}{2} - p \epsilon(s) \right) + \ln 2.
\]

Its time derivative now reads

\[
- \frac{d}{ds} I(\Lambda(t) \otimes V_{s,t}(\rho_{AS})) |_{t=s} = \frac{d \epsilon(t)}{dt} p (\ln(1+p) - \ln(1-p)).
\]

Note that \( p(\ln(1+p) - \ln(1-p)) > 0 \) for \( 0 < p < 1 \). Therefore, for \( \frac{d \epsilon(t)}{dt} > 0 \) the time derivative of the mutual information is positive for \( \rho_{AS} \). Moreover, since \( \epsilon(t) \) is assumed to be analytic \( \frac{d \epsilon(t)}{dt} > 0 \) implies that the map \( V_{s,t} \) is non-\( P \) for a sufficiently small \( \delta \). Since \( \Lambda_t \) is bijective and unital, there always exists a sufficiently small \( p \) such that \( \rho_{AS} \) is in the image of \( \Lambda_t \).

\[\square\]

B. Non-Markovian random unitary dynamics and maximally entangled states

We provide a condition for the mutual information not to show backflows when a maximally entangled state is evolved by a random unitary dynamical map. Thereafter, we formulate a version of this condition that applies to qubits, where the \( P \)-divisibility of the dynamical map is implied.

We consider the bipartite scenario where the systems \( S \) and \( A \) are qudits \( (d_S = d_A = d) \), \( S \) is evolved by a random unitary dynamical map \( \Lambda_t \) (see Eq. (9)) and the ancillary system
$A$ is left untouched. We study the evolution of a maximally entangled state $\phi = |\phi\rangle_{A} |\phi\rangle_{S}$, where

$$|\phi\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |a_{i}\rangle_{A} \otimes |s_{i}\rangle_{S},$$

and $\{|s_{i}\rangle_{S}\} (\{|a_{i}\rangle_{A}\})$ is an orthonormal basis of $\mathcal{H}_{S}$ ($\mathcal{H}_{A}$).

First, we show that the evolved state is diagonal in a Bell basis with eigenvalues given by the same probability distribution $\{p_{k}(t)\}_{k}$ that defines $\Lambda_{t}$ (see Eq. (9)). Indeed

$$\phi(t) = \sum_{k=1}^{N} p_{k}(t)(|\Lambda_{k}\rangle \otimes \sigma_{k})|\phi\rangle_{A} \otimes \phi_{k},$$

where $N = d^{2} - 1$. The set of states $\{|\phi_{k}\rangle\}_{k} \equiv \{|\Lambda_{k}\rangle \otimes \sigma_{k}\} |\phi\rangle_{k}$ define a Bell basis and are orthonormal since:

$$\langle \phi_{j}|\phi_{k}\rangle = \text{Tr} \left[ |\phi_{j}\rangle \langle \phi_{k}| \otimes \sigma_{j}\sigma_{k} \right] = \frac{1}{2} \text{Tr} \left[ \sigma_{j}\sigma_{k} \right] = \delta_{jk}.$$  

It follows that the von Neumann entropy of $\phi(t)$ is defined by the distribution $p_{k}(t)$

$$S(\phi(t)) = - \sum_{k=0}^{N} p_{k}(t) \ln p_{k}(t).$$  

The reduced states of $\phi(t)$ to the subsystems $S$ and $A$ are maximally mixed: $\rho_{S}(t) = Tr_{A}[\phi(t)] = \mathbb{1}_{S}/d$ and $\rho_{A}(t) = Tr_{S}[\phi(t)] = \mathbb{1}_{A}/d$. Thus, the only evolving component of the mutual information of $\phi(t)$ is given by (38): $I(\phi(t)) = 2 \log_{2} d - S(\phi(t))$. The time derivative of this quantity is

$$\frac{d}{dt} I(\phi(t)) = \sum_{k=0}^{N} \frac{dp_{k}(t)}{dt} \ln p_{k}(t) + 1$$

$$= \frac{dp_{0}(t)}{dt} \ln p_{0}(t) + 1 + \sum_{k=1}^{N} \frac{dp_{k}(t)}{dt} \ln p_{k}(t) + 1$$

$$= \sum_{k=1}^{N} \frac{dp_{k}(t)}{dt} \ln p_{k}(t) \frac{p_{k}(t)}{p_{0}(t)}.$$  

It follows that, if $\frac{d}{dt} p_{k}(t) \geq 0$ and $p_{k}(t) \geq p_{0}(t)$ for $k = 1, 2, \ldots, N$, we cannot witness any backflow of mutual information with a maximally entangled state since $\frac{d}{dt} I(\phi(t)) \leq 0$.

Finally, we consider the qubit case, namely when $d = 2$. From Eqs. (17) and (13) it follows that the conditions $p_{0}(t) \geq p_{k}(t)$, for $k = x, y, z$ and $t \geq 0$, are always satisfied. We conclude that, when $S$ and $A$ are qubits, if $\frac{d}{dt} p_{k}(t) \geq 0$ for any $k = x, y, z$ and $t \geq 0$, we cannot obtain any backflow of mutual information if the initial state is maximally entangled.

We notice that in a time interval where $\frac{d}{dt} p_{k}(t) \geq 0$ the dynamics is P-divisible, but not necessarily CP-divisible. Thus, there are some cases of non-Markovian P-divisible qubit dynamics that cannot be witnessed by the mutual information of an evolved maximally entangled state. In order to prove this result, we write the time derivative of $p_{k}(t)$

$$\frac{d}{dt} p_{k}(t) = \frac{1}{2} \left( (\gamma_{x}(t) + \gamma_{y}(t)) A_{xy}(t) + (\gamma_{z}(t) + \gamma_{y}(t)) A_{xz}(t) \right.$$

$$\left.- (\gamma_{z}(t) + \gamma_{x}(t)) A_{yz}(t) \right).$$  

Similarly, we can write $\frac{d}{dt} p_{y}(t)$ and $\frac{d}{dt} p_{z}(t)$. We notice that $\frac{d}{dt} p_{x}(t) + \frac{d}{dt} p_{y}(t) = (\gamma_{x}(t) + \gamma_{y}(t)) A_{xy}(t)$. Therefore, given the positivity of $A_{xy}(t), A_{yz}(t)$ and $A_{xz}(t)$ (see Eq. (13)), if $\frac{d}{dt} p_{k}(t) \geq 0$ for $k = x, y, z$, the dynamics is P-divisible (in this case the conditions given in Eqs. (15) are automatically satisfied). However, in general the converse is not true. Indeed, in Sec. [V B 1] we study two similar P-divisible evolutions for qubits where in the first the conditions $\frac{d}{dt} p_{k}(t) \geq 0$ for $k = x, y, z$ are not satisfied, while in the second they are.

In order to obtain an intuitive meaning of the conditions presented in this section, we look at the dynamics given in Eq. (39) for random unitary evolutions. We notice that $p_{0}(t)$ represents the fraction of $\Lambda_{t}$ that acts as the identity map on $\rho_{S}(0)$. Therefore, if the value of $p_{0}(t)$ is increasing for some $t$, i.e., $\frac{d}{dt} p_{0}(t) > 0$, it is reasonable to expect that at time $t$ the system $S$ is getting closer to its initial configuration $\rho_{S}(0)$ and therefore evolving under a non-Markovian evolution that can be witnessed. Conversely, since $\sum_{k=0}^{N} p_{k}(t) = 1$, if $\frac{d}{dt} p_{0}(t) \geq 0$ for any $k \neq 0$, it follows that $\frac{d}{dt} p_{0}(t) \leq 0$. We expect that in this situation, where the overlap of $\rho_{S}(t)$ with its initial configuration decreases, $S$ undergoes an evolution that cannot be distinguished from a Markovian one.

1. Example: the quasi-eternal non-Markovian model

In the last section we gave a set of conditions for non-Markovian random unitary dynamics such that, if satisfied, cannot be witnessed evolving maximally entangled states. In the case of qubits, these conditions are given in terms of the time derivative of the probability distribution $\{p_{k}(t)\}_{k}$ that defines $\Lambda_{t}$ (see Eq. (9)). In this section we consider two examples. First, we consider a model that does not satisfy the conditions given in Section [V B] and we check if, apart from the maximally entangled states, the evolution of random pure states can provide any backflow. Secondly, we consider a model that satisfies these conditions, i.e., cannot be witnessed by any initially maximally entangled state, and we notice that the evolution of random pure initial states do not provide any backflow of mutual information.

First of all, we consider the qubit random unitary dynamics given in Section [V C 2] where $a = 2/5$ and $t_{0} = 1$. This example satisfies the condition of physicality given by Eq. (21). Indeed, $\lim_{t \to \infty} B_{xyz}^{(2/5,1)}(t) \approx 0.146 > 0$. Moreover, the physicality of this model is shown by the positivity of the time-dependent distribution $\{p_{k}(t)\}_{k}$, for $k = x, y, z$, that defines the corresponding random unitary evolution

$$p_{x}(t) = p_{y}(t) = \frac{1}{4} \left( 1 - e^{-4t/5} \right),$$

$$p_{z}(t) = \frac{1}{4} \left( 1 + e^{-4t/5} - 2e^{-2i/5} \frac{\cosh(t-1)}{\cosh(1)} \right)^{2/5},$$

where $p_{0}(t) = 1 - p_{x}(t) - p_{y}(t) - p_{z}(t) \geq 0$ and $p_{0}(0) = 1$. We define $t_{NM}(\rho_{SA}(0))$ the time when the backflow of mutual information starts if the initial state considered is $\rho_{SA}(0)$. 

We evaluated $t_{NM}(\rho_{SA}(0))$ for $2 \cdot 10^4$ pure random states
$
\rho_{SA}(0) = |\psi_{SA}\rangle\langle\psi_{SA}|$
of the form
\[
|\psi_{SA}\rangle = a_0|00\rangle_S + a_1|01\rangle_S + a_2|10\rangle_S + a_3|11\rangle_S,
\]
where the parameters $a_i$ are normalized complex random numbers. The minimum value of $t_{NM}(\rho_{SA}(0))$ obtained has been $t_{NM}(\rho_{SA}(0)) \approx 2.404$, where the values of the parameters that generates $\rho_{SA}(0)$ are characterized by: $|a_1| \approx |a_2| \approx |a_4| \approx 0$ and $|a_3| \approx 1$ up to local unitary operations on $A$. Our numerical analysis does not give any insight about the possible existence of a class of initial states for which the corresponding $t_{NM}(\rho_{SA}(0))$ is arbitrarily close to $t_0 = 1$, i.e., the earliest time for which the intermediate map $V_{n,e}$ is P but not CP. If there exist pure states with $t_{NM}$ closer to $t_0 = 1$, they must belong to a small subset that we did not sample. We point out that for this model, while $\frac{d}{dt}p_x(t) \geq 0$ and $\frac{d}{dt}p_y(t) \geq 0$ for any $t \geq 0$, $\frac{d}{dt}p_z(t) < 0$ for $t > 1.3254$. Indeed, the evolution of a maximally entangled state $|\phi^+\rangle_{SA} = (|00\rangle_S + |11\rangle_S)/\sqrt{2}$ shows a backflow of mutual information with $t_{NM}(|\phi^+\rangle\langle\phi^+|) \approx 2.741$.

Finally, we have studied the eternal non-Markovian model given by $\alpha = 1$ and $t_0 = 0$. Interestingly, in this case $\frac{d}{dt}p_x(t) \geq 0$, $\frac{d}{dt}p_y(t) \geq 0$ and $\frac{d}{dt}p_z(t) \geq 0$ for any $t \geq 0$. Therefore, this model satisfies the conditions given in Section IV B i.e., maximally entangled states that are evolved by this dynamics do not show backflows of mutual information. In our numerical simulations no backflow of mutual information has been observed evolving $10^3$ random pure initial states.

C. Non-maximally entangled states help to detect non-Markovian dynamics

The purpose of this section is to examine, through a concrete example, how the use of non-maximally entangled states can be of benefit for the detection of non-Markovian effects. This example also serves as an illustration of the witnessing potential of the mutual information for non-Markovian dynamics that are non-unital and not P-divisible.

The model we consider is a generalized amplitude damping channel (GADC) with two time dependent parameters, defined by the following set of Kraus operators
\[
K_1(t) = \sqrt{s(t)} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{r(t)} \end{pmatrix},
K_2(t) = \sqrt{s(t)} \begin{pmatrix} 0 & \sqrt{1-r(t)} \\ 0 & 0 \end{pmatrix},
K_3(t) = \sqrt{1-s(t)} \begin{pmatrix} \sqrt{r(t)} & 0 \\ 0 & 1 \end{pmatrix},
K_4(t) = \sqrt{1-s(t)} \begin{pmatrix} 0 & \sqrt{1-r(t)} \\ \sqrt{1-r(t)} & 0 \end{pmatrix},
\]
where $s(t) = \cos^2(5t)$ and $r(t) = e^{-t}$. The evolution induced by these operators is equivalent to that described by a generator $L'_t$, of the form given in Eq. (2) with $G_- = 3\gamma_-$ and $G_+ = 3\gamma_+$ and the respective time-dependent rates
\[
\gamma_-(t) = \cos^2(5t) - 5(1-e^{-t})\sin(10t),
\gamma_+(t) = \sin^2(5t) + 5(1-e^{-t})\sin(10t),
\]
for which the following equality holds
\[
\gamma_-(t) + \gamma_+(t) = 1.
\]

In order to understand the non-Markovian behaviour of this model, it is possible to calculate the function $g(t)$ introduced in [8]
\[
g(t) = \lim_{dt \to 0} \frac{\| (I_A \otimes V_{i\sigma}(t)) (\phi_{AS}) \|_1 - 1}{dt},
\]
where $\phi_{AS}$ is the maximally entangled state $|\phi^+\rangle_{SA}$ and $\| \cdot \|_1$ represents the trace norm. Indeed it is positive if and only if $V_{i\sigma}(t)$ is not CP. In our case $g(t) > 0$ if and only if either $\gamma_-(t)$ or $\gamma_+(t)$ is negative
\[
g(t) = \frac{1}{2} \sum_{i=\pm} |\gamma_i(t)| - \gamma_i(t) = \begin{cases} \gamma_-(t) & t \in T^- \\ \gamma_+(t) & t \in T^+ \\ 0 & \text{otherwise} \end{cases},
\]
where $T^- \equiv \{ t : \gamma_-(t) < 0 \}$ are two non-overlapping sets of time intervals. We can define $T^+$ as the union of the time intervals $T^+_k \equiv (t_{f_{in,k}}, t_{f_{in,k}})$ when the rate $\gamma^-(t)$ is negative. If we do the same with $T^+$, we can write
\[
T^+ \equiv \bigcup_{k=1}^{\infty} T^+_k = \bigcup_{k=1}^{\infty} (t_{f_{in,k}}, t_{f_{in,k}}).
\]
In Ref. [19] the authors compare the ability of mutual information and entanglement of formation to witness non-Markovianity when a maximally entangled state is shared between $S$ and $A$ for the considered dynamics. They note that the mutual information does not show any backflow during the first time interval where dynamics is not CP-divisible, i.e., for $t \in T^-$, while the entanglement of formation shows a backflow in a time interval that is a proper subset of $T^-$. However, in order to fairly compare the witnessing potential of two different correlation measure, we must consider any possible initial state.

From now on we will focus on detecting backflows of mutual information for times $t \in T^- = (0.13437, 0.31416)$: we give numerical results indicating that the mutual information can detect backflows for any $t \in T^-$. We consider as initial states the pure states
\[
|\psi^-\rangle(e) \equiv \sqrt{1-e^2} |00\rangle_S + e |11\rangle_S,
\]
and provide strong evidence that these states provide backflows of mutual information for any time $t \in T^-$ for $e \to 0^+$. More precisely, we have observed backflows in the mutual information for $t \in T^- \subset T^-$, where $T^- \equiv (t_{f_{in,k}}, t_{f_{in,k}})$, $\delta T^-$, and $\delta T^- \equiv 10^{-10}$.

We consider $\rho^-(0, e) = |\psi^-\rangle(e)\langle\psi^-|e(e)$ as the initial state of our complete system and we study its evolution $\rho^-(t, e) = I_A \otimes \Lambda_T(\rho^-(0, e))$, where $\Lambda_T$ represents the GADC described above. In Fig. we show the behaviour of $I(\rho^-(t, e))$ for several values of $e$. We notice that as $e$ approaches zero, the time interval where $I(\rho^-(t, e))$ is increasing widens and approaches
The leading term of the expansion is positive. For \( \delta \tau > 0 \), the leading order term of the series expansion of \( I(\rho^\tau(t, \epsilon)) \) is negative for any \( \beta \), and the resulting sign of the leading order term of the expansion is negative for any \( \beta \). For \( \delta \tau > 0 \) both \( \beta(\delta \tau) \) and \( \alpha(\delta \tau) \) are negative and therefore, for sufficiently small \( \epsilon \), the leading order term of the expansion is positive. For \( \delta \tau = 0 \) the value of \( \beta(\delta \tau) \) is zero to within numerical precision.

\[ \delta \tau = T^{-1} , \] while the amplitude of the mutual information decreases. The latter effect, and the increasingly small values of \( \epsilon \), makes it difficult to numerically verify the possibility to witness a backflow of mutual information for \( \tau \in T^{-1} \) arbitrarily close to \( t_{in,1} \) and \( t_{in,1}' \).

To better understand the behaviour of \( I(\rho^\tau(t, \epsilon)) \) when \( \tau = t_{in,1} \), we consider a series expansion in \( \epsilon \) of the time derivative of this quantity for times close to the beginning of \( T^{-1} \), e.g., for \( |\delta \tau| \equiv |\tau - t_{in,1}'| \leq 10^{-6} \). We find

\[ \frac{d}{dt} I(\rho^\tau(t_{in,1} + \delta \tau, \epsilon)) = (\alpha(\delta \tau) + \beta(\delta \tau) \ln(\epsilon)) \epsilon^2 + O(\epsilon^3) . \]

This expansion (Fig. 5) is characterized by \( \alpha(\delta \tau) < 0 \) and the relation \( \mathrm{sign}(\beta(\delta \tau)) = -\mathrm{sign}(\delta \tau) \), which has been verified up to \( \delta \tau = \pm 10^{-10} \). An analogous result is obtained when \( \tau = t_{in,1}' \). For this case \( \alpha(\delta \tau) \) is negative and \( \mathrm{sign}(\beta(\delta \tau)) = \mathrm{sign}(\delta \tau) \).

In summary, the numerical analysis indicates that for any value of \( \epsilon < 0 \), the fraction of the time interval of \( T^{-1} \) for which \( \frac{d}{dt} I(\rho^\tau(t, \epsilon)) > 0 \) is smaller than the one that we have for \( T^{-1} \) when \( \epsilon > 0 \). Similar results are obtained for \( T^{-1} \) if we consider the initial state \( \rho^\tau(0, \epsilon) \equiv |\psi^\tau(\epsilon)\rangle \langle \psi^\tau(\epsilon)| \) for \( \epsilon > 0 \) that approaches to zero, where \( |\psi^\tau(\epsilon)\rangle = e(0)_{SA} + \sqrt{1 - \epsilon^2}|11\rangle_{SA} \). We underline that we cannot perform this numerical analysis for each \( T^{-1} \) since there is an infinite number of such intervals.

