Quantum Non-Markovianity: Characterization, Quantification and Detection

Ángel Rivas\textsuperscript{1}, Susana F. Huelga\textsuperscript{2,3} and Martin B. Plenio\textsuperscript{2,3}

\textsuperscript{1}Departamento de Física Teórica I, Facultad de Ciencias Físicas, Universidad Complutense, 28040 Madrid, Spain.
\textsuperscript{2}Institut für Theoretische Physik, Universität Ulm, Albert-Einstein-Allee 11, 89073 Ulm, Germany.
\textsuperscript{3}Center for Integrated Quantum Science and Technologies, Albert-Einstein-Allee 11, 89073 Ulm, Germany.

Abstract. We present a comprehensive and updated review on the paradigm of quantum non-Markovianity, a central theme in the theory of open quantum systems. We motivate the introduction of a concept of quantum Markovian process on the basis of a divisibility property that generalizes the classical definition of Markovianity and relate this notion to the intuitive idea that links non-Markovianity with the persistence of memory effects. A detailed comparison with other definitions presented in the literature is provided. We then discuss several existing proposals to quantify the degree of non-Markovianity of quantum dynamics and to witness non-Markovian behavior, the latter providing sufficient conditions to detect deviations from strict Markovian behavior. Finally, we conclude by enumerating some timely open problems in the field and provide an outlook on possible research directions.
Contents

1. Introduction

In recent years, renewed attention has been paid to the characterization of quantum non-Markovian processes. Different approaches have been followed and several methods have been proposed which in some cases yield inequivalent conclusions. Given the considerable amount of literature that has built up on the subject, we believe that the time is right to summarize most of the existing results in a review article that clarifies both the underlining structure and the interconnections between the different approaches.

On the one hand, we are fully aware of the risk we take by writing a review of a quite active research field, with new results arising during the completion of the work. Although, on the other hand, we hope that possible shortcoming will be well balanced by the potential usefulness of such a review in order to, hopefully, clarify some misconceptions and generate further interest in the field.

Essentially, the subject of quantum non-Markovianity splits into two main questions, namely,

I. What is a quantum Markovian process and hence what are non-Markovian processes? (characterization problem).

II. If a given process deviates from Markovianity, by how much does it deviate? (quantification problem).

In this work we examine both questions in detail. More specifically, concerning the characterization problem, we adopt the so-called called divisibility property as a definition of quantum Markovian processes. Nonetheless in section ?? we make a comparison and comment on other proposed definitions. In this regard, we should stress that it is not our intention to make an unquestionable claim of what should be called quantum Markov process. What we hope is to convince the reader about the strong analogy between the definition taken here and the classical definition of Markov processes, as well as emphasizing the resulting good mathematical properties which will allow us to address the characterization problem in simple terms. Concerning the quantification problem, we have analyzed most of the quantifiers present in the literature, which we classify into measures and witnesses of non-Markovianity, depending on whether they are able to detect every non-Markovian dynamics or just a subset. Given the large body of literature that explores the application of these methods to different physical realizations, we have opted for keeping the presentation on the abstract level and not including specific examples in our discussion, but provide a detailed list of references instead.

Thus, this work is organized as follows. In section 2, we recall the classical concept of Markovian process and some of its main properties. This is crucial in order to understand why the divisibility property provides a good definition of quantum Markovianity. In
section 3 we introduce the concept of quantum Markovian process by establishing a step by step parallelism with the classical definition, and explain in detail why these quantum processes can be considered as memoryless. Section 4 gives a detailed review of different measures of non-Markovianity and section 4 describes different approaches in order to construct witnesses able to detect non-Markovian dynamics. Finally, section 6 is devoted to conclusions and to outline some of the problems which remain open in this field and possible future research lines.

2. Markovianity in Classical Stochastic Processes

In order to give a good definition of a Markov process in the quantum regime, it is essential to understand the concept of Markov process in the classical setting. Thus, this section is devoted to revise the definition of classical Markov processes and sketch the most interesting properties for our purposes without getting too concerned with mathematical rigor. More detailed explanations on the foundations of stochastic processes can be found in [?, ?, ?, ?, ?, ?].

2.1. Definition and properties

Consider a random variable \( X \) defined on a classical probability space \((\Omega, \Sigma, \mathbb{P})\), where \( \Omega \) is a given set (the sample space), \( \Sigma \) (the possible events) is a \( \sigma \)-algebra of subsets of \( \Omega \), containing \( \Omega \) itself, and the probability \( \mathbb{P}: \Sigma \rightarrow [0,1] \) is a \( \sigma \)-additive function with the property that \( \mathbb{P}(\Omega) = 1 \), (cf. [?, ?, ?, ?, ?, ?]). In order to avoid further problems when considering conditional probabilities (see for example the Borel-Kolmogorov paradox [?]) we shall restrict attention from now on to discrete random variables, i.e. random variables which take values on a finite set denoted by \( \mathcal{X} \).

A classical stochastic process is a family of random variables \( \{X(t), t \in I \subset \mathbb{R}\} \). Roughly speaking, this is nothing but a random variable \( X \) depending on a parameter \( t \) which usually represents time. The story starts with the following definition.

**Definition 2.1 (Markov process).** A stochastic process \( \{X(t), t \in I\} \) is a Markov process if the probability that the random variable \( X \) takes a value \( x_n \) at any arbitrary time \( t_n \in I \), provided that it took the value \( x_{n-1} \) at some previous time \( t_{n-1} < t_n \in I \), is uniquely determined, and not affected by the possible values of \( X \) at previous times to \( t_{n-1} \). This is formulated in terms of conditional probabilities as follows

\[
\mathbb{P}(x_n, t_n|x_{n-1}, t_{n-1}; \ldots; x_0, t_0) = \mathbb{P}(x_n, t_n|x_{n-1}, t_{n-1}), \quad \text{for all } \{t_n, t_{n-1}, \ldots, t_0\} \subset I,
\]

and informally it is summarized by the statement that “a Markov process does not have memory of the history of past values of \( X \)”. This kind of stochastic processes are so-named after the Russian mathematician A. Markov [?].

From the previous definition (??) it is possible to work out further properties of Markov processes. For instance, it follows immediately from (??) that for a Markov
random variable $X$ is a stochastic process. Thus, consider a linear map $T$

In this section we shall focus on the evolution of one-point probabilities $P$

2.2. Transition matrices

process 

$$E(x_n, t_n|x_{n-1}, t_{n-1}; \ldots; x_0, t_0) = E(x_n, t_n|x_{n-1}, t_{n-1}), \quad \text{for all } \{t_n, t_{n-1}, \ldots, t_0\} \subset I,$$

where $E(x|y) = \sum_{x \in X} xP(x|y)$ denotes the so-called conditional expectation.

In addition, Markov processes satisfy another remarkable property. If we take the joint probability for any three consecutive times $t_3 > t_2 > t_1$ and apply the definition of conditional probability twice we obtain

$$P(x_3, t_3; x_2, t_2; x_1, t_1) = P(x_3, t_3|x_2, t_2; x_1, t_1)P(x_2, t_2; x_1; t_1)$$

$$= P(x_3, t_3|x_2, t_2; x_1, t_1)P(x_2, t_2|x_1, t_1)P(x_1, t_1).$$

Since the Markov condition (??) implies that $P(x_3, t_3|x_2, t_2; x_1, t_1) = P(x_3, t_3|x_2, t_2)$, by taking the sum over $x_2$ and dividing both sides by $P(x_1, t_1)$ we arrive at

$$P(x_3, t_3|x_1, t_1) = \sum_{x_2 \in X} P(x_3, t_3|x_2, t_2)P(x_2, t_2|x_1, t_1),$$

which is called Chapman-Kolmogorov equation. Moreover, the next theorem gives an answer to the converse statement.

**Theorem 2.1.** A family of conditional probabilities $P(x_n, t_n|x_{n-1}, t_{n-1})$ with $t_n > t_{n-1}$ satisfying (??) can always be seen as the conditional probabilities of a Markov process $\{X(t), t \in I\}$.

**Proof.** The proof is by construction. Take some probabilities $P(x_n, t_n)$ and define the two-point joint probabilities by

$$P(x_n, t_n; x_{n-1}, t_{n-1}) := P(x_n, t_n|x_{n-1}, t_{n-1})P(x_{n-1}, t_{n-1}).$$

Then, set

$$P(x_n, t_n|x_{n-1}, t_{n-1}; \ldots; x_0, t_0) := P(x_n, t_n|x_{n-1}, t_{n-1}), \quad \text{for all } \{t_n, t_{n-1}, \ldots, t_0\} \subset I.$$

and construct higher joint probabilities by using expressions analogous to Eq. (??). This construction is always possible as it is compatible with (??), which is the presupposed condition satisfied by $P(x_n, t_n|x_{n-1}, t_{n-1})$. \hfill $\Box$

2.2. Transition matrices

In this section we shall focus on the evolution of one-point probabilities $P(x, t)$ during a stochastic process. Thus, consider a linear map $T$ that connects the probability of a random variable $X$, at different times $t_0$ and $t_1$:

$$P(x_1, t_1) = \sum_{x_0 \in X} T(x_1, t_1|x_0, t_0)P(x_0, t_0).$$
Since $\sum_{x_1 \in X} P(x_1, t_1) = 1$ and $P(x_1, t_1) \geq 0$ for every $P(x_0, t_0)$, we conclude that

$$\sum_{x_1 \in X} T(x_1, t_1| x_0, t_0) = 1,$$

$$T(x_1, t_1| x_0, t_0) \geq 0, \quad x_1, x_0 \in X. \quad (7)\quad (8)$$

Matrices $T$ fulfilling these properties are called *stochastic matrices.*

Consider $t = t_0$ to be the initial time of some (not necessarily Markovian) stochastic process $\{X(t), t \in I\}$. From definition of conditional probability,

$$P(x_2, t_2; x_0, t_0) = P(x_2, t_2| x_0, t_0)P(x_0, t_0) \Rightarrow P(x_2, t_2) = \sum_{x_0 \in X} P(x_2, t_2| x_0, t_0)P(x_0, t_0), \quad (9)$$

and therefore $T(x_2, t_2| x_0, t_0) = P(x_2, t_2| x_0, t_0)$ for every $t_2$. This relation is not valid in general for $t_1 > t_0$, $T(x_2, t_2| x_1, t_1) \neq P(x_2, t_2| x_1, t_1)$. The reason is that $P(x_2, t_2| x_1, t_1)$ is not fully defined for a general stochastic process; we need to know the value of $X$ for previous time instants as $P(x_2, t_2| x_1, t_0, t_0)$ could be different from $P(x_2, t_2| x_1, t_1; x_0, t_0)$ for $x_0 \neq x'_0$. However that is not the case for Markov processes which satisfy the following result.

**Theorem 2.2.** Consider a Markov process $\{X(t), t \in I\}$. Given any two time instants $t_1$ and $t_2$ we have

$$T(x_2, t_2| x_1, t_1) = P(x_2, t_2| x_1, t_1). \quad (10)$$

**Proof.** It follows from the fact that we can write $P(x_2, t_2; x_1, t_1) = P(x_2, t_2| x_1, t_1)P(x_1, t_1)$, as $P(x_2, t_2| x_1, t_1)$ is well defined for any $t_1$ and $t_2$. \qed

From this theorem and the Chapman-Kolmogorov equation (??) we obtain the following corollary.

**Corollary 2.1.** Consider a Markov process $\{X(t), t \in I\}$, then for any $t_3 \geq t_2 \geq t_1 \geq t_0$, the transition matrix satisfies the properties

$$\sum_{x_2 \in X} T(x_2, t_2| x_1, t_1) = 1,$$

$$T(x_2, t_2| x_1, t_1) \geq 0,$$

$$T(x_3, t_3| x_1, t_1) = \sum_{x_2 \in X} T(x_3, t_3| x_2, t_2)T(x_2, t_2| x_1, t_1). \quad (11)\quad (12)\quad (13)$$

In summary, for a Markov process the transition matrices are the two-point conditional probabilities and satisfy the composition law Eq. (12). Essentially, property (12) states that the evolution from $t_1$ to $t_3$ can be written as the composition of the evolution from $t_1$ to some intermediate time $t_2$, and from this $t_2$ to the final time $t_3$.

In case of non-Markovian processes, $T(x_2, t_2| x_1, t_1)$ might be not well defined for $t_1 \neq t_0$. Nevertheless, if the matrix $P(x_1, t_1| x_0, t_0)$ is invertible for every $t_1$, then $T(x_2, t_2| x_1, t_1)$ can be written in terms of well-defined quantities. Since the evolution
from $t_1$ to $t_2$ (if it exists) has to be the composition of the backward evolution to the initial time $t_0$ and the forward evolution from $t_0$ to $t_2$, we can write

$$T(x_2, t_2|x_1, t_1) = \sum_{x_0 \in \mathcal{X}} T(x_2, t_2|x_0, t_0)T(x_0, t_0|x_1, t_1)$$

$$= \sum_{x_0 \in \mathcal{X}} \mathbb{P}(x_2, t_2|x_0, t_0)[\mathbb{P}(x_1, t_1|x_0, t_0)]^{-1}. \quad (14)$$

In this case the condition Eq. (15) is satisfied and Eq. (16) also holds. However, condition Eq. (17) may be not fulfilled, which prevents from any interpretation of $T(x_2, t_2|x_1, t_1)$ as a conditional probability and therefore manifests the non-Markovian character of such a stochastic process.

Definition 2.2 (Divisible Process). A stochastic process $\{X(t), t \in I\}$ for which their transition matrices satisfy Eqs. (15), (16) and (17) is called divisible.

There are divisible processes which are non-Markovian. As an example (see [? , ?]), consider a stochastic process $\{X(t), t \in I\}$ with two possible results $\mathcal{X} = \{0, 1\}$, and just three discrete times $I = \{t_1, t_2, t_3\}$ ($t_3 > t_2 > t_1$). Define the joint probabilities as

$$\mathbb{P}(x_3, t_3; x_2, t_2; x_1, t_1) := \frac{1}{4} (\delta_{x_3,0}\delta_{x_2,0}\delta_{x_1,1} + \delta_{x_3,0}\delta_{x_2,1}\delta_{x_1,0} + \delta_{x_3,1}\delta_{x_2,0}\delta_{x_1,0} + \delta_{x_3,1}\delta_{x_2,1}\delta_{x_1,1}). \quad (15)$$

By computing the marginal probabilities we obtain $\mathbb{P}(x_3, t_3; x_2, t_2) = \mathbb{P}(x_2, t_2; x_1, t_1) = \mathbb{P}(x_3, t_3; x_1, t_1) = 1/4$, and then

$$\mathbb{P}(x_3, t_3|x_2, t_2; x_1, t_1) = \frac{\mathbb{P}(x_3, t_3; x_2, t_2; x_1, t_1)}{\mathbb{P}(x_2, t_2; x_1, t_1)} = (\delta_{x_3,0}\delta_{x_2,0}\delta_{x_1,1} + \delta_{x_3,1}\delta_{x_2,0}\delta_{x_1,0} + \delta_{x_3,1}\delta_{x_2,1}\delta_{x_1,1}). \quad (16)$$

Therefore the process is non-Markovian as, for example, $\mathbb{P}(1, t_3|0, t_2; 0, t_1) = 1$ and $\mathbb{P}(1, t_3|0, t_2; 1, t_1) = 0$. However the transition matrices can be written as

$$T(x_3, t_3|x_2, t_2) = \frac{\mathbb{P}(x_3, t_3; x_2, t_2)}{\mathbb{P}(x_2, t_2)} = \frac{1}{2},$$

and similarly $T(x_2, t_2|x_1, t_1) = T(x_3, t_3|x_1, t_1) = 1/2$. Hence the conditions (17), (17) and (17) are clearly fulfilled. Other examples of non-Markovian divisible processes can be found in [? , ? , ? , ? , ?].

