

# Equilibration of multitime quantum processes in finite time intervals

Neil Dowling<sup>1\*</sup>, Pedro Figueroa-Romero<sup>2</sup>, Felix A. Pollock<sup>1</sup>,  
Philipp Strasberg<sup>3</sup> and Kavan Modi<sup>1†</sup>

<sup>1</sup> School of Physics & Astronomy, Monash University, Victoria 3800, Australia

<sup>2</sup> Hon Hai Quantum Computing Research Center, Taipei, Taiwan

<sup>3</sup> Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física,  
Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

\* [neil.dowling@monash.edu](mailto:neil.dowling@monash.edu), † [kavan.modi@monash.edu](mailto:kavan.modi@monash.edu)

## Abstract

A generic non-integrable (unitary) out-of-equilibrium quantum process, when interrogated across many times, is shown to yield the same statistics as an (non-unitary) equilibrated process. In particular, using the tools of quantum stochastic processes, we prove that under loose assumptions, quantum processes equilibrate within finite time intervals. Sufficient conditions for this to occur are that multitime observables are coarse grained in both space and time, and that the initial state overlaps with many different energy eigenstates. These results help bridge the gap between (unitary) quantum and (non-unitary) statistical physics, i.e., when all multitime properties and correlations are well approximated by stationary quantities, which includes non-Markovianity and temporal entanglement. We discuss implications of this result for the emergence of classical stochastic processes from multitime measurements of an underlying genuinely quantum system.



Copyright N. Dowling *et al.*

This work is licensed under the Creative Commons

[Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Published by the SciPost Foundation.

Received 23-12-2022

Accepted 08-05-2023

Published 14-06-2023

doi:[10.21468/SciPostPhysCore.6.2.043](https://doi.org/10.21468/SciPostPhysCore.6.2.043)



Check for  
updates

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Preliminaries</b>	<b>4</b>
2.1	Single-time instruments	4
2.2	Tensor representation of quantum processes	5
2.3	The diamond norm distance	7
2.4	Underlying continuous process	8
2.5	Equilibration of processes over infinite times	9
<b>3</b>	<b>Equilibration of processes over finite times</b>	<b>10</b>
3.1	Main result	10
3.2	Proof of Eq. (24)	12

3.2.1	Diagonal terms	13
3.2.2	Cross terms	14
3.2.3	Off-diagonal terms	15
3.2.4	Final Bound for $k = 2$	15
3.2.5	Extension to arbitrary $k$	16
<b>4</b>	<b>Equilibration of geometric measures</b>	<b>16</b>
<b>5</b>	<b>Conclusions and discussion</b>	<b>18</b>
<b>A</b>	<b>Proof of Eq. (38)</b>	<b>19</b>
<b>B</b>	<b>Proof of Eq. (40)</b>	<b>21</b>
<b>C</b>	<b>Eq. (49) extended to <math>k = 3</math> time instruments</b>	<b>22</b>
<b>D</b>	<b>Proof of result 1</b>	<b>25</b>
	<b>References</b>	<b>26</b>

---

## 1 Introduction

Quantum processes allow the description of multitime statistics in quantum systems, providing information on the dynamics beyond the single time (quantum state) limit. Any quantum process can be represented as a single positive tensor  $\Upsilon_{\mathbf{k}}$  [1, 2], that can be used to compute the expectation value of a multitime measurement,

$$\text{tr}[\mathcal{A}_k \mathcal{U}_k \cdots \mathcal{A}_1 \mathcal{U}_1(\rho)] = \text{tr}[\Upsilon_{\mathbf{k}} \mathbf{A}_{\mathbf{k}}^T] =: \langle \mathbf{A}_{\mathbf{k}} \rangle_{\Upsilon}. \quad (1)$$

Here,  $\mathbf{A}_{\mathbf{k}}$  too is a tensor encoding the sequence of  $k$  singletime measurement operators  $\{\mathcal{A}_k, \mathcal{A}_{k-1}, \dots, \mathcal{A}_1\}$  (in the Schödinger picture) and  $\Upsilon_{\mathbf{k}}$  is called the process tensor [1–3], which encapsulates the unitary dynamics ( $\mathcal{U}_i$ ) and initial state ( $\rho$ ). Both of the tensors have free indices at the times  $\mathbf{k} := \{t_k, t_{k-1}, \dots, t_1\}$ , and are quantum combs [4]. These are depicted graphically in Fig. 1 (a), and will be more formally constructed later in this work. Such a description of quantum processes allows the characterization of temporal features such as the degree of non-Markovianity of a process [5, 6], the genuine multipartite entanglement in time [7], or when the statistics look classical [8, 9].

One can ask then, when such quantum processes look equilibrated? This is a foundational question of quantum statistical mechanics, concerned with how a thermal, or more generally a steady state, can arise from the picture of isolated quantum mechanics. There are a number of approaches to this research program, such as the celebrated Eigenstate Thermalization Hypothesis (ETH) which assumes that matrix elements of single time expectation values agree with their thermal value [10–18]. Equilibration, on the other hand, relies on rather minimal assumptions to show when expectation values look stationary on average. More specifically, for a quantum state in a large enough Hilbert space with many significant populations in the energy eigenbasis, realistic (coarse) observables look stationary on average [19–22].

Previous approaches to the task of deriving statistical mechanics from pure quantum mechanics predominately lie firmly within the single-time picture of quantum mechanics [13, 14, 23, 24]. However, this does not offer a complete picture, as generally there is a wealth of

hidden informational content in a quantum process  $\Upsilon_k$  compared to a state  $\rho$ , in the form of encoded temporal correlations. Our previous work Ref. [25] takes the first steps in addressing this problem, where it is shown which conditions lead analytically to a quantum process  $\Upsilon$  equilibrating on average (over infinite times) to an equilibrium process  $\Omega$ , where all unitary dynamics is replaced with dephasing (as summarized in Fig. 1). This means that equilibration is stable to perturbation, and that multitime correlations look stationary on average over long times. Much like the single time results, this occurs for an effectively large initial state and a spatially coarse measurement, but interestingly it additionally requires that the multitime observable is coarse in *time* (i.e., that the number of measurements  $k$  is not too large). This result, however, does not give any information on the time scales necessary for process equilibration, or the emergence of multitime features such as stochastic classicality.

Our results here and in Ref. [25] approach the question of equilibration from a new perspective. The advantage of considering the oft-neglected multitime setting, is that it exposes a wide plethora of phenomena of interest to the overarching question of the foundations of statistical mechanics. This includes why Markovianity is so prevalent in nature, and more generally why any multitime correlations should be well-approximated by stationary or thermal ones. For example, on the (microscopic) quantum level, a molecule will ‘remember’ if one applies laser pulse to it. However, chemically such a process is highly Markovian in practice. Yet, one theory clearly underlies the other. So how can one get from the microscopic theory to the (relatively) macroscopic? Measurement and interaction generally perturb quantum systems, so it is non-trivial to ask, what is the mechanism of the emergence of these multitime statistical properties, and how quickly do they appear? While quantum processes account for invasiveness of measurements, classicality emerges which is distinctly non-invasive [8, 9, 26]. Our work therefore helps address foundational questions *beyond thermalization and single-time equilibration*.

In this work we extend the results of Ref. [25] to show the equilibration of processes in finite time intervals and with degenerate energy levels. This means that, considering only the times between measurements within the interval  $\Delta t_i \in [0, T_i]$ , for large enough  $T_i$  the processes  $\Upsilon_k$  and  $\Omega_k$  are indistinguishable. That is, under this constraint for coarse multitime observables  $\mathbf{A}_k$ ,

$$\langle \mathbf{A}_k \rangle_{\Upsilon} \approx \langle \mathbf{A}_k \rangle_{\Omega} . \quad (2)$$

These time intervals  $T_i$  will be shown to be typically much smaller than recurrence times. Indeed, it is important that this is the case for the equilibration time scales to be meaningful. The infinite time intervals result of Ref. [25] readily implies the approximate equilibration for finite times equal to recurrence times, up to arbitrary accuracy. This is for any quantum processes evolving according to a finite dimensional, time-independent Hamiltonian, with arbitrary energy degeneracies and energy gap degeneracies. This is a generalization of the infinite time results, which relied on a technical assumption about energy gap degeneracies. Additionally, we also here give a general theorem of the equilibration of multitime geometric measures of quantum processes, showing the power of this result in comparison to single time equilibration. This may prompt further bounds on the equilibration of multitime properties, such as the time scales for a generic process to look Markovian or classical.