The results in this section demonstrate that it appears difficult to fully determine when a given non-Markovian dynamics experiences correlation backflow in terms of the mutual information, as one needs to consider all possible initial states. In fact, to our knowledge, it is not even clear if one can restrict the study to pure states. Despite all these difficulties, in the next sections we construct examples of non-Markovian dynamics where it can be proven that no backflow in the mutual information takes place.

D. Non-Markovian dynamics with no backflows of mutual information

We have analyzed several situations where correlations backflow as measured by the mutual information detect non-Markovianity, including explicit examples of non-P-divisible, P-divisible, unital and non-unital non-Markovian evolutions. In this section we construct, using the ideas introduced in [28], examples of non-Markovian evolutions for qubits that cannot be witnessed by backflows of mutual information. The key feature for the construction of these examples is that for large enough times all evolved states are contained in a neighborhood of the stationary state, that is, by choosing a large enough evolution time, the image of the dynamical map can be made as close as desired to its stationary state. Moreover, there are examples of evolutions in which it is possible to tune when non-Markovian effects take place. In particular, they can be made arbitrarily close to the stationary state. To understand the presence of correlation backflows for these maps it is enough to perform a perturbative study of the chosen correlation measure, the mutual information in this case, around the stationary state. This hugely simplifies the analysis and allows us to conclude the absence of correlation backflow in the mutual information for some non-Markovian dynamics. Examples of such maps are for instance shown in Sections V, VI and VII for P-divisible evolutions of qubits.
1. Example: the quasi-eternal non-Markovian model

We start by presenting examples of channels in which it is possible to tune when non-Markovian effects take place. In Section II C 2 we introduced a class of non-Markovian random unitary evolutions \([7] \) where the rates that define the corresponding \(P\)-divisible dynamical maps, see Eqs. (10), are given by Eq. (18), i.e.,

\[
\{\gamma_1(t), \gamma_2(t), \gamma_3^{(h)}(t)\} = \frac{1}{5}\{1, 1, -\tanh(t - t_0)\}. \tag{46}
\]

where we fixed \(\alpha = 2/5\).

We define \(\Lambda^{(h)}_{06}\) to be the qubit dynamical map induced by the rates given in Eq. (46). Moreover, let \(\nu^{(h)}_{\alpha_1,\alpha_2}\) be the intermediate map of \(\Lambda^{(h)}_{06}\) in the time interval \((t_1, t_2)\). We define the image of the dynamical map \(\Lambda^{(h)}_{06}\), i.e., the set of the accessible states of the evolution at time \(t\), to be

\[
\text{Im}(\Lambda^{(h)}_{06}) = \{\sigma \in S(\mathcal{H}_S) : \sigma = \Lambda^{(h)}_{06}(\rho) \text{ for some } \rho \in S(\mathcal{H}_S)\}. \tag{47}
\]

In Section V B 1 we showed that \(t_0 = 1\) and \(\alpha = 2/5\) define a physical evolution. Indeed, from Eq. (21), \(\Lambda^{(h)}_{06}\) is a CPTP map for any \(t \geq 0\) and \(t_0 \geq 7^{2/5} \approx 0.769\).

The structure of this section is the following. First, we pick \(t_0 = t_0'\) to define \(\Lambda^{(G)}_{s}\) and we consider an intermediate map \(\overline{V}\) that occurs in the time interval \((t_1, t_2)\), where \(t_1 > t_0'\). Therefore, we consider a larger value of \(t_0\), e.g., \(t_0'' > t_0\), and we show that the same \(\overline{V}\) occurs in a time interval which is delayed by the factor \(\Delta t_0 = t_0' - t_0''\), i.e., in \((t_1 + \Delta t_0, t_2 + \Delta t_0)\).

Secondly, once we understand that the value of \(t_0\) the same intermediate map \(\overline{V}\) can be delayed as much as needed, we study the image of the dynamics for arbitrary large values of \(t_0\). The final purpose is to show that, increasing enough the value of \(t_0\), \(\text{Im}(\Lambda^{(h)}_{06})\) is contained in an arbitrarily small neighborhood of the stationary state.

Following the steps we have described, we start considering the dynamical maps \(\Lambda^{(G)}_{t_0'}\) and \(\Lambda^{(G)}_{t_0''}\), where \(1 \leq t_0' < t_0''\) and \(\Delta t_0 \equiv t_0' - t_0'' > 0\). From Eqs. (10) and (11), since the rates that define the two evolutions differ by a simple time-shift (see Eq. (46)), we can express the intermediate maps of \(\Lambda^{(G)}_{t_0''}\) starting from time \(\Delta t_0\) in terms of the dynamical map \(\Lambda^{(G)}_{t_0'}\). Indeed, for \(\Delta t_0 < t\)

\[
\Lambda^{(G)}_{t_{0''},t_{0'}} = V^{(G)}_{t_{0''},t_{0'}}. \tag{48}
\]

As a consequence, if \(\Delta t_0 < t\), the dynamical map \(\Lambda^{(G)}_{t_0''}\) itself can be expressed as the composition of \(\Lambda^{(G)}_{\Delta t_0}\) and \(\Lambda^{(G)}_{t_0''-\Delta t_0}\)

\[
\Lambda^{(G)}_{t_0''} = \Lambda^{(G)}_{t_0''-\Delta t_0} \Lambda^{(G)}_{\Delta t_0}. \tag{49}
\]

Indeed, composing Eqs. (48) and (49), we obtain the definition of the intermediate map \(V^{(G)}_{t_{0''},t_{0'}}\), i.e., \(\Lambda^{(G)}_{t} = V^{(G)}_{t_{0''},t_{0'}} \Lambda^{(G)}_{\Delta t_0}\).

Now, we consider a time interval \((t_1', t_2')\) such that \(t_0' < t_1' < t_0''\) in this time interval, \(\gamma^{(G)}_{t_1'}(t) < 0\) and the intermediate map \(\overline{V} \equiv V^{(G)}_{t_{1'},t_{1''}}\) is P but not CP (see Section II C). From the results obtained above, it is clear that the action of \(\overline{V}\) can also be obtained as the intermediate map of \(\Lambda^{(G)}_{t_0''}\) that occurs in the shifted time interval \((t_1', t_2' + \Delta t_0)\). Therefore, we have the identity

\[
\overline{V} = V^{(G)}_{t_{1'},t_{1''}} = V^{(G)}_{t_{1'},t_{1''} + \Delta t_0}. \tag{50}
\]

Therefore, if we want to witness the non-Markovian effect of \(\overline{V}\) while we consider \(\Lambda^{(G)}_{t_0''}\), its action occurs from time \(t_1'\) instead, considering \(\Lambda^{(G)}_{t_1'}\), \(\overline{V}\) occurs from the delayed time \(t_1'' = t_1' + \Delta t_0\).

Having established that, for both \(\Lambda^{(G)}_{t_1'}\) and \(\Lambda^{(G)}_{t_0''}\), the intermediate map of the evolution that follows for a time interval that lasts \(t_2 - t_1\) is \(\overline{V}\), we proceed checking the images of the respective preceding evolutions, i.e., \(\text{Im}(\Lambda^{(G)}_{t_0''})\) and \(\text{Im}(\Lambda^{(G)}_{t_0'})\). The difference between the images of the two maps is given by the contractive action of \(\Lambda^{(G)}_{\Delta t_0}\).

In order to understand the effect of this map, since \(\Delta t_0 < t_0''\), the rates that define \(\Lambda^{(G)}_{\Delta t_0}\) through Eqs. (10) are positive in the time interval \((0, \Delta t_0)\). Therefore, the action of \(\Lambda^{(G)}_{\Delta t_0}\) is a "global" contraction, i.e., using Eq. (16) for \(\Lambda^{(G)}_{\Delta t_0}\),

\[
\Lambda^{(G)}_{\Delta t_0}(\sigma_i) = \lambda^{(G)}_{\Delta t_0}(\Delta t_0)\sigma_i, \tag{51}
\]

where \(\lambda^{(G)}_{\Delta t_0}(\Delta t_0) < 1\) for \(i = x, y, z\). Moreover, since \(\Lambda^{(G)}_{t_1'}\) is CP, we can write

\[
\Lambda^{(G)}_{t_1'}(\sigma_i) = \lambda^{(G)}_{t_1'}(t_1')\sigma_i, \tag{52}
\]
where $\lambda^{(i)}(t_i') < 1$ for $i = x, y, z$. Considering Eqs. [51] and [53] in Eq. [49], we obtain

$$\Lambda^{(G)}_{t_1'}(\sigma_i) = \lambda^{(i)}(t_1')\sigma_i = \lambda^{(i)}(t_1')\lambda^{(G)}_{t_1'}(\Delta t_0)\sigma_i,$$  \hspace{1cm} (53)

where we remember that $t_1' = t_1' + \Delta t_0$. Since $\lambda^{(i)}(t_1') < \min\{\lambda^{(i)}(t_1'), \lambda^{(G)}_{t_1'}(\Delta t_0)\}$, we conclude that

$$\text{Im}(\Lambda^{(G)}_{t_1'}) = \text{Im}(\Lambda^{(G)}_{t_1'}(\Delta t_0)) \subset \text{Im}(\Lambda^{(G)}_{t_1')},$$  \hspace{1cm} (54)

The next logical step is to fix the value of $t'_0 = 1$ and increase $t'_0$ (and consequently $\Delta t_0 = t'_0 - t'_0$) in order to, at the same time, delay the occurrence of $V$ for $\Lambda^{(G)}_{t_1'}$ and show that we can make the evolution that precedes its action more and more contractive (see Fig. 6). Indeed, as showed above, if $t_1'$ is the time from which $V$ occurs for $\Lambda^{(G)}_{t_1'}$, the same intermediate map (see Eq. [50]) occurs for $\Lambda^{(G)}_{t_1'}$ from time $t'_0 = t_1' + \Delta t_0$. Therefore, increasing $t'_0$, we delay more and more the time of occurrence $t_0'$ of the intermediate $P$ map $V = V^{(G)}_{t_1'}$.

Now, we prove that for any $\epsilon > 0$ there exists a value of $t'_0$ such that the action of $\Lambda^{(G)}_{t_1'}$ is characterized by $\lambda^{(i)}(t'_0) < \epsilon$ for each $i = x, y, z$. From Eqs. [10] and [46], it is easy to check that $\lambda^{(i)}(t) = \lambda^{(i)}(t') > \lambda^{(i)}(t)$ for any $t > 0$. Indeed,

$$\lambda^{(i)}(t') = \left(2e^{-2\gamma(t)}/\cosh(\gamma(t))/\cosh(\gamma(t))\right)^{1/3}.$$  \hspace{1cm} (55)

Therefore, for any $\epsilon > 0$, if the following condition is satisfied

$$t'_0 > \log \sqrt{2/\epsilon^3},$$  \hspace{1cm} (57)

we have $\lambda^{(i)}(t'_0) < \lambda^{(i)}(t'_0) = \lambda^{(i)}(t'_0) < \epsilon$.

We want to understand the effects on the set of accessible states of $\Lambda^{(G)}_{t_1'}$ that we obtain when $t'_0$ is increased over the bound given by Eq. [57]. Therefore, we consider a generic initial state $\rho_S(0) = (I_S + V(0)\sigma)/2$, represented by the Bloch vector $v(0) = (v_x(0), v_y(0), v_z(0))$, where in the vector $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ we collect the Pauli matrices. We evolve this qubit state with the dynamical map $\Lambda^{(G)}_{t_1'}$, where the condition of Eq. [57] is satisfied for some $\epsilon > 0$. At time $t = t'_0$ the Bloch vector is evolved to $v(t'_0) = (v(x(t'_0), v_y(t'_0), v_z(t'_0))$. From Eq. [55] and [56], it is straightforward to show that $\max(v(t'_0)) < \epsilon$ max $v(t)$. In particular

$$\|\rho_S(t'_0) - I_S/2\| = \frac{1}{2} \sum_{i} v_i^2(t'_0) < \epsilon/2.$$  \hspace{1cm} (58)

In other words, $\text{Im}(\Lambda^{(G)}_{t_0})$ is inside a neighbor of radius of order $\epsilon$ centered in the maximally mixed state $I_S/2$, namely the stationary state of the evolution. We can pick arbitrary small values of $\epsilon$ by considering arbitrary large values of $t'_0$ (and consequently $t'_0 = t'_0 + t'_0 - t'_0$), i.e., the time after which $v_i(t'_0) < \epsilon$ and consequently $\Lambda^{(G)}_{t_1'}$ ceases to be CP-divisible.

Finally, we consider the mutual information $I$ of two qubits $A$ and $S$, where $A$ is an ancilla and $S$ is evolved by the dynamical map $\Lambda^{(G)}_{t_1'}$. In this scenario, we want to witness a backflow of $I$ in the time interval $(t'_1, t'_2)$, i.e., when the $P$ intermediate map $V$ evolves $S$. Therefore, we increase the value of $t'_0$ until the image $\text{Im}(\Lambda^{(G)}_{t_1'}(I_A \otimes I_S I_{t_0}))$ of the evolution that precedes the action of $I_A \otimes I_S$ is in a neighbor of radius $\epsilon$ of the stationary states of the dynamics, namely $\rho_A \otimes I_S/2$, where $\rho_A$ is any state of $S(S_A)$. At time $t'_1$ we have that

- The states in $\text{Im}(\Lambda^{(G)}_{t_1'})$ are $\epsilon$-close from the stationary states of the evolution.
- The evolution in the following time interval $(t'_1, t'_2)$ is described by the $P$, but not $CP$, intermediate map $I_A \otimes V$.
- The previous point holds for arbitrary small time intervals $(t'_2 - t'_1)$.

To sum up, there exist dynamics in which non-Markovian effects take place only arbitrarily close to the stationary state. For these cases, to understand the existence of correlation backflows in the mutual information it suffices to perform a perturbative analysis of this quantity around the stationary state. This is precisely what we do in what follows, Sections 3.6 and 4.7 for a random unitary dynamics for qubits. The calculations are lengthy and can be skipped by readers not interested in the technical details, as they are not needed for the rest of the article. However, this perturbative analysis allows us to conclude that the studied non-Markovian dynamics do not present any increase of the mutual information for any initial state in $S(S_A \otimes H_S)$.

E. Taylor expansion of the time derivative of the mutual information

To study the time dependence of the mutual information perturbatively we here outline how the time derivative of the mutual information $\frac{d}{dt} I(\rho(t)) \equiv \frac{d}{dt} I(\bar{a}, V(s), s)$ in a neighbourhood of a given state $\rho$ can be described by a Taylor expansion in the coordinates $a_i$. In particular we consider Taylor expansions as a tool to investigate the neighborhoods of stationary states.

The mutual information as a function on the set of states $S(S_A \otimes H_S)$ is analytic for all states $\rho$ of full rank, i.e., everywhere in the interior of the set of states, here denoted int$(S(S_A \otimes H_S))$. Thus, in any open neighborhood $U \subset \text{int}(S(S_A \otimes H_S))$ of a state $\rho$ the mutual information equals its Taylor series and we can use Taylor expansions to analyze it perturbatively around $\rho$. Moreover, if the dynamics is differentiable the time derivative of any analytic correlation measure is analytic as well.
Proposition 3. Let \( M \) be a correlation measure that is analytic at \( \bar{a} \). If \( \frac{d}{dt} M(\bar{a}, t) \bigg|_{t = s} \) is well defined it follows that \( \frac{d}{dt} M(\bar{a}, t) \bigg|_{t = s} \) is analytic at \( \bar{a} \).

Proof. The time derivative \( \frac{d}{dt} M(\bar{a}, t) \bigg|_{t = s} \) is expressed as \( \frac{d}{dt} M(\bar{a}, V_{s,t}) \bigg|_{s,t} \) can be written as \( \sum_{i,j} a_j \frac{dV_{i,s}(t)}{dt} \bigg|_{s,t} \frac{dM(\bar{a}, t)}{dt} \). Assume that \( \frac{dV_{i,s}(t)}{dt} \bigg|_{s,t} \) is well defined for each \( i, j \). Then, since products, linear combinations, and derivatives of analytic functions are analytic it follows that \( \frac{d}{dt} M(\bar{a}, t) \) is analytic as a function of \( \bar{a} \).

Thus, in particular, if \( V_{s,t} \) is differentiable \( \frac{d}{dt} I(\rho(t)) \) can be described perturbatively in any open neighborhood of a state in \( \text{int}[S(\mathcal{H}_A \otimes \mathcal{H}_B)] \) by a Taylor expansion. On the other hand, for states of less than full rank, i.e., states on the boundary of \( S(\mathcal{H}_A \otimes \mathcal{H}_B) \), the partial derivatives in the coordinates \( \alpha_i \) need not even be well defined to all orders.

Now consider a stationary state \( \rho_0 \) of a linear divisible dynamic described by \( \Lambda \), which is of full rank, i.e., \( \rho \in \text{int}[S(\mathcal{H}_A \otimes \mathcal{H}_B)] \). Since \( \frac{d}{dt} I(\rho_0(t)) = 0 \) the sign of \( \frac{d}{dt} I(\rho(t)) \) in a neighborhood of \( \rho_0 \) is determined by the terms of order greater than zero in the Taylor expansion of \( \frac{d}{dt} I(\rho(t)) \).

In general \( \frac{d}{dt} I(\rho(t)) \) may take both positive and negative values in a neighborhood of \( \rho_0 \). If a neighborhood of \( \rho_0 \) exists where \( \frac{d}{dt} I(\rho(t)) \) is everywhere non-negative, or alternatively everywhere non-positive, depends only on \( \rho_0 \) and on \( \frac{d}{dt} V_{s,t} \bigg|_{s,t} \). In particular, the properties of the neighborhood is independent of the previous dynamic \( \Lambda \) and the properties of the image of \( \Lambda \) since we assumed linear divisibility of the dynamics.

This last observation allows us to make the following two statements about the mutual information. If there is some neighborhood of \( \rho_0 \) where \( \frac{d}{dt} I(\rho(t)) \) is somewhere positive, and this neighborhood is contained in the image of \( \Lambda \), we can observe an increase of the mutual information. If there is some neighborhood of \( \rho_0 \) where \( \frac{d}{dt} I(\rho(t)) \) is non-positive, and the image of \( \Lambda \) is contained in this neighborhood, we can not observe any increase of the mutual information.

1. Neighbourhoods of critical points

In the case that one or more first derivatives are zero one must consider higher order terms of the Taylor expansion to study how the sign of \( \frac{d}{dt} I(\rho(t)) \) behaves in a neighborhood of a stationary state \( \rho_0 \). In particular this is true if all first derivatives with respect to the \( \alpha_i \) are zero, i.e., if \( \rho_0 \) is a critical point of \( \frac{d}{dt} I(\rho(t)) \). The relevance of considering critical points in relation to stationary states can be understood from the following two observations. For any continuously differentiable dynamics a product state in the interior of the set of states is a critical point of \( \frac{dM}{dt} \) if \( M \) is analytic.

