Equivalence between non-Markovian dynamics and correlation backflows

Dario De Santis and Markus Johansson
ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain

Keywords: non-Markovianity, open quantum systems, backflow, correlation

Abstract
The information encoded into an open quantum system that evolves under a Markovian dynamics is always monotonically non-increasing. Nonetheless, for a given quantifier of the information contained in the system, it is in general not clear if for all non-Markovian dynamics it is possible to observe a non-monotonic evolution of this quantity, namely a backflow. We address this problem by considering correlations of finite-dimensional bipartite systems. For this purpose, we consider a class of correlation measures and prove that if the dynamics is non-Markovian there exists at least one element from this class that provides a correlation backflow. Moreover, we provide a set of initial probe states that accomplish this witnessing task. This result provides the first one-to-one relation between non-Markovian quantum dynamics and correlation backflows. Finally, we introduce a measure of non-Markovianity.

1. Introduction
The study of open quantum systems dynamics [1, 2] is of central interest in quantum mechanics. Since there are no experimental scenarios where a quantum system can be considered completely isolated, this approach provides a more realistic description of quantum evolutions.

The interaction between an open quantum system $S$ and its environment $E$ leads to two possible regimes of evolution. The phenomena associated with the Markovian regime are characterized by the monotonic non-increase of the information contained in the open system. Instead, in the non-Markovian regime, part of the information lost is recovered in one or more subsequent time intervals. This phenomenon is called backflow of information. For some detailed reviews on non-Markovian evolutions, see references [3–6].

It is nonobvious what mathematical structure is better suited to reproduce the backflow phenomenology. A framework based on a notion of divisibility of dynamical maps, namely the operators describing the dynamical evolution of the system, has achieved a promising consensus [2, 7–14]. A characterizations of non-Markovian evolutions based on divisibility is proposed in reference [15], where the authors introduce a degree of non-Markovianity to classify evolutions. Many efforts are presently directed towards testing this mathematical definition by studying the characteristic backflows that different physical quantities show when the evolution is non-Markovian. Once we consider a quantity that is non-increasing under Markovian evolutions, we can study its potential to show a backflow when the dynamics is non-Markovian. Distinguishability between states [11–13], correlation measures [14, 16–18], channel capacities [19] and the volume of accessible states [20] and quantum Fisher information [21] are some quantities that have been studied in this scenario. The non-trivial point that has to be analyzed is if it is possible to obtain one-to-one relations between backflows of these quantities and non-Markovian evolutions. Indeed, this result would imply a correspondence between the phenomenological and the mathematical description of non-Markovianity that we have presented. In reference [9] it was suggested that for bijective evolutions there is a one-to-one correspondence between backflow of the distinguishability of two-state ensembles and non-Markovianity. This correspondence follows from the results of [22, 23] together with an addendum given in reference [13]. Later it was shown in reference [12] that for general evolutions a one-to-one correspondence exists between non-Markovianity and backflow of the guessing probability for some ensemble of states. Furthermore, it was shown in reference [13] that for evolutions that...
are non-bijective for at most a discrete set of times there is a one-to-one correspondence between backflows of the distinguishability of an equiprobable two-state ensemble and non-Markovianity.

In this work we focus on the connection between revivals of bipartite correlation measures and when the evolution of one subsystem $S$ is non-Markovian. Several measures have been considered in this scenario, e.g. quantum mutual information [14, 16] and entanglement measures [17]. Recently, a correlation that witnesses almost all non-Markovian dynamics has been introduced [18]. However, it is unknown if any of these correlations can witness all non-Markovian dynamics [16].

The main result of this work is the first one-to-one relation between correlation backflows and non-Markovian dynamics. We consider a class of bipartite correlation measures that provides backflows if and only if the dynamics is not Markovian. For this purpose, we exploit supplementary ancillary systems to define initial probe states that succeed in this witnessing task. Finally, by considering the maximum backflow that these correlation measures can show when bipartite states evolve, we introduce a class of non-Markovianity measures. We prove that for any non-Markovian evolution there exists at least one measure from this class that is positive.

2. Markovianity and divisibility

Given a generic finite-dimensional Hilbert space $\mathcal{H}$, we define $B(\mathcal{H})$ to be the set of linear bounded operators that act on $\mathcal{H}$ and $S(\mathcal{H})$ the set of positive semidefinite, Hermitian and trace-one operators on $\mathcal{H}$, namely the state space of $\mathcal{H}$.

We consider an open quantum system $S$ defined by the Hilbert space $\mathcal{H}_S$ that, at the initial time $t_0$, is uncorrelated with the surrounding environment. The evolution of $S$ from $t_0$ to $t \geq t_0$ is given by a dynamical map: a completely positive and trace-preserving (CPTP) linear operator $\Lambda_S(t, t_0) : S(\mathcal{H}_S) \to S(\mathcal{H}_S)$. Therefore, the complete evolution of $S$ from $t_0$ to any time $t \geq t_0$ is described by a family of dynamical maps $\{\Lambda_S(t, t_0)\}_t$, where $\Lambda_S(t, t_0)$ is CPTP for every $t \geq t_0$.

The property of evolutions needed to define Markovianity is the completely positive (CP) divisibility of $\{\Lambda_S(t, t_0)\}$, in terms of intermediate maps $\mathcal{V}_S(t, t')$.

**Definition 1.** The evolution $\{\Lambda_S(t, t_0)\}_t$ is called CP-divisible if, for any $t \geq t_0$, the dynamical map $\Lambda_S(t, t_0)$ can be decomposed as a sequence of CPTP linear maps $\Lambda_S(t, t_0) = \mathcal{V}_S(t, t')\Lambda_S(t', t_0)$, where $\mathcal{V}_S(t, t')$ is a CPTP linear map for any $t_0 \leq t' \leq t$.

CP-divisibility is commonly used to define Markovian dynamics and it is the definition considered in this work: $\{\Lambda_S(t, t_0)\}$ is Markovian if and only if it is CP-divisible. Likewise, we call an evolution non-Markovian if and only if for some $t_0 \leq t' \leq t$ there is no CPTP intermediate map $\mathcal{V}_S(t, t')$.

3. Measurements with fixed output probability distributions

Any measurement process on a quantum state $\rho \in S(\mathcal{H})$ is defined by a positive-operator valued measure (POVM), namely an indexed set of Hermitian and positive semi-definite operators $\{P_i\}_{i=1}^n$ of $B(\mathcal{H})$ such that $\sum_{i=1}^n P_i = 1$, where $\mathcal{I} \in B(\mathcal{H})$ is the identity operator on $\mathcal{H}$ and $n$ is the number of possible outcomes. The $i$th output of the measurement is represented by $P_i$, where $p_i = \text{Tr} \{\rho P_i\}$ is the corresponding occurrence probability.

We consider finite probability distributions $\mathcal{P} = \{p_i\}_{i=1}^n$, where $\sum_{i=1}^n p_i = 1$ and define the set of $n$-output POVMs that, if applied on $\rho \in S(\mathcal{H})$, provide $\mathcal{P}$-distributed outcomes.

**Definition 2.** Given the finite probability distribution $\mathcal{P} = \{p_i\}_{i=1}^n$, the $n$-output POVM $\{P_i\}_{i=1}^n$ on $\mathcal{H}$ is a $\mathcal{P}$-POVM for $\rho \in S(\mathcal{H})$ if and only if it belongs to

$$
\Pi^\mathcal{P}(\rho) = \big\{ \{P_i\}_{i=1}^n : \text{Tr} \{\rho P_i\} = p_i, \forall i = 1, \ldots, n \big\}.
$$

Now consider a bipartite scenario where Alice and Bob share a state $\rho_{AB} \in S(\mathcal{H}_A \otimes \mathcal{H}_B)$. If Alice applies a POVM $\{P_{A,i}\}_{i=1}^n$ on her side of $\rho_{AB}$, an output ensemble $E(\rho_{AB}, \{P_{A,i}\}_{i=1}^n) = \{p_i, \rho_{B,i}\}_{i=1}^n$ is generated on Bob’s side, where each state $\rho_{B,i} \in S(\mathcal{H}_B)$ occurs with probability $p_i$ and

$$
p_i = \text{Tr} \{\rho_{AB} P_{A,i} \otimes \mathcal{I}_B\}, \quad \rho_{B,i} = \text{Tr}_A \left[\rho_{AB} P_{A,i} \otimes \mathcal{I}_B\right] / p_i.
$$

In particular, with probability $p_i$, Alice obtains the $i$th outcome of the measurement and Bob’s side of the shared state is transformed into $\rho_{B,i}$. We call $\{p_i\}_{i=1}^n$ and $\{\rho_{B,i}\}_{i=1}^n$ respectively the output probability distribution and the output states of the measurement.

Similarly to $\Pi^\mathcal{P}(\rho)$, we define the measurements that Alice can perform on $\rho_{AB}$ to generate $\mathcal{P}$-distributed output ensembles on Bob’s side (see figure 1).
Given the finite probability distribution \( \mathcal{P} \) and an \( n \)-output POVM \( \{P_{A,i}\}_{i=1}^{n} \) on \( \mathcal{H}_{A} \), \( \mathcal{P} \)-POVM on \( A \) for \( \rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB}) \) if and only if it belongs to

\[
\Pi_{A}^{\mathcal{P}}(\rho_{AB}) \equiv \{ \{P_{A,i}\}_{i=1}^{n} : \text{Tr}_{B} [\rho_{AB} P_{A,i} \otimes I_{B}] = p_{i}, \quad \forall i = 1, \ldots, n \}.
\]

Analogously, we can define \( \Pi_{B}^{\mathcal{P}}(\rho_{AB}) \). We notice that \( \Pi_{A}^{\mathcal{P}}(\rho_{AB}) = \Pi_{B}^{\mathcal{P}}(\rho_{A}) \) for any \( \mathcal{P} \) and \( \rho_{AB} \), where \( \rho_{A} = \text{Tr}_{B} [\rho_{AB}] \). Moreover, \( \Pi_{A}^{\mathcal{P}}(\rho) (\Pi_{A}^{\mathcal{P}}(\rho_{AB})) \) is a non-empty convex set for any \( \rho (\rho_{AB}) \) and \( \mathcal{P} \).

### 4. Witnessing non-Markovianity with distinguishability of ensembles

We consider the task of identifying a state that we randomly extract from a known ensemble \( \mathcal{E} = \{\rho_{i}, \rho_{i}'\}_{i=1}^{n} \) of states of \( \mathcal{S}(\mathcal{H}) \). The guessing probability \( P_{g}(\mathcal{E}) \) is the average probability to successfully identify the extracted state with an optimal measurement, that is,

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

where the maximization is performed over the \( n \)-output POVMs of \( \mathcal{B}(\mathcal{H}) \). We say that the larger is \( P_{g}(\mathcal{E}) \), the more \( \mathcal{E} \) is distinguishable. Indeed, the maximum value \( P_{g}(\mathcal{E}) = 1 \) is obtained for orthogonal state ensembles.

Now we describe how guessing probability can be used to witness non-Markovianity. We consider a finite-dimensional system \( \mathcal{H}_{S} \otimes \mathcal{H}_{A'} \), where the open quantum system \( \mathcal{S} \) is evolved by a generic \( \{A_{S}(t_{0})\}_{t} \), and \( A' \) is an ancilla that is left untouched. Given an initial ensemble \( \mathcal{E}_{SA}(t_{0}) = \{\rho_{i}, \rho_{i}'\}_{i=1}^{n} \), we consider its evolution:

\[
\mathcal{E}_{SA}(t_{0}) \rightarrow \mathcal{E}_{SA}(t) = \{\rho_{i}, A_{S}(t_{0}) \otimes I_{A'}(\rho_{i}'(\rho_{i})))\},
\]

where \( I_{A'} : \mathcal{S}(\mathcal{H}_{A'}) \rightarrow \mathcal{S}(\mathcal{H}_{A'}) \) is the identity map on \( \mathcal{S}(\mathcal{H}_{A'}) \). For any CPTP map \( \Lambda' \) acting on the states of \( \mathcal{E} = \{\rho_{i}, \rho_{i}'\}_{i} \), \( P_{g}(\mathcal{E}) \) is non-increasing: \( P_{g}(\{\rho_{i}, \rho_{i}'\}_{i}) \geq P_{g}(\{\rho_{i}, \Lambda'(\rho_{i}))_{i}\}) \). Therefore, if \( \{A_{S}(t_{0})\}_{t} \) is Markovian,

\[
P_{g}(\mathcal{E}_{SA}(\tau + \Delta \tau)) - P_{g}(\mathcal{E}_{SA}(\tau)) \leq 0
\]

for every \( \tau \geq t_{0} \) and \( \Delta \tau \geq 0 \).