In section 2 we will introduce the process tensor formalism that describes quantum processes, together with the notion of instruments which correspond to arbitrary multitime measurements. Additionally we will recap the infinite times result of Ref. [25]. In section 3 we give our main result on the equilibration of processes in finite time intervals, together with an analysis of what time scales this will occur. Finally, in section 4 we show that process equilibration implies the general equilibration of any geometric measure of the multitime properties of a process, such as non-Markovianity, classicality, etc.

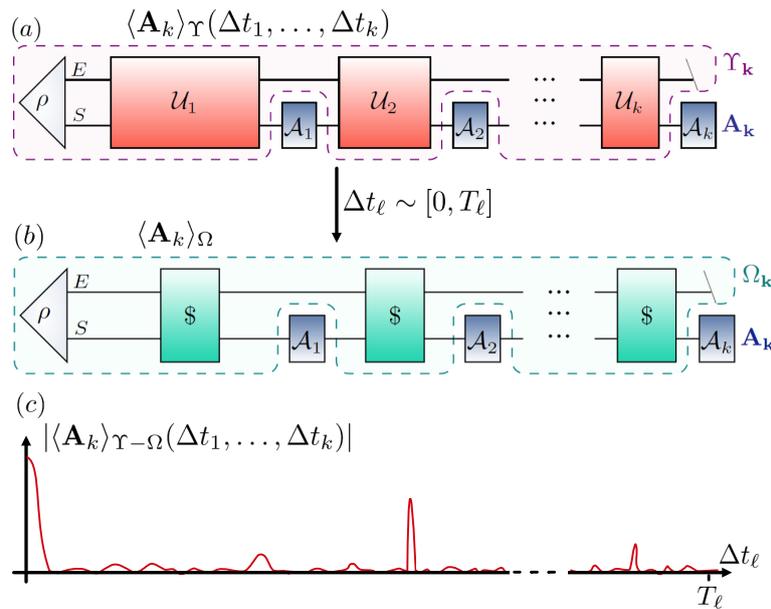


Figure 1: (a) An expectation value of the multitime instrument  $\mathbf{A}_k$  over an arbitrary quantum process  $\Upsilon$ , which is dependent on times  $\Delta t_\ell$  of unitary evolution  $U_\ell$ . (b) If these times are sampled from a large enough ranges of time,  $[0, T_\ell]$ , then this expectation value is likely to be approximately time-independent, equal to the expectation value over the equilibrated process  $\Omega$ . (c) A visualization of the effect of process equilibration. Deviations from the stationary value of multitime correlations are rare and small, across the time ranges  $\Delta t_\ell \sim [0, T_\ell]$  for  $1 \leq \ell \leq k$ . The conditions on  $k$  and  $T_\ell$  for this to occur can be seen in Eq. (29).

## 2 Preliminaries

Here we introduce the process tensor formalism from which our results are constructed from, and restate a main theorem from Ref. [25] from which this work generalizes to finite times and degenerate energy gaps.

### 2.1 Single-time instruments

Physical quantum transformations, including measurements, are in full generality described by a linear and completely positive (CP) map  $\mathcal{A}$  that takes an input quantum state  $\sigma$ , to an output quantum state  $\sigma'$ . Both the input and the output states are density operators, but the latter is not necessarily normalized except when the map is deterministic, i. e., when  $\mathcal{A}$  is further specified to be trace-preserving (TP).

Such a map admits a number of explicit representations, each useful in different circumstances [27, 28]. The operator sum or Kraus representation is given by

$$\mathcal{A}(\sigma) := \sum_{\alpha=1}^n K_\alpha \sigma K_\alpha^\dagger, \tag{3}$$

where if the representation is ‘minimal’,  $n$  is equal to the rank of the map. A second representation that will be essential below, is the *Choi* state. For a map  $\mathcal{A} : \mathcal{H}_b \rightarrow \mathcal{H}_c$  acting on an input  $\sigma$ , this is a matrix  $A$  such that

$$\mathcal{A}(\sigma) = \text{tr}_b[(\mathbb{1}_c \otimes \sigma^T)A]. \tag{4}$$

Note that the typewriter font will be used in the remainder of this work for Choi states of a single time map ( $A$  and  $U$ ), whereas a capital Greek letter will be used for the Choi state of a process ( $\Upsilon$  and  $\Omega$ ), and a capital boldfont Latin letter for a multitime instrument Choi state ( $\mathbf{A}_k$ ), which we define below.

Expanding on the above, we can construct the Choi state of a composition of two maps  $A \circ B$ , for  $B: \mathcal{H}_a \rightarrow \mathcal{H}_b$  and  $A$  as above, via the link product [29],

$$A * B := \text{tr}_b[(\mathbb{1}_c \otimes A^{\text{T}_b})(B \otimes \mathbb{1}_a)], \tag{5}$$

where  $B \in \mathcal{L}(\mathcal{H}_b \otimes \mathcal{H}_a)$  and  $A \in \mathcal{L}(\mathcal{H}_c \otimes \mathcal{H}_b)$ . That is, we append identity matrices to  $A$  and  $B$  such that they live on the same Hilbert space,  $\mathcal{L}(\mathcal{H}_a \otimes \mathcal{H}_b \otimes \mathcal{H}_c)$ , and then trace over the shared space  $\mathcal{H}_b$ . Here, the superscript  $A^{\text{T}_b}$  means the partial transpose of the matrix  $A$  over the space  $\mathcal{H}_b$ . The Link product essentially describes multiplication entirely within the Choi representation, via matrix multiplication on the shared space and tensor product on the independent. This is a key tool with which we can concisely and explicitly define the process tensor.

An operator norm of instruments that will be relevant to our results is the POVM (positive operator valued measured) norm,<sup>1</sup> which is the largest singular value of the POVM element  $A := \sum_{\alpha} K_{\alpha}^{\dagger} K_{\alpha} \leq \mathbb{1}$ ,

$$\|A\|_p := \max_{\|\psi\|_2=1} \|A|\psi\rangle\|_2 = \max_{\sqrt{\langle\psi|A^2|\psi\rangle}=1} \sqrt{\langle\psi|A^2|\psi\rangle}. \tag{6}$$

When a map is trace preserving, we have  $A = \sum_{\alpha} K_{\alpha}^{\dagger} K_{\alpha} = \mathbb{1}$  and the POVM norm is then equal to unity.

## 2.2 Tensor representation of quantum processes

Consider a partition of an isolated quantum system, initially in the state  $\rho$ , into a system ( $S$ ) of interest and the rest, which we call an environment ( $E$ ).  $SE$  then evolves unitarily up until some point  $t_1$ , according to the time-independent Hamiltonian

$$H = \sum E_n P_n, \tag{7}$$

i.e., via the supermap

$$\mathcal{U}_1(\cdot) := e^{-iH\Delta t_1}(\cdot)e^{iH\Delta t_1}. \tag{8}$$

An instrument is then applied to the  $S$  state at time  $t_1$ , described by a CP map  $\mathcal{A}_1 \equiv \mathcal{A}_1 \otimes \mathbb{1}_E$ . Note that this  $SE$  decomposition is consistent with the notion of a coarse (or fine) measurement on an isolated system. For example, measuring the total magnetization of a spin system is highly coarse-grained, and one can appropriately couple the system of spins ( $E$ ) to an ancilla spin ( $S$ ), such that measuring this ancilla will determine the total magnetization. Now, this dynamics followed by measurement is repeated, with a variable time of unitary evolution  $\Delta t_i$ , and choice of single time instruments  $\mathcal{A}_i$ , as depicted in Fig. 1 (a). The expectation value of the sequence of instruments is given by

$$\text{tr}[\mathcal{A}_k \mathcal{U}_k \cdots \mathcal{A}_1 \mathcal{U}_1(\rho)] = \text{tr}[\Upsilon_k \mathbf{A}_k^{\text{T}}] =: \langle \mathbf{A}_k \rangle_{\Upsilon}, \tag{9}$$

where we have introduced the process tensor  $\Upsilon_k$  and the multitime instrument  $\mathbf{A}_k$ , defined for the times  $\mathbf{k} := \{t_1, t_2, \dots, t_k\}$ . Note that adjacent calligraphic characters mean the composition of maps,  $\mathcal{A}_i \mathcal{U}_j := \mathcal{A}_i \circ \mathcal{U}_j$ .