Proposition 4. Let \( M \) be a correlation measure that is analytic at a state \( \rho \). Assume that \( V_{s,t} \) is continuously differentiable. Then if \( \rho \in \text{int}[S(\mathcal{H}_A \otimes \mathcal{H}_B)] \) and is a product state it is a critical points of \( \frac{dM}{dt} \).

Proof. See Appendix B.

Thus, in particular, all product states in \( \text{int}[S(\mathcal{H}_A \otimes \mathcal{H}_B)] \) are critical points of \( \frac{d}{dt} I(\rho(t)) \).

For qubit evolutions a stationary state in the interior of the set of states is a critical point of \( \frac{dM}{dt} \) if \( M \) is analytic.

Proposition 5. Let \( M \) be a correlation measure that is analytic at \( \rho \) and let \( \{\Lambda_i\} \) be a continuously differentiable dynamical qubit evolution. If \( \rho \) is in \( \text{int}[S(\mathcal{H}_A \otimes \mathcal{H}_B)] \) is a stationary state of \( \{\Lambda_i\} \) it is a critical point of \( \frac{dM}{dt} \).

Proof. See Appendix C.

Thus, for qubit evolutions all stationary states in \( \text{int}[S(\mathcal{H}_A \otimes \mathcal{H}_B)] \) are critical points of \( \frac{d}{dt} I(\rho(t)) \).

The nature of a critical point \( \rho_0 \) can be analyzed by obtaining the eigenvalues of the Hessian matrix, i.e., the matrix of second derivatives \( H_{i,j} = \frac{d^2}{d\alpha_i d\alpha_j} \frac{d}{dt} I(\rho(t)) \). However, for any stationary state, the Hessian \( H_{i,j} \) is of less than full rank since \( \frac{d}{dt} I(\rho(t)) = 0 \) on the set of stationary states of \( V_{s,t} \), denoted \( S \), and on the set of product states, denoted \( S.P \). From this follows that any eigenvector of the Hessian that is tangent to \( S \cup S.P \) corresponds to a zero eigenvalue. Thus, the sign of \( \frac{d}{dt} I(\rho(t)) \) in the part of the neighborhood of \( \rho_0 \) that coincides with the zero-eigenspace \( E_0 \) of \( H_{i,j} \) cannot be determined from the Hessian matrix alone since it depends on higher order derivatives.

On the overlap of the neighborhood of \( \rho_0 \) with the complement of \( E_0 \), i.e., with \( E_0^C = B(\mathcal{H}_A \otimes \mathcal{H}_B) \setminus E_0 \), the Hessian does describe the sign of \( \frac{d}{dt} I(\rho(t)) \) if the neighborhood is sufficiently small. In particular, if all non-zero eigenvalues of the Hessian, which correspond to eigenvectors tangent to \( E_0^C \), are negative there exists a neighborhood \( U_{\rho_0} \) of \( \rho_0 \) where \( \frac{d}{dt} I(\rho(t)) \) is negative in \( U_{\rho_0} \cap E_0^C \). If all the non-zero eigenvalues of the Hessian are positive there exists a neighborhood \( U_{\rho_0}^+ \) of \( \rho_0 \) where \( \frac{d}{dt} I(\rho(t)) \) is positive in \( U_{\rho_0}^+ \cap E_0^C \).

2. Calculating partial derivatives

Directly calculating the derivatives of \( \frac{d}{dt} I(\rho(t)) \) with respect to the coordinates \( \alpha_i \) can be demanding since the eigenvalues of \( \rho \) are the roots of a polynomial with degree equal to \( \text{dim}(\mathcal{H}_A \otimes \mathcal{H}_B) \). To circumvent this difficulty we calculate the derivatives and second derivatives at a state \( \rho \) using a method adapted from Ref. [20]. The method described in this reference is valid for real symmetric matrices but generalizing it to Hermitian complex matrices is very straightforward. We present the version of this method that works for Hermitian matrices in the following paragraphs.

Let \( f \) be a spectral function defined on a set of \( n \times n \) Hermitian matrices \( A \) parametrized by real numbers \( \alpha_i \). By spectral function we refer to a function that only depends on the eigenvalues \( \{\lambda_k\}_{k=1}^n \) of \( A \) but not on the ordering of the eigenvalues. Furthermore, assume that \( f \) is analytic in a given point \( \bar{a} \) and let \( \lambda_k(\bar{a}) \) be the eigenvalue of \( A(\bar{a}) \) with corresponding normalized eigenvector \( u_k(\bar{a}) \). Then the first and second order partial
derivatives of \( f \) with respect to the parameters \( a_i \) at the point \( \bar{a} \) can be expressed as
\[
\frac{\partial f(\bar{a})}{\partial a_i} = \sum_k \frac{\partial f[\lambda(\bar{a})]}{\partial \lambda_k} h_i^k(\bar{a}),
\]
(59)
and
\[
\frac{\partial^2 f(\bar{a})}{\partial a_i \partial a_j} = \sum_{k,l} \frac{\partial^2 f[\lambda(\bar{a})]}{\partial \lambda_k \partial \lambda_l} h_i^k(\bar{a}) h_j^l(\bar{a})
+ \sum_k \frac{\partial f[\lambda(\bar{a})]}{\partial \lambda_k} h_i^k(\bar{a}) + \eta_{ij}(\bar{a}),
\]
(60)
respectively, where
\[
h_i^k(\bar{a}) = u_k^i \frac{\partial A(\bar{a})}{\partial a_i} - u_k,
\]
\[
h_{ij}^k(\bar{a}) = u_k^i \frac{\partial^2 A(\bar{a})}{\partial a_i \partial a_j} u_k + \sum_{l,k \neq k} \frac{\partial A(\bar{a})}{\partial a_l} \left( \frac{\partial A(\bar{a})}{\partial a_k} - \frac{\partial A(\bar{a})}{\partial a_l} \right),
\]
\[
\alpha_{ij}^k(\bar{a}) = \left( u_k^i \frac{\partial A(\bar{a})}{\partial a_i} u_k \right) \left( u_j^i \frac{\partial A(\bar{a})}{\partial a_j} u_k \right)
+ \left( u_k^j \frac{\partial A(\bar{a})}{\partial a_j} u_k \right) \left( u_i^j \frac{\partial A(\bar{a})}{\partial a_i} u_k \right),
\]
\[
\eta_{ij}(\bar{a}) = \sum_{k,l,k \neq k} \alpha_{ij}^l(\bar{a}) \frac{\partial^2 f[\lambda(\bar{a})]}{\partial \lambda_k \partial \lambda_l}.
\]
(61)

When two or more eigenvalues coincide, the corresponding eigenvectors are not uniquely defined. Nevertheless, the method can still be used since while the expressions given in Eq. 61, e.g. \( \bar{h}_i \), may depend on the choice of eigenvectors, the partial derivatives themselves are independent and can be evaluated using any such choice.

If the diagonal form of \( A \) in the point \( \bar{a} \) and the corresponding eigenvectors \( u_k(\bar{a}) \) are known the method described here can greatly simplify the computation of the partial derivatives.

F. Random unitary dynamics that the mutual information cannot witness

We now show that the mutual information is non-increasing for some cases of random unitary qubit dynamics that are \( P \)-divisible but not \( CP \)-divisible by studying a neighborhood of the stationary states using the methods described in Sect. V.E.

We consider an ancilla that is also a qubit and explicitly introduce coordinates \( a_i \) for \( B(\mathcal{H}_A \otimes \mathcal{H}_S) \) with respect to an orthonormal basis \( |e_i\rangle \) defined by
\[
e_0 = \mathbb{1} \otimes \mathbb{1},
\]
\[
e_1 = \mathbb{1} \otimes \sigma_x,
\]
\[
e_2 = \mathbb{1} \otimes \sigma_y,
\]
\[
e_3 = \mathbb{1} \otimes \sigma_z,
\]
\[
e_4 = \sigma_x \otimes \mathbb{1},
\]
\[
e_5 = \sigma_y \otimes \mathbb{1},
\]
\[
e_6 = \sigma_z \otimes \mathbb{1},
\]
\[
e_7 = \sigma_x \otimes \sigma_x,
\]
\[
e_8 = \sigma_y \otimes \sigma_x,
\]
\[
e_9 = \sigma_y \otimes \sigma_y,
\]
\[
e_{10} = \sigma_y \otimes \sigma_z,
\]
\[
e_{11} = \sigma_y \otimes \sigma_z,
\]
\[
e_{12} = \sigma_z \otimes \mathbb{1},
\]
\[
e_{13} = \sigma_z \otimes \sigma_x,
\]
\[
e_{14} = \sigma_z \otimes \sigma_y,
\]
\[
e_{15} = \sigma_z \otimes \sigma_z.
\]
(62)
where all operators are of the form \( \chi_A \otimes \chi_S \) for \( \chi_A \in B(\mathcal{H}_A) \) and \( \chi_S \in B(\mathcal{H}_S) \). A state \( \rho \) is represented as
\[
\rho = \frac{1}{4} \mathbb{1} \otimes \mathbb{1} + \sum_{i=1}^{15} a_i |e_i\rangle \langle e_i|,
\]
(63)
where \( a_i = \frac{1}{\pi} \text{Tr}(\rho |e_i\rangle \langle e_i|) \).

The analysis of \( \frac{d^2 I(\rho(t))}{dt^2} \) in the neighborhood of the stationary states is done by first considering the states of full rank, i.e., the states in \( \text{int}[S(\mathcal{H}_A \otimes \mathcal{H}_S)] \), where \( \frac{d^2 I(\rho(t))}{dt^2} \) is analytic. There we calculate the second derivatives of \( \frac{d^2 I(\rho(t))}{dt^2} \) at the stationary states and find the eigenvalues of the Hessian matrix. On the subset of states that fall in the zero eigenspace of the Hessian we then directly evaluate \( \frac{d^2 I(\rho(t))}{dt^2} \). Finally, we consider the states of less than full rank and describe the neighborhood of the intersection of the stationary states with the boundary of the set of states.

The stationary states are of the form \( \rho_A \otimes \mathbb{1} / 2 \) for arbitrary \( \rho_A \). Since the stationary states in \( \text{int}[S(\mathcal{H}_A \otimes \mathcal{H}_S)] \) are critical by Propositions 4 and 5 and such that \( \frac{d^2 I(\rho(t))}{dt^2} = 0 \), there exists some sufficiently small neighborhood of the set of stationary states where the second order terms of the Taylor expansion in the \( a_i \) determine the sign of \( \frac{d^2 I(\rho(t))}{dt^2} \) in all directions where the second derivative is non-zero. To simplify the calculation of these derivatives we note that unitary transformations on the ancilla do not change the mutual information and it is therefore sufficient to consider diagonal \( \rho_A \). In other words, the purity of the state of the ancilla is the only parameter that is relevant for our analysis. The diagonal stationary states are of the form \( \frac{1}{2} \mathbb{1} \otimes \mathbb{1} + a_{12} \sigma_z \otimes \mathbb{1} \) for \(-1/4 \leq a_{12} \leq 1/4 \). The states for which \(-1/4 < a_{12} < 1/4 \) are in \( \text{int}[S(\mathcal{H}_A \otimes \mathcal{H}_S)] \) and the states with coordinates \( a_{12} = \pm 1/4 \) are at the boundary of the set of states.

The second derivatives at the diagonal stationary states of full rank were calculated using the method described in Sect. V.E. and the eigenvalues of the Hessian matrix were obtained. The Hessian has six eigenvalues that are zero for all stationary states in \( \text{int}[S(\mathcal{H}_A \otimes \mathcal{H}_S)] \), and for all values of the parameters \( \gamma(t) \). The remaining nine eigenvalues are func-
tions of the parameters \( \gamma_i(t) \) and of \( a_{12} \), and are given by
\[
32[\gamma_i(t) + \gamma_j(t)] \left( \frac{16a_{12}^2 + 1}{16a_{12}^2 - 1} \right),
32[\gamma_i(t) + \gamma_j(t)] \left( \frac{16a_{12}^2 + 1}{16a_{12}^2 - 1} \right),
32[\gamma_i(t) + \gamma_j(t)] \left( \frac{16a_{12}^2 + 1}{16a_{12}^2 - 1} \right),
\]
\[-8[\gamma_i(t) + \gamma_j(t)] \left( \frac{16a_{12}^2 + 1}{16a_{12}^2 - 1} \right) \frac{\tanh(4a_{12})}{a_{12}},
\]
\[-8[\gamma_i(t) + \gamma_j(t)] \frac{\tanh(4a_{12})}{a_{12}},
\]
\[-8[\gamma_i(t) + \gamma_j(t)] \frac{\tanh(4a_{12})}{a_{12}},
\]
\[-8[\gamma_i(t) + \gamma_j(t)] \frac{\tanh(4a_{12})}{a_{12}},
\]
\[-8[\gamma_i(t) + \gamma_j(t)] \frac{\tanh(4a_{12})}{a_{12}},
\]
\[-8[\gamma_i(t) + \gamma_j(t)] \frac{\tanh(4a_{12})}{a_{12}}. \quad (64)
\]

These nine eigenvalues are all non-positive if and only if the conditions in Eq. (15) are satisfied, i.e., if and only if the dynamics is P-divisible. In particular, they are all strictly negative if \( \gamma_i(t) + \gamma_j(t) > 0 \) for all \( i, j \). In this case there thus exists a sufficiently small neighborhood of the stationary states where \( \frac{d}{dt} I(\rho(t)) \) is negative in the intersection of the neighborhood with the complement of the zero eigenspace of the Hessian.

Next, we need to investigate \( \frac{d}{dt} I(\rho(t)) \) on the intersection of a neighborhood around a stationary state with the zero eigenspace of the Hessian. Here we would need higher order terms in the Taylor expansion to determine the sign of \( \frac{d}{dt} I(\rho(t)) \), however on this eigenspace we can evaluate it directly. The zero eigenspace \( E_0(a_{12}) \) as a function of \( a_{12} \), is spanned by the six eigenvectors \((\mathbb{1} + a_{12}\sigma_i) \otimes \sigma_j\) and \( \sigma_i \otimes \mathbb{1} \) for \( i = x, y, z \). These eigenvectors are tangent to the set of product states for all \( a_{12} \), but the plane they span, i.e. \( E_0(a_{12}) \), also contains correlated states. For a given stationary state \( \rho_0 = \frac{1}{4} \mathbb{1} \otimes \mathbb{1} + a_0 \sigma_i \otimes \mathbb{1} \) we can parametrize \( E_0(a_{12}) \). The states in the \( E_0(a_{12}) \) are of the form
\[
\frac{1}{4} \mathbb{1} \otimes \mathbb{1} + \left( \mathbb{1} + 4a_0 \sigma_i \right) \otimes (a_{12}\sigma_x + a_{12}\sigma_y + a_{12}\sigma_z)
+ (a_{12}\sigma_y + a_{12}\sigma_x + a_{12}\sigma_z) \otimes \mathbb{1}. \quad (65)
\]

Note that \( E_0(a_{12}) \) is an invariant subspace of \( V_{1,1} \) for all \( a_0 \) since the dynamics is unital. Thus any state in \( E_0(a_{12}) \) is mapped into a state also belonging to \( E_0(a_{12}) \). Furthermore, \( a_4, a_5 \) and \( a_{12} \) are time independent. Therefore the time derivative of the mutual information \( I \) as a function on \( E_0(a_{12}) \) depends only on the coordinates \( a_4, a_5, a_{12} \). Since the mutual information is independent of unitary transformations on the system we can diagonalize \( a_{12}\sigma_x + a_{12}\sigma_y + a_{12}\sigma_z \) without changing its value. Let \( \pm \lambda(s) = \pm \sqrt{a_4^2(s) + a_5^2(s) + a_{12}^2(s)} \) be the corresponding eigenvalues as functions of time where
\[
\begin{align*}
 a_1(s) &= a_1 \exp \left[ - \int_{t_0}^t (\gamma_i(\tau) + \gamma_j(\tau)) d\tau \right],
 a_2(s) &= a_2 \exp \left[ - \int_{t_0}^t (\gamma_i(\tau) + \gamma_j(\tau)) d\tau \right],
 a_3(s) &= a_3 \exp \left[ - \int_{t_0}^t (\gamma_i(\tau) + \gamma_j(\tau)) d\tau \right]. \quad (66)
\end{align*}
\]

Since the only time dependence of \( I \) is its dependence on \( \lambda(t) \) we can express the time derivative of the mutual information as
\[
\frac{dI[\rho(t)]}{dt} = \frac{dI[\rho(s)]}{ds} \frac{d\lambda(s)}{ds} \bigg|_{s=t}
\]
for any \( \rho(t) \in E_0(a_{12}) \), where \( \frac{d\lambda(s)}{ds} \bigg|_{s=t} \) is given by
\[
\lambda_1^2 \left[ \gamma_i(t) + \gamma_j(t) \right] + \lambda_2 \left[ \gamma_i(t) + \gamma_j(t) \right] + \lambda_3 \left[ \gamma_i(t) + \gamma_j(t) \right].
\]

When the conditions in Eq. (15) are satisfied, i.e., when the dynamics is P-divisible, \( \frac{d\lambda(s)}{ds} \bigg|_{s=t} \) is non-positive for all \( a_1, a_2, a_3 \). This is equivalent to stating that the length of the Bloch vector of the reduced state of the system does not increase when the dynamics is P-divisible. Moreover, we see that for all states in \( E_0(a_{12}) \) except those of the form \( \rho_A \otimes \mathbb{1} \), for which \( a_1 = a_2 = a_3 = 0 \), there exists some CP-divisible random unitary dynamics such that \( \frac{d\lambda(s)}{ds} \bigg|_{s=t} < 0 \). Since we know that \( \frac{d\lambda(s)}{ds} \bigg|_{s=t} \leq 0 \) for all \( \rho(t) \in E_0(a_{12}) \) when the dynamics is CP-divisible it follows that \( \frac{d\lambda(s)}{ds} \bigg|_{s=t} \) is non-negative for all \( \rho(s) \in E_0(a_{12}) \) not of the form \( \rho_A \otimes \mathbb{1} \). Therefore we can conclude that \( \frac{d\lambda(s)}{ds} \bigg|_{s=t} \leq 0 \) for all \( \rho(t) \in E_0(a_{12}) \) when \( V_{1,1} \) is P-divisible.