Given any evolution \( \{A_{S}(t_{0})\}_{t} \) and time interval \( [\tau, \tau + \Delta \tau] \), there exist an ancillary system \( A' \) and an initial ensemble \( \mathcal{E}_{SA}(t_{0}) \) of separable states of \( \mathcal{S}(\mathcal{H}_{S} \otimes \mathcal{H}_{A'}) \)

\[
\mathcal{E}_{SA}(t_{0}) \equiv \{\rho_{0}, \mathcal{E}_{SA}(t_{0})\}_{i=1}^{\pi},
\]

such that we have a backflow

\[
P_{g}(\mathcal{E}_{SA}(\tau + \Delta \tau)) - P_{g}(\mathcal{E}_{SA}(\tau)) > 0,
\]

if and only if there is no CPTP intermediate map \( V_{S}(\tau + \Delta \tau, \tau) \), as shown in reference [12]. Moreover,

\( \mathcal{E}_{SA}(t_{0}) \) depends on \( \{A_{S}(t_{0})\}_{t} \) and \( [\tau, \tau + \Delta \tau] \). The result of reference [12] is general and applies to any finite-dimensional system evolution.
5. A class of correlation measures

Let $\mathcal{P} \equiv \{p_i\}$ be a generic probability distribution and $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ a generic bipartite system state. We consider the correlation measure

$$C_A^\mathcal{P}(\rho_{AB}) \equiv \max_{\{P_{A,i}\} \subseteq \mathcal{P}(\rho_{AB})} P_\mathcal{P}\{\rho_{AB}, \{P_{A,i}\}\} - p_{\text{max}},$$

(7)

where the maximization is performed over the $\mathcal{P}$-POVMs on $A$ for $\rho_{AB}$ and we defined $P_\mathcal{P}(\rho_{AB}, \{P_{A,i}\}) \equiv P_A(\mathcal{E}(\rho_{AB}, \{P_{A,i}\}))$ and $p_{\text{max}} \equiv \max p_i$ (see figure 1). Therefore, we can consider the class of correlation measures where each element is defined by a different distribution $\mathcal{P}$. Notice that, if $\mathcal{P}_1$ can be transformed into $\mathcal{P}_2$ by a permutation and an addition or removal of one or more zero-valued probabilities, $C_A^\mathcal{P}_1(\rho_{AB}) = C_A^\mathcal{P}_2(\rho_{AB})$ for any $\rho_{AB}$.

The operational meaning of this correlation measure for a given $\mathcal{P}$ is the following. The value of $C_A^\mathcal{P}(\rho_{AB})$ (modulo $p_{\text{max}}$) corresponds to the guessing probability of the most distinguishable $\mathcal{P}$-distributed state ensembles of $B$ that Alice can generate measuring her side of $\rho_{AB}$. Therefore, $C_A^\mathcal{P}_1(\rho_{AB}) > C_A^\mathcal{P}_2(\rho_{AB})$ implies that the largest distinguishability of the $\mathcal{P}$-distributed output ensembles of $B$ that Alice can generate measuring $\rho_{AB}$ is greater than the largest distinguishability of the $\mathcal{P}_2$-distributed output ensembles of $B$ that Alice can generate measuring $\rho_{AB}$.

To consider $C_A^\mathcal{P}_1$ a proper correlation measure, we have to show that it is: zero-valued for product states, non-negative and monotonically decreasing under local operations [16]. In order to prove the first property, given a generic product state $\rho_{AB} = \rho_A \otimes \rho_B$, the output ensemble $\mathcal{E}(\rho_A \otimes \rho_B, \{P_{A,i}\}) = \{p_i, p_B\}$, is made of identical states for any POVM $\{P_{A,i}\}$, and $P_\mathcal{P}(p_i, p_B) = p_{\text{max}}$. Therefore, while $C_A^\mathcal{P}(\rho_{AB}) \geq 0$ is now trivial, the proof for the monotonicity of $C_A^\mathcal{P}(\rho_{AB})$ under local operations is in appendix A.

Similarly, we define

$$C_B^\mathcal{P}(\rho_{AB}) \equiv \max_{\{P_{B,i}\} \subseteq \mathcal{P}(\rho_{AB})} P_\mathcal{P}\{\rho_{AB}, \{P_{B,i}\}\} - p_{\text{max}}.$$

(8)

Since in general $C_A^\mathcal{P}(\rho_{AB}) \neq C_B^\mathcal{P}(\rho_{AB})$, we can consider the symmetric class of measures

$$C_{AB}^\mathcal{P}(\rho_{AB}) \equiv \max \{C_A^\mathcal{P}(\rho_{AB}), C_B^\mathcal{P}(\rho_{AB})\}.$$

(9)

Finally, we notice that equations (7)–(9) are generalizations of the correlation measures introduced in reference [18]. In particular, a scenario where only uniform probability distributions $\mathcal{P} = \{p_i = 1/n\}_{i=1}^n$ was considered in reference [18]. Moreover, the authors proved that by considering $C_{AB}^\mathcal{P}$ with $\mathcal{P} = \{1/2, 1/2\}$ it is possible to witness any bijective or pointwise non-bijective evolution, while a proof for generic non-Markovian evolutions is not provided.

5.1. The probe states

The goal of this work is to prove a one-to-one correspondence between non-Markovianity and correlation backflows. Therefore, we consider the most general evolution $\{A_t(t, t_0)\}$, and we focus on a generic time interval $[\tau, \tau + \Delta \tau]$. We provide an initial probe state $\rho_{AB}$ shared between Alice and Bob and a distribution $\mathcal{P}$ for which the correlation measure $C_A^\mathcal{P}$ shows a backflow between $\tau$ and $\tau + \Delta \tau$ if and only if there is no CPTP intermediate map $V_S(\tau + \Delta \tau, \tau)$.

First, we introduce the bipartition and the state space needed to consider $C_A^\mathcal{P}$ and the initial probe state. We define the bipartite system $S(\mathcal{H}_A \otimes \mathcal{H}_B)$ such that $\dim(\mathcal{H}_A) = \pi$ and $\mathcal{H}_B \equiv \mathcal{H}_S \otimes \mathcal{H}_A' \otimes \mathcal{H}_{\mathcal{A}'\mathcal{A}''}$, where $\dim(\mathcal{H}_S) = \dim(\mathcal{H}_{\mathcal{A}'\mathcal{A}''}) = d_S = \pi + 1$. We fix the following orthonormal basis for $\mathcal{H}_A$ and $\mathcal{H}_{\mathcal{A}'\mathcal{A}''}$: $\mathcal{A}_A \equiv \{|I\rangle_{\mathcal{A}_A}\}_{i=1}^{\pi+1} = \{|1\rangle_{\mathcal{A}_A}, |2\rangle_{\mathcal{A}_A}, \ldots, |\pi\rangle_{\mathcal{A}_A}\}$ and $\mathcal{A}_{\mathcal{A}'\mathcal{A}''} \equiv \{|I\rangle_{\mathcal{A}'\mathcal{A}''}\}_{i=1}^{\pi+1} = \{|1\rangle_{\mathcal{A}'\mathcal{A}''}, |2\rangle_{\mathcal{A}'\mathcal{A}''}, \ldots, |\pi + 1\rangle_{\mathcal{A}'\mathcal{A}''}\}$. Notice that the ancillas $A'$ and $A''$ can be considered as a single ancilla with Hilbert space $\mathcal{H}_{\mathcal{A}'\mathcal{A}''}$ (see figure 2).

We define $\mathcal{P}_{B,i} \equiv \mathcal{P}_{\mathcal{S}A'}(\mathcal{A}_A) \otimes (\pi + 1) |i\rangle |i\rangle_{\mathcal{A}''} \in S(\mathcal{H}_B)$, for $i = 1, \ldots, \pi$, where we made use of the elements of $\mathcal{E}(\mathcal{S}_A(t_0)) = \{\mathcal{P}_i \mathcal{P}_{\mathcal{S}A'}(\mathcal{A}_A)\}_{i=1}^\pi$ (see equation (5)). We introduce a class of initial probe states $\rho_{AB}^{(i)}(t_0) \in S(\mathcal{H}_A \otimes \mathcal{H}_B)$ parametrized by $\lambda \in [0, 1)$$

$$\rho_{AB}^{(i)}(t_0) \equiv \sum_{i=1}^\pi p_i |i\rangle |i\rangle_{\mathcal{A}_A} \otimes (\lambda \sigma_{\mathcal{S}A'} \otimes |i\rangle |i\rangle_{\mathcal{A}''} + (1 - \lambda) \mathcal{P}_{B,i}),$$

(10)

where $\sigma_{\mathcal{S}A'}$ is a generic state of $S(\mathcal{H}_S \otimes \mathcal{H}_{\mathcal{A}'\mathcal{A}''})$. Notice that the index $i$ runs from 1 to $\pi$ and $\dim(\mathcal{H}_{\mathcal{A}'\mathcal{A}''}) = \pi + 1$. Since the ancillary systems do not evolve, the action of the dynamical map of the
evolution on the probe state preserves the initial classical–quantum separable structure for any \( t \geq t_0 \)

\[
\rho_{\text{AB}}^{(i)}(t) = \sum_{i=1}^{\pi} \mathbf{P}_i |i⟩⟨i|_A \otimes (λ |i⟩⟨i|_{A′} + (1 − λ) \mathbf{P}_{\text{B}}(t)) ,
\]

where \( \mathbf{P}_{\text{B}}(t) = \Lambda_0(t, t_0) \otimes I_{A′A″} \otimes \mathbf{P}_{\text{B}}(t) \) and \( \sigma_{\text{A′A″}}(t) = \Lambda_S(t, t_0) \otimes I_{A′A″}(\sigma_{\text{A′A″}}) \). Finally, since \( \mathbf{P}_{\text{B}}(t) = \sum_{i=1}^{\pi} \mathbf{P}_i |i⟩⟨i|_A \), the set of \( \mathcal{P} \)-POVMs \( \Pi_{\text{A}}^{\pi}(\rho_{\text{AB}}^{(i)}(t)) = \Pi_{\text{B}}^{\pi}(\mathbf{P}_{\text{B}}(t)) \) does not depend on \( t \) and \( λ \).

5.2. Witnessing non-Markovianity with correlations

We provide a procedure that witnesses any non-Markovian dynamics with a correlation backflow. In the case of bijective or pointwise non-bijective \( \{\Lambda_{S}(t, t_0)\}_t \), this scenario has been studied in references [17, 18]. Moreover, the negativity entanglement measure witnesses any non-Markovian qubit evolution [17].

In order to witness non-Markovianity through backflows of \( C_{\Lambda}^{\pi} \), the evolution of the initial state \( \rho_{\text{ASA′}} = \sum_{i=1}^{\pi} \mathbf{P}_i |i⟩⟨i|_A \otimes \mathbf{P}_{\text{A′A″}} |i⟩⟨i|_{A′A″} \) is an intuitive choice. Indeed, \( |i⟩⟨i|_A \otimes |i⟩⟨i|_{A′A″} ∈ \Pi_{\text{A}}^{\pi}(\rho_{\text{ASA′}}(t)) \) for all \( t \geq t_0 \) and \( \mathbf{P}_{\text{B}}(\Lambda_{S}(t, t_0)) = \sum_{i=1}^{\pi} \mathbf{P}_i |i⟩⟨i|_A \), the set of \( \mathcal{P} \)-POVMs \( \Pi_{\text{A}}^{\pi}(\rho_{\text{AB}}^{(i)}(t)) = \Pi_{\text{B}}^{\pi}(\mathbf{P}_{\text{B}}(t)) \) (see equation (6))..

Nonetheless, in general \( |i⟩⟨i|_A \) is not selected by the maximization that defines \( C_{\Lambda}^{\pi}(\rho_{\text{ASA′}}(t)) \) [16].

We are now able to present the main result of this work.

**Theorem 1.** For any evolution \( \{\Lambda_{S}(t, t_0)\}_t \) defined on a finite-dimensional system \( S \) and time interval \( [τ, τ + Δτ] \) there exist at least one ancillary system \( \mathcal{H} \), one bipartite system \( \mathcal{H}_A \otimes \mathcal{H}_B \), where \( \mathcal{H}_B = \mathcal{H}_S \otimes \mathcal{H}_{A′} \), a correlation measure \( \mathcal{C}_{\text{AB}} = \mathcal{C}_{\Lambda}^{\pi} \), and an initial state \( \rho_{\text{AB}}(t_0) = \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B) \) such that a backflow

\[
\mathcal{C}_{\text{AB}}(\rho_{\text{AB}}(τ + Δτ)) − \mathcal{C}_{\text{AB}}(\rho_{\text{AB}}(τ)) > 0,
\]

occurs if and only if there is no CPTP intermediate map \( \mathcal{V}_{\mathcal{S}}(τ + Δτ, τ) \), where \( S \) is the only system that evolves during the evolution.

**Proof.** We consider the ancillary system \( \mathcal{H} = \mathcal{H}_{A′} \otimes \mathcal{H}_{A″} \), the correlation measure \( \mathcal{C}_{\text{AB}} = \mathcal{C}_{\Lambda}^{\pi} \) and the initial probe states \( \rho_{\text{AB}}^{(i)}(t_0) \). We prove that, for wisely chosen values of \( λ \),

\[
\Delta \mathcal{C}_{\Lambda}^{\pi} \equiv \mathcal{C}_{\Lambda}^{\pi}(\rho_{\text{AB}}(τ + Δτ)) − \mathcal{C}_{\Lambda}^{\pi}(\rho_{\text{AB}}(τ)) > 0,
\]

if and only if there is no CPTP intermediate map \( \mathcal{V}_{\mathcal{S}}(τ + Δτ, τ) \). As stated before, in appendix A we prove that \( \mathcal{C}_{\Lambda}^{\pi} \) is monotonically decreasing under local operations. It follows that \( \mathcal{C}_{\Lambda}^{\pi}(\rho_{\text{AB}}(τ)) \) cannot increase in case of CPTP intermediate maps \( \mathcal{V}_{\mathcal{S}}(τ + Δτ, τ) \) acting on \( \rho_{\text{AB}}^{(i)}(τ) \). Moreover, any correlation measure \( \mathcal{C}_{\text{AB}} \) is by definition monotonically decreasing under local operations. Therefore, in order to prove theorem 1, we follow by studying the cases where there is no CPTP intermediate map \( \mathcal{V}_{\mathcal{S}}(τ + Δτ, τ) \).