---

<sup>1</sup>We name it the POVM norm so as not to get mixed up with the operator norm of the instrument itself, that is the largest singular value of the Choi state of the instrument.

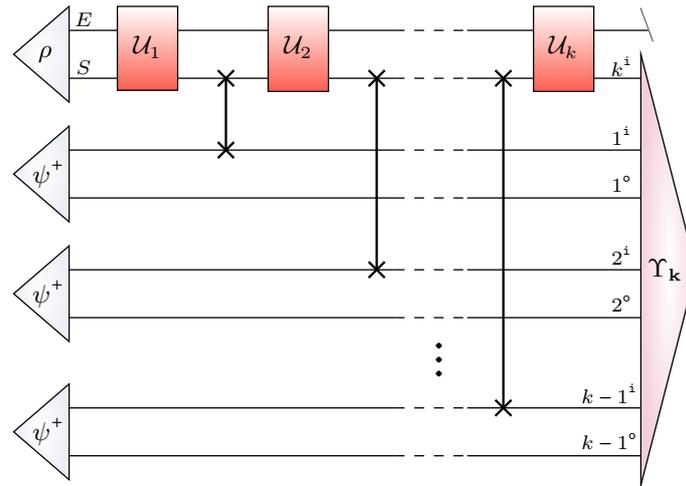


Figure 2: Circuit diagram of the construction of the Choi state of a process tensor through the generalized Choi-Jamiołkowski isomorphism [1, 28]. An ancilla system composed of  $k$  (unnormalized) Bell states  $\psi^+$  is appended to  $SE$ , with one half of a pair swapped with the  $S$  state before each unitary evolution  $\mathcal{U}_i$ , and the  $E$  space being discarded (traced over) at the end. Here the independent Hilbert spaces are labeled such that  $\ell^i$  ( $\ell^o$ ) is the input (output) index at time  $t_\ell$ , showing that the final tensor  $\Upsilon_k$  corresponds to a  $2k - 1$  body density matrix.

Recalling the definition of the Link product Eq. (5), we can directly convert the individual CP maps to their Choi representation, to prove Eq. (9): First, we have  $\rho \in B(\mathcal{H}_{S,0}^o \otimes \mathcal{H}_{E,0}^o)$ ,  $\mathcal{A}_k : B(\mathcal{H}_{S,k}^i) \rightarrow B(\mathcal{H}_{S,k}^o)$ ,  $\mathcal{U}_{k:k-1} : B(\mathcal{H}_{S,k-1}^o \otimes \mathcal{H}_{E,k-1}^o) \rightarrow B(\mathcal{H}_{S,k}^i \otimes \mathcal{H}_{E,k-1}^o)$ . The notation  $\mathcal{H}_{S(E),j}^{i/o}$  means the input/output Hilbert space of the system (environment) at measurement time  $j$ , which we give in order to specify the independent Hilbert spaces so it is clear which tensor indices contract in the following; see Fig. 2. We then write each of these in terms of their matrix indices to get the L.H.S. of Eq. (9)

$$\sum_{\text{all}} \delta_{x_3, y_3} \delta_{\alpha_3, \beta_3} A_{x_3 \alpha_3, y_3 \beta_3} U_{\alpha_3 x_2, \beta_3 y_2}^{\alpha_2 \alpha_1, \beta_2 \beta_1} A_{x_2 \alpha_2, y_2 \beta_2} U_{\alpha_2 x_1, \beta_2 y_1}^{\alpha_1 \alpha_0, \beta_1 \beta_0} A_{x_1 \alpha_1, y_1 \beta_1} U_{\alpha_1 x_0, \beta_1 y_0}^{\alpha_0, \beta_0} \rho_{x_0, y_0}^{\alpha_0, \beta_0}. \quad (10)$$

We have used Greek indices for the environment, and Latin for the system. Upon separating all terms with Greek indices from those that only have Latin ones we have

$$\sum_{\text{latin}} \left( \sum_{\text{greek}} U_{\alpha_3 x_2, \beta_3 y_2}^{\epsilon \alpha_2, \epsilon \beta_2} U_{\alpha_2 x_1, \beta_2 y_1}^{\alpha_2 \alpha_1, \beta_2 \beta_1} U_{\alpha_1 x_0, \beta_1 y_0}^{\alpha_1 \alpha_0, \beta_1 \beta_0} \rho_{x_0, y_0}^{\alpha_0, \beta_0} \right) \left( \sum_e A_{e \alpha_3, e \beta_3} A_{x_2 \alpha_2, y_2 \beta_2} A_{x_1 \alpha_1, y_1 \beta_1} \right). \quad (11)$$

Relabelling the two terms yields Eq. (9)

$$\text{tr}[\Upsilon_k \mathbf{A}_k^T] = \sum_{\text{latin}} (\Upsilon_k)_{\alpha_3 x_2 \alpha_2 x_1 \alpha_1, \beta_3 y_2 \beta_2 y_1 \beta_1} (\mathbf{A}_k)_{\alpha_3 x_2 \alpha_2 x_1 \alpha_1, \beta_3 y_2 \beta_2 y_1 \beta_1}. \quad (12)$$

While we have chosen a  $k = 3$  step process as an example, a more general procedure will yield the explicit definitions

$$\begin{aligned} \Upsilon_k &:= \text{tr}_E[U_k * \dots * U_1 * \rho], \\ \mathbf{A}_k &:= \mathbf{A}_k * \dots * \mathbf{A}_1. \end{aligned} \quad (13)$$

An alternative circuit construction of  $\Upsilon_k$  based on the generalized Choi-Jamiołkowski isomorphism can be seen in Fig. 2. Both  $\Upsilon_k$  and  $\mathbf{A}_k$  are examples of quantum combs that possess well-behaved positivity and trace properties [1, 2, 5, 29]. The former guarantees the positivity of probabilities and the latter is crucial for ensuring the causality of a process - that tracing over a final output leg of a process means the preceding input has no influence on the remainder of the process; ensuring future instruments cannot affect the statistics of past outcomes. In the derivation Eq. (10) it is assumed that the  $\mathcal{A}_i$  are not correlated, in which case the Link product definition of  $\mathbf{A}_k$  in Eq. (13) reduces to a tensor product. We rather allow instruments to carry quantum memory, which can be operationally realized by appending an ancilla space  $W$  such that each  $\mathcal{A}_i$  acts instead on the combined space  $SW$ , as shown in Fig. 3. In such a case  $\mathbf{A}_k$  is called a tester, and is the most general way to measure a process.

For our purposes,  $\Upsilon_k$  is a universal descriptor for any multitime quantum process [2, 30] and, in particular, central for describing non-Markovian processes [1, 3, 5]. This is because all dynamics and correlations that define a process are stored in the single object  $\Upsilon_k$ , which can be probed by the in-principle experimentally implementable instruments encoded in  $\mathbf{A}_k$ . Eq. (9) is then the multitime generalization of the Born rule [4, 31, 32], where  $\Upsilon_k$  plays the role of a state and  $\mathbf{A}_k$  that of a measurement. This definition allows the computation of any temporal correlation functions on a process, accounting for the invasive nature of measurements in quantum mechanics [33].