In the above analysis we have seen that there exists random unitary non-Markovian P-divisible dynamics for which no increase of the mutual information occurs in a sufficiently small neighborhood of the stationary states in \( \text{int}[S(H_1 \otimes H_2)] \). It remains to consider the neighbourhood of the stationary states of full rank, i.e., of stationary states in the boundary of the set of states. For these stationary states \( a_{12} = \pm 1/4 \) and they are thus of the form \( 1/4(\mathbb{1} \pm \sigma_i) \otimes \mathbb{1} \) and \( 1/4(\mathbb{1} - \sigma_i) \otimes \mathbb{1} \). It is sufficient to consider restricted neighborhoods of these states where the coordinate \( a_{12} \) is held fixed at \( 1/4 \) or \(-1/4\) respectively. Any other point in their neighborhoods either belongs to a neighbourhood of a stationary state in \( \text{int}[S(H_1 \otimes H_2)] \), or is unphysical. The physical states in these restricted neighbourhoods for which \( a_{12} = \pm 1/4 \) are product states of the form \( 1/4(\mathbb{1} \pm \sigma_i) \otimes \rho \), where \( \rho \in B(H_2) \). This can be seen by noting that if \( a_{12} = \pm 1/4 \), one must choose \( a_4 = a_8 = 0 \) to ensure non-negative eigenvalues of the reduced state on \( H_A \). Therefore, for all physical states in the restricted neighbourhoods the reduced state of the ancilla is pure and of the form \( 1/4(\mathbb{1} \pm \sigma_i) \), which implies that all such states are product states. Since any product state has zero mutual information and remains a product state during the evolution it follows that \( \frac{d\lambda(s)}{ds} \bigg|_{s=t} \) is zero for all states in any neighborhood of \( 1/4(\mathbb{1} \pm \sigma_i) \otimes \mathbb{1} \) where \( a_{12} = \pm 1/4 \).
Finally we can conclude that for random unitary evolutions such as the quasi-eternal model described in Sect. 1, where the non-Markovian dynamics is $P$-divisible and all initial states have been mapped to a sufficiently small neighborhood of the stationary states by the preceding Markovian evolution no increase in the mutual information occurs. Moreover, the neighborhood where no increase of the mutual information occurs only depends on $V_{t,t}$ and is independent of the preceding dynamics. Therefore, for any random unitary $P$-divisible evolution subsequent to time $t$ it is always possible to find a CP-divisible random unitary $\Lambda_t$ such that $\text{Im}(\mathcal{I}_t \otimes \Lambda_t)$ is contained in this neighborhood by appropriately choosing the rates $\gamma_k(\tau) > 0$ for $0 \leq \tau \leq t$.

VI. CORRELATION MEASURES DETECTING ALMOST-ALL NON-MARKOVIAN DYNAMICS

After seeing that the mutual information and entanglement measures are unable to detect all non-Markovian dynamics, in this section we review and expand the results in [23], where it was proven that correlation backflows can be observed for all non-Markovian dynamics that are at most point-wise non-bijective. To do so, we need to (i) move to the enlarged scenario in which two auxiliary particles are employed and (ii) consider a new correlation measure based on the distinguishability of quantum states. To present this correlation measure we need to introduce the notion of maximally entropic measurement and review the definition of guessing probability for an ensemble of quantum states.

1. Maximally entropic measurements

Consider a quantum state $\rho$. For this state, we say that a measurement, defined by a Positive Operator Valued Measure (POVM), is maximally entropic, and denote by $\text{ME-POVM}$, if, when applied on the state $\rho$, each outcome has the same probability of occurrence $p_i = 1/n$, where $n$ is the generic number of outputs of the POVM considered. Indeed, if $\mathcal{S}([p_i]_i^n) = -\sum p_i \log p_i$ is the Shannon entropy of the resulting $n$-outcome probability distribution, where we take as the basis of the logarithm the number of probabilities $\mathcal{S}([p_i]_i^n) = 1$ if and only if $p_i = 1/n$. We define the set of ME-POVMs for $\rho$ as

$$\Pi(\rho) \equiv \{ [p_i]_i^n : \mathcal{S}([p_i]_i^n) = 1 \}. \quad (67)$$

For any state $\rho$, this collection is non-empty and contains measurements with any number of outputs (see Appendix 1). It is trivial to adapt this notion to the case when the number of outputs is fixed to $n$, having the set of measurements $\Pi_n(\rho)$.

2. Guessing probability of an ensemble

Consider an ensemble of states $\mathcal{E} = \{ p_i, \rho_i \}_{i=1}^n$ (with $i = 1, \ldots, n$) defined on a finite dimensional state space $S(H)$. Assume that we know the composition of the ensemble $\mathcal{E}$, and we want to answer the question: What is the average probability to correctly identify a state extracted from $\mathcal{E}$, maximized over all possible measurements? This quantity is called the guessing probability of the ensemble

$$P_g(\mathcal{E}) \equiv \max_{[p_i]} \sum_{i=1}^n p_i \mathcal{I}[\rho_i, P_i], \quad (68)$$

where the maximization is performed over the space of the $n$-output POVMs. It is clear that $P_g(\rho) \geq \rho_1$, where $\rho_1 = \max_i [p_i]_i^n \geq 1/n$ and $P_g(\mathcal{E}) = \rho_1$ if the states of the ensemble are identical: $\rho_i = \rho_1$ for any $i = 1, \ldots, n$. It means that, if we fix the number of states of an ensemble to $n$, the minimum value of the guessing probability, i.e., $1/n$, is obtained if, and only if, the distribution is uniform and the states are identical.

Note that when the ensemble is composed by two equiprobable states, i.e., $\mathcal{E}^n = \{ [1/2], [\rho_1, \rho_2] \}$, $P_g(\mathcal{E}^n)$ can be expressed in terms of the distinguishability between $\rho_1$ and $\rho_2$

$$P_g(\mathcal{E}^n) = \frac{1}{4} (2 + ||\rho_1 - \rho_2||), \quad (69)$$

where $|| \cdot ||$ is the trace norm. The distinguishability is defined as $D(\rho_1, \rho_2) = ||\rho_1 - \rho_2|| + 1/2 = D(\rho_1, \rho_2)/2$.

A. The 2-output ME-POVMs’ scenario

We now have all the ingredients needed to define a correlation measure that witnesses almost-all non-Markovian evolutions.

Let $\rho_{AB}$ be a bipartite state acting on a finite dimensional state space of a composed system $S(H_A \otimes H_B)$. Now, consider the effect of a measurement in one of the two systems, say $A$. This measurement has $n$ possible outcomes and is represented by a POVM, i.e., a collection of positive semi-definite operators $[P_{A|i}]_i^n$ where the condition $\sum_{i=1}^n P_{A|i} = 1_A$ holds. Each $P_{A|i}$ represents a possible outcome with the probability $p_i$. After the measurement, depending on the outcome $i$, the state of the $B$ part is $\rho_{B|i}$. The output states and probabilities are

$$p_i = \mathcal{I}[\rho_{AB}, P_{A|i} \otimes 1_B], \quad \rho_{B|i} = \frac{\mathcal{I}[\rho_{AB}, P_{A|i} \otimes 1_B]}{p_i}. \quad (70)$$

We define by $\mathcal{E}(\rho_{AB}, [P_{A|i}]_i^n)$ the output ensemble made by the probabilities $p_i$ and the states $\rho_{B|i}$ obtained by the procedure described by Eq. (70).

Let’s now restrict the previous analysis to ME-POVMs of $n$ outcomes for the reduced state on system $A$. $\Pi_n(\rho_A)$, where $\rho_A = \mathcal{I}[\rho_{AB}]$. If $[P_{A|i}]_i^n \in \Pi_n(\rho_A)$, from Eqs. (70) and (67) it follows that,

$$\mathcal{E}(\rho_{AB}, [P_{A|i}]_i^n) = \left\{ p_i = \frac{1}{n}, \quad \rho_{B|i} = n \mathcal{I}[\rho_{AB}, P_{A|i} \otimes 1_B] \right\}_{i=1}^n \quad (71)$$

Therefore, Alice measures $\rho_{AB}$ with $[P_{A|i}]_i^n \in \Pi(\rho_A)$. Bob obtains an equiprobable ensemble of states (EES), namely an
In Section [14A] we show that the computation of $C^{(2)}(\rho_{AB})$ can be done efficiently using SDP for any $\rho_{AB}$. Operationally, $C^{(2)}_A(\rho_{AB})$ corresponds to the largest distinguishability between the pairs of equiprobable states of $A$ (A) and that we can obtain from $\rho_{AB}$ by performing measurements on $A$ (B).

We give two examples that provide an intuitive idea of the meaning of the correlation measures $C^{(2)}_A$, $C^{(2)}_B$ and $C^{(2)}$. First, we consider a generic product state $\rho_A \otimes \rho_B$. In this case, if Alice measures her side of $\rho_{AB}$ with a 2-output ME-POVMs $\{P_{A,i}\}_{i=1}^2$, the ensemble generated on Bob’s side consists of the two states: $\rho_{B,i} = 2 Tr_A [\rho_A \otimes \rho_B \cdot P_{A,i} \otimes 1_B] = \rho_{AB}$. which are identical and equal to $\rho_{AB}$. The corresponding guessing probability is $P_g = 1/2$ and therefore $C^{(2)}_A(\rho_A \otimes \rho_B) = 0$. In fact, the $1/2$ factor is chosen just to make the value of the correlation measure equal to zero for product states.

The second example is given by the maximally entangled state $\phi^{+}_{AB} = |\phi^+_A\rangle \langle \phi^+_A|$, where $|\phi^+_A\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. In order to evaluate $C^{(2)}_A(\phi^{+}_{AB})$, it is easy to realize that the projection measurement $\{P^+_{AB}\}_{i=1}^2 = \{|00\rangle_i \langle 00|_i, |11\rangle_i \langle 11|_i\}$ is the ME-POVM obtained by the maximization of Eq. (73). Indeed, in this case, Alice generates on Bob’s side an orthogonal ensemble of two states: $\rho_{B,i} = 2 Tr_A [\phi^{+}_{AB} \cdot i|\psi_A\rangle \langle \psi_A| \otimes 1_B] = i|\psi_A\rangle \langle \psi_A|_B$, which is perfectly distinguishable: $P_g (|\psi_A\rangle = 1/2, i|\psi_A\rangle^2_{e_{1,2}}) = 1$. It follows that, since the guessing probability of an ensemble cannot be greater than 1, maximally entangled states are maximally correlated states with respect to $C^{(2)}_A$ and, as it is straightforward to prove, also to $C^{(2)}_B$ and $C^{(2)}$. Note however that the same maximum value can be obtained by a maximally correlated classical bit, defined by the equal mixture of states $|00\rangle$ and $|11\rangle$.

For these quantities to define a proper correlation measure, we need to demonstrate that they are non-increasing under local operations. This is proven in Appendix [14] therefore showing that $C^{(2)}_A$, $C^{(2)}_B$ and $C^{(2)}$ are proper correlation measures. Moreover, in Appendix [D] we show that the correlation measure $C^{(2)}$ is continuous on the set of states $S(\mathcal{H}_A \otimes \mathcal{H}_B)$.

1. Computation of the correlation measure using semi-definite programming

A useful property of the introduced correlation measure is that its computation can be done by means of semi-definite programming (SDP). Without loss of generality, see Eqs. (73), (74) and (75), we present this result for the case where Alice performs the ME-POVM.

The main reason why we were able to express $C^{(2)}_A$ using SDP is that for the case of two equiprobable states, $\rho_1$ and $\rho_2$, the guessing probability can be expressed as a function of the trace distance between the two states which, in turn, can be computed through SDP as follows:

$$||\rho_1 - \rho_2|| = \max Tr [P_+] + Tr [P_-] ,$$

s.t. $\rho_1 - \rho_2 = P_+ - P_- , P_+ , P_- \geq 0 .

(76)

It now suffices to replace $\rho_1$ and $\rho_2$ by the two states prepared
with probability 1/2 on Bob’s side by Alice’s measurement on her part of $\rho_{AB}$. The final SDP to compute $C^{(2)}_A$ reads

$$C^{(2)}_A(\rho_{AB}) = \max \frac{1}{4} \left( \text{Tr} [P_+] + \text{Tr} [P_-] + 2 \right),$$

s.t. $\rho_{B,1} - \rho_{B,2} = P_+ - P_-$,

$$\text{Tr} [\rho_A P_{A,1}] = \text{Tr} [\rho_A P_{A,2}],$$

$$\rho_{B,1} = 2 \text{Tr}_A [\rho_{AB} P_{A,1} \otimes I_B],$$

$$\rho_{B,2} = 2 \text{Tr}_A [\rho_{AB} P_{A,2} \otimes I_B],$$

$$P_{A,1} + P_{A,2} = I_A,$$

$$P_+ + P_- + P_{A,1} + P_{A,2} \geq 0.$$  \hfill (77)

where again $\rho_A = \text{Tr}_B [\rho_{AB}]$ is the reduced state on $A$.

2. Witnessing non-Markovian dynamics

We now show how to use the correlation measure introduced above to detect non-Markovian evolutions. We prove that for any evolution that is at most point-wise non-bijective, we can find an initial state $\rho^{(\tau)}_{AB}(t)$ such that $C^{(2)}_A(\rho^{(\tau)}_{AB}(t))$ increases between time $t = \tau$ and $t = \tau + \Delta t$ if and only if there is no CP intermediate map $V_{\tau + \Delta t \tau}$, where the evolution of the system $B$ is given by the dynamical map $I_A \otimes \Lambda$, where $I_A$ is the identity map on $A$.

The particular bipartite separable states $\rho^{(\tau)}_{AB}(t)$ for which we examine the correlation $C^{(2)}_A$ are classical-quantum states. Our “probe” state is

$$\rho^{(\tau)}_{AB}(t) \equiv \frac{1}{2} \left( |0\rangle_\Lambda \otimes \rho^{(\tau)}_B(t) + |1\rangle_\Lambda \otimes \rho^{(\tau)}_B(t) \right),$$

where and $\rho^{(\tau)}_B(t)$ and $\rho^{(\tau)}_B(t)$ are the states that appear in Eq. (78) and $\rho^{(\tau)}_B(t)$ is an orthonormal basis for $H_A$. Since only the system $B$ is involved in the evolution, $\rho^{(\tau)}_{AB}(t)$ is given by Eq. (79) for any $t \geq 0$. Note that from Eq. (79) it follows that $\rho^{(\tau)}_{AB}(t)$ does not contain any entanglement. Moreover, the state can be chosen arbitrarily close to an uncorrelated state since, as shown in [11], one can always choose states $\rho^{(\tau)}_B(0)$ and $\rho^{(\tau)}_B(0)$ arbitrarily close to each other.

4. Detecting the correlation backflow

In this section we show how the correlation measure $C^{(2)}_A(\rho^{(\tau)}_{AB}(t))$, and later $C^{(2)}(\rho^{(\tau)}_{AB}(t))$, witnesses bijective or pointwise non-bijective non-Markovian dynamics.

To evaluate $C^{(2)}_A(\rho^{(\tau)}_{AB}(t))$, given the maximization used in Eq. (73), we have to find a ME-POVM $\{P_{A,1}, P_{A,2}\}$ that, applied on $\rho^{(\tau)}_{AB}(t)$, generates the output ensemble $\{p_{1,2} = 1/2, \rho_{B,1}(t), \rho_{B,2}(t)\}$ with the largest value of $\|\rho^{(\tau)}_{AB}(t) - \rho^{(\tau)}_{B,2}(t)\|$. Let $\lambda \in [0,1]$ and $\eta \in [0,1]$ be the diagonal elements of $P_{A,1}$ in the basis $\mathcal{B}_A = \{ |0\rangle_\Lambda, |1\rangle_\Lambda \}$. It is easy to show that $\lambda + \eta = 1$ for ME-POVMs. The corresponding output states are

$$\rho_{B,1}(t) = \lambda \rho^{(\tau)}_B(t) + \eta \rho^{(\tau)}_B(t),$$

$$\rho_{B,2}(t) = (1 - \lambda) \rho^{(\tau)}_B(t) + (1 - \eta) \rho^{(\tau)}_B(t).$$

It follows that

$$\|\rho_{B,1}(t) - \rho_{B,2}(t)\|_1 = |\lambda - \eta| \|\rho^{(\tau)}_B(t) - \rho^{(\tau)}_B(t)\|_1.$$  \hfill (82)

Since $0 \leq |\lambda - \eta| \leq 1$, the maximum is obtained when either $\lambda$ or $\eta$ is equal to 1. In both cases the output states are $\rho^{(\tau)}_B(t)$ and $\rho^{(\tau)}_B(t)$ and

$$C^{(2)}_A(\rho^{(\tau)}_{AB}(t)) = \frac{\|\rho^{(\tau)}_B(t) - \rho^{(\tau)}_B(t)\|_1}{4}. $$

\hfill (83)

In Appendices [11] and [22, 23] we prove that $C^{(2)}(\rho^{(\tau)}_{AB}(t)) = C^{(2)}(\rho^{(\tau)}_{AB}(t)) \geq C^{(2)}(\rho^{(\tau)}_{AB}(t))$. Therefore, using Eqs. (78) and (83), we obtain a correlation backflow

$$C^{(2)}(\rho^{(\tau)}_{AB}(t + \Delta t)) > C^{(2)}(\rho^{(\tau)}_{AB}(t)),$$

if and only if there is no CP intermediate map $V_{\tau + \Delta t \tau}$.
5. Example: the quasi-eternal non-Markovian model

For the sake of clarity, we illustrate the previous results through a specific dynamics. Let us consider the example introduced in Section II C 2 namely the class of dynamical maps \( \Lambda_{\alpha}(t_0) \) where \( \alpha > 0 \) and \( t_0 \geq 0 \) satisfies the relation given in Eq. (21). Recall that in Section V D 1 we showed that the mutual information fails to detect some non-Markovian dynamics that belong to this class.