First, \( \Pi_{\text{A}}^{\pi} = \Pi_{\text{A}}^{\pi}(\rho_{\text{AB}}^{(i)}(t)) \) does not depend on \( λ \) or \( t \), and \( |i⟩⟨i|_A \) \( \otimes |i⟩⟨i|_{A′A″} \in \Pi_{\text{A}}^{\pi} \). In the following, if not specified otherwise, the index \( i \) runs from 1 to \( \pi \). Moreover, we omit the dependence on \( τ \) of some quantities to increase readability.

By applying \( |i⟩⟨i|_A \), on \( \rho_{\text{AB}}^{(i)}(t) \), we get \( \mathcal{E}(\rho_{\text{AB}}^{(i)}(t), |i⟩⟨i|_A) = \{\mathbf{P}_i, λ \sigma_{\text{A′A″}}(t) \otimes |i⟩⟨i|_{A′A″} + (1 − λ) \mathbf{P}_{\text{B}}(t)\} \), and (see appendix B)

\[
P_{\mathbf{E}}(\rho_{\text{AB}}^{(i)}(t), |i⟩⟨i|_A) = λ + (1 − λ) P_{\mathbf{E}}(\mathbf{P}_{\text{B}}(t)) .
\]
For a \( \{P_{\lambda,i}\}_i \) different from \( \{|i\rangle\langle i_A|\} \), we obtain (see appendix B):
\[
\mathcal{E}(\rho^{(\lambda)}_{AB}(t), \{P_{\lambda,i}\}_i) = \left\{ P_{\lambda,i} \lambda \sigma^{(\lambda)}_{AB}(t) + (1 - \lambda) \sigma^{(\lambda)}_{i,i} \right\}_i. 
\]
where \( \sigma^{(\lambda)}_{AB} \) is a convex combination of \( \{|i\rangle\langle i_A|\} \). Analogously, \( \sigma^{(\lambda)}_{ij} \) is a convex combination of \( \{|i\rangle\langle i_A|\} \) (see appendix B). The corresponding guessing probability is
\[
P_g(\rho^{(\lambda)}_{AB}(t), \{P_{\lambda,i}\}_i) = \lambda P_g \left( \left\{ P_{\lambda,i} \rho^{(\lambda)}_{i,i} \right\}_i \right) + (1 - \lambda) P_g \left( \left\{ P_{\lambda,i} \rho^{(\lambda)}_{ij} \right\}_i \right). \tag{14}
\]

Next consider the following lower bound for equation (12)
\[
\Delta C^{\tau}_{\lambda} \geq P_g(\rho^{(\lambda)}_{AB}(\tau + \Delta \tau), \{i\rangle\langle i_A|\}_i) - P_g(\rho^{(\lambda)}_{AB}(\tau), \{P_{\lambda,i}\}_i), \tag{15}
\]
where we define \( \{P_{\lambda,i}\}_i \) to be one of the optimal \( \mathcal{F} \)-POVMs for the maximization that defines \( C^{\tau}_{\lambda}(\rho^{(\lambda)}_{AB}(t)) \) at \( t = \tau \), namely \( C^{\tau}_{\lambda}(\rho^{(\lambda)}_{AB}(\tau)) = P_g(\rho^{(\lambda)}_{AB}(\tau), \{P_{\lambda,i}\}_i) - \mathcal{F}_{\text{max}} \). Consider equation (B.2) for \( \{P_{\lambda,i}\}_i \) and define \( \mathcal{E}^\perp \left( \{P_{\lambda,i}\}_i \right) \) and \( \mathcal{E}^\parallel \left( \{P_{\lambda,i}\}_i \right) \), such that
\[
P_g(\rho^{(\lambda)}_{AB}(\tau), \{P_{\lambda,i}\}_i) = \lambda P_g(\mathcal{E}^\perp \left( \{P_{\lambda,i}\}_i \right)) + (1 - \lambda) P_g(\mathcal{E}^\parallel \left( \{P_{\lambda,i}\}_i \right)). \tag{16}
\]

In order to evaluate \( \Delta C^{\tau}_{\lambda} \), we distinguish the two possible scenarios for \( P_g(\rho^{(\lambda)}_{AB}(\tau), \{P_{\lambda,i}\}_i) \): (A) \( \{i\rangle\langle i_A|\}_i \) is an optimal \( \mathcal{F} \)-POVM for some \( \lambda \in [0, 1] \) and (B) \( \{i\rangle\langle i_A|\}_i \) is not an optimal \( \mathcal{F} \)-POVM for any \( \lambda \in [0, 1] \).

Case (A): if \( \{i\rangle\langle i_A|\}_i \) is an optimal \( \mathcal{F} \)-POVM for some \( \lambda^* \), then the same is true for any \( \lambda \in (\lambda^*, 1) \) (see appendix C). From equations (6), (B.1) and (15), for \( \lambda \in (\lambda^*, 1) \)
\[
\Delta C^{\tau}_{\lambda} \geq (1 - \lambda) \left( P_g(\mathcal{E}_{\text{SA}}(\tau + \Delta \tau)) - P_g(\mathcal{E}_{\text{SA}}(\tau)) \right) > 0, \tag{17}
\]
if and only if there is no CPTP intermediate map \( V_5(\tau + \Delta \tau, \tau) \).

Case (B): we start by noting that \( P_g(\rho^{(\lambda)}_{AB}(\tau), \{P_{\lambda,i}\}_i) \) is Lipschitz continuous in \( \lambda \), \( V_5(\rho^{(\lambda)}_{AB}(\tau), \{P_{\lambda,i}\}_i) \) is Lipschitz continuous in \( \{P_{\lambda,i}\}_i \) and the unique optimal \( \mathcal{F} \)-POVM for \( \lambda = 1 \) is \( \{i\rangle\langle i_A|\}_i \) (see appendix D). Therefore, the set of optimal \( \mathcal{F} \)-POVMs is contained in a neighbourhood of \( \{i\rangle\langle i_A|\}_i \), with size decreasing towards zero as \( \lambda \rightarrow 1 \). This in turn implies that the values of \( P_g(\mathcal{E}^\parallel \left( \{P_{\lambda,i}\}_i \right)) \) for different \( \{P_{\lambda,i}\}_i \) are contained in an interval that converges on \( P_g(\mathcal{E}_{\text{SA}}(\tau)) \) (see appendix D). If we define \( P_g^{(\lambda)}(\tau) \equiv \max_{\{i\rangle\langle i_A|\}_i} P_g(\mathcal{E}^\parallel \left( \{P_{\lambda,i}\}_i \right)) \) and \( P_g^{(\lambda)}(\tau) \equiv \max_{\{i\rangle\langle i_A|\}_i} P_g(\mathcal{E}^\parallel \left( \{P_{\lambda,i}\}_i \right)) \), it holds that
\[
\forall \delta > 0, \exists \lambda_\delta > 0 : P_g^{(\lambda)}(\tau) - P_g(\mathcal{E}_{\text{SA}}(\tau)) < \delta, \quad \forall \lambda \in (\lambda_\delta, 1). \tag{18}
\]
Hence, for \( \delta \equiv P_g(\mathcal{E}_{\text{SA}}(\tau + \Delta \tau)) - P_g(\mathcal{E}_{\text{SA}}(\tau)) > 0 \), there exists \( \lambda \in [0, 1] \) such that \( P_g^{(\lambda)}(\tau) - P_g(\mathcal{E}_{\text{SA}}(\tau)) < P_g(\mathcal{E}_{\text{SA}}(\tau + \Delta \tau)) - P_g(\mathcal{E}_{\text{SA}}(\tau)) \) for any \( \lambda \in (\lambda, 1) \), namely
\[
P_g(\mathcal{E}_{\text{SA}}(\tau + \Delta \tau)) - P_g^{(\lambda)}(\tau) > 0, \quad \forall \lambda \in (\lambda, 1). \tag{19}
\]
We consider inequalities (15) and (19) for \( \lambda \in (\lambda, 1) \) and obtain a backflow
\[
\Delta C^{\tau}_{\lambda} \geq \lambda (1 - P_g^{(\lambda)}(\tau)) + (1 - \lambda) \left( P_g(\mathcal{E}_{\text{SA}}(\tau + \Delta \tau)) - P_g^{(\lambda)}(\tau) \right) > 0, \tag{20}
\]
if and only if there is no CPTP intermediate map \( V_5(\tau + \Delta \tau, \tau) \).

We proved that any non-Markovian evolution can be witnessed with backflows of \( C^{\tau}_{\lambda} \) by considering the probe states \( \rho^{(\lambda)}_{AB}(\tau) \). These backflows are robust, namely, if we add sufficiently small perturbations to \( \rho^{(\lambda)}_{AB}(\tau) \) and the optimal \( \mathcal{F} \)-POVMs obtained by the maximization in equation (7) for any given non-Markovian dynamics, we still obtain a backflow of \( C^{\tau}_{\lambda} \). Hence, there exists a set of initial states with the same dimension as \( S(H_A \otimes H_B) \) that provide a backflow of \( C^{\tau}_{\lambda} \) in the scenario described above (see appendix E for more details).

Since there are no assumptions for the structure of \( \mathcal{E}_{\text{SA}}(\rho^{(\lambda)}_{AB}(t)) \) [12], it is straightforward to adapt our results to any other ensemble. In particular, if the evolution of an initial ensemble \( \{p_i, \phi^{(\lambda)}_{AB}(t)\}_{i=1}^n \) provides a backflow of \( P_g(\{p_i, \phi^{(\lambda)}_{AB}(t)\}_{i=1}^n) \) in a time interval \( [\tau, \tau + \Delta \tau] \), we can consider \( C^{\tau}_{\lambda}(\phi^{(\lambda)}_{AB}(t)) \), where \( \mathcal{P} \equiv \{p_i\}_{i=1}^n \) and \( \phi^{(\lambda)}_{AB}(t) = \sum_{i=1}^n p_i |i\rangle\langle i_A| \otimes (\lambda \phi_{SA} \otimes |i\rangle\langle i_A| + (1 - \lambda) \phi_{SB} \otimes |n + 1\rangle\langle n + 1|) \) and obtain a backflow of \( C^{\tau}_{\lambda}(\phi^{(\lambda)}_{AB}(t)) \) in \( [\tau, \tau + \Delta \tau] \). We make some examples of ensembles (different from
that can be considered to witness particular classes of non-Markovian evolutions. A constructive method that provides ensembles of two equiprobable states that witness any bijective or pointwise non-bijective non-Markovian dynamics is given in reference [13]. The existence of two-state ensembles that detect any image non-increasing evolution, i.e., such that \( \text{Im}(\Lambda_t) \subseteq \text{Im}(\Lambda_s) \) for any \( s < t \), is proven in reference [25]. Finally, two-state ensembles suffice to witness any non-Markovian qubit evolution [24].

Similarly to prior measures of non-Markovianity that catch increases of quantities that are monotonically decreasing under Markovian evolutions [10, 11, 14, 19, 20], we define the class

\[
N^P(\{A_\lambda(t, t_0)\}) := \sup_{\rho_{\Lambda A}(t_0)} \int_{(C_\Lambda^P(\rho_{\Lambda A}(t))) > 0} \frac{dC_\Lambda^P(\rho_{\Lambda A}(t))}{dt},
\]

where the sup is over the possible ancillary systems (A and A') and the initial states \( \rho_{\Lambda A}(t_0) \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_S \otimes \mathcal{H}_{A'}) \). As a consequence of theorem 1, if \( C_\Lambda^P(\rho_{\Lambda A}(t)) \) is differentiable, \( N^P(\{A_\lambda(t, t_0)\}) > 0 \) if and only if the evolution is non-Markovian (see appendix F for details and a discussion of the non-differentiable case). Indeed, for any time interval where the evolution cannot be described by a CPTP intermediate map, we proved the existence of a set of initial states that show an increase of \( C_\Lambda^P \) in the same time interval. We notice that \( N^P \) with \( P = \{1/2, 1/2\} \) is non-zero for any bijective or pointwise non-bijective non-Markovian evolution [18]. The measure of non-Markovianity \( NRHP \) proposed in [10] and the class \( N^P \) are the only measures that are proved to be positive for any non-Markovian evolution. Notice that the value of \( N^P \), differently from \( NRHP \), represents the backflow of a physical quantity, namely \( C_\Lambda^P \), shown by a state that undergoes the target evolution. Nonetheless, while \( NRHP \) is easy to compute in many different cases, the computation of the class \( N^P \) may be difficult in the general case, since it involves a supremum over initial states.

6. Discussion

In this work we showed that any non-Markovian dynamics can be witnessed through backflows of the correlation measure \( C_\Lambda^P \). For this purpose, we introduced a class of initial probe states \( \rho_{\Lambda A}(t_0) \) that allows to accomplish this task. Hence, we proved the first one-to-one correspondence between CP-divisibility of evolutions, namely Markovianity, and the absence of correlation backflows.

It would be useful to obtain a constructive method that provides the elements of \( C_\Lambda^P(\rho_{\Lambda A}(t)) \) that we used to define the initial probe state. Moreover, since the class of bipartite correlations that we studied does not consider the subsystems A and B symmetrically, an open question is to understand if also \( C_\Lambda^{AP} \) (see equation (9)) is able to witness any non-Markovian evolution.