It is worth pointing out that quantum combs arise naturally in many areas of modern quantum mechanics, including: channels with operational memory [34–37], operational quantum gravity [38–40], spatiotemporal density matrix [41], causally indefinite processes [31], quantum stochastic thermodynamics [42, 43], and the quantum-to-classical transition [8, 9]. They are, of course, central to the studies of multitime correlations in open quantum systems.

### 2.3 The diamond norm distance

In the previous section we have shown that there exists a representation for quantum processes that yields  $k$ -time correlations as a  $(2k - 1)$ -body quantum state, i.e. the Choi state of the process. This allows us to define distances between two quantum processes. This is very much akin to defining distance between two probability distributions, which may represent two (classical) stochastic processes [2].

Considering two processes  $\Upsilon_k$  and  $\Omega_k$ , the most natural way to compute a distance between them is to ask how well one can distinguish them using the optimal multitime measurement. The most general measurement we can perform is a tester, that is we allow the multitime instrument to carry quantum memory and so be correlated in time. We therefore define the generalized diamond norm distance [2, 29, 44] between two processes as the maximum norm difference in the expectation value of any tester on them (with normalization 1/2),

$$D_{\blacklozenge}(\Upsilon, \Omega) := \frac{1}{2} \max_{\mathbf{A}_k} \sum_{\vec{x}} |\langle \mathbf{A}_{\vec{x}} \rangle_{\Upsilon - \Omega}|, \tag{14}$$

where we have defined the shorthand  $\langle \mathbf{A}_{\vec{x}} \rangle_{\Upsilon - \Omega} := \langle \mathbf{A}_{\vec{x}} \rangle_{\Upsilon} - \langle \mathbf{A}_{\vec{x}} \rangle_{\Omega}$ . This definition is motivated by the fact that a particular outcome of an instrument generates a probability distribution,  $\langle \mathbf{A}_{\vec{x}} \rangle_{\Upsilon} = \mathbb{P}(x_k, x_{k-1}, \dots, x_1)$ , and then Eq. (14) is simply a trace difference between probability distributions, maximized over all possible distributions that can be generated on the process  $\Upsilon$ . In practice, however, one does not generally have access to the optimal tester. Instead, consider a restricted set of instruments  $\mathcal{M}_k = \{\mathbf{A}_k\}$  which a hypothetical experimenter has access to, that probe a process at most  $k$  times. We define the operational diamond norm to be

$$D_{\mathcal{M}_k}(\Upsilon, \Omega) := \frac{1}{2} \max_{\mathbf{A}_k \in \mathcal{M}_k} \sum_{\vec{x}} |\langle \mathbf{A}_k \rangle_{\Upsilon - \Omega}|. \tag{15}$$

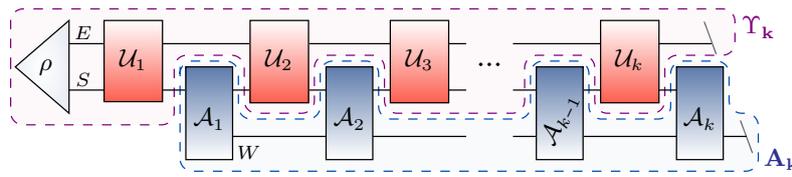


Figure 3: Allowing instruments to transmit quantum memory is equivalent to appending an ancilla space ( $W$ ) such that  $\mathbf{A}_k$  is now a quantum comb. So-called ‘testers’ are the most general way to probe a process  $\Upsilon$  [2, 29].

In the limit of  $\mathcal{M}_k$  containing all possible testers, we obtain the generalized diamond norm Eq. (14), and so  $0 \leq D_{\mathcal{M}_k} \leq D_{\blacklozenge} \leq 1$ . Intuitively, the operational diamond norm is how well one can distinguish between two processes in the best possible case, using only instruments available. This stems from the fact that we assume one has access to a limited set of measurements (for example, whatever is reasonably implementable in a given experimental setup), and measures the maximum difference in measurement statistics using the “most distinguishing” instrument available.

This is essential to our concept of process equilibration, as it allows us to describe a spatial coarse graining in the restriction of accessible instruments to a number that is operationally realistic. From now we drop the subscript  $k$  on  $\mathcal{M}_k$ . In the single-time measurement case, where  $\Upsilon \equiv \rho$  and the expectation values are the usual quantum mechanical ones,  $D_{\mathcal{M}}$  is the analogue of the distinguishability and  $D_{\blacklozenge}$  the trace distance [22, 45]. Note, however, that the inclusion of quantum memory in the multitime case is a non-trivial extension.

### 2.4 Underlying continuous process

So far we have considered a discrete process  $\Upsilon_k$ , with free indices at exactly the times  $\mathbf{k}$  where a chosen instrument  $\mathbf{A}_k$  measures. However, in principle there exists an underlying continuous process  $\Upsilon$ , with an infinite set of ‘free’ indices at all times, with implied identity operators at each. Then, when a discrete times instrument  $\mathbf{A}_k$  is chosen to measure this process, the marginal process  $\Upsilon_k \subset \Upsilon$  which we have derived above is used to compute the expectation value.<sup>2</sup> The existence of the underlying  $\Upsilon$  is ensured by the Generalized Extension Theorem, the quantum generalization of the Kolmogorov Extension Theorem for classical stochastic processes [2, 30].

Considering that the underlying  $\Upsilon$  can hypothetically be measured with a multitime instrument at any set of times  $\mathbf{k}$ , this motivates the definition of an equilibrium process, as the process generated by averaging over these times. The dynamics of such a process will be independent of times  $\mathbf{k} = \{t_k, t_{k-1}, \dots, t_1\}$ , with  $\mathbf{k}$  specifying only the times where the process is measured. The  $k$ -time equilibrium process  $\Omega_k$  is defined as the average marginal process over all possible (ordered) sets of times  $\mathbf{k}$ ,

$$\begin{aligned} \Omega_k &:= \overline{\Upsilon_k}^\infty = \left( \prod_{i=1}^k \lim_{T_i \rightarrow \infty} \frac{1}{T_i} \int_0^{T_i} d(\Delta t_i) \right) \Upsilon \\ &= \text{tr}_E [\hat{\$} * \dots * \hat{\$} * \rho], \end{aligned} \tag{16}$$

where  $\hat{\$}$  is the Choi state of the dephasing map with respect to the energy eigenbasis,

$$\$(\cdot) := \sum_n P_n(\cdot) P_n. \tag{17}$$

<sup>2</sup>For  $\mathbf{a} := \{t_1, t_2, \dots, t_a\}$  and  $\mathbf{b} := \{t_1, t_2, \dots, t_b\}$  with  $a < b$ , the notation  $\Upsilon_a \subset \Upsilon_b$  means that the  $a$ -time process  $\Upsilon_a$  is related to the  $b$ -time process  $\Upsilon_b$  by insertion of identity operators at the times  $\mathbf{b} \setminus \mathbf{a}$ .

This means that for a given dynamic, marginal process  $\Upsilon_{\mathbf{k}}$ , the corresponding equilibrium process  $\Omega_{\mathbf{k}}$  corresponds to replacing all global (SE) unitary dynamics with dephasing; see Fig. 1. We then define *process equilibration* as the time-average indistinguishability of multitime expectation values on this equilibrium process in comparison to an arbitrary non-equilibrium process  $\Upsilon_{\mathbf{k}}$ , via the operational diamond norm distance (15):

$$\overline{D_{\mathcal{M}_k}(\Upsilon_{\mathbf{k}}, \Omega_{\mathbf{k}})} \ll 1. \tag{18}$$

This is the key definition of this work. Note that in comparison to [25], here we consider the average in Eq. (18) to be over finite time intervals.

From now on we will generally drop the subscript  $\mathbf{k}$  on processes, with the understanding that the multitime instrument  $\mathbf{A}_{\mathbf{k}}$  dictates the times  $\mathbf{k} = (t_1, t_2, \dots, t_k)$  at which the underlying process  $\Upsilon$  marginalizes to.