Despite the existence non-Markovian divisible processes, we can establish the following key theorem.

Theorem 2.3. A family of transition matrices $T(x_2, t_2|x_1, t_1)$ with $t_2 > t_1$ which satisfies (17), (17) and (17) can always be seen as the transition matrices of some underlaying Markov process $\{X(t), t \in I\}$. 
Proof. Since the matrices $T(x_2, t_2|x_1, t_1)$ satisfy (??) and (??), they can be understood as conditional probabilities $\mathbb{P}(x_2, t_2|x_1, t_1) = T(x_2, t_2|x_1, t_1)$, and since (??) is also satisfied, the process fulfils Eq. (??). Then the final statement follows from Theorem ??.

Thus, we conclude that:

**Corollary 2.2.** *At the level of one-point probabilities, divisible and Markovian processes are equivalent. It is required to know the complete hierarchy of time-conditional probabilities to make any distinction.*

### 2.3. Contractive property

There is another feature of Markov processes that will be useful in the quantum case. Consider a vector $v(x)$, where $x$ denotes its different components. Then its $L_1$–norm is defined as

$$\|v(x)\|_1 := \sum_x |v(x)|. \quad (17)$$

This norm is particularly useful in hypothesis testing problems. Namely, consider a random variable $X$ which is distributed according to either probability $p_1(x)$ or probability $p_2(x)$. We know that, with probability $q$, $X$ is distributed according to $p_1(x)$, and, with probability $1 - q$, $X$ is distributed according to $p_2(x)$. Our task consists of sampling $X$ just once with the aim of inferring the correct probability distribution of $X$ [[$p_1(x)$ or $p_2(x)$]]. Then the minimum (averaged) probability to give the wrong answer turns out to be

$$\mathbb{P}_{\text{min}}(\text{fail}) = \frac{1 - \|w(x)\|_1}{2}, \quad (18)$$

where $w(x) := qp_1(x) - (1 - q)p_2(x)$. The proof of this result follows the same steps as in the quantum case (see section ??). Thus the $L_1$-norm of the vector $w(x)$ gives the capability to distinguish correctly between $p_1(x)$ and $p_2(x)$ in the two-distribution discrimination problem.

Particularly, in the unbiased case $q = 1/2$, we have

$$\|w(x)\|_1 = \frac{1}{2}\|p_1(x) - p_2(x)\|_1,$$

which is known as the *Kolmogorov distance*, $L_1$-distance, or variational distance between $p_1(x)$ and $p_2(x)$.

In order to identify non-divisible process, the $L_1$-norm also plays an important role.

**Theorem 2.4.** *Let $T(x_2, t_2|x_1, t_1)$ be the transition matrices of some stochastic process. Then such a process is divisible if and only if the $L_1$-norm does not increase when $T(x_2, t_2|x_1, t_1)$ is applied to every vector $v(x)$, $x \in \mathcal{X}$, for all $t_1$ and $t_2$,

$$\left\| \sum_{x_1 \in \mathcal{X}} T(x_2, t_2|x_1, t_1)v(x_1) \right\|_1 \leq \|v(x_2)\|_1, \quad t_1 \leq t_2. \quad (19)$$
Proof. The “only if” part follows from the properties (??) and (??):

$$\left\| \sum_{x_1 \in X} T(x_2, t_2 | x_1, t_1) v(x_1) \right\|_1 = \sum_{x_2 \in X} \left| \sum_{x_1 \in X} T(x_2, t_2 | x_1, t_1) v(x_1) \right|$$

$$\leq \sum_{x_1, x_2 \in X} T(x_2, t_2 | x_1, t_1) \left| v(x_1) \right|$$

$$= \sum_{x_1 \in X} \left| v(x_1) \right| = \sum_{x_2 \in X} \left| v(x_2) \right| = \left\| v(x_2) \right\|.$$  (20)

For the “if” part, as we aforementioned, if $T(x_2, t_2 | x_1, t_1)$ does exist, it always satisfies Eqs. (??) and (??). Take a vector to be a probability distribution $v(x) = p(x) \geq 0$ for all $x \in X$, because of Eq. (??) we have

$$\left\| p(x_1) \right\|_1 = \sum_{x_1 \in X} p(x_1) = \sum_{x_1, x_2 \in X} T(x_2, t_2 | x_1, t_1) p(x_1).$$  (21)

Since, by hypothesis, Eq. (??) holds for any vector, we obtain the following chain of inequalities

$$\left\| p(x_1) \right\|_1 = \sum_{x_1, x_2 \in X} T(x_2, t_2 | x_1, t_1) p(x_1) \leq \sum_{x_2 \in X} \left| \sum_{x_1 \in X} T(x_2, t_2 | x_1, t_1) p(x_1) \right|$$

$$\leq \sum_{x_2 \in X} \left| p(x_2) \right| = \sum_{x_1 \in X} \left| p(x_1) \right| = \left\| p(x_1) \right\|.$$  (22)

Therefore,

$$\sum_{x_2 \in X} \left| \sum_{x_1 \in X} T(x_2, t_2 | x_1, t_1) p(x_1) \right| = \sum_{x_1, x_2 \in X} T(x_2, t_2 | x_1, t_1) p(x_1),$$

for any probability $p(x_1)$, which is only possible if $\sum_{x_1 \in X} T(x_2, t_2 | x_1, t_1) p(x_1) \geq 0$. Then Eq. (??) has to be satisfied. □

Because of this theorem and Eq. (??), for a divisible process $P_{\min}(\text{fail})$ increases monotonically with time. In this regard, if the random variable $X$ undergoes a Markovian process, the best chance to rightly distinguish between the two possible distributions $p_1(x)$ and $p_2(x)$ is to sample $X$ at time instants as close as possible to the initial time $t_0$. However that is not the case if $X$ is subject to a non-divisible (and then non-Markovian) process. Then, in order to decrease the error probability, it could be better to wait until some time, say $t_1$, where $\left\| w(x_1, t_1) \right\|_1 = \left\| qp_1(x_1, t_1) - (1 - q)p_2(x_1, t_1) \right\|_1$ increases again. The fact that the error probability may decrease for some time $t_1$ posterior to the initial $t_0$ can be understood as a trait of underlying memory in the process. That is, the system keeps in some information about the probability of $X$ at $t_0$, which arises at a posterior time in the evolution.

In summary, classical Markovian processes are defined via multi-time conditional probabilities, Eq. (??). However, if the experimenter only has access to one-point probabilities, Markovian processes become equivalent to divisible processes. The latter
are more easily characterized, as they only depend on properties of transition matrices and the $L_1$-norm.

3. Markovianity in Quantum Processes

After the succinct review of classical Markovian processes in the previous section, here we shall try to adapt those concepts to the quantum case. By the adjective “quantum” we mean that the system undergoing evolution is a quantum system. Our aim is to find a simple definition of a quantum Markovian process by keeping a close analogy to its classical counterpart. Since this is not straightforward, we comment first on some points which make a definition of quantum Markovianity difficult to formulate. For the sake of simplicity, in the following we shall consider finite dimensional quantum systems unless otherwise stated.

3.1. Problems of a straightforward definition

Since the quantum theory is a statistical theory, it seems meaningful to ask for some analogue to classical stochastic processes and particularly Markov processes. However, the quantum theory is based on non-commutative algebras and this makes its analysis considerably more involved. Indeed, consider the classical definition of Markov process Eq. (??), to formulate a similar condition in the quantum realm we demand a way to obtain $P(x_n, t_n|x_{n-1}, t_{n-1}; \ldots; x_0, t_0)$ for quantum systems. The problem arises because we can sample a classical random variable without affecting its posterior statistics; however, in order to “sample” a quantum system, we need to perform measurements, and these measurements disturb the state of the system and affect the subsequent outcomes. This well-known fact is on the basis of the quantum cryptography theory (cf. [?]). Thus, $P(x_n, t_n|x_{n-1}, t_{n-1}; \ldots; x_0, t_0)$ does not only depend on the dynamics but also on the measurement process, and a definition of quantum Markovianity in terms of it, even if possible, does not seem very appropriate. Actually, in such a case the Markovian character of a quantum dynamical system would depend on which measurement scheme is chosen to obtain $P(x_n, t_n|x_{n-1}, t_{n-1}; \ldots; x_0, t_0)$. This is very inconvenient as the definition of Markovianity should be independent of what is required to verify it.

3.2. Definition in terms of one-point probabilities: divisibility

Given the aforementioned problems to construct $P(x_n, t_n|x_{n-1}, t_{n-1}; \ldots; x_0, t_0)$ in the quantum case, a different approach focuses on the study of one-time probabilities $P(x, t)$. For these ones, the classical definition of Markovianity reduces to the concept of divisibility (see Definition ??), and a very nice property is that divisibility may be defined in the quantum case without any explicit mention to measurement processes. To define quantum Markovianity in terms of divisibility may seem to lose generality, nevertheless Theorem ?? and Corollary ?? assert that this loss is innocuous, as divisibility and Markovianity are equivalent properties for one-time probabilities. These probabilities
are the only ones which can be constructed in the quantum case avoiding the difficulties associated to measurement disturbance.

Let us consider a system in a quantum state given by some (non-degenerate) density matrix $\rho$, the spectral decomposition yields

$$
\rho = \sum_x p(x) |\psi(x)\rangle \langle \psi(x)|.
$$

Here the eigenvalues $p(x)$ form a classical probability distribution, which may be interpreted as the probabilities for the system to be in the corresponding eigenstate $|\psi(x)\rangle$.

$$
P(|\psi(x)\rangle) := p(x).
$$

Consider now some time evolution of the quantum system such that the spectral decomposition of the initial state is preserved,

$$
\rho(t_0) = \sum_x p(x, t_0) |\psi(x)\rangle \langle \psi(x)|, \\
\rho(t) = \sum_x p(x, t) |\psi(x)\rangle \langle \psi(x)| \in S,
$$

where $S$ denotes the set of quantum states with the same eigenvectors as $\rho(t_0)$. Since this process can be seen as a classical stochastic process on the variable $x$, which labels the eigenstates $|\psi(x)\rangle$, we consider it to be divisible if the evolution of $p(x, t)$ satisfies the classical definition of divisibility (Definition ??). In such a case, there are transition matrices $T(x_1, t_1 | x_0, t_0)$, such that

$$
p(x_1, t_1) = \sum_{x_0 \in X} T(x_1, t_1 | x_0, t_0) p(x_0, t_0),
$$

fulfilling Eqs. (??), (??) and (??). This Eq. (??) can be written in terms of density matrices as

$$
\rho(t) = \mathcal{E}(t, t_0) [\rho(t_0)].
$$

Here, $\mathcal{E}(t, t_0)$ is a dynamical map that preserves the spectral decomposition of $\rho(t_0)$ and satisfies

$$
\mathcal{E}(t, t_0) \rho(t_0) = \sum_{x_0 \in X} p(x_0, t_0) \mathcal{E}(t, t_0) |\psi(x_0)\rangle \langle \psi(x_0)|
$$

$$
= \sum_{x_0 \in X} T(x_1, t_1 | x_0, t_0) p(x_0, t_0) |\psi(x_1)\rangle \langle \psi(x_1)|.
$$

Furthermore, because of Eqs. (??), (??) and (??), $\mathcal{E}(t, t_0)$ preserves positivity and trace of any state in $S$ and obeys the composition law

$$
\mathcal{E}(t_3, t_1) = \mathcal{E}(t_3, t_2) \mathcal{E}(t_2, t_1), \quad t_3 \geq t_2 \geq t_1.
$$

On the other hand, since the maps $\{\mathcal{E}(t_2, t_1), t_2 \geq t_1 \geq t_0\}$ are supposed to describe some quantum evolution, they are linear (there is not experimental evidence against this fact [?, ?, ?]). Thus, their action on another set $S'$ of quantum states with different spectral projectors to $S$ is physically well defined provided that the positivity of the states of $S'$ is preserved. Hence, by consistence, we formulate the following general definition of a P-divisible process.
**Definition 3.1 (P-divisible process)**. We say that a quantum system subject to some time evolution characterized by the family of linear maps \( \{ \mathcal{E}(t_2,t_1), t_2 \geq t_1 \geq t_0 \} \) is *P-divisible* if, for every \( t_2 \) and \( t_1 \), \( \mathcal{E}(t_2,t_1) \) is a positive map (preserve the positivity of any quantum state) and fulfills Eq. (?). The reason to use the terminology “P-divisible” (which stands for positive-divisible) instead of “divisible” comes from the difference between positive and complete positive maps which is essential in quantum mechanics. It is well-known that the requirement of positivity alone for a dynamical map presents difficulties [?]. Concretely, in order to keep the positivity of density matrices in presence of entanglement with another extra system we must impose complete positivity instead of positivity. Recall that a completely positive (linear) map \( \mathcal{E} \) is a map such that \( \mathcal{E} \otimes \mathbb{I}_d \) is a positive map for any dimension \( d \) of the identity map \( \mathbb{I} \). Incidentally, the detection of positive maps is typically very complicated, while the complete positivity is much more easy to check [?]. Thus, now we are able to give a definition of quantum Markovian process.

**Definition 3.2 (Markovian Quantum Process)**. We shall say that a quantum system subject to a time evolution given by some family of linear maps \( \{ \mathcal{E}(t_2,t_1), t_2 \geq t_1 \geq t_0 \} \) is *Markovian* (or *divisible* [?]) if, for every \( t_2 \) and \( t_1 \), \( \mathcal{E}(t_2,t_1) \) is a complete positive map and fulfills the composition law Eq. (?).

For the sake of comparison, the following table shows the clear parallelism between classical transition matrices and quantum evolution families in a Markovian process.