Different approaches that manipulate and evolve two-state ensembles defined over S and particular ancillary systems are proved to witness any bijective or alternatively at most point-wise non-bijective non-Markovian evolution [13, 17, 18]. On the other hand, methods that allow to witness any non-Markovian evolution, e.g. [12] and this work, make use of ensembles that in general are made by more than two states. We therefore find it interesting to know if the use of larger ensembles in [12] and in this work is necessary to witness any non-Markovian evolution. Finding an example of a non-Markovian evolution that two-state ensembles cannot witness would prove this in the positive. Such an example, if it exists, could perhaps also help elucidate how to explicitly construct the elements of \( C_\Lambda(\rho_{\Lambda A}(t)) \).

We consider interesting the possibility to formulate simplified versions of the measures of non-Markovianity \( N^P \) that permit simplified computations while still being positive for any non-Markovian evolution.

Acknowledgments

We would like to thank A Acín, B Bylicka and M Lostaglio for insightful discussions and comments on a previous draft. Support from the Government of Spain (FIS2020-TRANQI and Severo Ochoa CEX2019-000910-S), the Fundació Privada Cellex and the Generalitat de Catalunya (CERCA Program and SGR1381) is acknowledged. DDS acknowledges support from the ICFOstepstone programme, funded by the Marie Skłodowska-Curie COFUND action (GA665884).

Appendix A. Monotonic behavior of \( C_\Lambda^P \) under local operations

We consider a general bipartite finite-dimensional quantum system with Hilbert space \( \mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B \). Therefore, the states that we consider are \( \rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB}) \). We consider a generic finite probability
distribution $\mathcal{P} = \{ p_i \}_{i=1}^n$ and we prove that $C^q_\Lambda$ is monotone under local operations of the form $\Lambda_A \otimes I_B$, and $I_A \otimes \Lambda_B$ on $\rho_{AB}$, where $\Lambda_A$ ($\Lambda_B$) is a CPTP map on $A$ ($B$) and $I_A$ ($I_B$) is the identity map on $S(H_A)$ ($S(H_B)$).

In order to show the effect of the application of a local operation of the form $\Lambda_A \otimes I_B$ on $C^q_\Lambda(\rho_{AB})$, we look at $\Pi^q_\Lambda(\rho_{AB})$ in a different way. Each element of this collection is a $\mathcal{P}$-POVM for $\rho_{AB}$, i.e., they generate output ensembles where the output probability distribution is $\mathcal{P} = \{ p_i \}$. In fact, we can consider $C^q_\Lambda(\rho_{AB})$ as the maximization over all the possible output ensembles with output probability distribution $\mathcal{P}$ that we can generate measuring the subsystem $A$ of $\rho_{AB}$.

The effect of the first local operation that we consider is: $\tilde{\rho}_{AB} = \Lambda_A \otimes I_B (\rho_{AB}) = \sum_i (E_i^A \otimes I_B) \rho_{AB} (E_i^A \otimes I_B)$, where $\{ E_i \}$ is a set of Kraus operators that corresponds to $\Lambda_A$. Now we analyze the relation between $\Pi^q_\Lambda(\tilde{\rho}_{AB})$ and $\Pi^q_\Lambda(\rho_{AB})$. Given a $\mathcal{P}$-POVM for $\tilde{\rho}_{AB}$, i.e., $\{ P_{A,i} \} \subseteq \Pi^q_\Lambda(\tilde{\rho}_{AB})$, the probabilities and the states of the output ensemble $\mathcal{E} (\tilde{\rho}_{AB}, \{ P_{A,i} \})$ are

$$\text{Tr} [ \rho_{AB} \Lambda^*_A (P_{A,i} \otimes I_B) ] = \text{Tr} [ \rho_{AB} \tilde{\Lambda}^*_A (\tilde{P}_{A,i} \otimes I_B) ] = \text{Tr} [ \rho_{AB} \tilde{P}_{A,i} \otimes I_B ] ,$$

where we have defined the operators $\tilde{P}_{A,i} \equiv \Lambda^*_A (P_{A,i}) = \sum_j (E_j^A \otimes I_B) P_{A,j} (E_j^A \otimes I_B)$. Similarly, we can write $\tilde{\rho}_{iB} = \text{Tr}_A [ \tilde{\rho}_{AB} P_{A,i} ] / p_i = \text{Tr} [ \rho_{AB} \tilde{P}_{A,i} ] / p_i$.

Therefore, since $p_i = \text{Tr} [ \rho_{AB} \tilde{P}_{A,i} ]$ and $\tilde{\rho}_{iB} = \text{Tr} [ \rho_{AB} \tilde{P}_{A,i} ] / p_i$, if we apply $\{ P_{A,i} \}$ on $\rho_{AB}$ we obtain the same $\mathcal{P}$-distributed output ensemble $\{ p_i, \tilde{\rho}_{iB} \}$, that we obtain applying $\{ P_{A,i} \}$ on $\tilde{\rho}_{AB}$. Next we show that: $\{ \tilde{P}_{A,i} \} = \{ \Lambda^*_A (P_{A,i}) \} = \{ \sum_k (E_k^A \otimes I_B) \tilde{P}_{A,i} (E_k^A \otimes I_B) \}$, is a proper $n$-output POVM. First, the elements of $\{ \tilde{P}_{A,i} \}$ sum up to the identity:

$$\sum_k \tilde{P}_{A,i} = \sum_k (E_k^A \otimes I_B) \tilde{P}_{A,i} (E_k^A \otimes I_B) = \sum_k (E_k^A \otimes I_B) \tilde{P}_{A,i} = \sum_k (E_k^A \otimes I_B) = I_B. \quad \text{Moreover, we show that they are positive semi-definite operators. Indeed, for any } | \psi \rangle \in \mathcal{H}_A, \text{ we have } \langle \psi | \tilde{P}_{A,i} | \psi \rangle = \sum_k ( \langle \psi | E_k^A \rangle \tilde{P}_{A,i} (E_k^A | \psi \rangle) = \sum_k ( \langle \psi | E_k^A \rangle \langle E_k^A | \psi \rangle) \geq 0, \text{ where each element of the last sum is non-negative because } P_{A,i} \text{ is positive semi-definite.} \quad \text{It follows that } \{ \tilde{P}_{A,i} \} \text{ is a POVM and in particular a } \mathcal{P}\text{-POVM for } \tilde{\rho}_{AB} \text{, i.e., } \{ \tilde{P}_{A,i} \} \subseteq \Pi^q_\Lambda(\tilde{\rho}_{AB}).$$

Thus, for every $\mathcal{P}$-POVM $\{ P_{A,i} \} \subseteq \Pi^q_\Lambda(\rho_{AB})$ for $\rho_{AB}$, there is a $\mathcal{P}$-POVM $\{ \tilde{P}_{A,i} \} \subseteq \Pi^q_\Lambda(\tilde{\rho}_{AB})$ for $\tilde{\rho}_{AB}$, such that the output ensembles are identical: $\mathcal{E} (\rho_{AB}, \{ P_{A,i} \}) = \mathcal{E} (\tilde{\rho}_{AB}, \{ \tilde{P}_{A,i} \})$. Hence, any $\mathcal{P}$-distributed ensemble of $B$ that can be generated from $\rho_{AB}$ can also be obtained from $\tilde{\rho}_{AB}$. Therefore, we obtain the following inclusion

$$\bigcup_{\{ P_{A,i} \} \subseteq \Pi^q_\Lambda(\rho_{AB})} \mathcal{E} (\rho_{AB}, \{ P_{A,i} \}) \subseteq \bigcup_{\{ \tilde{P}_{A,i} \} \subseteq \Pi^q_\Lambda(\tilde{\rho}_{AB})} \mathcal{E} (\tilde{\rho}_{AB}, \{ \tilde{P}_{A,i} \}) .$$

Finally, since as we said above $C^q_\Lambda(\rho_{AB})$ is the maximum guessing probability of the $\mathcal{P}$-distributed output ensembles that can be generated from $\rho_{AB}$, from equation (A.2) we conclude that $C^q_\Lambda(\rho_{AB})$ is defined as a maximization over a set that includes the set over which maximization defines $C^q_\Lambda(\tilde{\rho}_{AB})$. Hence, for any state $\rho_{AB}$ and CPTP map $\Lambda_A$, we obtain

$$C^q_\Lambda(\rho_{AB}) \geq C^q_\Lambda (\Lambda_A \otimes I_B (\rho_{AB})) .$$

Next we show that $C^q_\Lambda(\rho_{AB})$ is monotonic under local operations of the form $I_A \otimes \Lambda_B$. We find that the collection of the $\mathcal{P}$-POVMs for $\tilde{\rho}_{AB} = I_A \otimes \Lambda_B (\rho_{AB})$, namely $\Pi^q_{\Lambda_B}(\rho_{AB})$, coincides with $\Pi^q_{\Lambda_B}(\rho_{AB})$. In order to prove this, we apply a general POVM $\{ P_{A,i} \}$ on $\rho_{AB}$ and $\tilde{\rho}_{AB}$ and we show that the respective output ensembles are defined by the same probability distribution. Indeed, being $\text{Tr} [ \rho_{AB} P_{A,i} ]$ the probability for the $i$-th output of the POVM considered when it is applied on $\rho_{AB}$ ($\tilde{\rho}_{AB}$), we have $\text{Tr} [ I_A \otimes \Lambda_B (\rho_{AB}) P_{A,i} ] = \text{Tr} [ \rho_{AB} P_{A,i} ]$, where this identity uses the trace-preserving property of the superoperator $I_A \otimes \Lambda_B$. Consequently, if $\{ P_{A,i} \}$ is a $\mathcal{P}$-POVM for $\rho_{AB}$, which means that $\text{Tr} [ \rho_{AB} P_{A,i} ] = p_i$ in the same way $\text{Tr} [ I_A \otimes \Lambda_B (\rho_{AB}) P_{A,i} ] = p_i$. Hence, $\{ P_{A,i} \} \subseteq \Pi^q_\Lambda(\rho_{AB})$ if and only if $\{ P_{A,i} \} \subseteq \Pi^q_\Lambda(\tilde{\rho}_{AB})$, i.e.,

$$\Pi^q_{\Lambda_B}(\rho_{AB}) = \Pi^q_{\Lambda_B}(\tilde{\rho}_{AB}).$$

Given a $\mathcal{P}$-POVM $\{ P_{A,i} \}$, both for $\rho_{AB}$ and $\tilde{\rho}_{AB}$, we compare the corresponding output states

$$\tilde{\rho}_{iB} = \Lambda_B (\text{Tr}_A [ \rho_{AB} P_{A,i} \otimes I_B ] / p_i) = \Lambda_B (\rho_{iB}).$$

(5)
From equation (A.5) and the definition of the guessing probability, it follows that

\[ P_g \left( \{p_i, \rho_{B_i}\}_i \right) \geq P_g \left( \{p_i, \Lambda_B(\rho_{B_i})\}_i \right). \tag{A.6} \]

The consequence of the last relation is that for any \( P \)-distributed output ensemble ensemble that we can generate from \( \hat{\rho}_{AB} \) there exists at least one \( P \)-distributed output ensemble that we can generate from \( \rho_{AB} \) for which the guessing probability is equal or greater. Hence, considering the definition of \( C^P_A \), equations (A.4) and (A.6), we conclude that

\[ C^P_A (\rho_{AB}) \geq C^P_A (I_A \otimes \Lambda_B (\rho_{AB})), \tag{A.7} \]

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

**Appendix B. Performing \( \mathcal{P} \)-POVMs on the probe state: the orthogonal and the parallel components**

In this section we prove that, if we apply the projective \( \mathcal{P} \)-POVM \( \{|i\rangle \langle i|\}_i \) on \( A \) for \( \rho_{AB}(t) \), we obtain

\[ P_g \left( \rho_{AB}(t) \otimes \{P_{\Lambda_i}\}_i \right) = \lambda + (1 - \lambda) P_g \left( \rho_{SA}(t) \right). \tag{B.1} \]

Moreover, for a general \( \mathcal{P} \)-POVM on \( A \) for \( \rho_{AB}(t) \) different from \( \{|i\rangle \langle i|\}_i \), we have

\[ P_g \left( \rho_{AB}(t), \{P_{\Lambda_i}\}_i \right) = \lambda P_g \left( \rho_{SA}(t) \right) + (1 - \lambda) P_g \left( \hat{\rho}_{SA}(t) \right), \tag{B.2} \]

for some \( \{|\hat{\rho}_{SA_i}(t)\}_i \) and \( \{|\rho_{SA_i}(t)\}_i \) that we define. First, we notice that the projective measurement \( \{|i\rangle \langle i|\}_i \) is a \( \mathcal{P} \)-POVM on \( A \) for \( \rho_{AB}(t) \) for any \( t \) and \( \lambda \). We consider \( E(\rho_{AB}(t), \{|i\rangle \langle i|\}_i) \), namely the ensemble of \( B \) that we obtain measuring \( \rho_{AB}(t) \) with \( \{|i\rangle \langle i|\}_i \):

\[ E(\rho_{AB}(t), \{|i\rangle \langle i|\}_i) = \left\{ \hat{\rho}_{SA}, \lambda \sigma_{SA}(t) \otimes |i\rangle \langle i|_{A^e} + (1 - \lambda) \rho_{B_i}(t) \right\}_{i=1}^\pi, \tag{B.3} \]

where \( \rho_{B_i} = \sigma_{SA_i} \otimes |i\rangle \langle i|_{A^e} \). We evaluate the guessing probability of this ensemble and we obtain

\[ P_g(\rho_{AB}(t), \{|i\rangle \langle i|\}_i) = \max \sum_{\{P_{\Lambda_i}\}_i} P_{\text{Tr}_B} \left[ (\lambda \sigma_{SA}(t) \otimes |i\rangle \langle i|_{A^e} + (1 - \lambda) \rho_{B_i}(t) \otimes |i\rangle \langle i|_{A^e}) P_{\text{Tr}_B} \right]. \tag{B.4} \]