### 2.5 Equilibration of processes over infinite times

The results of this work are an extension of the infinite time results from the related work [25], which we will now summarize. Specifically, the following bound was proven, in terms of the expectation value of any  $k$ -time instrument  $\mathbf{A}_{\mathbf{k}}$  applied to a process  $\Upsilon$  and its corresponding equilibrium  $\Omega$ ,

$$\overline{|\langle \mathbf{A}_{\mathbf{k}} \rangle_{\Upsilon - \Omega}|^2}^{\infty} \leq \max_{j \in [0, k-1]} \frac{(2^k - 1) \|\mathbf{A}_{k:\dots:(j+1)}\|_{\text{p}}^2}{d_{\text{eff}}[\mathcal{A}_j(\omega_j)]}. \tag{19}$$

Here,  $\mathbf{A}_{k:\dots:j}$  is the POVM element of the composition of CP maps

$$\mathcal{A}_k \mathcal{A}_{k-1} \cdots \mathcal{A}_{j+1} \mathcal{A}_j, \tag{20}$$

and we have defined the intermediate equilibrium state for  $j \in [0, k-1] := \{0, 1, \dots, k-1\}$ ,

$$\omega_j := \mathcal{A}_j \mathcal{A}_{j-1} \cdots \mathcal{A}_1 (\rho). \tag{21}$$

The crucial term of the right hand side is the effective dimension, which is defined as

$$d_{\text{eff}}[\sigma] := \frac{1}{\text{tr}[\mathcal{E}(\sigma)^2]}. \tag{22}$$

In generic physical situations the effective dimension scales exponentially with system size  $N$  [20, 24, 46], and is considered a quantifier for the validity of a statistical description of a many-body system. Therefore, with  $\|\mathbf{A}_{k:\dots:(j+1)}\|_{\text{p}}^2 \leq 1$  acting as a scale and  $k \ll N$  restricted by physical considerations, the right hand side of Eq. (19) is extremely small in typical many-body systems, leading directly to physical results on the equilibration of quantum processes over infinite time intervals, without energy gap degeneracies [25]. However, taking the infinite time average means that nothing can be said from this result about the time scales necessary to witness equilibration. It corresponds to averaging up to the recurrence time, which is typically doubly exponential in system size and so even for relatively small many body system these times can be longer than the age of the universe [47–49]. Instead, we here ask if similar equilibration results apply when averaging over finite time intervals, that are less than the recurrence times? and therefore can this give any additional insights into the open question of equilibration time scales [46, 50–55]? We will now explore this, first extending Eq. (19) to finite time intervals between instruments while allowing arbitrary energy gap degeneracies, in the spirit of Refs. [46, 50].

### 3 Equilibration of processes over finite times

Consider an isolated quantum process  $\Upsilon$  as described in Section 2.2 and represented by the purple dashed comb in Fig. 1 (a). This encompasses a finite quantum system evolving according to a time-independent Hamiltonian  $H = \sum E_n P_n$ , which we allow to have arbitrarily degenerate energy levels. Between the global system plus environment ( $SE$ ) unitary evolution, consider repeated, local measurements on the space  $S$  alone, at the times  $t_1, t_2, \dots, t_k$  by instruments  $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_k$ . Together this represents the expectation value of a  $k$ -time instrument  $\mathbf{A}_k$  on a process  $\Upsilon$ , shown in Fig. 1 (a). We then wish to investigate how the expectation value of this instrument varies when measured with respect to  $\Upsilon$  in comparison to the corresponding equilibrated  $\Omega$ , over finite time intervals  $T_\ell$  (such that  $\Delta t_\ell := t_\ell - t_{\ell-1} \leq T_\ell$ ). In contrast to the infinite time result Eq. (19), we additionally allow energy gap degeneracies and this will be included quantitatively in our results.

We will first state our main result on the equilibration of quantum processes over finite times, discuss its implications, and then detail the proof of this result in section 3.2. A reader not interested in the details of the proof should skip this section and read on to section 4, where we show how this result readily implies the equilibration of arbitrary geometric measures of quantum processes.

#### 3.1 Main result

Consider the difference in expectation values of some multitime instrument  $\mathbf{A}_k$ , acting at the set of times  $\mathbf{k} := \{t_1, t_2, \dots, t_k\}$ , between a process  $\Upsilon$  and corresponding equilibrium process  $\Omega$  and averaged over the time intervals  $\Delta t_\ell \in [0, T_\ell]$  for each  $\ell \in [1, k]$ . Define finite time-averaging as,

$$\overline{X}^{\vec{T}} = \left( \prod_{i=1}^k \frac{1}{T_i} \int_0^{T_i} d(\Delta t_i) \right) X, \quad (23)$$

where  $X$  is some function of  $\Delta t_1, \Delta t_2, \dots, \Delta t_k$ . Then in full generality, any process and multitime instrument satisfy the following bound, for time intervals  $T_\ell \geq 0$  and  $\epsilon > 0$ ,

$$\overline{|\langle \mathbf{A}_k \rangle_{\Upsilon - \Omega}|^2}^{\vec{T}} \leq \frac{2^{3k-1} \mathfrak{g}^k}{d_{\text{eff}}[\rho]_{\min}} =: \frac{C_k(\epsilon, T_{\min})}{d_{\text{eff}}[\rho]_{\min}}, \quad (24)$$

where

$$\mathfrak{g} := N(\epsilon) \left( 1 + \frac{8 \log_2 d_H}{\epsilon T_{\min}} \right) \quad \text{and} \quad d_{\text{eff}}[\rho]_{\min} := \min_{\sigma \in \{\rho, \omega, \mathcal{A}_1(\rho), \mathcal{A}_1(\omega), \dots\}} d_{\text{eff}}[\sigma]. \quad (25)$$

Here,  $d_H$  is the number of non-degenerate energy levels,  $T_{\min} := \min_{\ell \in [1, k]} T_\ell$  and, following Ref. [50],  $N(\epsilon)$  is the maximum number of energy gaps in an interval of size  $\epsilon > 0$ ,

$$N(\epsilon) := \max_E |\{(m, n) : E \leq E_m - E_n \leq E + \epsilon\}|. \quad (26)$$

When  $\epsilon \rightarrow 0^+$ , this reduces to the maximum degeneracy of any single energy gap,  $D_G$ . Eq. (24) is the finite time generalization of Eq. (19) and the main mathematical result from which our physical results are derived.

The key feature of the bound Eq. (24) is that it scales with the smallest effective dimension at any stage of either of the processes  $\Upsilon$  and  $\Omega$ . This means that the right hand side will typically scale exponentially with system size (when there are many significantly interacting energy eigenstates), as long as no instrument ‘knocks’ the total  $SE$  state into a small energy subspace. Also note that while  $\mathfrak{g} \geq 1$  can be very large in general, in the bound (24) it scales

logarithmically with total dimension, whereas the effective dimension typically scales linearly with  $d_H$  in a many body system. It will be discussed below what time scales are needed for the right hand side of the bound to be small. This will follow argumentation of Ref. [50].

Using this bound, we arrive at our main physical result on the distinguishability of the processes  $\Upsilon$  and  $\Omega$ .

**Result 1.** *For any quantum process  $\Upsilon$  with corresponding equilibrated  $\Omega$ , and given a set of multitime measurements  $\mathcal{M}$ , then for  $T_\ell > 0$ ,  $\epsilon > 0$  where  $\ell \in [1, k]$ ,*

$$\overline{D_{\mathcal{M}}(\Upsilon, \Omega)^{\vec{T}}} \leq \frac{\mathcal{S}_{\mathcal{M}} \sqrt{C_k(\epsilon, T_{\min})}}{2\sqrt{d_{\text{eff}}[\rho]_{\min}}}. \quad (27)$$

Here,  $\mathcal{S}_{\mathcal{M}}$  is the total combined number of outcomes for all instruments in the set  $\mathcal{M}$ ,

$$\mathcal{S}_{\mathcal{M}} = \sum_{M \in \mathcal{M}} \text{card}(M), \quad (28)$$

where  $\text{card}(M)$  is the cardinality of the set  $M$ .