<table>
<thead>
<tr>
<th>Normalization</th>
<th>Classical</th>
<th>Quantum</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sum_{x_2 \in X} T(x_2,t_2</td>
<td>x_1,t_1) = 1 )</td>
<td>( \mathcal{E}(t_2,t_1) ) trace-preserving</td>
</tr>
<tr>
<td>Positivity</td>
<td>( T(x_2,t_2</td>
<td>x_1,t_1) \geq 0 )</td>
</tr>
<tr>
<td>Composition Law</td>
<td>( T(x_3,t_3</td>
<td>x_1,t_1) = \sum_{x_2 \in X} T(x_3,t_3</td>
</tr>
</tbody>
</table>

Before we move on, it is worth to summarize the argument leading to the definition of Markovian quantum process, as it is the central concept of this work. Namely, since a direct definition from the classical condition Eq. (?) is problematic because of quantum measurement disturbance, we focus on one-time probabilities. For those, classical Markovian processes and divisible processes are equivalent, thus we straightforward formulate the divisibility condition for quantum dynamics preserving the spectral decomposition of certain set of states \( \mathcal{S} \). Then the Markovian (or divisibility) condition for any quantum evolution follows by linearity taking into account the completely positive requirement in the quantum evolution. We have sketched this reasoning in the scheme presented in figure 1.
Finally, we review a fundamental result regarding differentiable quantum Markovian processes (i.e. processes such that the limit \( \lim_{\epsilon \to 0} [E_{(t+t, t)} - 1]/\epsilon := L_t \) is well-defined). In this case, there is a mathematical result which is quite useful to characterize Markovian dynamics.

**Theorem 3.1 (Kossakowski-Lindblad form).** An operator \( L_t \) is the generator of a quantum Markov (or divisible) process if and only if it can be written in the form

\[
\frac{d\rho(t)}{dt} = L_t[\rho(t)] = -i[H(t), \rho(t)] + \sum_k \gamma_k(t) \left[ V_k(t)\rho(t)V_k^\dagger(t) - \frac{1}{2}\{V_k^\dagger(t)V_k(t), \rho(t)\} \right],
\]

(31)

where \( H(t) \) and \( V_k(t) \) are time-dependent operators, with \( H(t) \) self-adjoint, and \( \gamma_k(t) \geq 0 \) for every \( k \) and time \( t \).

This theorem is a consequence of the pioneering work by A. Kossakowski (and co-workers) [16] and independently G. Lindblad [17], who analyzed the case of time-homogeneous equations, i.e. time-independent generators \( L_t \equiv L \). For a completed proof including possible time-dependent \( L_t \) see [18, 19].

3.3. Where is the memoryless property in quantum Markovian processes?

As mentioned before, our motivation in the Definition ?? of a quantum Markovian process has been to keep formal analogy with the classical case. However, it is not immediate to appreciate that the classical memoryless property in the system evolution is also present in the quantum domain. There are at least two ways to visualize this (apparently hidden) property in Definition ??: One is based on the contractive properties of the completely positive maps and the other resorts to a collisional model of system-environment interactions.

3.3.1. **Contractive property of a quantum Markovian process** Similarly to the classical case (see section ??), memoryless properties of quantum Markovian processes become quite clear in hypothesis testing problems [20]. In the quantum case, we consider a system, with associated Hilbert space \( \mathcal{H} \), whose state is represented by the density matrix \( \rho_1 \) with probability \( q \), and \( \rho_2 \) with probability \( 1 - q \). We wish to determine which density matrix describes the true state of the quantum system by performing a measurement. If we consider some general positive operator valued measure (POVM) \( \{\Pi_x\} \) (cf. [21]), where \( x \in \mathcal{X} \) is the set of possible outcomes, we may split this set in two complementary subsets. If the outcome of the measurement is inside some \( A \subset \mathcal{X} \), then we say that the state is \( \rho_1 \). Conversely, if the result of the measurement belongs to the complementary set \( A^c \) such that \( A \cup A^c = \mathcal{X} \), we say that the state is \( \rho_2 \). Let us group the results of this measurement in another POVM given by the pair \( \{T, \mathbb{I} - T\} \), with \( T = \sum_{x \in A} \Pi_x \).

Thus, when the true state is \( \rho_1 \) (which happens with probability \( q \)) we erroneously
identify the state as \( \rho_2 \) with probability

\[
\sum_{j \in A^c} \text{Tr}[\rho_1 \Pi_x] = \text{Tr} \left[ \rho_1 \left( \sum_{x \in A^c} \Pi_x \right) \right] = \text{Tr} [\rho_1 (\mathbb{I} - T)].
\] (32)

On the other hand, when the true state is \( \rho_2 \) (which happens with probability \( 1 - q \)), we erroneously identify the state as \( \rho_1 \) with probability

\[
\sum_{j \in A} \text{Tr}[\rho_2 \Pi_x] = \text{Tr} \left[ \rho_2 \left( \sum_{x \in A} \Pi_x \right) \right] = \text{Tr} [\rho_2 T].
\] (33)

The problem in one-shot two-state discrimination is to examine the trade-off between the two error probabilities \( \text{Tr} [\rho_2 T] \) and \( \text{Tr} [\rho_1 (\mathbb{I} - T)] \). Thus, consider the best choice of \( T \) that minimizes the total averaged error probability

\[
\min_{0 \leq T \leq \mathbb{I}} \{ (1 - q) \text{Tr} [\rho_2 T] + q \text{Tr} [\rho_1 (\mathbb{I} - T)] \} = \min_{0 \leq T \leq \mathbb{I}} \{ q + \text{Tr} [(1 - q)\rho_2 T - q\rho_1 T] \}
= q - \max_{0 \leq T \leq \mathbb{I}} [\text{Tr} (\Delta T)],
\] (34)

where \( \Delta = q\rho_1 - (1 - q)\rho_2 \) is a Hermitian operator, with trace \( \text{Tr} \Delta = 2q - 1 \) vanishing in the unbiased case \( q = 1/2 \). \( \Delta \) is sometimes called Helstrom matrix [?]. We have the following result.

**Theorem 3.2.** With the best choice of \( T \), the minimum total error probability in the one-shot two-state discrimination problem becomes

\[
P_{\min}(\text{fail}) = \min_{0 \leq T \leq \mathbb{I}} \{ (1 - q) \text{Tr} [\rho_2 T] + q \text{Tr} [\rho_1 (\mathbb{I} - T)] \} = \frac{1 - \|\Delta\|_1}{2}, \]

where \( \|\Delta\|_1 = \text{Tr} \sqrt{\Delta^\dagger \Delta} \) is the trace norm of the Helstrom matrix \( \Delta \).

Thus, note that when \( q = 0 \) or \( q = 1 \) we immediately obtain zero probability to identify wrongly the true state.

**Proof.** The proof follows the same steps as for the unbiased \( q = 1/2 \) case (see [?, ?]). The spectral decomposition allows us to write \( \Delta = \Delta^+ - \Delta^- \), with positive operators \( \Delta^\pm = \pm \sum_k \lambda^\pm_k |\psi^\pm_k \rangle \langle \psi^\pm_k| \) where \( \lambda^+_k \) are the positive eigenvalues of \( \Delta \) and \( \lambda^-_k \) the negative ones. Then it is clear that for \( 0 \leq T \leq \mathbb{I} \)

\[
\text{Tr} (\Delta T) = \text{Tr} (\Delta^+ T) - \text{Tr} (\Delta^- T) \leq \text{Tr} (\Delta^+ T) \leq \text{Tr} (\Delta^+) ,
\] (36)

so that

\[
P_{\min}(\text{fail}) = q - \max_{0 \leq T \leq \mathbb{I}} [\text{Tr} (\Delta T)] = q - \text{Tr} (\Delta^+). \] (37)

On the other hand, because \( |\psi^+_j \rangle \langle \psi^+_j| \) are orthogonal projections (in other words as \( \|\Delta\|_1 = \sum_j |\lambda_j| \)), the trace norm of \( \Delta \) is

\[
\|\Delta\|_1 = \|\Delta^+\|_1 + \|\Delta^-\|_1 = \text{Tr}(\Delta^+) + \text{Tr}(\Delta^-). \] (38)
Since
\[ \text{Tr}(\Delta^+) - \text{Tr}(\Delta^-) = \text{Tr}(\Delta) = 2q - 1, \]  
we have
\[ \|\Delta\|_1 = 2\text{Tr}(\Delta^+) + (1 - 2q). \]  
(40)
Using this relation in (??) we straightforwardly obtain the result (??).

Thus the trace norm of \( \Delta = q\rho_1 - (1 - q)\rho_2 \) gives our capability to distinguish correctly between \( \rho_1 \) and \( \rho_2 \) in the one-shot two-state discrimination problem.

On the other hand, the following theorem connects trace-preserving and positive maps with the trace norm. It was first proven by Kossakowski in references [?, ?], while Ruskai also analyzed the necessary condition in [?].

**Theorem 3.3.** A trace preserving linear map \( \mathcal{E} \) is positive if and only if for any Hermitian operator \( \Delta \) acting on \( \mathcal{H} \),
\[ \|\mathcal{E}(\Delta)\|_1 \leq \|\Delta\|_1. \]  
(41)
Proof. Assume that \( \mathcal{E} \) is positive and trace preserving, then for every positive semidefinite \( \Delta \geq 0 \) the trace norm is also preserved, \( \|\mathcal{E}(\Delta)\|_1 = \|\Delta\|_1 \). Consider \( \Delta \) not to be necessarily positive semidefinite, then by using the same decomposition as in the proof of Theorem ??, \( \Delta = \Delta^+ - \Delta^- \), we have
\[ \|\mathcal{E}(\Delta)\|_1 = \|\mathcal{E}(\Delta^+) - \mathcal{E}(\Delta^-)\|_1 \leq \|\mathcal{E}(\Delta^+)\|_1 + \|\mathcal{E}(\Delta^-)\|_1 = \|\Delta^+\|_1 + \|\Delta^-\|_1 = \|\Delta\|_1, \]  
(42)
where the penultimate equality follows from the positivity of \( \Delta^\pm \). Therefore, \( \mathcal{E} \) fulfils Eq. (??).

Conversely, assume that \( \mathcal{E} \) satisfies (??) and preserves the trace, then for a positive semidefinite \( \Delta \) we have the next chain of inequalities:
\[ \|\Delta\|_1 = \text{Tr}(\Delta) = \text{Tr}[\mathcal{E}(\Delta)] \leq \|\mathcal{E}(\Delta)\|_1 \leq \|\Delta\|_1, \]  
for \( \Delta \geq 0 \), hence \( \|\mathcal{E}(\Delta)\|_1 = \text{Tr}[\mathcal{E}(\Delta)] \). Since \( \|\Delta\|_1 = \text{Tr}(\Delta) \) if and only if \( \Delta \geq 0 \), we obtain that \( \mathcal{E}(\Delta) \geq 0 \).

There is a clear parallelism between this theorem and Theorem ?? for classical stochastic processes. As a result, quantum Markov processes are also characterized in the following way.

**Theorem 3.4.** A quantum evolution \( \{\mathcal{E}(t_2,t_1), t_2 \geq t_1 \geq t_0 \} \) is Markovian if and only if for all \( t_2 \) and \( t_1 \), \( t_2 \geq t_1 \),
\[ \left\| \left[ \mathcal{E}(t_2,t_1) \otimes 1 \right] \tilde{\Delta} \right\|_1 \leq \|\tilde{\Delta}\|_1, \]  
(43)
for any Hermitian operator \( \tilde{\Delta} \) acting on \( \mathcal{H} \otimes \mathcal{H} \).
Figure 2. Illustration of the quantum two-state discrimination problem. Under Markovian dynamics (blue line) the trace norm of the Helstrom matrix \( \Delta(t) = q\rho_1(t) - (1 - q)\rho_2(t) \) decreases monotonically from its initial value at \( t_0 \). On the contrary, for a non-Markovian dynamics (red line) there exist revivals at some time instants, say \( t_2 \), where the trace norm of \( \Delta(t) \) is larger than the previously attained values, for example at \( t_1 \). Thus, the memoryless property of a Markovian process implies that information about the initial state is progressively lost as time goes by. That is not the case for non-Markovian dynamics.

Proof. Since for a quantum Markovian process \( \mathcal{E}(t_2,t_1) \) is completely positive for any \( t_2 \geq t_1 \), the map \( \mathcal{E}(t_2,t_1) \otimes 1 \) is positive, and the results follows from Theorem ??.

Therefore, similarly to the classical case, a quantum Markovian process increases in a monotonic way the averaged probability \( P_{\min(\text{fail})} \), Eq. (??), to give the wrong answer in one-shot two-state discrimination problem. More concretely, consider a quantum system “S” which evolves from \( t_0 \) to the current time instant \( t_1 \), through some dynamical map \( \mathcal{E}(t,t_0) \). This system S was prepared at \( t_0 \) in the state \( \rho_1 \) with probability \( q \) and \( \rho_2 \) with probability \( (1 - q) \), and we aim at guessing which state was prepared by performing a measurement on S at the present time \( t_1 \). If the dynamics \( \mathcal{E}(t,t_0) \) is Markovian the best we can do is to measure at the present time \( t_1 \), however for non-Markovian processes it may be better to wait for some posterior time \( t_2 > t_1 \) where the trace norm of the Helstrom matrix \( \Delta(t) = \mathcal{E}(t,t_0)(\Delta) = q\rho_1(t) - (1 - q)\rho_2(t) \) increases with respect to its value at time \( t_1 \) (see illustration in figure 2).

Moreover, the same result applies if we make measurements including a static ancillary system “A”, in such a way that the evolution of the enlarged system “S + A” is given by \( \mathcal{E}(t_2,t_1) \otimes 1 \). That is not the case for a P-divisible dynamics where just the positivity of \( \mathcal{E}(t_2,t_1) \) is required instead of the complete positivity.

The fact that for a quantum non-Markovian process \( P_{\min(\text{fail})} \) can decrease for some time period may again be interpreted as a trait of the underlying memory in the process. The system seems to retain information about its initial state, that arises at some posterior time \( t_2 \).

3.3.2. Memoryless environment. A different way to visualize the memoryless properties characteristic of a quantum Markovian process is in the context of system-environment interactions. Since for a closed quantum system the evolution is given by some two-parameter family of unitary operators \( U(t_1,t_0) \), which fulfill \( U(t_2,t_0) = U(t_2,t_1)U(t_1,t_0) \) (see for instance [?]), the evolution of a closed quantum system is trivially Markovian. However, the situation changes regarding the time evolution of open quantum systems. Despite the most studied models to describe such a dynamics result in Markovian master equations of the form of Eq. (??) [??,??,??,??], it is well-known that the exact dynamics of an open quantum system is essentially non-Markovian [??,??,??,??].