We note that, for any \( i = 1, \ldots, \pi \), every state that belongs to the set \( \{|i\rangle \langle i|\}_i \) is orthogonal to every state of the set \( \{|i\rangle \langle i|\}_i \otimes |i\rangle \langle i|_{A^e} \} \), if it follows that, for any \( i = 1, \ldots, \pi \), the value of \( P_{\text{Tr}_B} \left[ (\lambda \sigma_{SA}(t) \otimes |i\rangle \langle i|_{A^e}) P_{\text{Tr}_B} \right] \) depends only on the components of \( P_{\text{Tr}_B} \) that belong to span\( (\{|j\rangle \langle j|\}_B) \), where \( |j\rangle_B \) and \( |j\rangle_B \) belong to the tensor product between the elements of \( M_{A^e} \), i.e., an orthonormal basis of \( H_A \otimes H_{A^e} \), and \( \{|i\rangle \langle i|\}_i \) (notice that \( \dim(H_{A^e}) = \pi + 1 \)). Similarly, for any \( i = 1, \ldots, \pi \), the value of \( P_{\text{Tr}_B} \left[ (\lambda \sigma_{SA}(t) \otimes |i\rangle \langle i|_{A^e}) P_{\text{Tr}_B} \right] \) depends only on the components of \( P_{\text{Tr}_B} \) that belong to span\( (\{|f\rangle \langle f|\}_B) \), where \( |f\rangle_B \) and \( |f\rangle_B \) belong to the tensor product between the elements of \( M_{A^e} \) and \( \{|i\rangle \langle i|\}_i \) (notice that \( \dim(H_{A^e}) = \pi + 1 \)). We further note that no operator defined on span\( (\{|i\rangle \langle j|\}_B) \) span\( (\{|i\rangle \langle j|\}_B) \) span\( (\{|f\rangle \langle f|\}_B) \) span\( (\{|f\rangle \langle f|\}_B) \) without affecting the optimal value. Since span\( (\{|i\rangle \langle j|\}_B) \) is orthogonal to span\( (\{|f\rangle \langle f|\}_B) \), the maximization in equation (B.4) can be divided in two independent maximizations

\[ P_g \left( \rho_{AB}(t), \{|i\rangle \langle i|\}_i \right) = \lambda \max \sum_{\{P_{\Lambda_i}\}_i} P_{\text{Tr}_B} \left[ (\lambda \sigma_{SA}(t) \otimes |i\rangle \langle i|_{A^e}) P_{\text{Tr}_B} \right] + (1 - \lambda) \max \sum_{\{P_{\Lambda_i}\}_i} P_{\text{Tr}_B} \left[ (\lambda \sigma_{SA}(t) \otimes |i\rangle \langle i|_{A^e}) P_{\text{Tr}_B} \right]. \tag{B.5} \]
where we have used $P_k(\{\rho_i, |i⟩⟨i|_A\}_i) = 1$, namely the possibility to perfectly distinguish ensembles of orthonormal states, $P_k(\{\rho_i, |i⟩⟨i|_A\}_i) = P_k(\{\rho_i, |i⟩⟨i|_A\}_i)$ and $P_k(\{\rho_i, |i⟩⟨i|_A\}_i) = P_k(\{\rho_i, |i⟩⟨i|_A\}_i)$ for orthonormal states

The output ensemble that we obtain applying a generic POVM $\{P_{\lambda A}\}$ on $A$ for $\rho^{(\lambda)}_{AB}(t)$ different from $\{|i⟩⟨i|_A\}_i$ is $E(\rho^{(\lambda)}_{AB}(t), \{P_{\lambda A}\})$. The $k$th state of this ensemble is

$$\rho^{(\lambda)}_{\lambda k}(t) = \frac{\text{Tr}_A \left[ \rho^{(\lambda)}_{AB}(t) P_{\lambda k} \otimes 1_B \right]}{P_k}$$

$$= \sum_{i=1}^{\pi} \frac{P_k}{P_k} \text{Tr}_A \left[ |i⟩⟨i|_A P_{\lambda A} \right] (\lambda |\sigma_{SA}(t)⟩⟨\sigma_{SA}|_A + (1 - \lambda) |\bar{\sigma}_{SA}|_A)$$

$$= \sum_{i=1}^{\pi} \frac{P_k (P_{\lambda A})_{ik}}{P_k} (\lambda |\sigma_{SA}(t)⟩⟨\sigma_{SA}|_A + (1 - \lambda) |\bar{\sigma}_{SA}|_A),$$

where $(P_{\lambda A})_{ik} = (|i⟩⟨i|_A P_{\lambda A})_{ik} \geq 0$ is the $i$th diagonal element of $P_{\lambda A}$ in the basis $M_A = \{|i⟩⟩_A \}_i$. Keeping in mind that $\rho^\pi$ is a finite probability distribution and $P_k > 0$ for any $k$, we define the parameters $e_k = \text{Pr}(\lambda |\sigma_{SA}(t)⟩⟨\sigma_{SA}|_A + (1 - \lambda) |\bar{\sigma}_{SA}|_A)$ and the states $\lambda |\sigma_{SA}(t)⟩⟨\sigma_{SA}|_A + (1 - \lambda) |\bar{\sigma}_{SA}|_A$ are trace one operators for any $i = 1, \ldots, \pi$. Therefore, $\{e_k\}_i$ is an $\pi$-element probability distribution for any value of $k = 1, \ldots, \pi$. We write:

$$\rho^{(\lambda)}_{\lambda k}(t) = \sum_{i=1}^{\pi} e_k \lambda |\sigma_{SA}(t)⟩⟨\sigma_{SA}|_A + (1 - \lambda) |\bar{\sigma}_{SA}|_A)$$

$$= \lambda \sum_{i=1}^{\pi} e_k = \lambda \sum_{i=1}^{\pi} e_k |\sigma_{SA}(t)⟩⟨\sigma_{SA}|_A + (1 - \lambda) \sum_{i=1}^{\pi} e_k |\bar{\sigma}_{SA}|_A$$

$$= \lambda |\sigma_{SA}(t)⟩⟨\sigma_{SA}|_A + (1 - \lambda) |\bar{\sigma}_{SA}|_A, (B.7)$$

where we have used the definitions

$$\sigma^{\perp}_{\lambda k}(t) = \sum_{i=1}^{\pi} e_k |i⟩⟨i|_A$$

$$\sigma^{∥}_{\lambda k}(t) = \sum_{i=1}^{\pi} e_k |\sigma_{SA}(t)⟩⟨\sigma_{SA}|_A + (1 - \lambda) |\bar{\sigma}_{SA}|_A(t)$$

Each state $\rho^{\perp}_{\lambda k}(t)$ is a convex combination of the states $\{|i⟩⟨i|_A\}_i = \{|i⟩⟨i|_A⟩\} = \{|\sigma_{SA}(t)⟩⟨\sigma_{SA}|_A⟩\}$ that does not depend on $\lambda$ but depends on the POVM $\{P_{\lambda A}\}_i$ chosen. From equation (B.7) it follows that, if we consider a generic POVM $\{P_{\lambda A}\}_i$ for $\rho^{(\lambda)}_{AB}(t)$, we obtain

$$E(\rho^{(\lambda)}_{AB}(t), \{P_{\lambda A}\}) = \left\{ |i⟩⟨i|_A⟩, (B.10) \right\}$$

and, therefore, similarly to equation (B.5), now we can write

$$P_k(\rho^{(\lambda)}_{AB}(t), \{P_{\lambda A}\}) = \Lambda P_k \left( \{\rho_{\lambda k}(t)\}_i\right) + (1 - \lambda) P_k \left( \{\rho_{\lambda k}^\perp(t)\}_i\right).$$

**Appendix C. Analysis of case (A)**

Let assume that for some $\alpha \in [0, 1]$ we have that $\{P^{(\alpha)}_{\lambda k}\} = \{|i⟩⟨i|_A\}_i$, i.e., this projective measurement is one of the optimal POVM that accomplishes the maximization for $\sigma^{\pi}(\rho^{(\alpha)}_{AB}(t))$, and that for some $\beta > \alpha$ instead we have that $\{i⟩⟨i|_A\}_i$ is not optimal. In this section we show that these two assumptions are incompatible and lead to a contradiction. The first condition implies that, when $\lambda = \alpha$ the optimal POVM that provides the greatest value of $P_k(\rho^{(\alpha)}_{AB}(t), \{P_{\lambda A}\})$ is $\{P^{(\alpha)}_{\lambda k}\} = \{|i⟩⟨i|_A\}_i$ and therefore
\[
\alpha + (1 - \alpha)P_g(\mathcal{E}_{SA}(\tau)) \geq \alpha P_g(\mathcal{E}^{(1)}(\beta)) + (1 - \alpha)P_g(\mathcal{E}^{(1)}(\beta)),
\]
\[
\alpha \left(1 - P_g(\mathcal{E}^{(1)}(\beta))\right) + (1 - \alpha) \left[ P_g(\mathcal{E}_{SA}(\tau)) - P_g(\mathcal{E}^{(1)}(\beta)) \right] \geq 0,
\]

(C.1)

where we also considered the cases where \(\{P_{\beta j}\}_j\) is optimal both for \(\lambda = \alpha\) and \(\lambda = \beta\). On the other hand, for \(\lambda = \beta > \alpha\) we have that \(|i\rangle\langle i|_A\) is not an optimal \(\overline{\mathcal{F}}\)-POVM for the maximization needed for \(C_\lambda(\rho^{(1)}_{AB}(\tau))\) and
\[
\beta P_g(\mathcal{E}^{(1)}(\beta)) + (1 - \beta)P_g(\mathcal{E}^{(1)}(\beta)) > \beta + (1 - \beta)P_g(\mathcal{E}_{SA}(\tau)),
\]
which can be written as
\[
\beta \left( P_g(\mathcal{E}^{(1)}(\beta) - 1 + P_g(\mathcal{E}_{SA}(\tau)) - P_g(\mathcal{E}^{(1)}(\beta)) \right) > P_g(\mathcal{E}_{SA}(\tau)) - P_g(\mathcal{E}^{(1)}(\beta)),
\]
and therefore, subtracting the quantity \(\alpha \left( P_g(\mathcal{E}^{(1)}(\beta) - 1 + P_g(\mathcal{E}_{SA}(\tau)) - P_g(\mathcal{E}^{(1)}(\beta)) \right)\) from each side of inequality (C.3), we obtain
\[
(\beta - \alpha) \left( P_g(\mathcal{E}^{(1)}(\beta) - 1 + P_g(\mathcal{E}_{SA}(\tau)) - P_g(\mathcal{E}^{(1)}(\beta)) \right) > \alpha \left( P_g(\mathcal{E}^{(1)}(\beta)) \right) + (1 - \alpha) \left( P_g(\mathcal{E}_{SA}(\tau)) - P_g(\mathcal{E}^{(1)}(\beta)) \right).
\]

(C.4)

If inequality (C.2) holds, then \(P_g(\mathcal{E}^{(1)}(\beta)) > P_g(\mathcal{E}_{SA}(\tau))\). Therefore, \(P_g(\mathcal{E}_{SA}(\tau)) - P_g(\mathcal{E}^{(1)}(\beta)) < 0\) and we conclude that the left-hand side of inequality (C.4) is negative. The right-hand side of the same inequality is instead non-negative for inequality (C.1). This contradiction shows that if for some value of the parameter \(\lambda\) the orthogonal measurement \(|i\rangle\langle i|_A\) maximizes \(P_g(\rho_{AB}(\tau), \{P_{\lambda j}\}_j)\), then it is also the case for any greater value of \(\lambda\). In conclusion, if one of the optimal measurement is \(|i\rangle\langle i|_A\) for \(\lambda = \alpha\), the same is true for any \(\beta \in [\alpha, 1)\).