*Proof.* A proof for this applies Eq. (24) to a result from Ref. [25], and is given in App. D.  $\square$

Assuming that process equilibration occurs in the infinite time intervals case [25], i.e. assuming that  $\mathcal{S}_{\mathcal{M}} \sqrt{2^k - 1} \ll \sqrt{d_{\text{eff}}[\rho]_{\min}}$ , we additionally have equilibration within finite time intervals  $\vec{T}$  given that

$$T_\ell \gtrsim \frac{\log_2 d_H}{\epsilon} \quad \text{and} \quad k \ll \frac{1}{3} \log_2 d_{\text{eff}}[\rho]_{\min}. \quad (29)$$

The first condition states that if no energy gap is hugely degenerate and each time interval  $T_\ell$  is big enough, the factor  $g$  is not too large. Physically this time scale is much smaller than the recurrence times, which typically scale exponentially with effective dimension (doubly exponentially with system size). However, in realistic physical examples  $g$  can blow up due to the  $N(\epsilon)$  factor defined in Eq. (26). We therefore expect tighter bounds than Eq. (24) to hold with additional physical assumptions, with strong evidence that realistic (locally interacting) models with generic initial states equilibrate significantly faster than our bound (and the bound of Ref. [50]) suggests [13, 14, 56, 57]. Although it should be noted there exists relevant physical examples of transitionally invariant lattice models where it takes exponentially long to equilibrate in the single time sense [24, 58]. The strength of the present work is that it holds analytically under minimal assumptions, where analogous results are difficult to obtain with increasing physical assumptions (although some progress in the multitime case has been made very recently, see Refs. [26, 59]).

The second condition ensures the number of times at which an instrument measures the system is far less than the number of components in the system. Clearly this is satisfied in typical many body situations, with for example  $\mathcal{O}(10^{23})$  particles, where measuring even  $\mathcal{O}(10^2)$  time correlations is experimentally unfeasible. Physically, choosing a (small) finite  $k$  is a kind of coarse graining in time; analogous to the assumption that  $\mathcal{S}_{\mathcal{M}}$  is small, which is a coarse graining in (Hilbert) space. This is essential in defining process equilibration. As a counter example, a perfectly fine grained instrument which continuously measures a process at all times can distinguish a dynamical out-of-equilibrium  $\Upsilon$  from a stationary  $\Omega$ .

A disadvantage of this result in comparison to the infinite time one, is that it cannot easily be interpreted in terms of a probability bound, such as Chebyshev's inequality [25]. This is because the bound Eq. (24) (the variance) is computed over a finite times uniform distribution, whereas the equilibrated process expectation value  $\langle A_k \rangle_\Omega$  (the mean) is computed over an infinite times one. Nonetheless, it offers additional insight when interpreted via the operational

diamond norm in Eq. (15), allowing us to show quantum process equilibration in finite time intervals that are generally much less than recurrence times.

Under the conditions (29), Result 1 indicates that processes equilibrate within finite time intervals. However, due to the generality of the setup, the time scales involved are still typically very large. Additional assumptions on the physical scenario would be needed for an estimate on realistic equilibration times. To this purpose, the process tensor formalism directly allows for physical assumptions on the dynamical details of a system. This will be immediately clear from a physical corollary to Result 1, which we describe in Section 4.

### 3.2 Proof of Eq. (24)

We will here derive our main result explicitly for  $k = 2$  instruments, and then motivate a generalization to arbitrary  $k$ , with further details set out in the Appendix. We wish to bound  $|\langle \mathbf{A}_k \rangle_{\Upsilon} - \langle \mathbf{A}_k \rangle_{\Omega}|^2$ , where

$$\overline{X}^{\overline{T}} := \frac{1}{T_1 \cdots T_k} \int_0^{T_k} \cdots \int_0^{T_1} X d(\Delta t_1) \cdots d(\Delta t_k) \tag{30}$$

is the finite-time average over all evolution time intervals  $\Delta t_i$  within the range  $[0, T_i]$ .

Recalling the multitime Born rule Eq. (9), we can expand unitary evolution operators in the energy eigenbasis for the following,

$$\langle \mathbf{A}_k \rangle_{\Upsilon} - \langle \mathbf{A}_k \rangle_{\Omega} = \sum_{n_i, m_i} \text{tr}[\mathcal{A}_k \mathcal{P}_{n_k m_k} \cdots \mathcal{A}_1 \mathcal{P}_{n_1 m_1}(\rho)] \left\{ \prod_{i=1}^k e^{-i\Delta t_i(E_{n_i} - E_{m_i})} - \prod_{i=1}^k \delta_{m_i n_i} \right\}, \tag{31}$$

where  $\mathcal{P}_{n_i m_j}(\cdot) := P_{n_i}(\cdot)P_{m_j}$  is the superoperator that projects onto the  $(n_i, m_j)$ th component in the energy eigenbasis.

Specifying to  $k = 2$  we have,

$$\begin{aligned} \langle \mathbf{A}_k \rangle_{\Upsilon - \Omega} & \stackrel{=:X}{=} \sum_{\substack{n_1 \neq m_1 \\ n_2 \neq m_2}} \text{tr}[\mathcal{A}_2 \mathcal{P}_{n_2 m_2} \mathcal{A}_1 \mathcal{P}_{n_1 m_1}(\rho)] \left[ e^{-i\Delta t_2(E_{n_2} - E_{m_2})} e^{-i\Delta t_1(E_{n_1} - E_{m_1})} \right] \\ & + \underbrace{\sum_{n_1 \neq m_1} \text{tr}[\mathcal{A}_2 \mathcal{P}_{n_2 m_2} \mathcal{A}_1 \mathcal{P}_{n_1 m_1}(\rho)] e^{-i\Delta t_1(E_{n_1} - E_{m_1})}}_{=:Y} + \underbrace{\sum_{n_2 \neq m_2} \text{tr}[\mathcal{A}_2 \mathcal{P}_{n_2 m_2} \mathcal{A}_1(\omega_1)] e^{-i\Delta t_2(E_{n_2} - E_{m_2})}}_{=:Z}. \end{aligned} \tag{32}$$

We will now multiply this with its complex conjugate and independently time-average over each  $\Delta t_\ell$  over the range  $[0, T_\ell]$ , in order to obtain the desired quantity  $|\langle \mathbf{A}_k \rangle_{\Upsilon} - \langle \mathbf{A}_k \rangle_{\Omega}|^2$ . We label the indices corresponding to complex conjugate parts with primes,  $n'_i$  and  $m'_i$ , and after taking the modulus square we obtain exponential multiplicative factors with exponents  $-i\Delta t_\ell(E_{n_\ell} - E_{n'_\ell} - E_{m_\ell} + E_{m'_\ell})$  and  $-i\Delta t_\ell(E_{n_\ell} - E_{m_\ell})$ . These factors contain all the time dependencies, and so we therefore define the tensors  $\mathcal{G}^{(\ell)}$  and  $G^{(\ell)}$  with components

$$\begin{aligned} \overline{\mathcal{G}}_{n_\ell m_\ell n'_\ell m'_\ell}^{(\ell)} & := \overline{\exp[-i\Delta t(E_{n_\ell} - E_{n'_\ell} - E_{m_\ell} + E_{m'_\ell})]}^{T_\ell}, \\ \overline{G}_{n_\ell m_\ell}^{(\ell)} & := \overline{\exp[-i\Delta t(E_{n_\ell} - E_{m_\ell})]}^{T_\ell} \end{aligned} \tag{33}$$

in the corresponding energy eigenbasis. For brevity we will also define the following multilinear function

$$f(\mathcal{X}, \mathcal{Y}) := \text{tr}[\mathcal{A}_2 \mathcal{X} \mathcal{A}_1 \mathcal{Y}(\rho)]. \tag{34}$$

So for example  $f(\mathcal{P}_{n_2 m_2}, \mathcal{P}_{n_1 m_1})$ ,  $f(\$, \mathcal{P}_{n_1 m_1})$ , and  $f(\mathcal{P}_{n_2 m_2}, \$)$  appear in Eq. (32).