To illustrate in what sense Markovian dynamics are memoryless in this context,
Figure 3. Schematic action of a memoryless environment as described by a collisional model. At every time $t_n$ the system interacts with the environmental state $\rho_E$ via some unitary operation $U$. At the following time step $t_{n+1}$, the system finds again the environment in the same state $\rho_E$, forgetting any correlation caused of the previous interaction at $t_n$. Provided the limit $\max_n |t_{n+1} - t_n| \to 0$ is well-defined, this process can be seen as a discrete version of quantum Markovian dynamics.

Consider the collisional model in the formulation proposed in [??, ??, ??], which is depicted in figure 3. In this model, the interaction between system and environment is made up of a sequence of individual collisions at times $t_1, t_2, \ldots, t_n$. Each collision produces a change in the state of the system $\rho_S$ given by

$$
\rho_S(t_{n+1}) = \text{Tr}_E[U(t_{n+1}, t_n)\rho_S(t_n) \otimes \rho_E U^\dagger(t_{n+1}, t_n)] = \mathcal{E}(t_{n+1}, t_n)[\rho_S(t_n)],
$$

(44)

where $\rho_E$ is the state of the environment assumed to be the same for every collision, and $U(t_{n+1}, t_n)$ is a unitary operator describing the system-environment interaction. Moreover, $\mathcal{E}(t_{n+1}, t_n)(\cdot) = \sum_{ij} K_{ij}(\cdot) K^\dagger_{ij}$ is a completely positive map whose Kraus operators are given by $K_{ij} = \sqrt{p_j^E} |\phi_j^E\rangle \langle \phi_E^j|$, for $\rho_E = \sum_j p_j^E |\phi_j^E\rangle \langle \phi_E^j|$. The successive concatenations of these collisions lead to a quantum Markovian process. Indeed, if we write

$$
\rho_S(t_{n+2}) = \mathcal{E}(t_{n+2}, t_n)[\rho_S(t_n)],
$$

(45)

as

$$
\rho_S(t_{n+2}) = \mathcal{E}(t_{n+2}, t_{n+1})[\rho_S(t_{n+1})] = \mathcal{E}(t_{n+2}, t_{n+1}) \mathcal{E}(t_{n+1}, t_n)[\rho_S(t_n)],
$$

(46)

we conclude that

$$
\mathcal{E}(t_{n+2}, t_n) = \mathcal{E}(t_{n+2}, t_{n+1}) \mathcal{E}(t_{n+1}, t_n),
$$

(47)

and since $\mathcal{E}$ are completely positive maps the process is Markovian. In addition, if the limit $\max_n |t_{n+1} - t_n| \to 0$ does exist, it is possible to obtain equations with the form of (??) for these models [??, ??].

Notably, any Markovian dynamics can be seen as a collisional model like this. This is a consequence of the following theorem [??, ??, ??].

**Theorem 3.5 (Stinespring).** A completely positive dynamics $\mathcal{E}(\rho_S)$ can be seen as the reduced dynamics of some unitary evolution acting on an extended state with the form $\rho_S \otimes \rho_E$, where $\rho_E$ is the same independently of $\rho_S$.

Thus, since for Markovian evolutions $\mathcal{E}(t_2, t_1)$ exists for all $t_2 \geq t_1$ and is completely positive, we may write it as

$$
\mathcal{E}(t_2, t_1)(\cdot) = \text{Tr}_E[U(t_2, t_1)(\cdot) \otimes \rho_E U^\dagger(t_2, t_1)],
$$

(48)

where $U(t_2, t_1)$ may depend on $t_2$ and $t_1$, but $\rho_E$ can be taken to be independent of time. Hence, Markovian evolutions may be thought to be made up of a sequence of memoryless
collisions, where the environmental state is the same and the total state of system and environment is uncorrelated in every collision as if there were no previous interaction (figure 3). Note that this does not mean that we must impose system and environment to be uncorrelated at any time to get Markovian evolutions \[\text{[?]}\]. Actually the total state of system and environment may be highly correlated even in a Markovian environment \[\text{[?]}\]. Rather, the conclusion is that the obtained evolution may also be thought as the result of memoryless system-environment infinitesimal collisions.

Interestingly, this kind of collisional models can be adapted to simulate non-Markovian dynamics by breaking the condition of uncorrelated collisions \[\text{[?]}\].

### 3.4. Comparison with other definitions of quantum Markovianity

Before moving on, it is worth to dedicate a brief section to comment on other definitions of quantum Markovianity that have been presented in the literature.

- **Semigroup Definition.** Historically, the absence of memory effects in quantum dynamics was commonly associated to the formulation of differential dynamical equations for \(\rho(t)\) with time-independent coefficients. In contrast, differential equations with time-dependent coefficients or integro-differential equations were linked to non-Markovian dynamics (see for instance \[\text{[?]}, \text{[?]}, \text{[?]}, \text{[?]}, \text{[?]}, \text{[?]}, \text{[?]}, \text{[?]}, \text{[?]}\]), and references therein). From this point of view, Markovian evolutions would be given only by quantum dynamical semigroups \[\text{[?]}\], i.e. families of trace preserving and completely positive maps, \(\mathcal{E}_\tau\), fulfilling the condition

\[
\mathcal{E}_\tau \mathcal{E}_\sigma = \mathcal{E}_{\tau+\sigma}, \quad \tau, \sigma \geq 0.
\]

(49)

It should be noted however, that this definition does not coincide with the definition adopted in this review and, in our view, suffers from certain drawbacks. The semigroup law Eq. \(\text{[??]}\) is just a particular case of the two-parameter composition law Eq. \(\text{[??]}\), which encompasses the case of time-inhomogeneous Markovian processes. In other words, this approach does not distinguish between Markovian and non-Markovian equations of motion with time-dependent coefficients. Moreover, this problem persists in the classical limit.

- **Algebraic Definition.** In the 1980s a rigorous definition of quantum stochastic process was introduced by using the algebraic formulation of quantum mechanics \[\text{[?], ?}\]. It is difficult to summarize in a few words those results, but we will try to sketch the main idea for those amongst the readership that are familiar with \(C^*\)-algebras. In this context, a quantum stochastic process on a \(C^*\)-algebra \(\mathcal{A}\) with values in a \(C^*\)-algebra \(\mathcal{B}\) is defined by a family \(\{j_t\}_{t \geq 0}\) of \(*\)-homomorphism \(j_t : \mathcal{B} \to \mathcal{A}\). To define a Markov property two ingredients are necessary. The first one is the following sub-algebra of \(\mathcal{A}\),

\[
\mathcal{A}_t = \vee\{j_s(b) : b \in \mathcal{B}, s \leq t\}
\]

(50)
which is called a *past filtration* or a *filtration* \([?]\). Here the symbol \(\lor S\) denotes the \(C^*-\)algebra generated by the subset \(S\) of \(A\). The second one is the introduction of the concept of conditional expectation on \(A\) \([?], \ldots\), which is a generalization of the usual conditional expectation, see Eq. \((??)\), to non-commutative algebras.

Mathematically, a *conditional expectation* of \(A\) on a sub-algebra \(A_0 \subset A\) is a linear map

\[
\mathbb{E}[\cdot|A_0] : A \to A_0,
\]

satisfying the properties:

(i) For \(a \in A\), \(\mathbb{E}[a|A_0] \geq 0\) whenever \(a \geq 0\).

(ii) \(\mathbb{E}[1|A_0] = 1\).

(iii) For \(a_0 \in A_0\) and \(a \in A\), \(\mathbb{E}[a_0a|A_0] = a_0\mathbb{E}[a|A_0]\).

(iv) For \(a \in A\), \(\mathbb{E}[a^*|A_0] = (\mathbb{E}[a|A_0])^*\).

(v) For \(a \in A_0\) and any state \(\varphi\) on \(A\), \(\varphi(a) = \varphi(\mathbb{E}[a|A_0])\).

Thus, the stochastic process \(\{j_t\}_{t \geq 0}\) is said to be Markovian if for all \(s, t \geq 0\) and all \(X \in A_0\) an analogous to Eq. \((??)\) is fulfilled,

\[
\mathbb{E}[j_{t+s}(X)|A_s] = \mathbb{E}[j_{t+s}(X)|j_s(A_0)].
\]

We will not go into further details here. What is important for our purposes is that, on the one hand, Accardi, Frigerio and Lewis proved in their seminal paper \([?]\) that this definition of Markovian process implies our Definition \((??)\) (rewritten in the Heisenberg picture). On the other hand, the opposite problem, namely to prove that any Markovian evolution according to Definition \((??)\) is also Markovian according to \((??)\) requires a technically complicated step known as the dilation problem (see \([?]\) and references therein). That is quite related to what was explained informally in section \(??\), but we do not enter into details. Fortunately, under well-chosen and reasonable conditions (boundedness of operators, fulfilment of Lipschitz conditions, etc.) \([?], \ldots\), it is possible to prove that Definition \((??)\) also implies \((??)\). Therefore, within the scope this paper, i.e. finite dimensional systems, we can consider the algebraic definition of Markovianity to be essentially equivalent to the one given here in terms of the divisibility condition.

• BLP Definition. Recently, Breuer, Laine and Piilo (BLP) proposed a definition of non-Markovian dynamics in terms of the behavior of the trace distance \([?], \ldots\). Concretely, they state that a quantum evolution, given by some dynamical map \(E_{(t,t_0)}\), is Markovian if the trace distance between any two initial states \(\rho_1\) and \(\rho_2\) decreases monotonically with time. This definition is a particular case of Definition \((??)\). As was explained in section \(??\), for any Markovian dynamics \(E_{(t,t_0)}\) and Hermitian operator \(\Delta\) acting on \(\mathcal{H} \otimes \mathcal{H}\), \(\| [E_{(t,t_0)} \otimes 1] \Delta \|_1\) monotonically decreases with time, and so does \(\|E_{(t,t_0)}(\Delta)\|_1\) for any Hermitian \(\Delta\) acting on \(\mathcal{H}\). Concretely, for \(\Delta = \frac{1}{2}(\rho_1 - \rho_2)\), which corresponds to the unbiased case in the two-state discrimination problem \(q = 1/2\), the property

\[
\|E_{(t_2,t_0)}(\Delta)\|_1 \leq \|E_{(t_1,t_0)}(\Delta)\|_1, \quad t_2 \geq t_1,
\]

\((53)\)
reduces to the BLP definition; this is, for all $\rho_1$ and $\rho_2$,

$$\|\rho_1(t_2) - \rho_2(t_2)\|_1 \leq \|\rho_1(t_1) - \rho_2(t_1)\|_1, \quad t_2 \geq t_1.$$  \hspace{1cm} (54)

However, the reverse implication fails to hold, i.e. not every dynamics fulfilling Eq. (54) satisfies Theorem ?? (e.g. [?, ?, ?]). Thus, we believe that it is more appropriate to consider the BLP definition as a particular case which arises in the study of memory properties in unbiased two-state discrimination problems. Note that the apparent lack memory in an unbiased case does not imply a general memoryless property; it only manifested in a general biased case (and taking into account possible ancillary systems). Nevertheless, from equation (54), it is possible to construct a very useful witness of non-Markovianity as we will see in section ??.

Remarkably, the previous different definitions of quantum Markovianity satisfy a hierarchical relation with our Definition ?? based on the divisibility condition. That is sketched in figure 4.

**Figure 4.** Relation between divisibility, semigroup, algebraic and BLP definitions of quantum Markovianity. The divisibility definition is essentially equivalent to the algebraic one (see main text). In addition, any Markovian dynamics according to the semigroup definition, it is also Markovian according to the divisibility definition, and hence Markovian according to BLP definition. However the converse implication does not hold.

- Markovianity in microscopic derivations. When deriving evolution equations for open quantum systems from microscopic models, the adjective “Markovian” is widely used to design master equations obtained under the so-called “Born-Markov” approximation. More concretely, if $\rho(t)$ is the state of the open quantum system, the Born approximation consists in shorten the perturbative expansion in the interaction Hamiltonian, $V = \sum_i A_i \otimes B_i$, at first non-trivial order. This leads to some differential equation of the form [?, ?, ?, ?, ?, ?]:

$$\frac{d\rho(t)}{dt} = -i[H_S, \rho(t)] + \sum_{i,j} \Omega_{ij}(t)\rho(t)A_j + A_j\rho(t)\Omega_{ij}^\dagger(t) - A_j\Omega_{ij}(t)\rho(t) - \rho(t)\Omega_{ij}^\dagger(t)A_j.$$  \hspace{1cm} (55)

where $H_S$ stands for the free Hamiltonian of the open system, and

$$\Omega_{ij}(t) = \int_0^t dsC_{ij}(s)e^{-iH_Ss}A_ie^{iH_Ss}.$$  \hspace{1cm} (56)

Here, $C_{ij}(s) = \text{Tr}(B_j e^{-iH_Bs} B_i e^{iH_Bs} \rho_B)$ are the correlation functions of the bath, which is in the state $\rho_B$ and has free Hamiltonian $H_B$. Eq. (56) is sometimes called Bloch-Redfield equation (e.g. [?]). Now, if the correlation functions of the bath $C_{ij}(s)$ are narrow in comparison to the typical time scale of $\rho(t)$ due to $V$,
the upper limit in the integral of $\Omega_{ij}(t)$ can be safely extended to infinity. This conforms what is sometimes called “Markov” approximation.

Two comments are pertinent regarding the connection of these dynamical equations with the Markovian processes as defined this work. Firstly, despite the “Born-Markov” approximation leads to master equations with time-independent coefficients, they do not always define a valid quantum dynamical semigroup [?, ?]. This is because they break complete positivity. Thus, these models should not be referred as “Markovian” in strict sense, as a Markovian processes must preserve the positivity of any state, or probability distribution in the classical limit. Secondly, if the “Born-Markov” approximation is combined with the secular approximation (i.e. neglecting fast oscillating terms in the evolution equation) a valid quantum dynamical semigroup is obtained [?, ?, ? , ?, ?], and then the dynamics can be certainly called Markovian. However, the fact that the “Born-Markov-secular” approximation generates Markovian dynamics, should not be understood as the only framework to obtain Markovian dynamics.

Despite this is representative for other definitions of quantum non-Markovianity different from Definition ??; the reader may also find proposals based on the behavior of multi-time correlation functions [?, ?, ?], initial-time-dependent generators [?, ?, ?], or properties of the asymptotic state [?]. See also [?, ?, ?, ?] for a definition of non-Markovianity in the context of stochastic Schrödinger equations.

4. Measures of Quantum non-Markovianity

After introducing the concept of quantum non-Markovianity in previous sections, we may ask about its quantification in terms of suitable measures and its detection in actual experiments. As we shall see, there have recently been several developments towards these goals, and we shall present them separately. Thus, the present section is devoted to the quantification problem whereas the detection of non-Markovian dynamics by witnesses is left to section ??.