Now, we show how to build the states \( \rho_B^{(\alpha)}(t_0) \) and \( \rho_B^{(\alpha)}(t_0) \) that appear in Eqs. (78) and (79) and therefore \( \rho_{AB}^{(\alpha)}(t_0) \) itself, when the evolution is given by \( \Lambda_{\alpha}(t_0) \). Picking \( \tau > t_0 \), the intermediate map \( V_{\alpha}(t_0) \), \( \tau \in \Delta_t \), is not CP for any \( \Delta_t > 0 \) (see Section VI A 2). The constructive method given in [11] suggests to consider, together with the qubit \( S \) evolved by \( \Lambda_{\alpha}(t_0) \), an ancillary qutrit \( A' \): \( S(\mathcal{H}_B) = S(\mathcal{H}_A \otimes \mathcal{H}_S) \). Now, being \( \{0\}_{A',1\}_{A',2\}_{A'} \) and \( \{0\}_{S},1\}_{S},2\}_{S} \) orthonormal basis respectively for \( \mathcal{H}_A \) and \( \mathcal{H}_S \), we have:
\[
\rho_{AS}^{(\alpha)}(\tau) = (1-p)\sigma_{AS} + p\phi_{AS}^{(\alpha)}\,,
\]
\[
\rho_{AS}^{(\alpha)}(\tau) = (1-p)\sigma_{AS} + p\rho_{AS}^{(\alpha)}\,.
\]
\[
\phi_{AS}^{(\alpha)} \equiv \phi_{A'S}^{(\alpha)} \text{ is the maximally entangled state, where } |\phi_{A'S}^{(\alpha)}\rangle = \{|0\rangle_{A'}|1\rangle_{A'} \rangle + |1\rangle_{A'}|0\rangle_{A'} \rangle / \sqrt{2} \text{ and } \sigma_{AS} \in \text{int}(\text{Im}(\Lambda_{\alpha}(t_0))).
\]

In order to define completely \( \rho_{AS}^{(\alpha)}(\tau) \) and \( \rho_{AS}^{(\alpha)}(\tau) \), we fix their free components: \( \sigma_{AS} = \langle 0|0\rangle_{A'} + |1\rangle_{1\rangle_{A'}} / 2 \otimes 1 / 2 \) and \( \rho_{AS} = 1 / 2 \) and we get:
\[
\rho_{AS}^{(\alpha)}(\tau) = \frac{(1-p)|0\rangle_{A'} |1\rangle_{1\rangle_{A'}} + |1\rangle_{1\rangle_{A'}} / 2 \otimes 1 / 2}{2} \otimes 1 / 2. \]

Consider the rates given in Eq. (46), the evolution induced by the dynamical map \( \Lambda_{\alpha}(t_0) \) (see Eqs. (10) that precedes the action of \( \Lambda_{\alpha}(t_0) \)), can be written as:
\[
\Lambda_{\alpha}(t_0)(\sigma_x) = \left( e^{-t_0} \cosh(t_0) \cosh\left( t_0 \right) \right)^{\alpha/2} \sigma_x = \Lambda_{\alpha}(t_0)(\tau) \sigma_x \,,
\]
\[
\Lambda_{\alpha}(t_0)(\sigma_y) = \left( e^{-t_0} \cosh(t_0) \cosh\left( t_0 \right) \right)^{\alpha/2} \sigma_y = \Lambda_{\alpha}(t_0)(\tau) \sigma_y \,,
\]
\[
\Lambda_{\alpha}(t_0)(\sigma_z) = e^{-t_0} \sigma_z = \Lambda_{\alpha}(t_0)(\tau) \sigma_z \,,
\]
\[
\Lambda_{\alpha}(t_0)(1) = 1 \,.
\]
where, for \( \tau > t_0 \), we have \( \Lambda_{\alpha}(t_0)(\tau) \). The state \( \rho_{AS}^{(\alpha)}(\tau) \), since it assumes the form \( \rho_{AS} \otimes 1 / 2 \) and the evolution is random unitary, is stationary for \( \mathcal{I}_A \otimes \Lambda_{\alpha}(t_0) \). Therefore, 
\[
\rho_{AS}^{(\alpha)}(\tau) = (I_A \otimes \Lambda_{\alpha}(t_0))^{-1} \rho_{AS}^{(\alpha)}(\tau) = \rho_{AS}^{(\alpha)}(\tau) \,.\]
Conversely, \( \rho_{AS}^{(\alpha)}(\tau) \) is not a stationary state and \( (I_A \otimes \Lambda_{\alpha}(t_0))^{-1} \rho_{AS}^{(\alpha)}(\tau) \) is not physical for every \( p \in [0,1] \). Indeed, we can write
\[
\rho_{AS}^{(\alpha)}(0) = \left( \Lambda_{\alpha}(t_0) \right)^{-1} \rho_{AS}^{(\alpha)}(\tau) = \left( |0\rangle \langle 0| + |1\rangle \langle 1| \right) \otimes 1 / 4 + \left( 2 \sigma_x \otimes \sigma_y \right) / 4 + \left( 2 \sigma_y \otimes \sigma_z \rangle \rangle \right) / 4 \,.
\]
which is a physical state for \( p / \Lambda_{\alpha}(t_0) < p / \Lambda_{\alpha}(\tau) \leq 1 \) (see Section V F). Therefore, if \( p \) satisfies the condition \( p < \Lambda_{\alpha}(\tau) \), Eq. (90) is the physical initial state \( \rho_{AS}^{(\alpha)}(0) \) which at time \( \tau \) evolved to the state given in Eq. (87), and therefore, together with \( \rho_{AS}^{(\alpha)}(0) \), fulfills the requirements of the constructive method given in [11].

The construction of the state \( \rho_{AS}^{(\alpha)}(0) \) is now straightforward (see Eq. (79)). Being \( A \) an ancillary qubit for which we adopt the orthonormal basis \( |0\rangle_{A},|1\rangle_{A},|2\rangle_{A} \), we have
\[
\rho_{AB}^{(\alpha)}(0) = \frac{1}{2} \left( |0\rangle \langle 0|_{A} + |1\rangle \langle 1|_{A} \right) \rho_{B}^{(\alpha)}(0) \,.
\]

B. Scenarios where ME-POVMs have more than two outputs

As mentioned, similar correlation measures \( C^{(n)} \) can be obtained by fixing the number of outputs of the ME-POVMs to any integer \( n \geq 3 \). Indeed, similarly to Eqs. (78), (74) and (75), we define:
\[
C_A^{(n)}(\rho_{AB}) = \max_{\{P_{A_i} \}_{i=1}^{n}} P_{\rho_{AB}} \left( \mathcal{E} \left( \rho_{AB}, \{P_{A_i} \}_{i=1}^{n}\right) \right) - \frac{1}{n} \quad (92)
\]
\[
C_B^{(n)}(\rho_{AB}) = \max_{\{P_{B_i} \}_{i=1}^{n}} P_{\rho_{AB}} \left( \mathcal{E} \left( \rho_{AB}, \{P_{B_i} \}_{i=1}^{n}\right) \right) - \frac{1}{n} \quad (93)
\]
and
\[
C^{(n)}(\rho_{AB}) = \max \left( C_A^{(n)}(\rho_{AB}), C_B^{(n)}(\rho_{AB}) \right) \quad (94)
\]
The measuring scenario considered by \( C_A^{(n)}(\rho_{AB}), \) and similarly \( C_B^{(n)}(\rho_{AB}), \) is described in Fig. [7]. Indeed, the value of \( C_A^{(n)}(\rho_{AB}), \) is given when Alice chooses an \( n \)-output ME-POVM that maximizes the guessing probability of the output ensemble generated on Bob’s side.

Moreover, we define \( C(\rho_{AB}) = \max \left( C_A(\rho_{AB}), C_B(\rho_{AB}) \right) \), where \( C_A(\rho_{AB}) = \max \left( P_{\rho_{AB}} \left( \mathcal{E} \left( \rho_{AB}, \{P_{A_i} \}_{i=1}^{n}\right) \right) \right) \), namely
\[
C(\rho_{AB}) = \max_{\{P_{A_i} \}_{i=1}^{n}} P_{\rho_{AB}} \left( \mathcal{E} \left( \rho_{AB}, \{P_{A_i} \}_{i=1}^{n}\right) \right) - \frac{1}{2} \quad (95)
\]
We define \( C_B(\rho_{AB}) \) similarly.

The details of the proof that shows that \( C^{(n)} \) (for any \( n \geq 3 \)) and \( C \) are valid correlation measures, i.e., satisfy the conditions given in Sec. III A can be found in Appendix C. Moreover, in Appendix D we show that the correlation measures \( C^{(n)} \) are continuous on the set of states for any \( n \geq 2 \) and in Appendix E we give a bound for the number of outcomes of
the ME-POVMs that Alice has to consider in order to evaluate \( C_{\rho_{AB}}^{\text{int}} \).

In Section VI A we gave a physical interpretation of the value of the measure \( C_{\rho_{AB}}^{(2)} \) in terms of the most distinguishable pair of output states that Alice, performing two-output ME-POVMs on her share of \( \rho_{AB} \), can produce on Bob’s side. Increasing to \( n \geq 3 \) the number of outputs of the ME-POVMs adopted by Alice, this interpretation can be easily adopted also for \( C_{\rho_{AB}}^{(3)} \).

The meaning of the correlation measure \( C_A \), and consequently of \( C_B \) and \( C \), is slightly different. Indeed, \( C_A(\rho_{AB}) \) represents the guessing probability of the most distinguishable ensemble of states that Alice can produce on Bob’s side applying any ME-POVM on \( \rho_{AB} \), namely without fixing the number of states of the output ensemble obtained by Bob (although the number of outputs should be larger than one). Note that for product states the best measurement consists of only two outputs.

Making use of the probe state \( |\rho_{AB}^{(t)}⟩ \) given in Eq. (79), also the correlation measure \( C_A \) is able to witness almost-all non-Markovian evolutions. Indeed, in Appendices I and II we prove that for this state \( C_B(\rho_{AB}^{(t)}(t)) = C_{\rho_{AB}}^{(2)}(\rho_{AB}^{(t)}(t)) \) and \( C_{\rho_{AB}}^{(3)}(\rho_{AB}^{(t)}(t)) \). From these results it follows that for this initial probe state \( C_{\rho_{AB}}^{(2)}(\rho_{AB}^{(t)}(t)) = C_{\rho_{AB}}^{(3)}(\rho_{AB}^{(t)}(t)) \) at any time \( t \geq 0 \). Therefore, for this state, \( C_{\rho_{AB}}^{(t)}(t) \) can also be calculated efficiently using SDP and shows the same correlation backflow given in Eq. (84).

### C. Problems to witness any non-Markovian dynamics

The construction of the correlation measure is inspired by the results in [11], valid for distinguishability measures for dynamics that at most point-wise non-bijective. However, in Ref. [10] it was shown that for any non CP-divisible intermediate map \( V_{ij} \) an ensemble of states \( \mathcal{E} = \{|p_i,\rho_i⟩\} \) exists such that the guessing probability \( P_g(\mathcal{E}) \) increases between time \( t \) and \( s \). How to construct such an ensemble is not known and there is no reason to believe that an equiprobable ensemble with this property exists for a general dynamics. A natural question is therefore if a correlation measure of the type described in Sect. VI can be defined that shows an increase when \( P_g(\mathcal{E}) \) increases.

A first attempt in this direction could go as follows. It is possible to generalize the correlation measure \( C_{\rho_{AB}}^{(2)} \) by performing a maximization of \( P_g \) over the set of POVMs \( \Pi_{\rho_{AB}}^{(2)} \) that produce some given fixed output probabilities \( \{p_i\} \) when acting on \( \rho_{AB} \), not necessarily fully entropic. The resulting correlation measure \( C_{\rho_{AB}}^{(2)} \) reads

\[
C_{\rho_{AB}}^{(2)} = \max_{\{p_i\} \mid |\rho_{AB}^{(t)}⟩} P_g(E(\rho_{AB}^{(t)},\{p_i\})).
\]

The idea would be to mimic the previous result by using as initial probe a state consisting of the ensemble states of Ref. [10] for systems \( S \) and \( A' \) correlated with orthogonal states on \( A \) according to the probabilities in the ensemble.

Unfortunately, this approach does not work. In fact, we show that for a classical-quantum state \( \rho_{AB} \) of the type

\[
\rho_{AB} = \sum_i p_i |i⟩⟨i| \otimes \rho_i.
\]

the ensemble that maximizes \( P_g \) is not defined by the projections on the states \( |i⟩ \) in system \( A \). We show this by constructing a counterexample. Consider the state

\[
\rho_{AB} = p_1 |1⟩⟨1| \otimes \rho_1 + p_2 |2⟩⟨2| \otimes \rho_2 + p_3 |3⟩⟨3| \otimes \rho_3,
\]

where \( 2p_3 > p_1 > 2p_2 \) and

\[
\rho_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}, \rho_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \rho_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.
\]

For this example we can directly find \( P_g \). To do so we use that \( P_g \) is the solution to a convex optimization problem where strong duality holds [33, 34]. The dual formulation of \( P_g \) is

\[
P_g = \min_{k} \text{Tr} K \quad \text{s. t.} \quad K \succeq p_k \rho_i, \forall i.
\]

For the output ensemble \( \mathcal{E} = \{|p_i,\rho_i⟩\} \), achieved by the p-POVM \( \{|1⟩⟨1|,|2⟩⟨2|,|3⟩⟨3|\} \), it can be seen that \( P_g = p_1/2 + p_3 \) by directly constructing the optimal \( K \).

However, for the output ensemble \( \{|p_i,\rho_i⟩\} \) achieved by a non-projective p-POVM \( \{|1⟩⟨1|,|2⟩⟨2|,|3⟩⟨3|\} \) the construction of the optimal \( K \) gives \( P_g = p_1/2 + p_2/2 + p_3 \). Thus, in general the maximization of \( P_g \) over \( \Pi_{\rho_{AB}}^{(2)} \) does not produce the desired ensemble for systems \( S \) and \( A' \). Therefore, it is not clear to us how to adapt the results in [10] into a correlation measure that detects all non-Markovian evolutions.

### VII. QUASI-CORRELATION MEASURES

We would like to conclude our study by discussing the use of what we call quasi-correlation measures in the context of non-Markovian detection. Note that while for a correlation measure it is demanded that it does not increase by local operations, for the detection of non-Markovianity it suffices to consider functions that do not increase under the action of operations by one of the parties (the one evolving through the dynamics). An increase on the value of these measures is enough to detect the failure of CP divisibility. We name quasi-correlation measures those functions of a bipartite state that do not increase when applying operations on only one half of the state.

An example of such measures is the version of the singlet fraction defined in Ref. [32]. Given a bipartite state \( \rho_{AB} \), it is defined as

\[
SF(\rho_{AB}) = \max_{\epsilon} |\langle \phi^+ | (1_\Lambda \otimes \epsilon)(\rho_{AB}) | \phi^+ \rangle|,
\]

that is the maximum of the fidelity with a maximally entangled state optimised over local operations by one of the parties. Other definitions of the singlet fraction allow operations by both parties and include classical communication. It is easy to see that this measure cannot increase by local operations on
the second system, as any further local processing can always be adsorbed in the optimization in (101).

It immediately follows from previous results that this measure detects all non-Markovian dynamics. In [32], it was shown that for classical-quantum correlated states of the form

\[ \rho_{\text{CQ}} = \sum_i p(i) |i\rangle\langle i| \otimes \rho_i, \]

(102)

the singlet fraction is equal to the guessing probability of the ensemble \( \{p(i), \rho_i\} \). Therefore, we can combine this with the results in [10], proving the existence of an ensemble with increasing guessing probability for any non-Markovian dynamics, to conclude that this version of the singlet fidelity also detects all such dynamics.

VIII. DISCUSSION

Understanding the operational consequences of non-Markovian effects in terms of information backflow is a fundamental question. In this work, we focus on correlations and study how they can be used to detect the failure of CP-divisibility. We have identified strengths and weaknesses of several known correlation measures. In particular, we have shown that:

- Non-Markovian effects in single-parameter dynamics, such as depolarization, dephasing or amplitude damping, always lead to correlation backflows for any continuously differentiable measure that is not time-independent on the image of the the preceding evolution;

- It is possible to detect backflows in the mutual information for any qubit unital non-P-divisible dynamics;

- Maximally entangled states are not necessarily optimal for observing backflows in the mutual information;

- There exist quasi-eternal non-Markovian dynamics with no backflow in the mutual information;

- There exist quasi-correlation measures that an be used for non-Markovianity detection and always show a backflow.

In addition to this we have revisited and expanded upon or exemplified several points from Ref. [28]:

- Measures of entanglement between the system and an ancilla cannot provide any backflow in those cases where the non-Markovian dynamics is P-divisible and appear only after the dynamics has become entanglement breaking;

- Measures based on state distinguishability that detect almost all non-Markovian dynamics can be constructed in a setting with two additional particles.

Our results clarify many issues but also point to several open questions. The most obvious one is to construct a correlation measures able to detect all dynamics that are not CP-divisible, either by adapting the results in [10] to our approach, or by considering a novel approach. A second open question is to understand if the use of the second additional particle can be of use for existing correlation measures, such as those based on entanglement or mutual information.

As mentioned, while completing this work, we became aware of the work [29] where it is shown that negativity, a computable entanglement measure, is able to detect all the bijective non-Markovian dynamics, and even non-bijective in the case of qubits, in the enlarged scenario using two auxiliary particles. Note that this second particle is also crucial to extend the detection power of our correlation measure and detect all at most point-wise non-bijective evolutions. Whether similar effects can be observed for other correlation measures, such as the mutual information, deserves further investigation.

ACKNOWLEDGMENTS

Support from the ERC CoG QITBOX, the AXA Chair in Quantum Information Science, Spanish MINECO (QIBEQI FIS2016-80773-P and Severo Ochoa SEV-2015-0522), Fundació Privada Cellex, and the Generalitat de Catalunya (SGR1381 and CERCA Program) is acknowledged. D.D.S. acknowledge support from the ICFOstepstone programme, funded by the Marie Skłodowska-Curie COFUND action (GA665884). B.B. acknowledge support by the ICFO-MPQ Fellowship. Fellowship. N.K.B. thanks the Brazilian funding agency CAPES.

[1] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford Univ. Press, Oxford, 2007).
[2] U. Weiss, *Quantum Dissipative Systems* (World Scientific, Singapore, 2000).
[3] A. Rivas and S. F. Huelga, *Open Quantum Systems. An Introduction*, (Springer, Heidelberg, 2011).
[4] A. Rivas, S. F. Huelga, and M. B. Plenio, Rep. Prog. Phys. 77, 094001 (2014).
[5] H.-P. Breuer, E.-M. Laine, J. Piilo, and B. Vacchini, Rev. Mod. Phys. 88, 021002 (2016).
[6] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, J. Math. Phys. 17, 821 (1976).
[7] G. Lindblad, Comm. Math. Phys. 48, 119 (1976).
[8] Á. Rivas, S. F. Huelga, and M. B. Plenio, Phys. Rev. Lett. 105, 050403 (2010).
[9] H.-P. Breuer, E.-M. Laine and J. Piilo, Phys. Rev. Lett. 103,
Appendix A: Nonzero time derivatives of initially zero eigenvalues of rank one matrices for non-unitary single parameter maps

Let $\rho$ be a positive semidefinite Hermitian matrix. Consider an eigenvalue $\lambda_k$ of $\rho$ and its corresponding normalized eigenvector $u_k$. By definition it holds that $\lambda_k = u_k^\dagger \rho u_k$. The time derivative of $\lambda_k$ is

$$\frac{d\lambda_k}{dt} = \frac{du_k^\dagger}{dt} \rho u_k + u_k^\dagger \frac{\rho}{dt} u_k + u_k^\dagger \frac{d\rho}{dt} u_k$$

$$= \lambda_k \left( \frac{du_k^\dagger}{dt} u_k + u_k^\dagger \frac{du_k}{dt} \right) + u_k^\dagger \frac{d\rho}{dt} u_k.$$ (A1)

Since $u_k$ is normalized it follows that its derivative is orthogonal to $u_k$, i.e., $\frac{du_k^\dagger}{dt} u_k = 0$. If the evolution of $\rho$ is described by a continuously differentiable family of dynamical maps so that $\frac{d\rho}{dt} = \frac{d}{dt} V_{\gamma}(\rho||\varnothing) = \mathcal{L}_\gamma(\rho)$ it follows that $\frac{d\lambda_k}{dt} = u_k^\dagger \mathcal{L}_\gamma(\rho) u_k$.