Then for  $k = 2$ , we obtain

$$\overline{|\langle \mathbf{A}_k \rangle_{\Gamma-\Omega}|^2}^{(k=2)} = \overline{X^2} + \overline{Y^2} + \overline{Z^2} + 2\text{Re}\{\overline{XY^*} + \overline{XZ^*} + \overline{YZ^*}\}. \quad (35)$$

We will address each of these terms in turn, and see that they are all bounded by a quantity that scales with an effective dimension, using results from Refs. [25, 50]. There are three different type of terms in Eq. (35): the first three only have double sums over  $n_\ell^{(\prime)} \neq m_\ell^{(\prime)}$  (where a prime with a bracket here means a sum over both prime and non-prime indices). We call these ‘diagonal’; the next two terms have both a double sum over  $n_\ell^{(\prime)} \neq m_\ell^{(\prime)}$  and an independent sum over  $n_j \neq m_j$  with  $j \neq \ell$ . We call these ‘cross terms’; and the final term contains only independent sums over  $n_\ell \neq m_\ell$ , which we call ‘off-diagonal’. These three type of terms require somewhat different methods to bound, but classify all the different terms that appear at higher  $k$ . The following methods of will therefore directly generalize to many time instruments.

### 3.2.1 Diagonal terms

Looking at the first term, we have

$$\overline{X^2} = \sum_{\substack{n_1^{(\prime)} \neq m_1^{(\prime)} \\ n_2^{(\prime)} \neq m_2^{(\prime)}}} \overline{\mathcal{G}}_{n_1 m_1 n'_1 m'_1}^{(1)} \overline{\mathcal{G}}_{n_2 m_2 n'_2 m'_2}^{(2)} f(\mathcal{P}_{n_2 m_2}, \mathcal{P}_{n_1 m_1}) f(\mathcal{P}_{n'_2 m'_2}, \mathcal{P}_{n'_1 m'_1})^*, \quad (36)$$

where  $\sum_{n_\ell^{(\prime)} \neq m_\ell^{(\prime)}} := \sum_{n_\ell \neq m_\ell} \sum_{n'_\ell \neq m'_\ell}$  (with  $\ell \in \{1, 2\}$  here). The indices with primes,  $(n'_\ell$  and  $m'_\ell)$ , come from the complex conjugate  $X^*$ , and the complex prefactor  $\mathcal{G}_{n_i m_i n'_i m'_i}^{(i)}$  (a tensor with four indices) is defined in Eq. (33). We can write this prefactor as a matrix, by gathering the indices as  $\alpha \equiv (n_\ell, m_\ell)$ . One can then see that  $\overline{\mathcal{G}}_{n_\ell m_\ell n'_\ell m'_\ell}^{(\ell)} \equiv \overline{\mathcal{G}}_{\alpha\alpha'}^{(\ell)}$  is Hermitian in  $\alpha$ , and similarly  $M_{\alpha\alpha'} := \overline{\mathcal{G}}_{n_1 m_1 n'_1 m'_1}^{(1)} \overline{\mathcal{G}}_{n_2 m_2 n'_2 m'_2}^{(2)}$  is Hermitian in the indices  $\alpha = (n_1, m_1, n_2, m_2)$ . We may therefore use that for Hermitian  $M$ ,  $\sum_{\alpha\alpha'} v_\alpha^* M_{\alpha\alpha'} v_{\alpha'} \leq \|M\| \sum_\alpha |v_\alpha|^2$ , where  $\|M\|$  is the usual operator norm of the matrix  $M$ .<sup>3</sup> Therefore,

$$\overline{X^2} \leq \|\overline{\mathcal{G}}^{(1)}\| \|\overline{\mathcal{G}}^{(2)}\| \sum_{\substack{n_1 \neq m_1 \\ n_2 \neq m_2}} |f(\mathcal{P}_{n_2 m_2}, \mathcal{P}_{n_1 m_1})|^2. \quad (37)$$

From here, we will use an identity used to obtain the infinite time result of Ref. [25],

$$\sum_{\substack{n_i \neq m_i \\ \dots \\ n_j \neq m_j}} |\text{tr}[\mathcal{A}_k \mathcal{D}_k \dots \mathcal{D}_{j+1} \mathcal{A}_j \mathcal{P}_{n_j m_j} \mathcal{A}_{j-1} \mathcal{S}_{j-1} \dots \mathcal{A}_j \mathcal{S}_1(\rho)]| \leq \|\mathbf{A}_{k:\dots:j}\|_{\text{p}}^2 d_{\text{eff}}^{-1}[\mathcal{A}_{j-1}(\omega_{j-1})], \quad (38)$$

where  $1 \leq i < j \leq k$ ,  $\mathcal{D}_\ell \in \{\$, \mathcal{I}\}$  for identity superoperator  $\mathcal{I}$ , and  $\mathcal{S}_\ell \in \{\$, \mathcal{P}_{n_\ell m_\ell}\}$ . The norm  $\|\mathbf{A}_{k:\dots:j}\|_{\text{p}}^2$  is the POVM norm of the composition of CP maps  $\mathcal{A}_k \mathcal{D}_k \mathcal{A}_{k-1} \dots \mathcal{D}_{j+1} \mathcal{A}_j$ , as defined in Eq. (6). The key thing to note is that the choice of each  $\mathcal{S}_\ell$  does not matter for the inequality (38), instead only the final (leftmost) projector is what determines the bound. A proof of this is given in App. A for completeness, reproduced from Ref. [25]. We then obtain

$$\overline{X^2} \leq g_1 g_2 \|\mathbf{A}_2\|_{\text{p}}^2 d_{\text{eff}}^{-1}[\mathcal{A}_1(\omega_1)], \quad (39)$$

<sup>3</sup>The operator norm of a matrix is its largest singular value; formally,  $\|M\| := \sup\{\|Mv\| : \|v\| = 1\}$ .

where we have also introduced a bound derived in Ref. [50],

$$\|\overline{\mathcal{G}}^{(\ell)}\| \leq N(\epsilon) \left( 1 + \frac{8 \log_2 d_H}{\epsilon T_\ell} \right) =: g_\ell, \tag{40}$$

where  $\epsilon > 0$ ,  $d_H$  is the number of distinct energies, and  $N(\epsilon)$  is maximum number of energy gaps in an interval of size  $\epsilon$ , as defined in Eq. (26). A proof for Eq. (40) is given in Appendix B for completeness. An equivalent method can be applied to bound the other two diagonal terms,

$$\overline{Y^2} \leq g_1 \|A_{2:1}\|_p^2 d_{\text{eff}}^{-1}[\rho] \quad \text{and} \quad \overline{Z^2} \leq g_2 \|A_2\|_p^2 d_{\text{eff}}^{-1}[\mathcal{A}_1(\omega_1)]. \tag{41}$$

### 3.2.2 Cross terms

Now, the next two cross terms of Eq. (35) will also be proportional to an effective dimension, but require a different treatment and obtain an additional multiplicity.

$$\begin{aligned} 2\text{Re}\{\overline{XY^*}\} &= 2\text{Re} \left\{ \sum_{\substack{n_1^{(\prime)} \neq m_1^{(\prime)} \\ n_2 \neq m_2}} \overline{\mathcal{G}}_{n_1 m_1 n_1' m_1'}^{(1)} \overline{\mathcal{G}}_{n_2 m_2}^{(2)} f(\mathcal{P}_{n_2 m_2}, \mathcal{P}_{n_1 m_1}) f(\$, \mathcal{P}_{n_1' m_1'})^* \right\} \\ &\leq 2\overline{G}_{\text{max}}^{(2)} \left| \sum_{\substack{n_1^{(\prime)} \neq m_1^{(\prime)} \\ n_2 \neq m_2}} \overline{\mathcal{G}}_{n_1 m_1 n_1' m_1'}^{(1)} f \left( \sum_{n_2 \neq m_2} \mathcal{P}_{n_2 m_2}, \mathcal{P}_{n_1 m_1} \right) f(\$, \mathcal{P}_{n_1' m_1'}) \right|, \end{aligned} \tag{42}$$