In order to quantify non-Markovianity, the so-called measures of non-Markovianity are introduced. Basically, a measure of non-Markovianity is a function which assigns a number (positive or zero) to each dynamics, in such a way that the zero value is obtained if and only if the dynamics is Markovian. We will also use the name degree of non-Markovianity for a normalized measure of non-Markovianity, with values between 0 and 1, although other normalizations may eventually be taken.

4.1. Geometric measures

Consider a dynamical map $\mathcal{E}_{(t,t_0)}$ describing the evolution from some initial time $t_0$. A first attempt to formulate a measure of non-Markovianity may be a distance-based approach. Here the measure of non-Markovianity is expressed as a distance between $\mathcal{E}_{(t,t_0)}$ and its closest Markovian dynamics (see figure 5). More concretely, let $\mathfrak{M}$ denote
Figure 5. Illustration of the geometric measure of non-Markovianity. At each \( t \), \( \mathcal{N}^{\text{geo}}_t[\mathcal{E}(t,t_0)] \) measures the distance between the map \( \mathcal{E}(t,t_0) \) and the non-convex set of Markovian maps \( \mathcal{M} \). For a time interval \( t \in I \), \( \mathcal{D}_{\text{NM}(g)}^I \) in Eq. (57) is the maximum of every value of \( \mathcal{N}^{\text{geo}}_t[\mathcal{E}(t,t_0)] \) for \( t \in I \).

the set of all Markovian dynamics, and \( \mathcal{D}(\mathcal{E}_1, \mathcal{E}_2) \in [0,1] \) be some (normalized) distance measure in the space of dynamical maps. We define the geometric non-Markovianity at time \( t \) as

\[
\mathcal{N}^{\text{geo}}_t[\mathcal{E}(t,t_0)] := \min_{\mathcal{E}_M \in \mathcal{M}} \mathcal{D}[\mathcal{E}(t,t_0), \mathcal{E}_M(t,t_0)],
\]

which is zero if and only if \( \mathcal{E}(t,t_0) \) belongs to the set of Markovian dynamics \( \mathcal{M} \).

The geometric measure of non-Markovianity in some time interval \( I \) may be defined as the maximum value of the geometric non-Markovianity for \( t \in I \),

\[
\mathcal{D}_{\text{NM}(g)}^I := \max_{t \in I} \mathcal{N}^{\text{geo}}_t[\mathcal{E}(t,t_0)].
\]

This quantity lies between 0 and 1 and is positive if and only if the process is non-Markovian, therefore it is a degree of non-Markovianity.

Despite the conceptually clear meaning of \( \mathcal{D}_{\text{NM}(g)} \), it suffers from an important drawback, as this quantity is very hard to compute in practice because the involved optimization process. In fact, note that the set of Markovian maps \( \mathcal{M} \) is non-convex \([3]\), what makes the problem computationally intractable as the dimension of the system grows.

This approach was originally proposed by Wolf and collaborators \([3, 4]\) to quantify the non-Markovianity of what is sometimes called a “snapshot” of the dynamics. This is a map \( \mathcal{E}_{(t_1, t_0)} \) for fixed \( t_0 \) and \( t_1 \). Instead, in the case considered throughout this paper, a family of maps \( \{\mathcal{E}_{(t,t_0)}, t \geq t_0\} \) is assumed to be known. These authors put forward the aforementioned practical problems of the geometric measure and introduced an alternative measure (see also \([5]\)). Later on, the problem to decide whether some map is a “snapshot” of a Markovian dynamics was shown to be very hard in general \([3, 5]\), however it has been analyzed for small size systems in \([5]\) and \([6]\). For a review about quantum channels with memory see \([7]\).

4.2. Optimization of the Helstrom matrix norm

Another approach to quantify non-Markovianity is based on the result of Theorem 35. Recall that the trace norm of a Helstrom matrix \( \Delta = q\rho_1 + (1-q)\rho_2 \) is a measure of the capability to distinguish between the states \( \rho_1 \) and \( \rho_2 \) given the outcome of some POVM, see Theorem 35. Thus, if the dynamics is such that for some \( t \) and \( \epsilon > 0 \), \( ||\Delta(t)||_1 < ||\Delta(t + \epsilon)||_1 \), the probability to distinguish whether the system was in state \( \rho_1 \) or \( \rho_2 \) at time \( t_0 \), is higher at \( t + \epsilon \) than it was at a time \( t \). As commented in section 35, this phenomenon denotes the existence of memory effects in the dynamics, as an increase
of information at time $t + \epsilon$ with respect to $t$ suggests that the system is “remembering” its original state at $t + \epsilon$. In fact, the intuitive understanding of the word “memory” demands that a memoryless process does not have the property to keep information, and that this always decreases with time.

This revival of information at $t + \epsilon$ may be understood as a positive flow of information from the environment to the system. Thus, for purely Markovian dynamics the flow of information goes always from the system to the environment. However, as pointed out in [??, ??], this interpretation in terms of information flowing between system and environment may be problematic if taken strictly, because it is possible to obtain quantum non-Markovian dynamics with the form

$$E(t,t_0)(\rho) = \sum_i p_i U_i(t,t_0) \rho U_i^\dagger(t,t_0).$$

This type of evolution can be generated simply by applying randomly the unitary evolutions $U_i(t,t_0)$ in accordance to the probabilities $p_i$. It is a fact that these probabilities can be generated independently of the dynamics of $\rho$ by some random (or pseudo-random) number generator.

On the other hand, since we cannot discard the presence of a decoupled and inert ancillary space “A”, we generally consider an enlarged Helstrom matrix $\tilde{\Delta} = q \rho_{1A} - (1 - q) \rho_{2A}$, where $\Delta = \text{Tr}_A(\tilde{\Delta})$, and $\|\tilde{\Delta}(t)\|_1 = \|[E(t,t_0) \otimes 1] \tilde{\Delta}(t_0)\|_1$. Thus, as an increment of information as accounted by $\|\tilde{\Delta}\|_1$ denotes non-Markovianity in the dynamics, we can take the maximum of information gained to assess how non-Markovian the evolution is. More concretely, we may write

$$\tilde{\sigma}(\tilde{\Delta},t) := \frac{d\|\tilde{\Delta}(t)\|_1}{dt} := \lim_{\epsilon \rightarrow 0^+} \frac{\|\tilde{\Delta}(t+\epsilon)\|_1 - \|\tilde{\Delta}(t)\|_1}{\epsilon},$$

by adding up every increment of information in some interval $I$:

$$\int_{t \in I, \tilde{\sigma} > 0} dt \tilde{\sigma}(\tilde{\Delta},t).$$

Then by maximizing over the initial Helstrom matrix $\tilde{\Delta}$ (i.e. maximizing over $\rho_{1A}, \rho_{2A}$ and the bias $q$) we define

$$\mathcal{N}^I_{\text{H}} := \max_{\Delta} \int_{t \in I, \tilde{\sigma} > 0} dt \tilde{\sigma}(\tilde{\Delta},t),$$

where the subindex “H” stands for Helstrom, as a measure of non-Markovianity. In virtue of Theorem ??, $\mathcal{N}^I_{\text{H}} = 0$ if and only if the process is Markovian in the interval $I$. The quantity $\mathcal{N}^I_{\text{H}}$ can be normalized via exponential or rational functions, for instance $\mathcal{D}^I_{\text{NM(exp-H)}} := 1 - e^{-\mathcal{N}^I_{\text{H}}}$ or $\mathcal{D}^I_{\text{NM(rat-H)}} := \mathcal{N}^I_{\text{H}}/(1 + \mathcal{N}^I_{\text{H}})$.

This proposal was first suggested in [??]. For the unbiased case $q = 1/2$ and without taking into account the possible presence of ancillary systems was previously formulated in [??]. As in the case of geometric measures, the main drawback of the quantity $\mathcal{N}^I_{\text{H}}$ is the difficult optimization process, which makes this measure rather impractical. For the restricted case of [??] there has been some progress along this line [??], see also section ?? and references therein.
4.3. Choi-Jamiolkowski based measure

As pointed out, despite that the two previous measures enjoy several nice geometric or informational interpretations, they are very difficult to compute in practice. A computationally simpler measure of non-Markovianity was introduced in \cite{?}. Given a family \( \{ \mathcal{E}(t,t_0), t \geq t_0 \} \), the basic idea of this measure is to quantify how much non-completely positive the intermediate dynamics \( \{ \mathcal{E}(t,t_1), t \geq t_1 \geq t_0 \} \) is for every time \( t_1 \). To obtain these partitions, by time-continuity we have

\[
\mathcal{E}(t,t_0) = \mathcal{E}(t,t_1) \mathcal{E}(t_1,t_0),
\]

after right-multiplication with the inverse of \( \mathcal{E}(t_1,t_0) \) on both sides we obtain the desired partitions

\[
\mathcal{E}(t,t_1) = \mathcal{E}(t,t_0) \mathcal{E}^{-1}(t_1,t_0).
\]

If these maps are completely positive (CP) for all \( t_1 \), the time evolution is Markovian (Definition ??). At the moment, we shall assume that \( \mathcal{E}^{-1}(t_1,t_0) \) does exist (we will come back to this point later on), so \( \mathcal{E}(t,t_1) \) is well-defined by Eq. (??). For non-Markovian dynamics there must be some \( t_1 \), such that \( \mathcal{E}(t,t_1) \) is not completely positive. Therefore by measuring how much the intermediate dynamics \( \{ \mathcal{E}(t,t_1), t \geq t_1 \geq t_0 \} \) depart from completely positive maps, we are measuring up to what extent the time evolution is non-Markovian. Note that \( \mathcal{E}(t,t_1) \) is always trace-preserving as it is a composition of two trace-preserving maps.

In order to quantify the degree of non-complete positiveness of the maps \( \{ \mathcal{E}(t,t_1), t \geq t_1 \geq t_0 \} \), we resort to the Choi-Jamiolkowski isomorphism \cite{?,?}. Consider the maximally entangled state between two copies of our system \( |\Phi\rangle = \frac{1}{\sqrt{d}} \sum_{n=0}^{d-1} |n\rangle |n\rangle \) (assumed here with finite dimension \( d \)), we associate the map \( \mathcal{E}(t,t_1) \) to a (Choi-Jamiolkowski) matrix constructed by the rule

\[
[\mathcal{E}(t,t_1) \otimes 1] (|\Phi\rangle \langle \Phi|).
\]

The Choi’s theorem asserts that \( \mathcal{E}(t,t_1) \) is completely positive if and only if the matrix Eq. (??) is positive semidefinite. In addition, note that since \( \mathcal{E}(t,t_1) \) is trace-preserving, the trace norm of matrix (??) provides a measure of the non-completely positive character of \( \mathcal{E}(t,t_1) \). More concretely,

\[
\| [\mathcal{E}(t,t_1) \otimes 1] (|\Phi\rangle \langle \Phi|) \|_1 = 1 \quad \text{iff} \quad \mathcal{E}(t,t_1) \text{ is CP},
\]

\[
> 1 \quad \text{otherwise.}
\]

Following \cite{?}, we define a function \( g(t) \) via the right derivative of the trace norm as

\[
g(t) := \lim_{\epsilon \to 0^+} \| [\mathcal{E}(t+t,\epsilon) \otimes 1] (|\Phi\rangle \langle \Phi|) \|_1 - \frac{1}{\epsilon},
\]

so that \( g(t) > 0 \) for some \( t \) if and only if the evolution is non-Markovian. Therefore the total amount of non-Markovianity in an interval \( t \in I \) will be given by

\[
N_{CJ}^I := \int_I g(t) dt,
\]
where “CJ” stands for Choi-Jamiołkowski.

The quantity $N^I_{CJ}$ may be normalized via exponential or rational methods, for instance $D^I_{NM(exp-CJ)} := 1 - e^{-N^I_{CJ}}$ or $D^I_{NM(rat-CJ)} := N^I_{CJ}/(1 + N^I_{CJ})$. However these normalizations turn out to be not very discriminative. Another way to obtain a more useful normalized measure was proposed in [?], we explain here a slightly modified method in terms of $g(t)$.

Define the function

$$\bar{g}(t) := \tanh[g(t)],$$

(69)

where $g(t)$ is given by Eq. (??). Therefore $1 \geq \bar{g}(t) \geq 0$ with $\bar{g}(t) = 0$ for all $t$ if and only if the evolution is Markovian. Then, for a bounded interval $t \in I$ we define its normalized degree of non-Markovianity as

$$D^I_{NM(CJ)} := \frac{\int_I \bar{g}(t)dt}{\int_I \chi[\bar{g}(t)]dt}, \quad \text{(with “0/0 = 0” by convention),}$$

(70)

where the indicator function $\chi(x)$ is defined as

$$\chi(x) := \begin{cases} 0 & \text{if } x = 0, \\ 1 & \text{otherwise.} \end{cases}$$

Thus, the degree Eq. (??) is basically the non-Markovianity accumulated for each $t \in I$, divided by the total length of the subintervals of $I$ where the dynamics is non-Markovian. It is easy to proof that $D^I_{NM(CJ)}$ is normalized. Let $I_n \subset I$ be the collection of subintervals such that $\bar{g}(t) > 0$ for $t \in I_n$. If $|I_n|$ denotes the length of the subinterval $I_n$, we have

$$D^I_{NM(CJ)} = \frac{\int_I \bar{g}(t)dt}{\int_I \chi[\bar{g}(t)]dt} = \frac{\sum_n \int_{I_n} \bar{g}(t)dt}{\sum_n \int_{I_n} \chi[\bar{g}(t)]dt} = \frac{\sum_n \int_{I_n} \bar{g}(t)dt}{\sum_n |I_n|} \leq \frac{\sum_n |I_n|}{\sum_n |I_n|} = 1,$$

(72)

because of the bound $\bar{g}(t) \leq 1$.

It is worth mentioning several points to have in mind when evaluating this measure of non-Markovianity.

(i) Note that, in general [see an exception in the point (iii) below], we need to know the complete dynamical map $\{\mathcal{E}(t_{\ell_0}), t \geq t_0\}$ to compute the function $g(t)$. The standard way to obtain it is resorting to process tomography. Thus, one considers the evolution for different final times $t$ of a complete set of states which span the space of dynamical maps. Then the dynamical map is reconstructed by tomography of the evolved final states (see for instance [?]). This is the only experimental way to proceed. However if we know the theoretical evolution, for example by mean of some model, there is another trick which is worth explaining. In that case, we may consider directly the evolution of the basis $\{|i\rangle\langle j|\}$ for different final times $t$. We write the resulting matrix as a (column) vector by stacking the columns on top of one another. This process is sometimes called vectorization and denoted by
“vec” [?, ?]. As a result, the dynamical map $E_{(t,t_0)}$ can be seen as a matrix $E_{(t,t_0)}$ acting on states written as (column) vectors, and moreover in the basis of $\{|i\rangle\langle j|\}$ such a matrix is given by

$$E_{(t,t_0)} = [v_{11}(t), \ldots, v_{1N}(t), v_{21}(t), \ldots, v_{2N}(t), \ldots, v_{NN}(t)]$$  \hspace{1cm} (73)

where the column vectors are $v_{ij}(t) = \text{vec}[|i\rangle\langle j|](t)$. These are the vectorization of the matrix $|i\rangle\langle j|(t)$, which denotes the matrix obtained by evolving the basis element $|i\rangle\langle j|$ from $t_0$ to $t$.