Next consider the special case where $\rho$ is a rank one positive semidefinite trace one $n \times n$ matrix and consider its block diagonal form.

$$\rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$ (A2)

where $1$ represents the $1 \times 1$ block corresponding to the nonzero eigenvalue $1$ and $0$ represents the $n \times 1$, $1 \times n$, and $n \times n$ zero blocks. We want to investigate $\frac{d}{dt} V_{\gamma}(\rho||\varnothing) = \mathcal{L}_\gamma(\rho)$ for the case of single parameter evolution. In particular we want to study the projection of $\mathcal{L}_\gamma(\rho)$ onto the zero-eigenspace of $\rho$.

First we consider the unitary part of $\mathcal{L}_\gamma$. Let $P_0$ be the projector onto the zero eigenspace of $\rho$. One easily finds that $P_0[H,\rho]P_0 = 0$ for any $H$.

Then we consider $\left( G_{k}\rho G^\dagger_{k} - \frac{1}{2} \{G^\dagger_{k}G_{k},\rho\} \right)$ and express the matrix $G_{k}$ on the same block form as $\rho$, i.e.,

$$G_{k} = \begin{pmatrix} A_k & B_k \\ C_k & D_k \end{pmatrix},$$ (A3)

where $A_k$ is the $1 \times 1$ block. The projection of $\left( G_{k}\rho G^\dagger_{k} - \frac{1}{2} \{G^\dagger_{k}G_{k},\rho\} \right)$ onto the zero eigenspace of $\rho$ is then

$$P_0 \left( G_{k}\rho G^\dagger_{k} - \frac{1}{2} \{G^\dagger_{k}G_{k},\rho\} \right) P_0 = C_{k} C^\dagger_{k}.$$ (A4)

The matrix $C_{k} C^\dagger_{k}$ is clearly Hermitian and positive semidefinite. It follows that

$$P_0 \mathcal{L}_{\gamma} P_0 = \gamma(t) \sum_k C_k C^\dagger_k,$$ (A5)

is also a positive semidefinite matrix if $\gamma(t) > 0$ and negative semidefinite if $\gamma(t) < 0$. Moreover, $P_0 \mathcal{L}_{\gamma} P_0$ is zero if and only if $C_k$ is zero for every $k$. Hence, if and only if for each $k$ the lower off-diagonal $n \times 1$ block of $G_k$ is zero in every basis will there be no rank one $\rho$ such that $P_0 \mathcal{L}_{\gamma} P_0$ is nonzero. In this case $G_k$ is proportional to the identity which implies $\left( G_{k}\rho G^\dagger_{k} - \frac{1}{2} \{G^\dagger_{k}G_{k},\rho\} \right) = 0$ for every $\rho$. Thus, for any $\mathcal{L}_{\gamma}$ with non-zero dissipative part there exist at least one rank one $\rho$ such that the time derivative of the initial zero-eigenspace is nonzero.
To analyse the special case when \( \rho = \phi_s^+ \) where \( \phi_s^+ \) is the maximally entangled state on \( \mathcal{H} \otimes \mathcal{H} \) and \( \dim(\mathcal{H}) = n \) we note that the condition \( G_k = 0 \) can be formulated as \( G_i \rho = \rho G_i \). We write \( \phi^+ = \frac{1}{n} \sum_i E_{ij} \otimes E_{ij} \) where \( E_{ij} \) is the matrix with the \( ij \)th element equal to 1 and all other elements equal to zero. We write \( G_k = I \otimes G \) where \( G \) is any matrix. Then

\[
G_i \rho = \frac{1}{n} \sum_{ij} E_{ij} \otimes GE_{ij}
\]

\[
\rho G_i \rho = \frac{1}{n^2} \sum_{ijkl} E_{ij} E_{kl} \otimes GE_{ij} = \frac{\text{Tr}(G)}{n^2} \sum_{ij} E_{ij} \otimes E_{ij}.
\]

(A6)

These two expressions are equal if and only if \( GE_{ij} = \text{Tr}(G)/n E_{ij} \) for each \( i \). Since the matrices \( E_{ij} \) form a basis for the matrix space it follows that \( G_i \rho = \rho G_i \) if and only if \( G_k \propto I \otimes I \). As noted before this implies \( \{ G_i G_k \} = 0 \) for every \( k \), and thus for any \( L \) with non-zero dissipative part there is an eigenvalue of \( V_s, (\phi_s^+ ) \) that is zero for \( s = t \) but has a non-zero time derivative.

Appendix B: Proof of Proposition [4]

Let \( M(\tilde{a}) \) be a correlation measure that is an analytic function of the coordinates \( a_i \) in a point \( \tilde{a} \) corresponding to a product state \( \rho \) and let \( V_{st} \) be a continuously differentiable intermediate map for all \( s, t \). Consider a family of states \( \rho_s = \rho + \epsilon \chi \) where \( \chi \) is Hermitian. The Taylor expansion of \( M(\rho_s) \) in \( \epsilon \) around \( \epsilon = 0 \) is

\[
M(\rho_s) = \left. \frac{\partial M(\rho_s)}{\partial \epsilon} \right|_{\epsilon=0} \epsilon + \left. \frac{\partial^2 M(\rho_s)}{2 \partial \epsilon^2} \right|_{\epsilon=0} \epsilon^2 + \ldots,
\]

(B1)

where we have used that \( M(\rho_s)|_{\epsilon=0} = 0 \). Since \( M \geq 0 \) on \( S(\mathcal{H}_A \otimes \mathcal{H}_S) \) it follows that the first order term of the expansion must be zero if \( \rho \in \text{int}[S(\mathcal{H}_A \otimes \mathcal{H}_S)] \). Otherwise there would be a sufficiently small \( \epsilon \) for which both \( \rho_+, \rho_- \in \text{int}[S(\mathcal{H}_A \otimes \mathcal{H}_S)] \) and either \( M(\rho_+) \) or \( M(\rho_-) \) was negative.

Assume that \( \rho \in \text{int}[S(\mathcal{H}_A \otimes \mathcal{H}_S)] \) and consider the Taylor expansion of \( M[I \otimes V_{st}(\rho_s)] \) in \( \epsilon \) around 0 in

\[
M[I \otimes V_{st}(\rho_s)] = \left. \frac{\partial M[I \otimes V_{st}(\rho_s)]}{\partial \epsilon} \right|_{\epsilon=0} \epsilon + \left. \frac{\partial^2 M[I \otimes V_{st}(\rho_s)]}{2 \partial \epsilon^2} \right|_{\epsilon=0} \epsilon^2 + \ldots,
\]

(B2)

where we have used that \( M[I \otimes V_{st}(\rho_s)]|_{\epsilon=0} = 0 \). Since \( M \geq 0 \) it follows again that the first order term of the expansion must be zero. Thus \( \left. \frac{\partial M[I \otimes V_{st}(\rho_s)]}{\partial \epsilon} \right|_{\epsilon=0} = 0 \) for all \( s \).

Next, consider the Taylor expansion of the derivative \( \frac{d}{d\epsilon} M[I \otimes V_{st}(\rho_s)] \) in \( \epsilon \)

\[
\frac{d}{d\epsilon} M[I \otimes V_{st}(\rho_s)]|_{\epsilon=0} = 0
\]

and

\[
\frac{d}{d\epsilon} M[I \otimes V_{st}(\rho_s)]|_{\epsilon=0} + \ldots
\]

(B3)

where we have used that \( \frac{d}{d\epsilon} M[I \otimes V_{st}(\rho_s)]|_{\epsilon=0} = 0 \). Due to the continuous differentiability of \( V_{st} \), it follows that \( \frac{d}{d\epsilon} \frac{d}{d\epsilon} M[I \otimes V_{st}(\rho_s)] \) is continuous, and therefore \( \frac{d}{d\epsilon} \frac{d}{d\epsilon} M[I \otimes V_{st}(\rho_s)] = \frac{d}{d\epsilon} \frac{d}{d\epsilon} M[I \otimes V_{st}(\rho_s)] \).

Appendix C: Proof of Proposition [5]

We begin by considering the following two propositions.

Proposition 6. Let \( \{ \Lambda_t \} : B(\mathcal{H}_S) \to B(\mathcal{H}_S) \) be a family of CP qubit maps. If the set of stationary states in \( S(\mathcal{H}_S) \) is of non-zero dimension, the \( \Lambda_t \) are unital.

Proof. Assume that \( \Lambda_t(1) = 1 + \theta \) and \( \Lambda_t(\rho_1) = \rho_1, \Lambda_t(\rho_2) = \rho_2 \) where \( \rho_1 \neq \rho_2 \). If \( \theta = 0 \) it follows that \( \Lambda_t(1 + x(\rho_1 - \rho_2)) = 1 + x(\rho_1 - \rho_2) + \theta \). Note that \( x \) can be chosen such that \( 1 + x(\rho_1 - \rho_2) \) is rank one. Since these rank one qubit states are antipodal points on the Bloch ball \( S(\mathcal{H}_S) \) it follows that unless \( \theta = 0 \) at least one of \( 1 + x(\rho_1 - \rho_2) + \theta \) and \( 1 - x(\rho_1 - \rho_2) + \theta \) is not positive semidefinite. Thus if the set of stationary states has dimension greater than zero, it follows that \( \Lambda_t \) is unital.

Proposition 7. Let \( \{ \Lambda_t \} : B(\mathcal{H}_S) \to B(\mathcal{H}_S) \) be a family of CP qubit maps. Then the set of stationary states of \( \{ \Lambda_t \} \) in \( S(\mathcal{H}_S) \) has a dimension different from 2.

Proof. Assume that the dimension of the set of stationary states in \( S(\mathcal{H}_S) \) is 2. From Prop. 6 follows that \( \Lambda_t \) is unital. Without loss of generality we can assume that \( \Lambda_t(\sigma_z) = \sigma_z \) and \( \Lambda_t(\sigma_y) = \sigma_y \) and \( \Lambda_t(\sigma_x) = a \sigma_x + b \sigma_y + c \sigma_z \). The Choi matrix of \( \Lambda_t \) has eigenvalues \( \pm \sqrt{1 - 2a + a^2 + b^2 + c^2} \) and \( \pm \sqrt{1 + 2a + a^2 + b^2 + c^2} \). Therefore \( \| \chi \| \) is CP if and only if \( a = 1 \) and \( b = c = 0 \), i.e., if an only if \( \Lambda_t = T \). In this case the the set of stationary states has dimension 3, contradicting the assumption.

Next, consider a family of continuously differentiable dynamical qubit maps \( \{ \Lambda_t \} \) and a correlation measure \( M \). If the set of stationary states in \( S(\mathcal{H}_S) \) is non-empty its dimension is either zero or non-zero. If the dimension is zero, the set of stationary states in \( S(\mathcal{H}_A \otimes \mathcal{H}_S) \) is a set of product states. Then it follows from Prop. 6 that such a stationary state is a critical point if \( M \) is analytical at the state and the state is in the interior of \( S(\mathcal{H}_A \otimes \mathcal{H}_S) \).

If the dimension is 3 all states are stationary points and \( \frac{dM}{dt} = 0 \) on all of \( S(\mathcal{H}_A \otimes \mathcal{H}_S) \). Thus all states are critical.
points. By Prop. [7] the dimension is never 2. The remain-
ing case is a one-dimensional set of stationary states. With-
out loss of generality we can express any state in this set as

\[ \rho_{AB} = \rho_A \otimes \mathbb{1} + \chi_A \otimes \sigma, \]

for some \( \rho_A, \chi_A \in B(\mathcal{H}_A) \). Now, assume that \( \rho_{AB} \) is
in the interior of \( S(\mathcal{H}_A \otimes \mathcal{H}_B) \) and that the correlation measure \( M \) is an analytic function at
\( \rho_{AB} \). Then consider the family of states \( \rho_{e} = \rho_{AB} + \epsilon \chi_A \otimes \chi_S \) parametrized by \( \epsilon \), where \( \chi_A \otimes \chi_S \) are Hermitian and \( \text{Tr} (\chi_A \otimes \chi_S) = 0 \). If \( \chi_S = \sigma_z \) or \( \chi_S = \mathbb{1} \) it follows that \( \rho_e \) is also a stationary state and thus \( \frac{d}{d\epsilon} \text{Tr}(\rho_e) \bigg|_{\epsilon=0} = 0 \). If \( \chi_S = \sigma_x \) or \( \chi_S = \sigma_y \), there exist a local unitary operation, \( I \otimes \sigma_z \), that commutes with \( \rho_{AB} \) but anticommutes with \( \chi_A \otimes \chi_S \). Since \( M(\tilde{\alpha}), \tilde{\alpha} \) is invariant under local unitary operations it follows that \( M(\rho_e) = M(\rho_{AB}) \). Thus \( M(\rho_e) \) is an even analytic function in \( \epsilon \) and it follows that \( \frac{d}{d\epsilon} M(\rho_{AB}) \bigg|_{\epsilon=0} = 0 \) for every \( \chi_A \otimes \chi_S \). Moreover, \( \frac{d}{d\epsilon} M(\rho_{AB}) \bigg|_{\epsilon=0} = 0 \) for any \( t \). Therefore we can conclude that \( \frac{d}{d\epsilon} M(\rho_{AB}) \bigg|_{\epsilon=0} = 0 \) holds for any \( t \). By the analyticity of \( M \) and the continuous differentiability of \( \Lambda_t \) it follows that \( \frac{d}{d\epsilon} M, \frac{d^2}{d\epsilon^2} M \) and \( \frac{d}{d\epsilon} M \) exist and that \( \frac{d^2}{d\epsilon^2} M \) is continuous as a function of \( \epsilon \) and \( t \). Therefore it follows that \( \frac{d}{d\epsilon} M \) exist and \( \frac{d^2}{d\epsilon^2} M = \frac{d}{d\epsilon} M \). We can conclude that all first derivatives of \( \frac{d}{d\epsilon} M(\tilde{\alpha}, t) \) with respect to \( \tilde{\alpha} \) equal zero for states in the interior of \( S(\mathcal{H}_A \otimes \mathcal{H}_B) \) that are stationary under a continuously differentiable \( \Lambda_t \).

**Appendix D: Continuity of \( C(\rho_{AB}) \)**

We here show that the measures \( C^n(\rho_{AB}) \) for any \( n \) as well as \( C(\rho_{AB}) \) are continuous functions on the set of states. Let \( \text{Tr}(\sum_i P_{A,i} \otimes P_{B,i} \rho_{AB}) \) be an optimal solution to the optimization \( \max_{\rho_{AB}} \text{Tr}(\rho_{AB} P_{A,i} \otimes P_{B,i} \rho_{AB}) \) (or alternatively \( \max_{\rho_{AB}} \text{Tr}(\rho_{AB} P_{A,i} \otimes P_{B,i} \rho_{AB}) \)). Let \( S_X \) be the set of traceless Hermitian operators \( X \) satisfying \( \text{Tr}(XX^\dagger) = 1 \) and consider a state \( \rho_{AB} = \rho_{AB} + \epsilon X \) where \( X \in S_X \). Let \( \text{Tr}(\sum_i P_{A,i} \otimes P_{B,i} \rho_{AB}) = \text{Tr}(\sum_i P_{A,i} \otimes P_{B,i} \rho_{AB}) + \epsilon \text{Tr}(\sum_i P_{A,i} \otimes P_{B,i} X) \) be an optimal solution to the optimization \( \max_{\rho_{AB}} \text{Tr}(\rho_{AB} P_{A,i} \otimes P_{B,i} \rho_{AB}) \) (or alternatively \( \max_{\rho_{AB}} \text{Tr}(\rho_{AB} P_{A,i} \otimes P_{B,i} \rho_{AB}) \)). Then we can construct the following upper and lower bounds

\[
\text{Tr} \left( \sum_i P_{A,i} \otimes P_{B,i} \rho_{AB} \right) - \epsilon \max_{X \in S_X} \text{Tr} \left( \sum_i P_{A,i} \otimes P_{B,i} X \right) \\
\leq \text{Tr} \left( \sum_i P_{A,i} \otimes P_{B,i} \rho_{AB} \right) + \epsilon \text{Tr} \left( \sum_i P_{A,i} \otimes P_{B,i} X \right) \\
\leq \text{Tr} \left( \sum_i P_{A,i} \otimes P_{B,i} \rho_{AB} \right) + \epsilon \text{Tr} \left( \sum_i P_{A,i} \otimes P_{B,i} X \right) \\
\leq \text{Tr} \left( \sum_i P_{A,i} \otimes P_{B,i} \rho_{AB} \right) + \epsilon \max_{X \in S_X} \text{Tr} \left( \sum_i P_{A,i} \otimes P_{B,i} X \right),
\]

where we have used that \( \text{Tr}(\sum_i P_{A,i} \otimes P_{B,i} \rho_{AB}) \leq \text{Tr}(\sum_i P_{A,i} \otimes P_{B,i} \rho_{AB}) \) and \( \text{Tr}(\sum_i P_{A,i} \otimes P_{B,i} \rho_{AB}) \leq \text{Tr}(\sum_i P_{A,i} \otimes P_{B,i} \rho_{AB}) \). Since any eigenvalues of an \( X \in S_X \) is smaller or equal to one, the lower and upper bounds in Eq. 1 converge to \( \text{Tr}(\sum_i P_{A,i} \otimes P_{B,i} \rho_{AB}) = \epsilon \) converges to zero. An analogous argument can be made for the case where the roles of \( A \) and \( B \) are interchanged. Thus \( C^n(\rho_{AB}) \) converges to \( C^n(\rho_{AB}) \) as \( \rho_{AB} \) converges to \( \rho_{AB} \) for any \( \rho_{AB} \). Likewise \( C(\rho_{AB}) \) converges to \( C(\rho_{AB}) \) as \( \rho_{AB} \) converges to \( \rho_{AB} \) for any \( \rho_{AB} \).

**Appendix E: Sufficient number of POVM elements**

We can express the correlation measure \( C(\rho_{AB}) \) as

\[
C(\rho_{AB}) = \max_{P_{A,i} \in \mathcal{P}(\rho_{AB})} \max_{P_{B,i} \in \mathcal{P}(\rho_{AB})} \frac{1}{2} \sum_{i} \text{Tr} \left( P_{A,i} \otimes \rho_{AB} \right) - \frac{1}{2} \sum_{j} \text{Tr} \left( P_{B,j} \otimes \rho_{AB} \right),
\]

where the first maximization is over constraining either \( \{P_{A,i}\} \) or \( \{P_{B,j}\} \) to be a ME-POVM and \( \max_{n \geq 2} \) denotes maximization over the number of POVM-elements running over all integers \( n \geq 2 \).

While the number of POVM elements in the maximization is not constrained, here we show that it is sufficient to maximize over a finite range of integers. We give a proof that it is sufficient to consider \( n = 2 \) and odd \( n \) in the range \( 2 < n < \tilde{n} \) where \( \tilde{n} \) is a function of \( d_A = \dim(\mathcal{H}_A) \) and \( d_B = \dim(\mathcal{H}_B) \) that satisfies \( \tilde{n} \leq \max(d_A, d_B) \). It is unclear what the necessary conditions for maximization over \( n \) are.