**Appendix D. Study of the limit \(\lambda \to 1\) in case (B)**

First, we notice that the set of \(\overline{\mathcal{F}}\)-POVMs on \(A\) for \(\rho^{(1)}_{AB}(t)\) is a set that does not depend on \(\lambda\) and \(t\). Indeed, we use the notation \(\Pi_{\overline{\mathcal{F}}} = \Pi_{\overline{\mathcal{F}}}(\rho^{(1)}_{AB}(\tau))\). Now we prove that the only optimal \(\overline{\mathcal{F}}\)-POVM for \(C_\lambda(\rho^{(1)}_{AB}(\tau))\) is the projective measurement \(|i\rangle\langle i|_A\). In the case of an optimal \(\{P_{\lambda j}\}_j \in \Pi_{\overline{\mathcal{F}}}\) for \(\rho^{(1)}_{AB}(\tau)\) we can write the output ensemble (see equation (B.8))
\[
\mathcal{E} \left( \rho^{(1)}_{AB}(\tau), \{P_{\lambda j}\}_j \right) = \left\{ P_{\lambda j} \sigma_{SA}(\tau) \otimes \sum_{j} \rho_{ji} |j\rangle\langle j|_{A'} \right\}_j,
\]

(D.1)

where \(\sum_j \rho_{ji} = 1\) for any \(i = 1, \ldots, \mathcal{F}\). Since \(P_{\lambda j} \rho^{(1)}_{AB}(\tau), \{P_{\lambda j}\}_j = 1\), an optimal \(\overline{\mathcal{F}}\)-POVM different from \(|i\rangle\langle i|_A\) must provide an output ensemble \(\mathcal{E}(\rho^{(1)}_{AB}(\tau), \{P_{\lambda j}\}_j)\) of orthogonal states. Given the identity \(P_{\lambda j} (\mathcal{E}(\rho^{(1)}_{AB}(\tau), \{P_{\lambda j}\}_j)) = P_{\lambda j} \left( \sum_{|j\rangle\langle j|_{A'}} \rho_{ji} |j\rangle\langle j|_{A'} \right)\), we have to check if, for some \(\rho_{ji}\) the ensemble \(\{P_{\lambda j} \sum_{|j\rangle\langle j|_{A'}} \rho_{ji} |j\rangle\langle j|_{A'}\} \) can be an orthogonal ensemble of states different from \(|i\rangle\langle i|_{A'}\). Each state \(\rho_{ij_{A'}} = \sum_{|j\rangle\langle j|_{A'}} \rho_{ji} |j\rangle\langle j|_{A'}\) is defined as a convex combination of the states \(|i\rangle\langle i|_{A'}\). Two such states are orthogonal only if the respective convex combinations do not have any element \(|i\rangle\langle i|_{A'}\) in common. Therefore, the only way to have \(n\) orthogonal output states is if for each \(i\) the state is of the form \(\rho_{ij_{A'}} = |j\rangle\langle j|_{A'}\) for some \(j = j(i)\) exclusively assigned to \(i\). Thus, each \(P_{\lambda j}\) has only one nonzero diagonal element \((P_{\lambda j})_{ij} = |j\rangle\langle j|_{A' P_{\lambda j} = 1}\) this is only possible if \(\{P_{\lambda j}\} = \{|i\rangle\langle i|_A\}\).

We proved that \(|i\rangle\langle i|_A\) is the only optimal \(\overline{\mathcal{F}}\)-POVM for the evaluation of \(C_\lambda(\rho^{(1)}_{AB}(\tau))\). Therefore, for any \(\overline{\mathcal{F}}\)-POVM \(\{P_{\lambda j}\}_j \neq |i\rangle\langle i|_A\) we have that \(P_{\lambda j}(\rho^{(1)}_{AB}(\tau), \{P_{\lambda j}\}_j) < 1\). We notice that the set \(\Pi_{\overline{\mathcal{F}}} \) is closed and bounded, i.e., it is compact. Indeed, it is a subset of \(B(H_A)\) that is defined through linear constraints involving identities and relations of semi-positivity. The guessing probability \(P_g(\rho^{(1)}_{AB}(\tau), \{P_{\lambda j}\}_j)\) is a continuous function on this compact set of \(\overline{\mathcal{F}}\)-POVMs.

We now show that \(P_g(\rho^{(1)}_{AB}(\tau), \{P_{\lambda j}\}_j)\) is Lipschitz continuous in \(\lambda\). In other words we construct a bound on the change of the guessing probability for a given change in \(\lambda\). To do so we first show that \(P_g(\rho_{AB}, \{P_{\lambda j}\}_j)\) is Lipschitz continuous on the set of states. Consider \(P_g(\rho_{AB}, \{P_{\lambda j}\}_j)\) as a function of \(\rho_{AB}\). We consider a pair \(\rho_{AB}^\lambda, \rho_{AB}^\beta\) and observe that
\[
\text{max}_{\{\rho_{ik}\}} \sum_i \text{Tr}[P_{Aj_i} \otimes P_{Bk_i}(\rho_{AB}^2 + (\rho_{AB}^2 - \rho_{AB}^2))] \\
\leq \frac{1}{\text{max}} \sum_i \text{Tr}[P_{Aj_i} \otimes P_{Bk_i}(\rho_{AB}^2 - \rho_{AB}^2)] \\
\leq \frac{1}{\text{max}} \sum_i \text{Tr}[P_{Aj_i} \otimes P_{Bk_i}(U^\dagger U \Delta_+ - \Delta_-)] U]
\]

Let $\Delta$ be a diagonal matrix such that $\Delta = U(\rho_{AB}^2 - \rho_{AB}^2)U^\dagger$ for a unitary $U$. Let $\Delta_+$ and $\Delta_-$ be the two diagonal positive semidefinite matrices such that $\Delta = \Delta_+ - \Delta_-$. Note that $U^\dagger \Delta_+ U$ and $U^\dagger \Delta_- U$ are positive semidefinite. This implies

\[
\text{max}_{\{\rho_{ik}\}} \sum_i \text{Tr}[P_{Aj_i} \otimes P_{Bk_i}(\rho_{AB}^2 - \rho_{AB}^2)] \\
= \text{max}_{\{\rho_{ik}\}} \sum_i \text{Tr}[P_{Aj_i} \otimes P_{Bk_i}(U^\dagger U \Delta_+ - \Delta_-)] U]
\]

Since POVM elements are positive semidefinite $\text{Tr}[P_{Aj_i} \otimes P_{Bk_i}(U^\dagger U \Delta_+ - \Delta_-)]$ is positive for each pair $P_{Aj_i}, P_{Bk_i}$. Therefore $\text{Tr}[\sum_i P_{Aj_i} \otimes P_{Bk_i}(U^\dagger U \Delta_+ - \Delta_-)] U] = \text{Tr}[\sum_i P_{Aj_i} \otimes P_{Bk_i}(U^\dagger U \Delta_+ - \Delta_-)] = \text{Tr}[\Delta_+ - \Delta_-]$. Likewise $\text{Tr}[\sum_i P_{Aj_i} \otimes P_{Bk_i}(U^\dagger U \Delta_+ - \Delta_-)] = \text{Tr}[\Delta_-].$ Thus, \n
\[
\text{max}_{\{\rho_{ik}\}} \sum_i \text{Tr}[P_{Aj_i} \otimes P_{Bk_i}(\rho_{AB}^2 - \rho_{AB}^2)] \leq \text{Tr}[\Delta_+ + \Delta_-] = ||\rho_{AB}^2 - \rho_{AB}^2||^2.
\]

Considering equations (D.2) and (D.4) we can now conclude that

\[
P_{Bk}(\rho_{AB}^2, \{P_{Aj_i}\}_i) - P_{Bk}(\rho_{AB}^2, \{P_{Aj_i}\}_i) \leq ||\rho_{AB}^2 - \rho_{AB}^2||^2.
\]

By exchanging the 1 and 2 in the above derivation we obtain

\[
P_{Bk}(\rho_{AB}^2, \{P_{Aj_i}\}_i) - P_{Bk}(\rho_{AB}^2, \{P_{Aj_i}\}_i) \leq ||\rho_{AB}^2 - \rho_{AB}^2||^2.
\]

Thus

\[
P_{Bk}(\rho_{AB}^2, \{P_{Aj_i}\}_i) - P_{Bk}(\rho_{AB}^2, \{P_{Aj_i}\}_i) \leq ||\rho_{AB}^2 - \rho_{AB}^2||^2.
\]

Note that this bound is independent of \{\{P_{Aj_i}\}_i\}. Thus we see that $P_{Bk}(\rho_{AB}^2, \{P_{Aj_i}\}_i)$ is Lipschitz continuous on the set of states. Next we consider the pair $\rho_{AB}^2(\tau), \rho_{AB}^2(\tau)$ and note that the trace norm \n
\[
||\rho_{AB}^2(\tau) - \rho_{AB}^2(\tau)|| = 2|\lambda_1 - \lambda_2|.
\]

Therefore,

\[
P_{Bk}(\rho_{AB}^2(\tau), \{P_{Aj_i}\}_i) - P_{Bk}(\rho_{AB}^2(\tau), \{P_{Aj_i}\}_i) \leq 2|\lambda_1 - \lambda_2|.
\]

Thus we see that $P_{Bk}(\rho_{AB}^2(\tau), \{P_{Aj_i}\}_i)$ is Lipschitz continuous in $\lambda$.}

We next consider how the set of optimal $\mathcal{F}$-POVMs converges to $\{\{\langle i | \langle i | \rangle \}_A\}$, as $\lambda \rightarrow 1$ using the bound in equation (D.8). Consider a semi-open neighbourhood $O_1$ of the projective $\mathcal{F}$-POVM $\{\{\langle i | \langle i | \rangle \}_A\}$ such that the set $S_1 \equiv \Pi_{\lambda} - O_1$ of $\mathcal{F}$-POVMs not in $O_1$ is closed. Since the set $S_1$ is closed and bounded and $P_{Bk}(\rho_{AB}^2(\tau), \{P_{Aj_i}\}_i)$ is a continuous function on $\Pi_{\lambda}$ there exists a maximum value $m_1 < 1$ of $P_{Bk}(\rho_{AB}^2(\tau), \{P_{Aj_i}\}_i)$ on $S_1$, i.e., $m_1 \equiv \max_{\{\rho_{AB}^2(\tau), \{P_{Aj_i}\}_i\} \in S_1} P_{Bk}(\rho_{AB}^2(\tau), \{P_{Aj_i}\}_i) < 1$. Then, due to equation (D.8), for $\epsilon > 0$ and $\lambda = 1 - \epsilon$ it holds that $P_{Bk}(\rho_{AB}^2(\tau), \{P_{Aj_i}\}_i) \leq m_1 + 2\epsilon$ on $S_1$ and the maximum value of $P_{Bk}(\rho_{AB}^2(\epsilon(1-\epsilon), \{P_{Aj_i}\}_i)$ on $O_1$ is larger or equal to $1 - 2\epsilon$. There exists a sufficiently small $\epsilon_1 > 0$ such that $1 - 2\epsilon_1 = m_1 + 2\epsilon_1$. For all $\epsilon < \epsilon_1$ the set of optimal $\mathcal{F}$-POVMs belongs to $O_1$.

We next consider a sequence of semi-open sets $O_i$ which all contain $\{\langle i | \langle i | \rangle \}_A\}$ and are such that $O_{i+1} \subset O_i$. There is a corresponding sequence of closed sets $S_i \equiv \Pi_{\lambda} - O_1$ and non-decreasing sequence of maximal values $m_i < 1$ of $P_{Bk}(\rho_{AB}^2(\tau), \{P_{Aj_i}\}_i)$ on $S_i$. For each $m_i$ there is an $\epsilon_i$ such that for all $\epsilon < \epsilon_i$ the optimal $\mathcal{F}$-POVMs, namely the POVMs that maximize $P_{Bk}(\rho_{AB}^2(\tau), \{P_{Aj_i}\}_i)$, belong to $O_i$. The sequence $\epsilon_i$ is non-increasing since the sequence of $m_i$ is non-decreasing.

Let us consider a distance measure $d(\cdot, \cdot)$ on $B(H_A)$ and define a sequence $O(\delta_i)$ of semi-open sets as the $\mathcal{F}$-POVMs $\{P_{Aj_i}\}_i$ such that $d(\rho_{Aj_i}, |i\rangle \langle i|_A) < \delta_i$ for any $i = 1, \ldots, \pi$, for a strictly decreasing sequence $\delta_{i+1} < \delta_i$ where $\delta_i \rightarrow 0$ as $i \rightarrow \infty$.

Then from the above argument we can conclude that, for any $\delta > 0$ there exists a value $\lambda_3 \in (0, 1)$ such that, if $\lambda \in (\lambda_3, 1)$, any optimal $\mathcal{F}$-POVM $\{\rho_{AB}^2(\lambda_i), \langle i | \langle i | \rangle \}_A\}$ for this $\lambda$ is such that $d(\rho_{Aj_i}, |i\rangle \langle i|_A) < \delta$ for any $i = 1, \ldots, \pi$. 