Once $E_{(t,t_0)}$ is known for some interval $t \in I$, we can compute the intermediate dynamics in $I$ accordingly to Eq. (??), $E_{(t,t_1)} = E_{(t,t_0)} E_{(t_1,t_0)}^{-1}$, where $E_{(t_1,t_0)}^{-1}$ is just the standard matrix inverse. Finally, $g(t)$ can be computed in the following way: first, construct the matrix $U_{2+3}[E_{(t_1,t_0)} \otimes I]U_{2+3}$ where $U_{2+3}$ is the commutation matrix between the “second” and the “third” subspace [?]; second, apply $U_{2+3}[E_{(t_1,t_0)} \otimes I]U_{2+3}$ on $\text{vec}(\Phi\langle\Phi|)$; third, write the result as a matrix, i.e. “devectorize”; forth, compute the trace norm of that matrix; and finally, evaluate the right limit of Eq. (??).

(ii) It may happen that for some $t_1$ the map $E_{(t_1,t_0)}$ is not bijective, so that the intermediate map $E_{(t_1,t_0)}$ given by (??) is ill-defined. There are several ways to deal with this problem. If the singularity in $t_1$ is isolated, and we know the dynamics in some neighborhood of $t_1$, one can evaluate the function $g(t)$ in this neighboring region of $t_1$. By taking the limit $t \rightarrow t_1$ we usually obtain a divergence, $\lim_{t \rightarrow t_1} g(t) \rightarrow \infty$. However, since the hyperbolic tangent removes the divergence $\lim_{t \rightarrow t_1} \tanh(t) = 1$, we can compute $D_{\text{NM(3)}t}$ without further problems. Another way to remove the singularity may be to compute indirectly the inverse of $E_{(t_1,t_0)}$ by finding the inverse of $I + E_{(t_1,t_0)}$, which always exist. Then, at the end of the computation of $\tilde{g}(t)$, we proceed by taking the limit $\epsilon \rightarrow 0$. Another more sophisticated (and in sometimes inequivalent) method has been proposed by using the Moore-Penrose pseudoinverse [?]. See also [?] for other considerations about singularities in dynamical maps.

(iii) There are cases where we know the dynamics fulfills some linear differential equation,

$$\frac{d\rho(t)}{dt} = \mathcal{L}_t \rho(t) = -i[H(t), \rho(t)]$$  \hspace{1cm} (74)

$$+ \sum_k \gamma_k(t) \left[ V_k(t) \rho(t) V_k^\dagger(t) - \frac{1}{2} \{V_k^\dagger(t)V_k(t), \rho(t)\} \right],$$

where the decay rates may be negative $\gamma_k(t) < 0$ for some $t$ and so describes non-Markovian evolutions. Then, there is a very practical way to obtain the function $g(t)$. Since for small enough $\epsilon$ we have [?]

$$E_{(t+\epsilon,t)} = \mathcal{T} \exp \left( \int_t^{t+\epsilon} \mathcal{L}_s ds \right) \simeq \exp(\mathcal{L}_t \epsilon) \simeq 1 + \epsilon \mathcal{L}_t,$$  \hspace{1cm} (75)
the function $g(t)$ can be computed directly from the generator $\mathcal{L}_t$:

$$g(t) = \lim_{\epsilon \to 0^+} \frac{\| [1 + \epsilon (\mathcal{L}_t \otimes 1)] |\Phi\rangle \langle \Phi|_1 - 1}{\epsilon}. \quad (76)$$

(iv) It is possible to extend the definition Eq. (76) to unbounded intervals, typically $I = [t_0, \infty)$. However, this extension must be carefully handled. It can be understood as a limiting procedure of bounded intervals $I_n$, such that $\lim_{n \to \infty} I_n = [t_0, \infty)$, for example $I_n = [t_0, n)$. Very crucially, this limit has to be taken at the last step in the computation:

$$\mathcal{D}_{\text{NM(CJ)}}^{[t_0, \infty)} := \lim_{n \to \infty} \mathcal{D}_{\text{NM(CJ)}}^{I_n}, \quad \text{with} \quad \lim_{n \to \infty} I_n = [t_0, \infty). \quad (77)$$

As aforementioned, the measure of non-Markovianity $\mathcal{I}_{\text{CJ}}$, Eq. (77), was first introduced in [?]. The more discriminative degree Eq. (77) is a variant based on the same ideas as [?], where the normalization problem was further analyzed. Examples where this measure is studied can be found in [?], [?], [?], [?], [?], [?], [?], [?] for qubits coupled to bosonic environments, and [?] for more general spin systems coupled to bosonic environments, in [?] for qubits coupled to other $d$-level systems, in [?] for qubits interacting with composite environments, in [?], [?], [?] for classical stochastic dynamics, and in [?] for the so-called semi-Markov quantum processes. In addition, the application of the Choi-Jamiołkowski criterion to study the complete positivity of intermediate dynamics for some specific examples is considered in [?] as well.

### 4.4. Other measures

#### 4.4.1. Negative decay rates

Since a Markovian dynamics is characterized by generators with the form of Eq. (76), in [?] Andersson, Cresser and Hall proposed a measure of non-Markovianity focused on properties of the generator. Let us consider some dynamical evolution given by its generator,

$$\frac{d\rho(t)}{dt} = \mathcal{L}_t \rho(t) = -i[H(t), \rho(t)] \quad (78)$$

In order to characterize its non-Markovianity, we may write $\mathcal{L}_t$ in an orthonormal basis $\{G_j\}_{j=0}^{d-1}$ with respect to the Hilbert-Schmidt product $\text{Tr}(G_m^* G_n) = \delta_{mn}$. More specifically, in [?] it is proposed to use a self-adjoint basis with $G_0 = 1/\sqrt{d}$, so that $\{G_j\}_{j=1}^{d-1}$ can be taken to be the (normalized) generators of the $\mathfrak{su}(d)$ algebra. Thus, by expanding every operator of the dissipative part of the generator,

$$V_k(t) = \sum_m v_{km}(t) G_m, \quad v_{km}(t) = \text{Tr}[G_m V_k(t)], \quad (79)$$

$$W_k(t) = \sum_n w_{kn}(t) G_n, \quad w_{kn}(t) = \text{Tr}[G_m W_k(t)]. \quad (80)$$
Introducing this in Eq. (82), one obtains
\[
  \mathcal{L}_t [\rho(t)] = -i[H(t), \rho(t)] + \sum_{m,n} \tilde{c}_{mn}(t) \left[ G_m \rho(t) G_n - \frac{1}{2} \{ G_n G_m, \rho(t) \} \right],
\]
where \( \tilde{c}_{mn}(t) = \sum_{k,l} w_{kn}^*(t) c_{kl}(t) v_{lm}(t) \) forms a Hermitian matrix, \( \tilde{c}_{mn}(t) = \tilde{c}_{nm}^*(t) \), because \( \mathcal{L}_t \) preserves the Hermiticity of \( \rho \). Therefore, this matrix is diagonalized via some unitary operation, \( \tilde{c}_{mn}(t) = \sum_j u_{mj}(t) \gamma_j(t) u^*_{nj}(t) \) and the generator can be rewritten in the form
\[
  \mathcal{L}_t [\rho(t)] = -i[H(t), \rho(t)] + \sum_{j=0}^{d-1} \gamma_j(t) \left[ L_j(t) \rho(t) L_j^\dagger(t) - \frac{1}{2} (L_j^\dagger(t) L_j(t), \rho(t)) \right]
\]
with \( L_j(t) = \sum_m u_{mj}(t) G_m \), keeping orthonormality \( \text{Tr}[L_i^\dagger(t) L_j(t)] = \delta_{ij} \). Note that since the eigenvalues \( \gamma_j(t) \) are independent of the basis, this form is unique (up to degeneracy). Now, Andersson, Cresser and Hall define some functions of the eigenvalues (decay rates) \( \gamma_j(t) \):
\[
  f_j(t) := \min\{ \gamma_j(t), 0 \}.
\]
Because of Theorem 2, every \( f_j(t) \) vanishes at any time if and only if the evolution is Markovian. Therefore the functions \( f_j(t) \) can be used to construct a measure of non-Markovianity. For example, defining \( f(t) = \sum_{j=0}^{d-1} f_j(t) \), for a (bounded) time interval \( I \),
\[
  \mathcal{N}_I^f := \int_I f(t) dt.
\]
Interestingly this approach also suggest a discrete measure, by computing \( F_j^I = \int_I f_j(t) dt \), a non-Markovianity index can defined by the rule
\[
  \mathcal{N}_I^{\text{index}} := \sum_{j=0}^{d-1} \chi(F_j^I),
\]
i.e. the number of non-zero \( F_j^I \)'s in the interval \( I \).

An example where these measures were applied can be found in [1]. See also [2] for an experimental proposal to probe non-Markovianity by negative decay rates.

4.4.2. Hierarchical \( k \)-divisibility degrees Recently, Chruściński and Maniscalco have proposed a hierarchical way to assess non-Markovianity [2]. Their approach, based on the concept of \( k \)-divisibility, is interesting as it provides a way to define some kind of maximally non-Markovian dynamics. Basically, a family of dynamical maps, \( \{ \mathcal{E}_{(t_2,t_1)} \}, t_2 \geq t_1 \geq t_0 \), is \( k \)-divisible, if \( \mathcal{E}_{(t_2,t_1)} \otimes \mathbb{1}_k \) is a positive operator for all \( t_2 \geq t_1 \geq t_0 \) (here \( \mathbb{1}_k \) denotes the identity map acting on the space of \( k \times k \) matrices). Therefore, if the dimension of the quantum system is \( d \), a \( k \)-divisible process with \( k \geq d \), is what in this work has been called divisible or Markovian process (see Definition 2). The 1-divisible processes are the P-divisible processes as introduced in Definition 2, and
the 0-divisible processes are processes where $E_{(t_2,t_1)}$ is not a positive operator for some $t_1$ and $t_2 \geq t_1$.

Moreover, analogously to Theorems 23 and 24, we have that a process is $k$-divisible if and only if $\tilde{\sigma}_k(\tilde{\Delta},t) := \frac{d}{dt} \left\| E_{(t,t_0)} \otimes 1_k \tilde{\Delta} \right\|_1 \leq 0$ for every Helstrom matrix $\tilde{\Delta} = q \rho_{1A} - (1 - q) \rho_{2A}$ with an ancillary space of dimension $k$. In similar fashion to Eq. (67), Chruściński and Maniscalco define a set of degrees to quantify departure from $k$-divisibility for $t \in I$,

$$D^I_{ND(k)} := \sup_{\Delta} \frac{N^+_k(\tilde{\Delta},I)}{|N^-_k(\tilde{\Delta},I)|},$$

where $N^+_k(\tilde{\Delta},t) := \int_{t \in I, \tilde{\sigma} \succ 0} dt \tilde{\sigma}(\Delta,t)$, and the subindex “ND” stands for non-divisibility.

Since $E_{(t,t_0)}$ is completely positive for any final time $t$, it is easy to prove that $N^+_k(\tilde{\Delta},I) \leq |N^-_k(\tilde{\Delta},I)|$ [67], therefore $D^I_{ND(k)} \in [0, 1]$ for all $k$. Moreover, as $k$ increases, so does the dimension of the space in the optimization problem Eq. (68), and hence it is clear that

$$0 \leq D^I_{ND(1)} \leq \ldots \leq D^I_{ND(d)} \leq 1.$$  

In this equation, only $D^I_{ND(d)}$ is a degree of non-Markovianity as defined in this work. The other quantities are zero for non-Markovian but $k$-divisible ($d > k$) dynamics. This hierarchy of degrees of non-divisibility is particularly useful to try a definition of maximally non-Markovian dynamics. Indeed, Chruściński and Maniscalco propose to call “maximally non-Markovian dynamics” to those that $D^I_{ND(1)} = 1$, and consequently $D^I_{ND(2)} = D^I_{ND(3)} = \ldots = D^I_{ND(d)} = 1$. A particular example of this kind dynamics for one qubit is the one generated by the master equation $\frac{d}{dt} \rho = \gamma(t)(\sigma_z \rho_{\sigma_z} - \rho)$, for an interval $I$ such that $\gamma(t)$ is periodic in $I$. For instance, $\gamma(t) = \sin(t)$ or $\gamma(t) = \tan(t)$ in $t \in [0, 2\pi]$. Notably, for these two examples the Choi-Jamiolkowski measure Eq. (68) provides different values. We obtain $D^I_{NM(CJ)} = 0.758$ and $D^I_{NM(CJ)} = 0.803$, for $\gamma(t) = \sin(t)$ and $\gamma(t) = \tan(t)$ respectively.

5. Witnesses of Quantum non-Markovianity

In this section we revise the different ways to detect non-Markovianity via witnesses. A witness of non-Markovianity is a quantity that vanishes for all Markovian dynamics (see also [? , ?, ?]), but it may also vanish for some non-Markovian dynamics. Thus, when a witness of non-Markovianity gives a non-zero value, we are sure that the dynamics is non-Markovian.