Moreover, for classical-quantum states it is sufficient to consider maximization for \( n = 2 \) and odd \( n \) in the range \( 2 < n < d_A \). For quantum-classical states it is sufficient to consider maximization for \( n = 2 \) and odd \( n \) in the range \( 2 < n < d_B \). In particular this implies that \( C(\rho_{AB}) = C_{B(\chi)}(\rho_{AB}) \) and that \( C(\rho_{AB}) = C_{A(\chi)}(\rho_{AB}) \).

To find these sufficient conditions we introduce the following notation. Denote by \( C(n)(\rho_{AB}) \) the correlation measure where the maximization is limited to \( n \)-outcome POVMs

\[
C(n)(\rho_{AB}) = \max_{P_{A,i} \in \mathcal{P}(\rho_{AB})} \max_{P_{B,i} \in \mathcal{P}(\rho_{AB})} \frac{1}{2} \sum_{i} \text{Tr} \left( P_{A,i} \otimes \rho_{AB} \right) - \frac{1}{2} \sum_{j} \text{Tr} \left( P_{B,j} \otimes \rho_{AB} \right),
\]

Let \( \{\tilde{P}_{A,i}\} \) and \( \{\tilde{P}_{B,j}\} \) be two POVMs that maximize \( C(n)(\rho_{AB}) \) where either \( \{\tilde{P}_{A,i}\} \) or \( \{\tilde{P}_{B,j}\} \) is a ME-POVM

\[
C(n)(\rho_{AB}) = \text{Tr} \left[ \rho_{AB} \left( \sum_{i=1}^{n} \tilde{P}_{A,i} \otimes \tilde{P}_{B,j} \right) \right] - \frac{1}{2}.
\]
1. The general case

If \( n \) is even we consider the following 2-output POVM:

\[
P_{A,1}^{(2)} = \sum_{i \in E_1} \mathcal{P}_{A,i}^{(n)}
\]

\[
P_{A,2}^{(2)} = \sum_{i \in E_2} \mathcal{P}_{A,i}^{(n)}
\]

(\text{E4})

where \( E_1 \) and \( E_2 \) are any two sets of \( n/2 \) indices such that \( E_1 \cup E_2 = \{1, 2, \ldots, n\} \). This structure guarantees that Eq. (13) is a 2-output ME-POVM for \( \rho_{AB} \) if \( \{\mathcal{P}_{A,i}^{(n)}\}_i \) is. Similarly we define the 2 element POVM \( \{\mathcal{P}_{B,j}^{(2)}\}_i \) as

\[
P_{B,1}^{(2)} = \sum_{i \in E_1} \mathcal{P}_{B,i}^{(n)}
\]

\[
P_{B,2}^{(2)} = \sum_{i \in E_2} \mathcal{P}_{B,i}^{(n)}
\]

(\text{E5})

We compare Eq. (12) with the guessing probability of the output ensemble that we obtain applying Eq. (13) on \( \rho_{AB} \):

\[
C^{(2)}(\rho_{AB}) = \max_{\{\rho_{AB}^{(n)}\}_{i \in \Pi(\rho_{AB})}} \max_{\{\rho_{B,x}\}} \sum_{i=1}^{2} \text{Tr} \left[ \rho_{AB} P_{A,i} \otimes P_{B,x} \right] - \frac{1}{2}
\]

where \( O_1^x \) and \( O_2^x \) are any two sets of \( (n-1)/2 \) indices such that \( O_1^x \cup O_2^x = \{1, 2, \ldots, n\} \setminus x \). There are thus \( n \) different ways to choose \( x \). This structure guarantees that \( \{\mathcal{P}_{A,k}^{(2)}\}_k \) is a 2-output ME-POVM for \( \rho_{AB} \) if \( \{\mathcal{P}_{A,k}^{(n)}\}_k \) is and likewise for \( B \).

We can relate \( C^{(2)}(\rho_{AB}) \) and \( C^{(n)}(\rho_{AB}) \) as

\[
C^{(2)}(\rho_{AB}) \geq \text{Tr} \left[ \rho_{AB} \left( \sum_{i \in O_1^x} \mathcal{P}_{A,i}^{(n)} + \frac{1}{2} \mathcal{P}_{A,x}^{(n)} \right) \otimes \sum_{i \in O_2^x} \mathcal{P}_{B,i}^{(n)} + \frac{1}{2} \mathcal{P}_{B,x}^{(n)} \right] - \frac{1}{2}
\]

\[
= \text{Tr} \left[ \rho_{AB} \left( \sum_{i=1}^{n} \mathcal{P}_{A,i}^{(n)} \otimes \mathcal{P}_{B,i}^{(n)} + \frac{1}{2} \mathcal{P}_{A,x}^{(n)} \otimes \mathcal{P}_{B,x}^{(n)} \right) \right] - \frac{1}{2}
\]

\[
= \text{Tr} \left[ \rho_{AB} \left( \sum_{i=1}^{n} \mathcal{P}_{A,i}^{(n)} \otimes \mathcal{P}_{B,i}^{(n)} - \frac{3}{2} \mathcal{P}_{A,x}^{(n)} \otimes \mathcal{P}_{B,x}^{(n)} + \frac{1}{2} \mathcal{P}_{A,x}^{(n)} \otimes \mathcal{P}_{B,x}^{(n)} \right) \right] - \frac{1}{2}
\]

\[
= C^{(n)}(\rho_{AB}) + \text{Tr} \left[ \rho_{AB} \left( \frac{3}{2} \mathcal{P}_{A,x}^{(n)} \otimes \mathcal{P}_{B,x}^{(n)} + \frac{1}{2} \mathcal{P}_{A,x}^{(n)} \otimes \mathcal{P}_{B,x}^{(n)} \right) \right].
\]

(\text{E6})

A sufficient condition for \( C^{(2)}(\rho_{AB}) \geq C^{(n)}(\rho_{AB}) \) is that \( \theta_x \equiv -3 \mathcal{P}_{A,x}^{(n)} \otimes \mathcal{P}_{B,x}^{(n)} + \mathcal{P}_{A,x}^{(n)} \otimes \mathcal{P}_{B,x}^{(n)} + \mathcal{P}_{A,x}^{(n)} \otimes \mathcal{P}_{B,x}^{(n)} \) is positive semidefinite. Let \( \lambda_{AB} \) and \( \lambda_{BA} \) be the eigenvalues of \( \mathcal{P}_{AB}^{(n)} \) and \( \mathcal{P}_{BA}^{(n)} \) respectively. Then \( \theta_x \) is positive-semidefinite if \( -3 \lambda_{AB} \mathcal{I}_A \otimes \mathcal{I}_B + \lambda_{BA} \mathcal{I}_B \otimes \mathcal{I}_A \) is positive semidefinite for every \( k, l \). If \( \lambda_{AB} \leq 1/2 \) or \( \lambda_{BA} \leq 1/2 \) for every \( k, l \) this inequality holds. If \( \lambda_{AB} \geq 1/2 \) or \( \lambda_{BA} \geq 1/2 \) the inequality holds if \( \lambda_{AB} \leq z \) and \( \lambda_{BA} \leq z/(3z - 1) \) for some \( 1/2 \leq z \leq 1 \), and every \( k, l \). A sufficient condition for the existence of a pair \( \mathcal{P}_{A,k}^{(n)} \) and \( \mathcal{P}_{B,l}^{(n)} \) satisfying this is \( n \geq d_B(3z - 1)/z \) and \( n \geq d_A/z \) since \( \text{Tr}(\mathcal{P}_{B,l}^{(n)}) = d_B \) and \( \text{Tr}(\mathcal{P}_{A,k}^{(n)}) = d_A \).

Thus, a sufficient condition for \( C^{(2)}(\rho_{AB}) \geq C^{(n)}(\rho_{AB}) \) is that \( n \geq \tilde{n} \) where \( \tilde{n} \)
In conclusion, for general states it is sufficient to consider maximization for \( n = 2 \) and odd \( n \) in the range \( 2 < n < \bar{n} \).

2. Classical-quantum and Quantum-classical states

For the case of classical-quantum or quantum-classical states we can derive a different bound on \( n \). To see this we derive an inequality similar to that of Eq. (E9).

\[
P_g^{(2)} \geq \text{Tr} \left[ \rho_{AB} \left( \sum_{i=1}^{n} P_{A,i} \otimes P_{B,i} + \frac{1}{2} \sum_{i=1}^{n} P_{A,i} \otimes P_{B,i} + \frac{1}{2} \sum_{i=1}^{n} P_{A,i} \otimes P_{B,i} + P^{mix}_{AB} \right) \right]
\]

\[
\geq \text{Tr} \left[ \rho_{AB} \left( \sum_{i=1}^{n} P_{A,i} \otimes P_{B,i} - \frac{1}{2} P_{A,\bar{i}} \otimes P_{B,\bar{i}} + \frac{1}{2} \sum_{i=1}^{n} P_{A,i} \otimes P_{B,i} \right) \right]
\]

\[
= P_g^{(n)} + \text{Tr} \left[ \rho_{AB} \left( -\frac{P_{A,\bar{i}}}{2} \otimes P_{B,\bar{i}} + \sum_{i \neq \bar{i}} P_{A,i} \otimes P_{B,i} \right) \right] = P_g^{(n')} + \text{Tr} \left[ \rho_{AB} \left( \frac{1}{2} - \frac{2P_{A,\bar{i}}}{2} \otimes P_{B,\bar{i}} \right) \right].
\]

By definition \( \{ P_{A,i} \}_{i=1,...,n} \in \Pi \) if the output ensemble \( E(\rho_{AB}, \{ P_{A,i} \}) = \{ p_i, P_{B,i} \} \) is characterized by \( p_i = 1/n \). In general, we have that \( p_i = \text{Tr} [ \rho_{AB} (P_{A,i}) ] = \text{Tr} \left[ \rho_{AB} P_{A,i} \right] \), where \( \rho_{AB} = \text{Tr}_B [ \rho_{AB} ] \). Using an orthogonal decomposition of \( \rho_A \), we can always write it as: \( \rho_A = \sum_{i=1}^{d_A} \alpha_i | i \rangle \langle i | \), where \( \{ | i \rangle \} \) is an orthonormal basis of \( \mathcal{H}_A \). The condition \( \sum_{i=1}^{d_A} \alpha_i = 1 \) implies that there exist an \( \bar{i} \) such that \( S(\bar{i}) \equiv \sum_{i=1}^{d_A} \alpha_i > 1/2 \) and \( S(\bar{i} - 1) \equiv \sum_{i=1}^{d_A} \alpha_i \leq 1/2 \). We consider the following class of 2-output POVM that depends on a real parameter \( \omega \in [0,1] \) \( P_{A,\bar{i}}(\omega) = \sum_{i=1}^{d_A} | i \rangle \langle i | + \omega | \bar{i} \rangle \langle \bar{i} | , P_{A,\bar{i}}(\omega) = | 1 - \omega \rangle \langle 1 - \omega | + \sum_{i=1, i \neq \bar{i}} | i \rangle \langle i | \). We evaluate \( p_i \) for a general value of \( \omega \) and we obtain \( p_1(\omega) = \sum_{i=1}^{d_A} \alpha_i + \omega \alpha_{\bar{i}} = S(\bar{i} - 1) + \omega \alpha_{\bar{i}} \). It is clear that, since \( p_1(0) = S(\bar{i} - 1) \leq 1/2 \) and \( p_1(1) = S(\bar{i}) > 1/2 \), the value \( \omega = \bar{\omega} \equiv (1/2 - S(\bar{i} - 1))/\pi \), gives the uniform distribution \( p_{1,2}(\bar{\omega}) = 1/2 \) and consequently \( \{ P_{A,i}(\bar{\omega}) \}_{i} \in \Pi \) if \( \rho_{AB} \) is a ME-POVM for \( \rho_{AB} \).

Appendix G: Monotonic behaviour of \( C \) and \( C^{(n)} \) under local operations

Firstly, we prove that \( C_A \) is monotone under local operations of the form \( \Lambda_A \otimes I_B \), and secondly we consider the case where the local operation is \( I_A \otimes \Lambda_B \), where \( \Lambda_A \) (\( \Lambda_B \)) is a CPTP map on \( A \) (\( B \)) and \( I_A \) (\( I_B \)) is the identity map on \( A \) (\( B \)). The proof for \( C_A \) easily generalizes to \( C_B \) and \( C \). Finally, we prove that the same monotonicity property holds for \( C^{(n)} \) for any \( n \geq 2 \). We denote the set of ME-POVMs acting on \( A \) for the state \( \rho_{AB} \) by \( \Pi_A(\rho_{AB}) \) and similarly for \( B \).
In order to show the effect of the application of a local operation of the form \( \Lambda_i \otimes I_B \) on \( C_A(\rho_{AB}) \), we look at \( \Pi_\Lambda(\rho_{AB}) \) in a different way. Each element of this collection is a ME-POVM for \( \rho_{AB} \), i.e., they generate sets of equiprobable ensembles of states (EES) from \( \rho_{AB} \). In fact

\[
C_A(\rho_{AB}) = \max_{\{P_i\}, \Pi_\Lambda(\rho_{AB})} \mathbb{P}_\rho \left( E \left( \rho_{AB}, \{P_i\} \right) \right) - \frac{1}{2}, \tag{G1}
\]

is a maximization over all the possible EES that we can generate from \( \rho_{AB} \) with a measurement procedure on \( A \).

The effect of the first local operation that we consider is: \( \bar{\rho}_{AB} = \Lambda_A \otimes I_B(\rho_{AB}) = \sum_k (E_k \otimes \mathbb{I}_B) \cdot \rho_{AB} \cdot (E_k \otimes \mathbb{I}_B)^\dagger \), where \( \{E_k\} \) is the set of Kraus operators that defines \( \Lambda_A \). What is the relation between \( \Pi_\Lambda(\rho_{AB}) \) and \( \Pi_\Lambda(\bar{\rho}_{AB}) \)? Given an \( n \)-output ME-POVM for \( \bar{\rho}_{AB} \), i.e., \( \{P_{ij}\} \subset \Pi_\Lambda(\bar{\rho}_{AB}) \), the probabilities and the states of the output ensemble \( E(\bar{\rho}_{AB}, \{P_{ij}\}) \) are \( \bar{p}_i = \operatorname{Tr} [\bar{\rho}_{AB} \cdot P_{ij}] = 1/n \) and \( \bar{\rho}_{Bj} = \operatorname{Tr} [\bar{\rho}_{AB} \cdot P_{ij}] / \bar{p}_i \). Now we look at the term

\[
\operatorname{Tr}_A [\bar{\rho}_{AB} \cdot P_{ij}] = \operatorname{Tr} [\Lambda_A \otimes I_B(\rho_{AB}) \cdot P_{ij}] =
\]

\[
= \operatorname{Tr} \left[ \sum_k (E_k \otimes \mathbb{I}_B) \cdot \rho_{AB} \cdot (E_k^\dagger \otimes \mathbb{I}_B) \cdot P_{ij} \right] =
\]

\[
= \operatorname{Tr} \left[ \rho_{AB} \sum_k (E_k^\dagger \otimes \mathbb{I}_B) \cdot P_{ij} \cdot (E_k \otimes \mathbb{I}_B) \right] =
\]

\[
= \sum_{k} \rho_{AB} \cdot (E_k^\dagger \otimes \mathbb{I}_B) \cdot P_{ij} \cdot (E_k \otimes \mathbb{I}_B)
\]

and we rewrite the probabilities and the output states as: \( \bar{p}_i = \operatorname{Tr} [\bar{\rho}_{AB} \cdot P_{ij}] = 1/n \) and \( \bar{\rho}_{Bj} = \operatorname{Tr} [\bar{\rho}_{AB} \cdot P_{ij}] / \bar{p}_i \).

This ensemble is an EES. Next we show that: \( \{P_{ij}\} \in \{\Lambda_A \cdot (P_{ij})\} = \{\sum_k E_k^\dagger \cdot P_{ij} \cdot E_k\} \), is a POVM. The elements of \( \{P_{ij}\} \) sum up to the identity: \( \sum_i \bar{P}_{ij} = \sum_{k,j} E_k \cdot P_{ij} \cdot E_k^\dagger = \sum_j E_j \cdot P_{ij} \cdot E_j^\dagger = \mathbb{I}_B \), and they are positive operators: \( \bar{P}_{ij} = \sum_k E_k \cdot P_{ij} \cdot E_k^\dagger \). This follows that, \( \bar{P}_{ij} \) is a ME-POVM for \( \rho_{AB} \), i.e., \( \{P_{ij}\} \subset \Pi_\Lambda(\rho_{AB}) \). Thus, for every ME-POVM \( \{P_{ij}\} \) in \( \Pi_\Lambda(\bar{\rho}_{AB}) \) for \( \bar{\rho}_{AB} \), there is a ME-POVM \( \{P_{ij}\} \subset \Pi_\Lambda(\rho_{AB}) \) for \( \rho_{AB} \), such that the output ensembles are identical: \( E(\bar{\rho}_{AB}, \{P_{ij}\}) = E(\rho_{AB}, \{P_{ij}\}) \).

Therefore, it follows that:

\[
C_A(\rho_{AB}) \geq C_A(\Lambda_A \otimes I_B(\rho_{AB})) \tag{G5}
\]

for any integer \( n \geq 2 \), state \( \rho_{AB} \) and CPTP map \( \Lambda_A \).

Next we show the property of monotonicity of \( C_A(\rho_{AB}) \) under the action of local operations of the form \( I_A \otimes \Lambda_B \). We find that the collection of the ME-POVMs for \( \bar{\rho}_{AB} = I_A \otimes \Lambda_B(\rho_{AB}) \), i.e., \( \Pi_\Lambda(\bar{\rho}_{AB}) \), coincides with \( \Pi_\Lambda(\rho_{AB}) \).

In order to prove this, we apply a general POVM \( \{P_{ij}\} \) on both \( \rho_{AB} \) and \( \bar{\rho}_{AB} \) and we show that the respective output ensembles are defined by the same probability distribution. We can write \( p_i = \operatorname{Tr} [\rho_{AB} \cdot P_{ij}] \) and \( \bar{p}_i = \operatorname{Tr} [\bar{\rho}_{AB} \cdot P_{ij}] = \operatorname{Tr} [\rho_{AB} \cdot P_{ij}] \), where the last step uses the trace-preserving property of the superoperator \( I_A \otimes \Lambda_B \). Consequently, \( p_i = 1/n \) if and only if \( \bar{p}_i = 1/n \) and \( \{P_{ij}\} \in \Pi_\Lambda(\rho_{AB}) \) if and only if \( \{P_{ij}\} \in \Pi_\Lambda(\bar{\rho}_{AB}) \).

\[
\Pi_\Lambda(\rho_{AB}) = \Pi_\Lambda(\bar{\rho}_{AB}) \tag{G6}
\]

Given a ME-POVM for both \( \rho_{AB} \) and \( \bar{\rho}_{AB} \), we relate the output states

\[
\bar{\rho}_{Bj} = \Lambda_B \cdot \operatorname{Tr} [\bar{\rho}_{AB} P_{ij}] / \bar{p}_i = \Lambda_B(\rho_{Bj}) \tag{G7}
\]

From Eq. (G7) and the definition of the guessing probability, it follows that

\[
P_\rho \left( \{p_i\}, \rho_{Bj} \right) \geq P_\rho \left( \{p_i\}, \Lambda_B(\rho_{Bj}) \right) \tag{G8}
\]

and, considering Eq. (G6), Eq. (G7) and Eq. (G8)

\[
C_A(\rho_{AB}) \geq C_A \left( I_A \otimes \Lambda_B(\rho_{AB}) \right) \tag{G9}
\]

that is true for any state \( \rho_{AB} \) and CPTP map \( \Lambda_B \).