12
Next we show that $P_g(\rho_{AB}, \{P_{A,i}\}_i)$ is Lipschitz continuous as a function of $\{P_{A,i}\}_i$. In other words, we construct a bound on the change of the guessing probability proportional to a distance measure quantifying the change of the POVM $\{P_{A,i}\}_i$, valid for any $\rho_{AB} \in S(\mathcal{H}_A \otimes \mathcal{H}_B)$. We select a pair $\{P_{A,i}^1\}_i$, $\{P_{A,i}^2\}_i$ and observe that

$$\max_{\{P_{A,i}\}_i} \sum_i \text{Tr}[P_{A,i}^1 \otimes P_{B,i}\rho_{AB}] = \max_{\{P_{A,i}\}_i} \sum_i \text{Tr}[P_{A,i}^2 \otimes P_{B,i}\rho_{AB} + (P_{A,i}^1 - P_{A,i}^2) \otimes P_{B,i}\rho_{AB}]$$

$$\leq \max_{\{P_{A,i}\}_i} \sum_i \text{Tr}[P_{A,i}^2 \otimes P_{B,i}\rho_{AB}] + \max_{\{P_{A,i}\}_i} \sum_i \text{Tr}[(P_{A,i}^1 - P_{A,i}^2) \otimes P_{B,i}\rho_{AB}].$$

(D.9)

Let $\Delta_i$ be a diagonal matrix such that $\Delta_i = U_i(\rho_{AB} - P_{A,i})U_i^\dagger$ for a unitary $U_i$. Let $\Delta_{i+}$ and $\Delta_{i-}$ be the two diagonal positive semidefinite matrices such that $\Delta_i = \Delta_{i+} - \Delta_{i-}$. Note that $U_i^\dagger \Delta_{i+} U_i$ and $U_i^\dagger \Delta_{i-} U_i$ are positive semidefinite. This implies

$$\max_{\{P_{A,i}\}_i} \sum_i \text{Tr}[(P_{A,i}^1 - P_{A,i}^2) \otimes P_{B,i}\rho_{AB}] \leq \sum_i \text{Tr}[(U_i^\dagger \Delta_{i+} - \Delta_{i-}) U_i \otimes P_{B,i}\rho_{AB}]$$

$$\leq \sum_i \text{Tr}[(U_i^\dagger \Delta_{i+} - \Delta_{i-} U_i \otimes 1_{B}\rho_{AB}]$$

$$= \max_{\{P_{A,i}\}_i} \sum_i \text{Tr}[(P_{A,i}^1 - P_{A,i}^2) \otimes P_{B,i}\rho_{AB}] = (\overline{\tau} + 1)d^2 \max_{\{P_{A,i}\}_i} \sum_i \left\| P_{A,i}^1 - P_{A,i}^2 \right\|_1 .$$

(D.10)

where we used that $\text{Tr}[1_B] = (\overline{\tau} + 1)d^2$ and for the second inequality we have used Von Neumann’s trace inequality and that the largest eigenvalue of $\rho_{AB}$ is smaller or equal to 1. By combining equations (D.9) and (D.10) we can now conclude that

$$P_g(\rho_{AB}, \{P_{A,i}^1\}_i) - P_g(\rho_{AB}, \{P_{A,i}^2\}_i) \leq (\overline{\tau} + 1)d^2 \sum_i \left\| P_{A,i}^1 - P_{A,i}^2 \right\|_1 .$$

(D.12)

By exchanging the $\{P_{A,i}\}_i$ and $\{P_{A,i}^2\}_i$ in the above derivation we obtain

$$P_g(\rho_{AB}, \{P_{A,i}^1\}_i) - P_g(\rho_{AB}, \{P_{A,i}^2\}_i) \leq (\overline{\tau} + 1)d^2 \sum_i \left\| P_{A,i}^1 - P_{A,i}^2 \right\|_1 .$$

(D.13)

Therefore

$$|P_g(\rho_{AB}, \{P_{A,i}^1\}_i) - P_g(\rho_{AB}, \{P_{A,i}^2\}_i) | \leq (\overline{\tau} + 1)d^2 \sum_i \left\| P_{A,i}^1 - P_{A,i}^2 \right\|_1 .$$

(D.14)

Thus we have shown that $P_g(\rho_{AB}, \{P_{A,i}\}_i)$ is Lipschitz continuous as a function of $\{P_{A,i}\}_i$ for any $\rho_{AB} \in S(\mathcal{H}_A \otimes \mathcal{H}_B)$. 

We now study the guessing probability of the ensemble that we obtain applying $\{P_{A,i}\}_i \in B^\times_\mathcal{A}$ on $\rho_{AB}^{(\lambda)}(t)$ given by

$$P_g(\rho_{AB}^{(\lambda)}(t), \{P_{A,i}\}_i) = \lambda P_g \left( \{P_{A,i}^{\lambda(\langle P_{A,i}\rangle_i)} \} \right) + (1 - \lambda)P_g \left( \{P_{A,i}^{\lambda(\langle P_{A,i}\rangle_i)} \} \right) .$$

(D.15)

We consider equation (D.15) when an optimal $\{P_{A,i}^{(\lambda)}\}_i$ is chosen. We define the corresponding ensembles that appear in this expression $\mathcal{E}^+(\{P^{(\lambda)}_{A,i}\}_i) \equiv \{\overline{P}_{j,i}^{\prime}, \rho_{A,j}^{(\lambda)}(t)\}_i$ and $\mathcal{E}^-(\{P^{(\lambda)}_{A,i}\}_i) \equiv \{\overline{P}_{j,i}^{\prime}, \rho_{A,j}^{(\lambda)}(t)\}_i$, so that

$$P_g \left( \rho_{AB}^{(\lambda)}(t), \{P_{A,i}^{(\lambda)}\}_i \right) = \lambda P_g \left( \mathcal{E}^+ \left( \{P_{A,i}^{(\lambda)}\}_i \right) \right) + (1 - \lambda)P_g \left( \mathcal{E}^+ \left( \{P_{A,i}^{(\lambda)}\}_i \right) \right) .$$

(D.16)
The ensembles $\mathcal{E}^i(\{P_{\lambda A}^i\})_i$ and $\mathcal{E}'(\{P_{\lambda A}^{1i}\})_i$ are functions on the set of optimal $\mathcal{P}$-POVMs $\{P_{\lambda A}^i\}_i$ for a given $\lambda$. Thus the image of the function $P_R(\mathcal{E}^i(\{P_{\lambda A}^i\}))$ over the set of optimal $\mathcal{P}$-POVMs $\{P_{\lambda A}^i\}_i$ for a given $\lambda$, denoted $\text{Im}(P_R(\mathcal{E}^i(\{P_{\lambda A}^i\})) \equiv \{P_R(\mathcal{E}^i(\{P_{\lambda A}^i\})) : \{P_{\lambda A}^i\}_i \text{ is optimal}, \}$, is a subset of the interval $[0,1]$, i.e., $\text{Im}(P_R(\mathcal{E}^i(\{P_{\lambda A}^i\}))) \subseteq [0,1]$. Likewise, the function $P_R(\mathcal{E}'(\{P_{\lambda A}^{1i}\}))$ takes values in a set $\text{Im}(P_R(\mathcal{E}'(\{P_{\lambda A}^{1i}\}))) \subseteq [0,1]$ for a given $\lambda$.

Using equation (D.14) we can now construct bounds on $\text{Im}(P_R(\mathcal{E}^i(\{P_{\lambda A}^i\})))$ and $\text{Im}(P_R(\mathcal{E}'(\{P_{\lambda A}^{1i}\})))$ for a given $\lambda$. First, based on the above argument we make the following observation: for any $\eta > 0$ there exists a value $\lambda_\eta \in (0,1)$ such that, if $\lambda \in (\lambda_\eta, 1)$, any optimal $\mathcal{P}$-POVM $\{P_{\lambda A}^i\}_i$ for this $\lambda$ is such that

$$
\|P_{\lambda A}^i - \tilde{\rho}_{\lambda A}\|_1 < \eta \text{ for any } i = 1, \ldots, n.
$$

Thus, by equation (D.14) the values in the image of $P_R(\mathcal{E}^i(\{P_{\lambda A}^i\}))$ for $\lambda \in (\lambda_\eta, 1)$ differ from $P_R(\mathcal{E}^i(\{\tilde{\rho}_{\lambda A}\})) = 1 - \eta$ by less than $n\eta$, i.e., $\|P_R(\mathcal{E}^i(\{P_{\lambda A}^i\})) - 1\| < n\eta$ for all optimal $\{P_{\lambda A}^i\}_i : \lambda \in (\lambda_\eta, 1)$. Likewise, the values in the range of $P_R(\mathcal{E}'(\{P_{\lambda A}^{1i}\}))$ for $\lambda \in (\lambda_\eta, 1)$ differ from $P_R(\mathcal{E}'(\{\tilde{\rho}_{\lambda A}\})) = 1 - \eta$ by less than $n\eta$, i.e., $\|P_R(\mathcal{E}'(\{P_{\lambda A}^{1i}\})) - 1\| < n\eta$ for all optimal $\{P_{\lambda A}^{1i}\}_i : \lambda \in (\lambda_\eta, 1)$. Using this we can state the following

$$
\forall \delta > 0, \exists \lambda_\delta > 0 : P_R(\mathcal{E}^i(\{P_{\lambda A}^i\})) - P_R(\mathcal{E}'(\{\tilde{\rho}_{\lambda A}\})) < \delta, \forall \lambda \in (\lambda_\delta, 1) .
$$

**Appendix E. Lipschitz continuity of $C^P_A$ on the set of states**

Consider a POVM $\{P_{\lambda A}\}_i$ and two states $\rho_{AB}$ and $\hat{\rho}_{AB}$. Let $p_i = \text{Tr}[P_{\lambda A}\rho_{AB}]$ and $\tilde{p}_i = \text{Tr}[P_{\lambda A}\hat{\rho}_{AB}]$. Let $\Delta$ be a diagonal matrix such that $U = \hat{\rho}_{AB} - \rho_{AB} U^U$ for a unitary $U$. Let $\Delta_+$ and $\Delta_-$ be the two diagonal positive semidefinite matrices such that $\Delta = \Delta_+ - \Delta_-$. Note that $U^U \Delta_+ U$ and $U^U \Delta_- U$ are positive semidefinite. Then

$$
\tilde{p}_i - p_i = \text{Tr}[P_{\lambda A}(\hat{\rho}_{AB} - \rho_{AB})] = \text{Tr}[P_{\lambda A}(U^U \Delta_+ U - U^U \Delta_- U)] \\
\leq \text{Tr}[P_{\lambda A}(U^U \Delta_+ U)] + \text{Tr}[P_{\lambda A}(U^U \Delta_- U)].
$$

(E.1)

Since POVM elements are positive semidefinite $\text{Tr}[P_{\lambda A}(U^U \Delta_+ U)]$ is positive for each $P_{\lambda A}$. Therefore

$$
\text{Tr}[P_{\lambda A}(U^U \Delta_+ U)] \leq \text{Tr}[\sum_{j} P_{\lambda A}(U^U \Delta_+ U)] = \text{Tr}[U^U \Delta_+ U] = \text{Tr}[\Delta_+].
$$

Likewise $\text{Tr}[P_{\lambda A}(U^U \Delta_- U)] \leq \text{Tr}[\Delta_-]$. Thus,

$$
\text{Tr}[P_{\lambda A}(U^U \Delta_+ U)] + \text{Tr}[P_{\lambda A}(U^U \Delta_- U)] \leq \text{Tr}[\Delta_+ + \Delta_-] = \|\hat{\rho}_{AB} - \rho_{AB}\|_1.
$$

(E.2)

It follows that

$$
\tilde{p}_i - p_i \leq \|\hat{\rho}_{AB} - \rho_{AB}\|_1.
$$

(E.3)

By exchanging $p_i$ and $\tilde{p}_i$ in the above derivation we obtain

$$
p_i - \tilde{p}_i \leq \|\hat{\rho}_{AB} - \rho_{AB}\|_1.
$$

(E.4)

From this we can conclude that

$$
|\tilde{p}_i - p_i| \leq \|\hat{\rho}_{AB} - \rho_{AB}\|_1.
$$

(E.5)

Assume now that $\{P_{\lambda A}\}_i$ is a $\mathcal{P}$-POVM for $\rho_{AB}$ but not necessarily for $\hat{\rho}_{AB}$. We can create a $\mathcal{P}$-POVM for $\hat{\rho}_{AB}$ from $\{P_{\lambda A}\}_i$ in the following way. If $\tilde{p}_i - p_i > 0$ we subtract $(1 - p_i/j\tilde{p}_i)P_{\lambda A}$ from $P_{\lambda A}$ to create a new element $P_{\lambda A} \equiv p_i/j\tilde{p}_iP_{\lambda A}$. Let $P_i \equiv \sum_{i \in \{\pm\}} (1 - p_i/j\tilde{p}_i)P_{\lambda A}$ where the $\{\pm\}$ is the set of all $i$ such that $\tilde{p}_i - p_i > 0$ and let $p_i \equiv \text{Tr}[P_i]\hat{\rho}_{AB} = \sum_{i \in \{\pm\}} \tilde{p}_i - p_i$. If $\tilde{p}_i - p_i < 0$ we add $(p_i/j\tilde{p}_i)P_{\lambda A}$ to $P_{\lambda A}$ to create a new element $P_{\lambda A} \equiv P_{\lambda A} + (p_i/j\tilde{p}_i)/(p_i)P_{\lambda A}$.