In general, we can classify the witnesses of non-Markovianity that have been presented in the literature according to two guiding principles. There are witnesses based on monotonic quantities under completely positive maps, and based on monotonic quantities under local completely positive maps. In the following we review several proposals in these two classes.
5.1. Witnesses based on monotonicity under completely positive maps

5.1.1. Trace distance

If we consider the unbiased situation in the two-state discrimination problem, \( q = 1/2 \) and the Helstrom matrix reads as \( \Delta = (\rho_1 - \rho_2)/2 \), where we have neglected the possible presence of ancillary systems. Thus, the trace norm of \( \Delta \) becomes the trace distance between states \( \rho_1 \) and \( \rho_2 \),

\[
D(\rho_1, \rho_2) := \frac{1}{2} \|\rho_1 - \rho_2\|_1.
\]

(88)

Analogously to Eqs. (??) and (??), we write,

\[
\sigma(\rho_1, \rho_2, t) := \frac{dD[\rho_1(t), \rho_2(t)]}{dt} := \lim_{\epsilon \to 0^+} \frac{D[\rho_1(t + \epsilon), \rho_2(t + \epsilon)] - D[\rho_1(t), \rho_2(t)]}{\epsilon},
\]

(89)

and for some interval \( I \),

\[
\int_{t \in (t_1, t_2), \sigma > 0} dt \sigma(\rho_1, \rho_2, t).
\]

(90)

If this quantity is not zero for some pair of states \( \rho_1 \) and \( \rho_2 \), we are sure the dynamics is non-Markovian in \( I \), as it is a lower bound the non-Markovianity measure \( N_{HI}^I \), Eq. (??). Particularly, we may be interested in finding the largest value of Eq. (??) in the time interval \( (0, \infty) \). To this end, Breuer, Laine and Piilo [?, ?] define the witness

\[
N := \max_{\rho_1, \rho_2} \int_{\sigma > 0} dt \sigma(\rho_1, \rho_2, t).
\]

(91)

The use of the trace distance to witness non-Markovianity was originally proposed in [?] and further analyzed in [?]. Due to its simplicity, it has been applied to detect non-Markovian features in dynamics of qubit [?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?] and qutrit systems [?] coupled to bosonic environments, qubits coupled to other finite dimensional systems [?, ?, ?], and to composite [?, ?, ?] and chaotic [?, ?, ?] environments. It has also been employed to analyze memory-kernel master equations [?, ?], quantum semi-Markov process [?], classical noise [?, ?, ?, ?, ?, ?], in fermionic systems [?] and to study exciton-phonon dynamics in energy transfer of photosynthetic complexes [?, ?]. Moreover this witness has been implemented experimentally within a linear optics set up [?, ?, ?]. On the other hand, some connections have been found between the non-monotonic behavior of the trace distance and geometric phases [?], Loschmidt echo [?], dynamical recovering of the quantum coherence by applying local operations [?], and the appearance of system-environment correlations [?]. In this regard, this witness has also been proposed as a tool to detect the presence of initial system-environment correlations [?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?].

While very efficient under certain conditions, the trace distance cannot witness non-Markovian processes where the non-Markovianity is encoded just in the “non-unital part” of the dynamics [?]. This part corresponds to the affine vector \( c_{(t,t_0)} \) when the dynamics is visualized in the Bloch space, see Eq. (??).
5.1.2. Fidelity

The fidelity $F(\rho_1, \rho_2)$ between two quantum states $\rho_1$ and $\rho_2$ is a generalization of the transition of probability $|\langle \psi_1 | \psi_2 \rangle|^2$ between two pure states $|\psi_1\rangle$ and $|\psi_2\rangle$, to density matrices. Specifically, the fidelity is defined \cite{uhlmann} by the equation

$$F(\rho_1, \rho_2) := \max_{|\Psi_1\rangle, |\Psi_2\rangle} |\langle \Psi_1 | \Psi_2 \rangle|^2. \quad (92)$$

Here, $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are two purifications of $\rho_1 = \text{Tr}_A(|\Psi_1\rangle\langle \Psi_1|)$ and $\rho_2 = \text{Tr}_A(|\Psi_2\rangle\langle \Psi_2|)$, where $\text{Tr}_A$ denotes the partial trace on some ancillary system $A$, and the maximum is taken over the all possible purifications (see \cite{uhlmann} for a pedagogical introduction \cite{uhlmann}). Uhlmann \cite{uhlmann} solved the optimization problem obtaining

$$F(\rho_1, \rho_2) = \left| \frac{\sqrt{\rho_1}}{\sqrt{\rho_2}} \right|^2 = \left[ \text{Tr} \left( \sqrt{\sqrt{\rho_1} \rho_2 \sqrt{\rho_1}} \right) \right]^2. \quad (93)$$

Among several properties (see \cite{uhlmann}), the fidelity is monotonic under complete positive maps $\mathcal{E}$,

$$F(\mathcal{E}(\rho_1), \mathcal{E}(\rho_2)) \geq F(\rho_1, \rho_2), \quad (94)$$

reaching the equality if and only if the completely positive map is unitary $\mathcal{E}(\cdot) = U(\cdot)U^\dagger$ \cite{uhlmann}. Thus, the fidelity is monotonically increasing for Markovian evolutions, and therefore it may be used to witness non-Markovian behavior. Concretely, in \cite{uhlmann} the approach of \cite{uhlmann,uhlmann}, originally proposed for the trace distance (see previous section), is reconsidered with the so-called Bures distance \cite{bures}:

$$D_B(\rho_1, \rho_2) := \sqrt{2 \left[ 1 - \sqrt{F(\rho_1, \rho_2)} \right]}. \quad (95)$$

Since the authors of \cite{uhlmann} aim at quantifying non-Markovianity in Gaussian states of harmonic oscillators, the use of the fidelity instead the trace distance is more convenient because a closed formula for the latter for Gaussian states is still lacking. See also \cite{uhlmann,uhlmann,uhlmann}.

A different witness in terms of fidelity was previously proposed in \cite{uhlmann}, however that was only able to detect deviations from time-homogeneous Markov processes, i.e. quantum dynamical semigroups where $\mathcal{E}_{(t_2,t_1)} = \mathcal{E}_{(t_2-t_1)}$ for every $t_1$ and $t_2$. See \cite{uhlmann} for another work focused in just in time-homogeneous dynamics.

5.1.3. Quantum relative entropies

Another similar witness is constructed with the (von Neumann) relative entropy between two quantum states $\rho_1$ and $\rho_2$,

$$S(\rho_1 \| \rho_2) := \text{Tr}(\rho_1 \log \rho_1) - \text{Tr}(\rho_1 \log \rho_2). \quad (96)$$

Despite the relative entropy neither being symmetric $S(\rho_1 \| \rho_2) \neq S(\rho_2 \| \rho_1)$, nor satisfying the triangle inequality, it is often intuited as a distance measure because $S(\rho_1 \| \rho_2) \geq 0$, vanishing if and only if $\rho_1 = \rho_2$ (Klein’s inequality \cite{uhlmann}). Moreover, if the intersection of the kernel of $\rho_2$ with the support of $\rho_1$ is non-trivial, then $S(\rho_1 \| \rho_2)$ becomes infinity.
Analogously to the Bures and the trace distance, the quantum relative entropy is monotonic under completely positive and trace preserving maps $\mathcal{E}$,

$$S[\mathcal{E}(\rho_1)\|\mathcal{E}(\rho_2)] \leq S(\rho_1\|\rho_2). \tag{97}$$

The proof of this result was first given by Lindblad [?] for finite dimensional systems, and Uhlmann [?] extend it to the general case (see also [?] and [?]). Therefore the quantum relative entropy between any two states is monotonically decreasing with time in a Markovian process, and any increment of it at some time instant reveals the non-Markovian character of the dynamics.

The use of the quantum relative entropy to witness non-Markovianity was originally proposed in [?]. In [?] it is suggested to use more general relative entropies due to Renyi [?] and Tsallis [?] for the same task. In this regard, [?] enumerates several distances fulfilling the monotonicity condition. Additionally [?] proposed to use the monotonicity of the relative entropy to look for initial system-environment correlations.

5.1.4. Quantum Fisher information

Following [?,?] (see also [?]) the quantum Fisher information can be defined as the infinitesimal Bures distance (??) between two quantum states. For simplicity, assume some one-parametric family of quantum states $\rho_\theta$, then

$$D^2_B(\rho_\theta, \rho_{\theta+\delta\theta}) = \frac{1}{4} \mathcal{F}(\rho_\theta)(\delta\theta)^2 + \mathcal{O}((\delta\theta)^3), \tag{98}$$

where $\mathcal{F}(\rho_\theta)$ is the so-called quantum Fisher information of the family $\rho_\theta$. Equivalently, we write

$$\mathcal{F}(\rho_\theta) := 4 \lim_{\delta\theta \to 0} \left[ \frac{D_B(\rho_\theta, \rho_{\theta+\delta\theta})}{\delta\theta} \right]^2. \tag{99}$$

Thus, the quantum Fisher information of $\rho_\theta$ measures the sensitivity of the Bures distance when $\theta$ is varied. In turn, this can be interpreted as the information about $\theta$ which is contained in the family $\rho_\theta$, in such a way that if $\rho_\theta$ does not depend on $\theta$, $\mathcal{F}(\rho_\theta) = 0$. We will come back to this point later.

Additionally, the quantum Fisher information admits other different but equivalent definitions [?,?,?,?]. For example [?], it can be defined as the maximum Fisher information of classical probabilities $p(x|\theta) = \text{Tr}(\Pi_x \rho_\theta)$, where the optimization is made over all possible POVMs $\{\Pi_x\}$,

$$\mathcal{F}(\rho_\theta) := \max_{\Pi_x} \mathcal{F}[p(x|\theta)]. \tag{100}$$

Recall that the Fisher information of a probability distribution $p(x|\theta)$ is defined as

$$\mathcal{F}[p(x|\theta)] := \int \frac{1}{p(x|\theta)} \left[ \frac{\partial p(x|\theta)}{\partial \theta} \right]^2 dx. \tag{101}$$

Another definition is given in terms of the so-called symmetric logarithmic derivative operator $L$, which is defined via the implicit equation

$$\frac{d\rho_\theta}{d\theta} := \frac{1}{2} (L\rho_\theta + \rho_\theta L), \tag{102}$$
and depends on the particular form of $\rho_\theta$, $L = L(\rho_\theta)$. The quantum Fisher information is given by the variance of this operator in the family $\rho_\theta$ [7],

$$ F(\rho_\theta) := \text{Tr} \left[ L^2(\rho_\theta) \rho_\theta \right]. \quad (103) $$

The equivalence between Eqs. (103) and (104) can be found explicitly proven in [7].

Going back to the problem of witnessing non-Markovianity, the quantum Fisher information is also monotonically decreasing under Markovian dynamics, as it cannot increase under completely positive maps. This can be showed directly from (104).

Because of (104), the Bures distance (and its square) is monotonically decreasing under a completely positive $E$,

$$ F[E(\rho_\theta)] = 4 \lim_{\delta \theta \to 0} \frac{D_B^2[E(\rho_\theta), E(\rho_\theta + \delta \theta)]}{(\delta \theta)^2} \geq 4 \lim_{\delta \theta \to 0} \frac{D_B^2(\rho_\theta, \rho_\theta + \delta \theta)}{(\delta \theta)^2} = F(\rho_\theta). \quad (104) $$

The use of the quantum Fisher information to witness non-Markovianity is originally due to Lu, Wang and Sun in [7]. These authors provided a proof of the monotonicity of the Fisher information by using the definition Eq. (103), and introduced a flow of quantum Fisher information by

$$ I_{\text{QFI}}(t) := \frac{\partial F[\rho_\theta(t)]}{\partial t}. \quad (105) $$

Thus if $I_{\text{QFI}}(t) > 0$ for some $t$, the time evolution is non-Markovian. Moreover, if the evolution is given by some master equation,

$$ \frac{d\rho(t)}{dt} = L_\rho(t) = -i[H(t), \rho(t)] + \sum_k \gamma_k(t) \left[ V_k(t) \rho(t) V_k^\dagger(t) - \frac{1}{2} \{ V_k^\dagger(t) V_k(t), \rho(t) \} \right], \quad (106) $$

the quantum Fisher information flow can be written as

$$ I_{\text{QFI}}(t) = \sum_k \gamma_k(t) J_k(t), \quad (107) $$

$$ J_k(t) := -\text{Tr} \left\{ \rho_\theta(t)[L(\rho_\theta, t), V_k(t)]^\dagger[L(\rho_\theta, t), V_k(t)] \right\} \leq 0. \quad (108) $$

Therefore, $I_{\text{QFI}}$ is negative if all $\gamma_k(t) \geq 0$ in accordance with Theorem 10.

Examples where the quantum Fisher information flow is computed can be found in the original reference [7] and in [?, ?, ?], where its possible relation with the Loschmidt echo was explored.

Notably, this witness of non-Markovianity may be relevant in the context of quantum parameter estimation. Specifically, the error (variance) of any (unbiased) estimation of the parameter $\theta$ is related to the quantum Fisher information through the quantum Cramer-Rao bound [?, ?, ?, ?]:

$$ (\Delta \theta)^2 \geq \frac{1}{F(\rho_\theta)} \quad (109) $$

Thus, an increment in $F(\rho_\theta)$ could be linked with a increment of information about the parameter $\theta$. Nevertheless note that the quantum Fisher information provide just a lower bound to the error on $\theta$, and in fact there are cases where this bound is not achievable.
5.1.5. Other proposals

Other quantities have been proposed to witness non-Markovianity based on the monotonicity property under completely positive maps. For example, in [?], Bylicka, Chruściński and Maniscalco have suggested to use channel capacities. More concretely, given a complete positive and trace-preserving map \( \mathcal{E} \) and some quantum state \( \rho \), we introduce the mutual information between \( \rho \) and \( \mathcal{E}(\rho) \) via

\[
I(\rho, \mathcal{E}) = S(\rho) + I_c(\rho, \mathcal{E}).
\]  

(110)

Here \( S(\rho) = -\text{Tr}(\rho \log \rho) \) is the von Neumann entropy and \( I_c(\rho, \mathcal{E}) \) is the so-called quantum coherent information, defined as [?],

\[
I_c(\rho, \mathcal{E}) = S[\mathcal{E}(\rho)] - S\{[\mathcal{E} \otimes 1]|\Psi_\rho\rangle\langle \Psi_\rho|\},
\]  

(111)

where \( |\Psi_\rho\rangle \in \mathcal{H} \otimes \mathcal{H}_A \) is a purification of \( \rho = \text{Tr}_A(|\Psi_\rho\rangle\langle \Psi_\rho|) \). Remarkably, the quantity \( S\{[\mathcal{E} \otimes 1]|\Psi_\rho\rangle\langle \Psi_\rho|\} \) does not depend on the particular choice of purification. The quantum coherent information is monotonic under completely positive maps [?]

\[
I_c(\rho, \mathcal{E}_2 \mathcal{E}_1) \leq I_c(\rho, \mathcal{E}_1),
\]  

(112)

and the same equation is satisfied for \( I(\rho, \mathcal{E}) \). Thus, in Ref. [?] the following two channel capacities are proposed to witness non-Markovianity:

\[
C_{ea}[\mathcal{E}(t,t_0)] = \sup_\rho I[\rho, \mathcal{E}(t,t_0)],
\]  

(113)

\[
Q[\mathcal{E}(t,t_0)] = \sup_\rho I_c[\rho, \mathcal{E}(t,t_0)],
\]  

(114)

which decrease monotonically with time for all Markovian dynamics. The entanglement assisted capacity \( C_{ea} \) sets a bound on the amount of classical information which can be transmitted along the quantum channel when sender and receiver are allowed to share an unlimited amount of entanglement. Similarly, the capacity \( Q \) provides the limit to the rate at which quantum information can be reliably sent by the quantum channel (for a single use). Examples of the use of these witnesses are found in [?] and [?] for a qubit interacting with a random classical field and with a bosonic environment, respectively.