From Eq. (G6) it follows the collection of the \( n \)-output ME-POVMs does not change if we apply a CPTP map \( \Lambda_B \) on \( \rho_{AB} \).

Therefore, since Eq. (G8) is true for any number of outputs:

\[
C_A^{(n)}(\rho_{AB}) \geq C_A^{(n)} \left( I_A \otimes \Lambda_B(\rho_{AB}) \right) \tag{G10}
\]

for any integer \( n \geq 2 \), state \( \rho_{AB} \) and CPTP map \( \Lambda_B \).

We underline that from this proof we automatically obtain the invariance under local unitary transformations of \( C \) and \( C^{(n)} \) for any \( n \geq 2 \).

**Appendix H**: \( C_A(\rho_{AB}) \geq C_B^{(2)}(\rho_{AB}) \)

In this section (where we omit the time dependence of \( \rho_{AB}(t), \rho_{B}^{\tau}(t) \) and \( \rho_B^{\tau}(t) \)) we show that \( C_A(\rho_{AB}) \geq C_B^{(2)}(\rho_{AB}) \), where \( C_B^{(2)}(\rho_{AB}) \) is defined by

\[
C_B^{(2)}(\rho_{AB}) \equiv \max_{\{P_{ij}\} \in \mathcal{C}_{\rho_{AB}}^{(2)}} \mathbb{P}_\rho \left( E(\rho_{AB}, \{P_{ij}\}) \right) - \frac{1}{2},
\]
where \( \Pi^{(2)}_{B}(\rho^{(r)}_{AB}) \) is the set of the 2-output ME-POVMs. In Appendix III we showed that \( C^{2}_{B}(\rho^{(r)}_{AB}) = C_{B}(\rho^{(r)}_{AB}) \) and this completes the proof that \( C_{A}(\rho^{(r)}_{AB}) \geq C_{B}(\rho^{(r)}_{AB}) \).

We apply a general but fixed 2-outcome ME-POVM for \( \rho^{(r)}_{AB} \), where now the measured system is \( B : \{ \rho^{(r)}_{B} \}_{i} = \{ P_{B}, \overline{P}_{B} \} \in \Pi_{B}(\rho^{(r)}_{AB}) \), where \( \overline{P}_{B} = I_{B} - P_{B} \). The output ensemble \( \mathcal{E}(\rho^{(r)}_{AB}, \{ \rho^{(r)}_{B} \}_{i}) = \{ \rho_{A,i}, \rho_{A,i} \} \) is composed by an uniform distribution (by definition of ME-POVM) and states in the following form:

\[
\rho_{A,1} = \frac{1}{2} \text{Tr}_{B} \left[ \rho^{(r)}_{B} + \rho^{(r)}_{B}' \right] P_{B},
\]

\[
\rho_{A,2} = \frac{1}{2} \text{Tr}_{B} \left[ \rho^{(r)}_{B} - \rho^{(r)}_{B}' \right] \overline{P}_{B},
\]

\[
\rho_{A,1} = |0\rangle \langle 0|_{A} \text{Tr}_{B} \left[ \rho^{(r)}_{B} P_{B} \right] + |1\rangle \langle 1|_{A} \text{Tr}_{B} \left[ \rho^{(r)}_{B}' P_{B} \right],
\]

\[
\rho_{A,2} = |0\rangle \langle 0|_{A} \text{Tr}_{B} \left[ \rho^{(r)}_{B} \overline{P}_{B} \right] + |1\rangle \langle 1|_{A} \text{Tr}_{B} \left[ \rho^{(r)}_{B}' \overline{P}_{B} \right].
\]

If we use Eq. (69) to get \( P_{g}(\mathcal{E}(\rho^{(r)}_{AB}, \{ \rho^{(r)}_{B} \}_{i})) \), firstly we have to evaluate \( \| \rho_{A,1} - \rho_{A,2} \| \). Hence, with Eqs. (H1)-(H4), we can write it as:

\[
\| |0\rangle \langle 0|_{A} \text{Tr}_{B} \left[ \rho^{(r)}_{B} P_{B} \right] + |1\rangle \langle 1|_{A} \text{Tr}_{B} \left[ \rho^{(r)}_{B}' P_{B} \right] \|_{1} = 
\| \text{Tr}_{B} \left[ \rho^{(r)}_{B} \Delta P_{B} \right] + \text{Tr}_{B} \left[ \rho^{(r)}_{B}' \Delta P_{B} \right] \|,
\]

where \( \Delta P_{B} = P_{B} - \overline{P}_{B} \). Hence:

\[
\| \rho_{A,1} - \rho_{A,2} \| = \max_{\rho} \| \text{Tr}_{B} \left[ \rho^{(r)}_{B} \pm \rho^{(r)}_{B}' \right] \Delta P_{B} \|.
\]

Using Eq. (H1) and Eq. (H2) we see that

\[
\| \text{Tr}_{B} \left[ \rho^{(r)}_{B} \pm \rho^{(r)}_{B}' \right] \Delta P_{B} \| = \| \text{Tr}_{B} \left[ \rho^{(r)}_{B} \pm \rho^{(r)}_{B}' \right] P_{B} \| = 2\| \rho_{A,1} - \rho_{A,2} \| = 0.
\]

Hence:

\[
\| \rho_{A,1} - \rho_{A,2} \| = \| \text{Tr}_{B} \left[ \rho^{(r)}_{B} - \rho^{(r)}_{B}' \right] (2P_{B} - \overline{P}_{B}) \| = 2\| \text{Tr}_{B} \left[ \rho^{(r)}_{B} - \rho^{(r)}_{B}' \right] \|,
\]

from which follows that \( C^{2}_{B}(\rho^{(r)}_{AB}) \) is equal to

\[
\max_{\{ P_{B}\}_{i} \in \Pi_{B}(\rho^{(r)}_{AB})} \frac{\| \text{Tr}_{B} \left[ \rho^{(r)}_{B} - \rho^{(r)}_{B}' \right] \|}{2}.
\]

To compare with \( C_{A}(\rho^{(r)}_{AB}) \), we write it as

\[
C_{A}(\rho^{(r)}_{AB}) = P_{g}(\{ \rho_{A,1,2} = 1/2 \}, \{ \rho^{(r)}_{B}, \rho^{(r)}_{B}' \}) - \frac{1}{2} = 
\max_{\{ P_{B}\}_{i} \in \Pi_{B}} \frac{\text{Tr}_{B} \left[ \rho^{(r)}_{B} P_{B} + \rho^{(r)}_{B}' \overline{P}_{B} \right] - 1}{2}.
\]

We have used the definition in Eq. (I1): \( \Pi_{B} \) is the collection of all the POVMs that we can perform on \( B \). It follows that the only difference between \( C^{2}_{B}(\rho^{(r)}_{AB}) \) and \( C_{A}(\rho^{(r)}_{AB}) \) is in the maximization procedure: in the former we maximize only over the ME-POVMs \( \Pi_{B}(\rho^{(r)}_{AB}) \), while in the latter we can pick any POVM: \( C_{A}(\rho^{(r)}_{AB}(i)) \geq C^{2}_{B}(\rho^{(r)}_{AB}(i)) \) follows as a natural consequence.

**Appendix I: Proof that \( C_{g}(\rho^{(r)}_{AB}) = C^{2}_{B}(\rho^{(r)}_{AB}) \)**

In this Appendix, in contrast to Appendix III, we consider the action of any ME-POVM over \( B \) for \( \rho^{(r)}_{AB} \). We want to show that for each ME-POVM \( \{ P_{B} \}_{i} \) that we can consider in \( C_{B}(\rho^{(r)}_{AB}) \), where \( i \) runs from 1 to \( n > 2 \), we can always find at least one 2-outcome ME-POVM acting on \( B \), i.e. \( \{ P_{B,1}, P_{B,2} \} \), that provides an ensemble with a higher value of \( P_{g}(\mathcal{E}) \). We recall that, if \( \mathcal{E} = \{ \rho_{1}, \rho_{2} \} \) is a generic ensemble of \( n \) states defined on \( \mathcal{H} \), where \( \mathcal{H} \) is a finite generic dimensional Hilbert space, the guessing probability of \( \mathcal{E} \) is

\[
P_{g}(\mathcal{E}) \equiv \max_{\{ P_{i} \}} \sum_{i=1}^{n} p_{i} \text{Tr}[\rho_{i} \cdot \rho_{i}],
\]

where the maximization is performed over the space of the \( n \)-outcome POVMs \( \{ P_{i} \} \), on \( \mathcal{H} \). Starting from a general \( n \)-outcome ME-POVM \( \{ P_{B,1}, P_{B,2} \} \), we construct the corresponding 2-output ME-POVM \( \{ P_{B,1}, P_{B,2} \} \) that accomplishes this task.

For every given \( n \)-outcome ME-POVM \( \{ P_{B} \}_{i} \) for \( \rho^{(r)}_{AB} \), we can generate an equiprobable ensemble of states (EES) of the form \( \mathcal{E}(\rho^{(r)}_{AB}(\{ P_{B} \}_{i})) = \{ \{ \rho_{i} = 1/n, \rho_{A,1} \} \} \). The guessing probability of this ensemble, which we denote by \( P^{(n)}_{g} = P_{g}(\mathcal{E}(\rho^{(r)}_{AB}(\{ P_{B} \}_{i}))) \), is

\[
P^{(n)}_{g} = \text{Tr} \left[ \rho^{(r)}_{AB} \cdot \sum_{i=1}^{n} \rho_{A,i} \otimes \rho_{B,i} \right],
\]

where \( \{ \rho_{A,i} \} \) is a POVM that provides the maximum in Eq. (I1). If \( n \) is even we consider the following 2-output POVM

\[
P^{(2)}_{B,1} = \sum_{i \in E_{1}} P^{(n)}_{B,i}, \quad P^{(2)}_{B,2} = \sum_{i \in E_{2}} P^{(n)}_{B,i},
\]

where \( E_{1} \) and \( E_{2} \) are any two sets of \( n/2 \) indices such that \( E_{1} \cup E_{2} = \{ 1, 2, \ldots, n \} \). This structure guarantees that Eq. (13) is a 2-output ME-POVM for \( \rho^{(r)}_{AB} \). We compare Eq. (12) with
the guessing probability of the output ensemble that we obtain by applying Eq. (13) on $\rho_{AB}^{(2)}$.

\[ P_g^{(2)} = \max_{\{P_{A,i}\}_{i=1}^2} \text{Tr} \left[ \rho_{AB}^{(2)} \left( \sum_{i=1}^2 P_{A,i} \otimes P_{B,i}^{(2)} \right) \right] \geq \text{Tr} \left[ \rho_{AB}^{(2)} \left( \sum_{i=1}^2 P_{A,i}^{(2)} \otimes P_{B,i}^{(2)} \right) \right], \quad (14) \]

where the POVM $\{P_{A,i}\}_{i=1}^2$ is defined by

\[ P_{A,1}^{(2)} = \sum_{i \in E_1} P_{A,i}^{(n)} \text{ and } P_{A,2}^{(2)} = \sum_{i \in E_2} P_{A,i}^{(n)}. \quad (15) \]

For the guessing probability $P_g^{(2)}$, we have

\[ P_g^{(2)} \geq \text{Tr} \left[ \rho_{AB}^{(2)} \left( P_{A,1}^{(2)} \otimes P_{B,1}^{(2)} + P_{A,2}^{(2)} \otimes P_{B,2}^{(2)} \right) \right] = \text{Tr} \left[ \rho_{AB}^{(2)} \left( \sum_{i=1}^n P_{A,i}^{(n)} \otimes P_{B,i}^{(n)} + P_{mix}^{(n)} \right) \right] = P_g^{(n)} + \text{Tr} \left[ \rho_{AB}^{(2)} \left( \frac{-2P_{A,x}}{2} \otimes P_{B,x}^{(n)} + \frac{\sum_{i \neq x} P_{A,i}^{(n)} \otimes P_{B,i}^{(n)}}{2} \right) \right] = P_g^{(n)} + \text{Tr} \left[ \rho_{AB}^{(2)} \left( \frac{1-A}{2} \otimes P_{B,x}^{(n)} \right) \right]. \quad (19) \]

where $P_{mix}^{(n)}$ represents terms that provide positive contributions to $P_g^{(2)}$. We have to find a value of $x$ that makes the second term of the last relation positive. Let $a_x$ and $b_x$ be the diagonal elements of $P_{A,x}$ in the orthonormal basis $\{|0\rangle_A, |1\rangle_A\}$. We recall that $\rho_{AB}^{(2)} = (|0\rangle_0 \otimes \rho_B^{(2)} + |1\rangle_1 \otimes \rho_B^{(2)})/2$ and we obtain

\[ P_g^{(2)} \geq P_g^{(n)} + \text{Tr} \left[ \rho_{AB}^{(2)} \left( \frac{-2a_x}{4} \otimes P_{B,x}^{(n)} + \frac{1}{4} \otimes \rho_B^{(2)} \right) \right] \geq P_g^{(n)}. \quad (19) \]

where the second term on the right-hand side of the inequality is definitely positive when $a_x, b_x \leq 1/2$. From $\sum_a P_{A,a}^{(n)} = 1$ follows that $\sum_a a_x = 1$ and $\sum_a b_x = 1$. Therefore, if $a_x > 1/2 (b_x > 1/2)$, then $a_x \leq 1/2 (b_x \leq 1/2)$ for any $y \neq x$. In order to fix the value of $x$, we must consider $a_x$ and $b_x$ could be bigger than 1/2 for two different values of $x$: let’s say $x_a$ and $x_b$. Even in this “worst-case” scenario we still have $n – 2$ other possible choices for $x$ such that $(1 - 2a_x), (1 - 2b_x) \geq 0$.

Equations (16) and (110) show that, when we evaluate $C_B(\rho_{AB}^{(2)})$, the guessing probability of the ensembles generated by the $n$-output ME-POVMs is never bigger than the one that we obtain if we only consider the 2-output ME-POVMs: $C_B^{(2)}(\rho_{AB}^{(2)}) = C_B(\rho_{AB}^{(2)})$. Thanks to this result we can finally say that $C_A(\rho_{AB}^{(2)}(t)) \geq C_B(\rho_{AB}^{(2)}(t))$ and $C(\rho_{AB}^{(2)}(t)) = C_A(\rho_{AB}^{(2)}(t))$. This result is valid if we consider $\rho_{AB}^{(2)}$, but in general it is not true.
Appendix J: Proof that $C_A(P_{AB}^{(r)}) = C_A^2(P_{AB}^{(r)})$

When we considered $C_A(P_{AB}^{(r)})$, we have seen that if the maximization over the ME-POVMs is considered only over the 2-output ones, the maximum is obtained for $\{P_{A,i}^{(n)}\}_i = \{|0\rangle|0\rangle, |1\rangle|1\rangle\}$. In order to complete the proof, we need to show that even if we consider general $n$-output ME-POVMs (as in the definition (G1)), we don’t get higher guessing probabilities of the corresponding output ensembles. In other words, if we use the definition

$$C_A^2(P_{AB}^{(r)}) = \max_{\{P_{A,i}\} \in \Pi_2^\epsilon(P_{AB}^{(r)})} P_g \left(\mathcal{E}(P_{A,i}^{(r)}, P_{A,i})\right) - \frac{1}{2},$$

where $\Pi_2^\epsilon(P_{AB}^{(r)})$ contains only the 2-output ME-POVMs of $P_{AB}^{(r)}$, then $C_A(P_{AB}^{(r)}) = C_A^2(P_{AB}^{(r)})$.

To see this we can make the same analysis as done in Appendix I for $C_A(P_{AB}^{(r)})$ but we switch the role of A and B in Eq. (13) and Eq. (15) when $n$ is even and Eq. (17) and Eq. (18) when $n$ is odd. The definitions for $P^{(n)}_g$, $P^{(2)}_g$, $E_{1,2}$ and $O_{1,2}$ are preserved.

The guessing probability of an EES generated by a ME-POVM $\{P_{A,i}^{(n)}\}_i$ with an even number of outputs is

$$P^{(n)}_g = \text{Tr} \left[ P_{AB}^{(r)} \cdot \left( \sum_{i=1}^{n} P_{A,i}^{(n)} \otimes P_{B,i}^{(n)} \right) \right],$$

where $\{P_{B,j}^{(n)}\}_j$ is a POVM that maximizes the guessing probability in Eq. (11). The 2-output ME-POVM that provides a higher guessing probability is

$$P^{(2)}_{A,1} = \sum_{i \in E_1} P_{A,i}^{(n)}, \quad P^{(2)}_{A,2} = \sum_{i \in E_2} P_{A,i}^{(n)}.$$  \hspace{1cm} (J1)

We define the following POVM on the system $B$

$$P^{(2)}_{B,1} = \sum_{i \in E_1} P_{B,i}^{(n)}, \quad P^{(2)}_{B,2} = \sum_{i \in E_2} P_{B,i}^{(n)}.$$  \hspace{1cm} (J2)

Consequently, we consider the following inequality

$$P^{(2)}_g \geq \text{Tr} \left[ P_{AB}^{(r)} \cdot \sum_{i=1}^{n} P_{A,i}^{(n)} \otimes P_{B,i}^{(n)} \right] = P^{(2)}_g + \text{Tr} \left[ P_{AB}^{(r)} \cdot \sum_{k=1}^{2} \sum_{i,j \in E_k} P_{A,i}^{(n)} \otimes P_{B,j}^{(n)} \right],$$

which shows that $P^{(2)}_g \geq P^{(n)}_g$. If $n$ is odd, we use again the technique from Appendix I, where we switch the role of A and B, to obtain the inequality

$$P^{(2)}_g \geq P^{(n)}_g + \text{Tr} \left[ P_{AB}^{(r)} \cdot \frac{1}{2} \frac{A_{i,x}}{A_{i,x}} \otimes \tilde{P}_{B,i}^{(n)} \right],$$

where the right-hand side is greater than $P^{(n)}_g$ if $x$ is suitably chosen.

We underline that the results given in this section and Appendix HI suffice to state that $C_A^2(P_{AB}^{(r)}) = C_A^2(P_{AB}^{(r)}) \geq C_B^2(P_{AB}^{(r)})$. 