Next consider the trace distance between $\{P_{\lambda A}\}_i$ and $\{P_{\lambda A}\}_i$.

$$
\sum_{i} \|P_{\lambda A} \neq P_{\lambda A}\|_1 = \sum_{i \in \{\pm\}} |\tilde{p}_i - p_i| \|P_{\lambda A}^i\|_1 + \sum_{i \in \{\pm\}} |p_i - \tilde{p}_i| \|P_{\lambda A}^i\|_1 = \sum_{i \in \{\pm\}} |\tilde{p}_i - p_i| \|P_{\lambda A}^i\|_1 + \|P_{\lambda A}^i\|_1 ,
$$

(E.6)

where we used that $\sum_{i \in \{\pm\}} p_i = \tilde{p}_i = p_i$. Since each $P_{\lambda A}$ is positive semidefinite with all eigenvalues less or equal to $1$ it follows that $\|P_{\lambda A}\|_1 \leq n_A$ where $n_A \equiv \text{dim}(H_A)$. Moreover,

$$
\|P_{\lambda A}\|_1 = \|\sum_{i \in \{\pm\}} (1 - p_i/j\tilde{p}_i)P_{\lambda A}\|_1 \leq \sum_{i \in \{\pm\}} |1 - p_i/j\tilde{p}_i| \|P_{\lambda A}\|_1.
$$

Therefore,

$$
\sum_{i} \|P_{\lambda A} \neq P_{\lambda A}\|_1 \leq 2 \sum_{i \in \{\pm\}} |\tilde{p}_i - p_i| \|P_{\lambda A}\|_1 \leq 2n_A \sum_{i \in \{\pm\}} |\tilde{p}_i - p_i| .
$$

(E.7)
We further note that $\hat{p}_i > p_i$ for $i \in \{i+\}$ and thus if $p_{\text{min}} \equiv \min_i p_i$ we have that $\hat{p}_i > p_{\text{min}}$ for $i \in \{i+\}$. It follows that $|\hat{p}_i - p_i|/p_{\text{min}} < |(\hat{p}_i - p_i)/p_{\text{min}}|$ for $i \in \{i+\}$. Hence,

$$\sum_i |\hat{P}_{Aj} - P_{Aj}| < \frac{2n_A}{p_{\text{min}}} \sum_i |\hat{P}_i - p_i| \leq \frac{2n_A}{p_{\text{min}}} \sum_i |\hat{\rho}_{AB} - \rho_{AB}| < \frac{2n_A |P|}{p_{\text{min}}} \|\hat{\rho}_{AB} - \rho_{AB}\|_1,$$

where $|P|$ is the number of elements of $P$ and we have used equation (E.5). Thus if $\{P_{Aj}\}$ is a $\mathcal{P}$-POVM for $\rho_{AB}$ the minimum trace distance between $\{P_{Aj}\}$ and a $\mathcal{P}$-POVM for $\rho_{AB}$ is upper bounded by $2n_A |P| \|\hat{\rho}_{AB} - \rho_{AB}\|_1/p_{\text{min}}$. By an analogous argument if $\{\hat{P}_{Aj}\}$ is a $\mathcal{P}$-POVM for $\hat{\rho}_{AB}$ the minimum trace distance between $\{\hat{P}_{Aj}\}$ and a $\mathcal{P}$-POVM for $\hat{\rho}_{AB}$ is upper bounded by $2n_A |P| \|\hat{\rho}_{AB} - \rho_{AB}\|_1/p_{\text{min}}$.

We now recall equations (D.7) and (D.14) from appendix D showing that the guessing probability $P_g(\rho_{AB}, \{P_{Aj}\})$ is Lipschitz continuous on the set of states for a fixed $\{P_{Aj}\}$, and how it can be extended to work for almost everywhere differentiable $\rho_{AB}$ such that all states in this neighbourhood show a backflow in the interval $[\tau, \tau + \Delta \tau]$, $\forall \tau \in [\tau_{\text{min}}, \tau_{\text{max}}]$, $\forall \Delta \tau > 0$. Thus a backflow can be seen also for evolution of a perturbed initial state $\rho_{AB}(t_0)$.

Thus $C^P_{\mathcal{A}}$ is Lipschitz continuous on the set of states.

Using equation (E.11) we can make some observations about the robustness of correlation backflows. If we have a backflow in the interval $[\tau, \tau + \Delta \tau]$ for an initial state $\rho_{AB}(t_0)$, i.e., $C^P_{\mathcal{A}}(\rho_{AB}(\tau + \Delta \tau)) - C^P_{\mathcal{A}}(\rho_{AB}(\tau)) > 0$, any state $\rho'_{AB}$ such that $\|\rho'_AB - \rho_{AB}(\tau + \Delta \tau)\|_1 < p_{\text{min}}/(p_{\text{min}} + 2n_A |P|)$ satisfies $C^P_{\mathcal{A}}(\rho'_{AB}) - C^P_{\mathcal{A}}(\rho_{AB}(\tau)) > 0$. Likewise, if $C^P_{\mathcal{A}}(\rho_{AB}(\tau + \Delta \tau)) - C^P_{\mathcal{A}}(\rho_{AB}(\tau)) > 0$ any state $\rho''_{AB}$ such that $\|\rho''_{AB} - \rho_{AB}(\tau)\|_1 < p_{\text{min}}/(p_{\text{min}} + 2n_A |P|)$ satisfies $C^P_{\mathcal{A}}(\rho''_{AB}) - C^P_{\mathcal{A}}(\rho_{AB}(\tau)) > 0$. Moreover, if $C^P_{\mathcal{A}}(\rho_{AB}(\tau + \Delta \tau)) - C^P_{\mathcal{A}}(\rho_{AB}(\tau)) > 0$ any pair of states $\rho_{AB}$ and $\rho''_{AB}$ such that $\|\rho_{AB} - \rho''_{AB}\|_1 + \|\rho_{AB} - \rho_{AB}(\tau)\|_1 < p_{\text{min}}/(p_{\text{min}} + 2n_A |P|)$ satisfies $C^P_{\mathcal{A}}(\rho_{AB}) - C^P_{\mathcal{A}}(\rho''_{AB}) > 0$.

Thus a backflow can be seen also for evolution of a perturbed initial state $\rho_{AB}(t_0)$ such that all states in this neighbourhood show a backflow in the interval $[\tau, \tau + \Delta \tau]$ and it includes all states $\rho_{AB}(t_0) + \chi$ such that $\|\chi\|_1 < p_{\text{min}}/(p_{\text{min}} + 2n_A |P|)$, $\forall \tau \in [\tau_{\text{min}}, \tau_{\text{max}}]$, $\forall \Delta \tau > 0$. Thus there is a neighbourhood of $\rho_{AB}(t_0)$ such that all states in this neighbourhood show a backflow in the interval $[\tau, \tau + \Delta \tau]$ and it includes all states $\rho_{AB}(t_0) + \chi$ such that $\|\chi\|_1 < p_{\text{min}}/(p_{\text{min}} + 2n_A |P|)$. Hence, this neighbourhood has the same dimension as $\mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$.

**Appendix F. Comments on the non-Markovianity measure: the case of non-differentiable $C^P_{\mathcal{A}}(\rho_{ASA'}(t))$**

Here we discuss the non-Markovianity measure $N^P(\{A_\Delta(t, t_0)\}_\Delta) \equiv \sup_{\rho_{ASA}(t_0)} \int_{\mathcal{P}(\rho_{ASA}(t_0)) \geq 0} \frac{d}{dt} C^P_{\mathcal{A}}(\rho_{ASA}(t)) \, dt,$ (F.1)

and how it can be extended to work for almost everywhere differentiable $C^P_{\mathcal{A}}(\rho_{ASA}(t))$. We also comment on how one may construct measures of non-Markovianity based on $C^P_{\mathcal{A}}(\rho_{ASA}(t))$ using finite differences.

First we consider the case where $C^P_{\mathcal{A}}(\rho_{ASA}(t))$ is differentiable. Consider the non-Markovianity measure introduced in equation (F.1) and let $[t_1, t_2]$ be a closed time interval for which it holds that
\( \frac{dC_A^P(\rho_{ASA}(t))}{dt} > 0 \). In equation (F.1) the type of integration used is not specified, but if the Henstock–Kurzweil integral is used it holds that
\[
\int_{t_1}^{t_2} \frac{d}{dt} C_A^P(\rho_{ASA}(t)) \, dt = C_A^P(\rho_{ASA}(t_2)) - C_A^P(\rho_{ASA}(t_1)),
\] (F.2)
if \( C_A^P(\rho_{ASA}(t)) \) is differentiable in \([t_1, t_2] \). If the Riemann or Lebesgue integral is used there would be the additional requirement that \( \frac{d}{dt} C_A^P(\rho_{ASA}(t)) \) is Riemann or Lebesgue integrable, respectively.

Next we consider the case where \( C_A^P(\rho_{ASA}(t)) \) is almost everywhere differentiable, i.e. \( C_A^P(\rho_{ASA}(t)) \) is non-differentiable for at most a countable set of times \( t_i \). At the times where \( C_A^P(\rho_{ASA}(t)) \) fails to be differentiable, it is either non-differentiable but continuous or has a discontinuity. Since \( C_A^P(\rho_{ASA}(t)) \) is a continuous function on the set of states it has a discontinuity only if the evolution of \( \rho_{ASA}(t) \) is discontinuous. To deal with these points of non-differentiability we can define a function \( \frac{d}{dt} C_A^P(\rho_{ASA}(t))^* \) that is equal to \( \frac{d}{dt} C_A^P(\rho_{ASA}(t)) \) for all \( t \) for which \( C_A^P(\rho_{ASA}(t)) \) is differentiable, and is equal to zero otherwise. If we use the Henstock–Kurzweil integral in the definition of the measure \( N^P\{\{A_S(t, t_0)\}\}_s \) it is insensitive to how we define \( \frac{d}{dt} C_A^P(\rho_{ASA}(t))^* \) in the countable set of \( t_i \) where \( C_A^P(\rho_{ASA}(t)) \) is not differentiable. Thus we can define the measure
\[
N^P\{\{A_S(t, t_0)\}\}_s \equiv \sup_{\rho_{ASA}(t_0) \in \mathcal{P}} \int_{t_0}^{t_1} \frac{d}{dt} C_A^P(\rho_{ASA}(t))^* \, dt + \sum_{t_i} \Delta_+(t_i),
\] (F.3)
where \( \Delta_+(t_i) \) is the value of a discontinuous increase of \( C_A^P(\rho_{ASA}(t)) \) at a time \( t_i \). This definition reduces to that of equation (F.1) when \( C_A^P(\rho_{ASA}(t)) \) is differentiable.

For the case when \( C_A^P(\rho_{ASA}(t)) \) is not almost everywhere differentiable the measure in equation (F.1) is not well defined. In this case one can resort to finite difference methods to estimate the amount of non-Markovianity in a given interval. A simple measure of this kind is
\[
N_{\text{finite}}^P\{\{A_S(t, t_0)\}\}_s \equiv \sup_{\rho_{ASA}(t_0) \in \mathcal{P}} \{0, C_A^P(\rho_{ASA}(t_1)) - C_A^P(\rho_{ASA}(t_0))\}
\] (F.4)
where \( t_0 \) and \( t_1 \) belong to the interval of interest. We know that if the evolution is non-Markovian there always exists at least one \( \mathcal{P}_0 \), some ancillas \( A \) and \( A' \), an initial state \( \rho_{ASA}(t_0) \) and a pair of times \( t_i \) and \( t_j \) such that \( C_A^P(\rho_{ASA}(t_j)) - C_A^P(\rho_{ASA}(t_i)) > 0 \) (See theorem 1). Therefore, \( N_{\text{finite}}^P\{\{A_S(t, t_0)\}\}_s \) > 0 if and only if the evolution \( \{A_S(t, t_0)\}_s \) is non-Markovian.

**ORCID iDs**

Dario De Santis  https://orcid.org/0000-0001-5666-1770

**References**

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

[2] Rivas A and Huelga S F 2011 *Open Quantum Systems. An Introduction* (Heidelberg: Springer)

[3] Breuer H-P, Laine E-M, Pillo J and Vacchini B 2016 Rev. Mod. Phys. 88 021002

[4] de Vega I and Alonso D 2017 Rev. Mod. Phys. 89 015001

[5] Rivas A, Huelga S F and Plenio M B 2014 Rep. Prog. Phys. 77 094001

[6] Li L, Hall M J W and Wiseman H M 2018 Phys. Rep. 759 1

[7] Teittinen J, Lyra H, Sokolov B and Maniscalco S 2018 New J. Phys. 20 073012

[8] Addis C, Bylicka B, Chrusciński D and Maniscalco S 2014 Phys. Rev. A 90 052103

[9] Chrusciński D, Kossakowski A and Rivas A 2011 Phys. Rev. A 83 052128

[10] Rivas A, Huelga S F and Plenio M B 2010 Phys. Rev. Lett. 105 050403

[11] Breuer H-P, Laine E-M and Pillo J 2009 Phys. Rev. Lett. 103 210401

[12] Buscemi F and Datta N 2016 Phys. Rev. A 93 012101

[13] Bylicka B, Johansson M and Acín A 2017 Phys. Rev. Lett. 118 120501

[14] Luo S, Fu S and Song H 2012 Phys. Rev. A 86 044101

[15] Chrusciński D and Maniscalco S 2014 Phys. Rev. Lett. 112 120404

[16] De Santis D, Johansson M, Bylicka B, Bernardes N K and Acín A 2020 Phys. Rev. A 102 022214

[17] Kobodzinski J, Rana S and Streltsov A 2020 Phys. Rev. A 101 020303(R)

[18] De Santis D, Johansson M, Bylicka B, Bernardes N K and Acín A 2019 Phys. Rev. A 99 012303

[19] Bylicka B, Chrusciński D and Maniscalco S 2014 Sci. Rep. 4 5720

[20] Lorenzo S, Plastina F and Paternostro M 2013 Phys. Rev. A 88 020102(R)

[21] Lu X-M, Wang X and Sun C P 2010 Phys. Rev. A 82 042103

[22] Kossakowski A 1972 Bull. Acad. Pol. Sci. Math. Ser. Math. Astron. 20 1021

[23] Ruskai M B 1994 Rev. Math. Phys. 06 1147

[24] Chakraborty S and Chrusciński D 2019 Phys. Rev. A 99 042105

[25] Chrusciński D, Rivas Á and Stormer E 2018 Phys. Rev. Lett. 121 080407