Another interesting proposal was suggested by Lorenzo, Plastina and Paternostro in [?]. These authors expand the state \( \rho \) in the basis \( \{G_j\}_{j=0}^{d-1} \) where \( G_0 = \mathbb{I}/\sqrt{d} \) and \( \{G_j\}_{j=1}^{d-1} \) are the (normalized) generators of the \( \mathfrak{su}(d) \) algebra,

\[
\rho = \frac{\mathbb{I}}{d} + \sum_{j=1}^{d-1} r_j G_j, \quad r_i = \text{Tr}(G_i \rho).
\]  

(115)

Then, it is well-know that the action of a dynamical map can be seen as an affine transformation of the Bloch vector \( r = (r_1, \ldots, r_{d-1})^t \),

\[
\rho(t) = \mathcal{E}(t,t_0)(\rho) \leftrightarrow r_t = M_{(t,t_0)} r_0 + c_{(t,t_0)}.
\]  

(116)
It can be proven that, since $E(t,t_0)$ is a composition of completely positive maps, the absolute value of the determinant of $M(t,t_0)$ decreases monotonically with time $[?]$. Interestingly $|\det[M(t,t_0)]|$ describes the change in volume of the set of states accessible through the evolution, so that Markovian evolutions reduce (or leave invariant) the volume of accessible states. Thus, this witness enjoys a nice geometrical interpretation. However, similarly to the trace distance, since the volume of accessible states is independent of the affine vector $c_{(t,t_0)}$, it is not sensitive to dynamics where the non-Markovianity is encoded in $c_{(t,t_0)}$.

For examples of the use of $|\det[M(t,t_0)]|$ to track non-Markovianity see the original reference $[?]$ and $[?,?,?]$.

5.2. Witnesses based on monotonicity under local completely positive maps

This second kind of witnesses are typically correlation measures between the dynamical system and some ancilla $A$, in such a way that they do not increase under the action of maps with the form $E \otimes 1_A$. Let us analyze three of them.

5.2.1. Entanglement From an operational point of view, entanglement can be defined as those correlations between different quantum systems which cannot be generated by local operations and classical communication (LOCC) procedures $[?]$. Thus, entanglement turns out to be a resource to perform tasks which cannot be done just by LOCC.

The degree of entanglement of a quantum state may be assessed by the so-called entanglement measures. These ones must fulfill a set of axioms in order to account for the genuine properties present in the concept of entanglement $[?,?,?,?]$. One of these requirements is the monotonicity axiom, which asserts that the amount of entanglement cannot increase by the application of LOCC operations. Actually, the quantifiers of entanglement that fulfill this axiom but do not coincide with the entropy of entanglement for pure states are simply called entanglement monotones. Since local operations are a particular example of LOCC, if some entanglement measure (or monotone) increases under a local map $E \otimes 1$, $E$ cannot be completely positive.

Thus, consider a system $S$ evolving according to some dynamical map $E_{(t,t_0)}$. We will study the evolution of an entangled state $\rho_{SA}$ between $S$ and some static ancillary system $A$, $\rho_{SA}(t) = [E_{(t,t_0)} \otimes 1] \rho_{SA}$. Then, an increment in the amount of entanglement of $\rho_{SA}(t)$ witnesses non-Markovianity. More specifically, consider initially a maximally entangled state $\rho_{SA} = |\Phi\rangle\langle\Phi|$, $|\Phi\rangle = \frac{1}{\sqrt{d}} \sum_{n=0}^{d-1} |n\rangle_S |n\rangle_A$. Provided that $E$ is an entanglement measurement (or monotone), the positive quantity

$$I^{(E)} := \Delta E + \int_{t_0}^{t_1} \left| \frac{dE[\rho_{SA}(t)]}{dt} \right| dt,$$

is different from zero only if $E_{(t,t_0)}$ is non-Markovian in the interval $(t_0,t_1)$. Here $\Delta E := E[\rho_{SA}(t_1)] - E[\rho_{SA}(t_0)]$. 

34
The use of the entanglement to witness non-Markovianity was first proposed in [?], where the expression (??) was suggested. This proposal has been theoretical addressed for cases of qubits coupled to bosonic environments [?, ?, ?, ?, ?], for a damped harmonic oscillator [?, ?, ?], and for random unitary dynamics and classical noise models [?, ?]. Experimentally this witness has been analyzed in [?, ?].

In addition, the relation between dynamical recovering of quantum coherence by applying local operations and entanglement generation has been studied in [?], and further connections between entanglement and non-Markovianity can be found in [?, ?].

5.2.2. Quantum mutual information The total amount of correlations (as classical as quantum) as measured by the quantum mutual information is another witness of non-Markovianity [?]. This quantity is defined as

\[ I(\rho_{SA}) := S(\rho_S) + S(\rho_A) - S(\rho_{SA}), \tag{118} \]

where \( S(\rho) := -\text{Tr}(\rho \log \rho) \) is the von Neumann entropy and \( \rho_{S,A} = \text{Tr}_{A,S}(\rho_{SA}) \). This expression can be rewritten as a relative entropy,

\[ I(\rho_{SA}) = S(\rho_{SA} \parallel \rho_S \otimes \rho_A). \tag{119} \]

Thus, if we apply a local operation, we have

\[
I[(E \otimes 1)\rho_{SA}] = S\{ (E \otimes 1)\rho_{SA} \parallel \text{Tr}_A[(E \otimes 1)\rho_{SA}] \otimes \text{Tr}_S[(E \otimes 1)\rho_{SA}] \}
= S[(E \otimes 1)\rho_{SA} \parallel (E \otimes 1)\rho_S \otimes \rho_A]
\leq S(\rho_{SA} \parallel \rho_S \otimes \rho_A) = I(\rho_{SA}),
\]

(120)

where have used that \( \text{Tr}_A[(E \otimes 1)\rho_{SA}] = E(\rho_S) \) and \( \text{Tr}_S[(E \otimes 1)\rho_{SA}] = \rho_A \), and the monotonicity of the relative entropy Eq. (??). Hence, the quantum mutual information is monotonic under local trace-preserving completely positive maps. See [?, ?, ?] for examples where the quantum mutual information has been used to study non-Markovianity.

5.2.3. Quantum discord Finally, quantum discord [?, ?, ?] can also be used to detect non-Markovian dynamics. Recall that the quantum discord between two quantum systems is a non-symmetric measure of correlations, so that it is not the same to measure the quantum discord between S and A, with respect to A, as with respect to S. For our purposes, we consider the quantum discord as measured by the ancillary system:

\[ D_A(\rho_{SA}) := S(\rho_A) - S(\rho_{SA}) + \min_{\{\Pi_j^A\}} \sum_j S(\rho_{S|\Pi_j^A}), \tag{121} \]

where the minimization is taken over all POVM \( \{\Pi_j^A\} \) on A, and \( \rho_{S|\Pi_j^A} \) is the system state after the outcome corresponding to \( \Pi_j^A \) has been detected,

\[ \rho_{S|\Pi_j^A} := \frac{\text{Tr}_A(\Pi_j^A \rho_{SA})}{\text{Tr}(\Pi_j^A \rho_{SA})}. \tag{122} \]
The quantum discord Eq. (123) is monotonic under local maps on the system $\mathcal{E} \otimes 1$ (see [?])
\[
D_A[\mathcal{E} \otimes 1(\rho_{SA})] \leq D_A(\rho_{SA}).
\]
(123)

However, note that this equation does not hold for local maps on the ancilla $1 \otimes \mathcal{E}$ because of the asymmetry of the quantum discord [?, ?]. Therefore, as long as we are certain that the ancilla does not evolve, we may use quantum discord to probe non-Markovianity.

The possible usefulness of the concept of quantum discord to witness non-Markovianity has been first discussed in [?]. Other relations between Markovianity and quantum discord can be found in [?, ?, ?]. Note however that the existence of quantitative connections between quantum discord and completely positive maps remain controversial [?, ?, ?, ?, ?, ?].

6. Conclusion and Outlook

In this work we have reviewed the topic of quantum non-Markovianity in the light of recent developments regarding its characterization and quantification. Quantum Markovian processes have been defined by taking the divisibility approach, which allows us to circumvent the problem of constructing a hierarchy of probabilities in quantum mechanics. We have also discussed the emergence of memoryless properties within this definition and compared the divisibility approach with other suggested ways to define Markovianity in the quantum realm. We have surveyed recently proposed measures and witnesses of non-Markovianity, explaining their foundations, as well as their motivation and interpretation. Each measure and witness of non-Markovianity has its pros and cons, and the ultimate question of which of them is preferable in practice strongly depends on the context. However, in the case of non Markovianity measures, up to date, just a few of them seem to be computationally efficient.

We hope that this article can be useful for future research in open quantum systems, and its implications for other areas such as quantum information, or statistical mechanics. Despite the tremendous quantity of new results in the characterization and quantification of non-Markovianity in recent years, there are still several important open questions that remain to be addressed. We conclude this review by providing a non-exhaustive list and some possible research directions.

▶ Classification of completely positive non-Markovian master equations. This is probably the most general open problem regarding non-Markovian evolutions. For instance, in Eq. (123) we have written a generic master equation without taking care of the completely positive character of its solutions. When the dynamics is non-Markovian, the structure of the generators which leads to completely positive dynamics is pretty much unknown, although a few partial results have been obtained [?, ?, ?]. The problem basically rests upon the difficult characterization of the generators of dynamical maps $\{\mathcal{E}_{(t_2,t_1)}, t_2 \geq t_1 \geq t_0\}$ under the weak assumption of complete positivity just for instants $t_2 \geq t_1 = t_0$, and not for any $t_2 \geq t_1 \geq t_0$ [?].
Computation of some measures of non-Markovianity. Despite their well grounded physical motivation, it would be desirable to provide efficient ways to compute some of the proposed measures of non-Markovianity. For instance, the geometric degree of non-Markovianity, $D^{I}_{\text{SM}(g)}$, Eq. (??). Similarly, the measure $\mathcal{N}_H^I$, Eq. (??), has been calculated just for unbiased problems or isolated cases without solving the complete optimization problem.

Performance of witnesses. Some questions may be posed regarding the performance of witnesses of non-Markovianity. For example, which kind of witnesses, be it the ones based on monotonicity or the ones based on local monotonicity, is more sensitive to non-Markovian dynamics? Moreover, which of them is more efficient to detect non-Markovian dynamics?

Witnesses of non-Markovianity without resorting to full-state tomography. A question of practical interest is to formulate ways to probe non-Markovianity avoiding full-state (or process) tomography. For example, if we manage to find good enough lower and upper bounds to properties like trace distance [?] or entanglement [?] in terms of simple measurements, we would be able to detect its non-monotonic behavior without resorting to expensive tomographic procedures.

Relation between different measures of non-Markovianity. Another fundamental question is to elucidate whether different measures of non-Markovianity induce the same order. Probably the answer is negative, but more progress has to be done on this line.

Non-Markovianity as a resource theory. Related to the previous point is the possible formulation of a resource theory for non-Markovianity. Similarly to other resource theories [?, ?, ?, ?, ?, ?, ?, ?], we may wonder if non-Markovianity can be seen as a resource to perform whatever tasks which cannot be done solely by Markovian evolutions. Then, an order relation follows, i.e. some dynamics $\mathcal{E}^{(1)}_{(t,t_0)}$ has smaller amount of Markovianity than another dynamics $\mathcal{E}^{(2)}_{(t,t_0)}$, if $\mathcal{E}^{(1)}_{(t,t_0)}$ can be constructed by $\mathcal{E}^{(2)}_{(t,t_0)}$ and Markovian evolutions. This approach also allows to introduce the notion of maximally non-Markovian evolutions, these would be the ones which cannot be generated in terms of other non-Markovian maps and Markovian evolutions. Would this maximally non-Markovian evolutions be the ones defined in section ???

Non-Markovianity and other properties. It will be very relevant to find systems where the presence of non-Markovianity is associated with other notable phenomena. For instance, some quantitative relations between non-Markovianity effects and criticality and phase transitions [?, ?], Loschmidt echo [?, ?], symmetry breaking [?], and Zeno and anti-Zeno effects [?] have already been described.

Potential applications of non-Markovianity. As final question, we may wonder is “what is a non-Markovian process good for in practice?” There are already studies showing its usefulness to prepare steady entangled states [?], to enhance the achievable resolution in quantum metrology [?] or to assist certain tasks in quantum
information and computation [? , ?]. However more research on this direction is required for the formulation of quantitative results [?].

Definitely, this list can be extended with other open questions. However, we think the enumerated points are representative enough to hopefully stimulate the readers into addressing some of these problems and shed further light on this remarkable phenomenon.

Acknowledgements

A. R. acknowledges to J. M. R. Parrondo for discussions about the classical definition of Markovianity and to L. Accardi and K. B. Sinha for their detailed explanations about the algebraic definition of stochastic quantum processes. Moreover, it has been a pleasure to share ideas about quantum non-Markovianity with D. Chruściński, A. Kossakowski, M. M. Wolf, F. Ticozzi and M. Paternostro. We acknowledge financial support from Spanish MINECO grant FIS2012-33152, CAM research consortium QUTEMAD S2009-ESP-1594, UCM-BS grant GICC-910758, EU STREP project PAPETS, the ERC Synergy grant BioQ and an Alexander von Humboldt Professorship.

References

[8] Andréi Andreyevich Márkov (Ryazan, 14 June 1856 Petrograd, 20 July 1922) was a Russian mathematician known because his contributions to number theory, analysis, and probability theory. For a reference about his life and work see G. P. Basharin, A. N. Langville and V. A. Naumov, Linear Algebra Appl. 386 3 (2004).


A permutation or commutation matrix (cf. [72,84]) is a matrix $U_P$ with the property that $U_P (A \otimes B) U_P = B \otimes A$. They are also called unitary swap matrices in quantum information literature. In order to compute them, we may use the relation $U_P \text{vec}(A) = \text{vec}(A^t)$. For the case
of $U_{2 \leftrightarrow 3}[E_{(t+\epsilon,t)} \otimes I]U_{2 \leftrightarrow 3}$, it is tacitly assumed that $E_{(t+\epsilon,t)} \otimes I$ acts in a tensor product of four spaces with the same dimension $H_1 \otimes H_2 \otimes H_3 \otimes H_4$. Then $U_{2 \leftrightarrow 3}$ denotes the permutation matrix interchanging the second and third subspace, i.e. $U_{2 \leftrightarrow 3} = I_1 \otimes U_P \otimes I_4$.

[181] Note that in [?] the fidelity is defined without square

\[ F(\rho_1, \rho_2) = \max_{\{\Psi_1, \Psi_2\}} |\langle \Psi_1 | \Psi_2 \rangle| \]

[201] This is sometimes called quantum data processing inequality, for a proof see [?].
[224] P. Haikka, T. H. Johnson and S. Maniscalco, Time Invariant Discord and Non-Markovianity,
Y.-J. Zhang, W. Han, C.-J. Shan and Y.-J. Xia, *Quantum correlations dynamics under different non-markovian environmental models*, arXiv:1111.2423.