where we have used that for  $z \in \mathbb{C}$ ,  $\text{Re}(z) \leq |z|$ , and defined the max value of the matrix  $\overline{\mathcal{G}}^{(\ell)}$ ,  $\overline{G}_{\text{max}}^{(\ell)} := \max_{m \neq n} |G_{nm}^{(\ell)}| \leq 1$ . Note also that the complex conjugate of the trace function  $f$  (Eq. (34)) corresponds only to a transpose of indices  $m$  and  $n$ . Noticing that  $\sum_{n \neq m} \mathcal{P}_{nm} \equiv \mathcal{I} - \$$ , we can use the linearity of  $f$  and the triangle inequality to expand the sum over  $n_2 \neq m_2$ . For the sum over  $n_1^{(\prime)} \neq m_1^{(\prime)}$ , we may again use that  $\mathcal{G}_{nmn'm'}^{(\ell)}$  is Hermitian, and so  $\sum_{\alpha\alpha'} u_\alpha M_{\alpha\alpha'} v_{\alpha'}$  defines an inner product. Applying the Cauchy-Schwarz inequality with respect to this, after the triangle inequality mentioned above, we arrive at

$$2\text{Re}\{\overline{XY^*}\} \leq 2\overline{G}_{\text{max}}^{(2)} \|\overline{\mathcal{G}}^{(1)}\| \sqrt{\sum_{n_1' \neq m_1'} |f(\$, \mathcal{P}_{n_1' m_1'})|^2} \left( \sqrt{\sum_{n_1 \neq m_1} |f(\mathcal{I}, \mathcal{P}_{n_1 m_1})|^2} + \sqrt{\sum_{n_1 \neq m_1} |f(\$, \mathcal{P}_{n_1 m_1})|^2} \right). \tag{43}$$

At this point we are left with terms of the form of the identity Eq. (38), and so arrive at the final bound for this term,

$$2\text{Re}\{\overline{XY^*}\} \leq 4s_2 g_1 \|A_{2:1}\|_p^2 d_{\text{eff}}^{-1}[\rho], \tag{44}$$

where  $\Delta E_{\text{min}}$  is the minimum energy gap, and we have computed

$$\begin{aligned} \overline{G}_{\text{max}}^{(\ell)} &= \max_{m \neq n} \frac{2|\sin[T_\ell(E_m - E_n)/2]|}{T_\ell |E_m - E_n|} \\ &= |\text{sinc}\left(\frac{T_\ell \Delta E_{\text{min}}}{2}\right)| =: s_\ell \leq 1. \end{aligned} \tag{45}$$

The inequality is saturated if and only if the time interval  $T_\ell$  is finite and the Hamiltonian  $H$  has a degenerate energy level. Note however, that we may in full generality consider only non-degenerate Hamiltonians as this generalizes to degenerate Hamiltonian's via an additional

inequality in convex mixtures of pure instruments and pure initial states.<sup>4</sup> Additionally considering that we here only consider finite  $T_\ell$ , the inequality in Eq. (45) is strict.

The fifth term of Eq. (35) can be upper bound in a similar manner,

$$\begin{aligned} 2\text{Re}\{\overline{XZ^*}\} &\leq 2\overline{G}_{\max}^{(1)} \|\overline{G}^{(2)}\| \|A_2\|_p^2 \left( \sqrt{d_{\text{eff}}^{-1}[\mathcal{A}_1(\omega_1)]} \sqrt{d_{\text{eff}}^{-1}[\mathcal{A}_1(\rho)] + d_{\text{eff}}^{-1}[\mathcal{A}_1(\omega_1)]} \right) & (46) \\ &\leq 4s_1 g_2 \|A_2\|_p^2 d_{\text{eff}}^{-1}[\rho]_{\min}, & (47) \end{aligned}$$

where we have used the minimum effective dimension at any stage of either process, as defined in Eq. (25).

### 3.2.3 Off-diagonal terms

We are finally left to bound the last term of Eq. (35), which is ‘off-diagonal in the sense that it contains no double-sums over  $n_\ell^{(\prime)} \neq m_\ell^{(\prime)}$ . The time averages for  $Y$  and  $Z^*$  are entirely independent, and so

$$\begin{aligned} 2\text{Re}\{\overline{YZ^*}\} &\leq 2|\overline{f(\$, \mathcal{U}_1 - \$)}^{T_1} \overline{f(\mathcal{U}_2 - \$, \$)}^{T_2}| \\ &\leq \sqrt{\overline{f(\$, \mathcal{U}_1 - \$)}^{2T_1} \overline{f(\mathcal{U}_2 - \$, \$)}^{2T_2}} \\ &\leq \sqrt{\frac{\|\overline{G}^{(1)}\| \|A_{2:1}\|_p^2 \|\overline{G}^{(2)}\| \|A_2\|_p^2}{d_{\text{eff}}[\rho] d_{\text{eff}}[\mathcal{A}_1(\omega)]}} & (48) \\ &\leq \frac{\sqrt{g_1 \|A_{2:1}\|_p^2 g_2 \|A_2\|_p^2}}{d_{\text{eff}}[\rho]_{\min}}. \end{aligned}$$

Here we have used the single time equilibration result of Ref. [50], which can be derived using the identity (38) together with Eq. (40).

### 3.2.4 Final Bound for $k = 2$

We can now combine Eqs. (39), (41), (44), (46), and (48) to obtain the full equilibration bound for  $k = 2$ ,

$$\begin{aligned} \overline{|\langle \mathbf{A}_k \rangle_{T-\Omega}|^2}^{\vec{T}} &\leq (g_1 g_2 \|A_2\|_p^2 + g_1 \|A_{2:1}\|_p^2 + g_2 \|A_2\|_p^2 \\ &\quad + 4g_1 s_2 \|A_{2:1}\|_p^2 + 4g_2 s_1 \|A_2\|_p^2 \\ &\quad + 2\sqrt{g_1 g_2 \|A_2\|_p^2 \|A_{2:1}\|_p^2}) d_{\text{eff}}^{-1}[\rho]_{\min}, & (49) \end{aligned}$$

where we have introduced an additional inequality to get the common factor  $d_{\text{eff}}^{-1}[\rho]_{\min}$  (which is defined in Eq. (25)).

A perceptive reader may notice that Eq. (49) does not exactly reduce to the infinite time bound Eq. (19). As when  $T_\ell \rightarrow \infty$ , then  $g_\ell \rightarrow 1$  and  $s_\ell \rightarrow 0$ , and so Eq. (49) reduces to,

$$\overline{|\langle \mathbf{A}_k \rangle_{T-\Omega}|^2}^\infty \leq (\|A_2\|_p^2 + \|A_{2:1}\|_p^2 + \|A_2\|_p^2 + 2\sqrt{\|A_2\|_p^2 \|A_{2:1}\|_p^2}) d_{\text{eff}}^{-1}[\rho]_{\min}, \quad (50)$$

<sup>4</sup>Via a similar argument to one given in Ref. [22], we can ensure that the evolution is equivalently according to a non-degenerate Hamiltonian by choosing a basis for any degenerate subspace such that the  $SE$  state at any particular time step may overlap only with one of the degenerate basis states for each distinct energy. This argument holds for pure states at all times, and so only for purity-preserving instruments  $\mathcal{A}_i$ . However,  $\overline{|\langle \mathbf{A} \rangle_{T-\Omega}|^2}^\infty$ , is convex in mixtures of pure instruments/states, so any bound for pure instruments directly implies a bound for mixed instruments/states.

which is a slightly looser bound than Eq. (19). This comes from the fact that the final term which we bounded,  $2\text{Re}\{\overline{YZ^*}\}$ , can in fact alternatively be bounded by a term  $\propto s_1 s_2$ , without any inverse effective dimension dependence. This then goes to zero in the infinite time limit. However, in the finite time case this term can be problematic:  $\Delta E_{\min}$  can be extremely small in a typical many body system (with a large  $d_{\text{eff}}[\rho]$ ), and so this term could cause the bound to become large in such a case, predicting larger equilibration times for systems with smaller  $\Delta E_{\min}$ . In summary, the final term of Eq. (49) can be replaced with,

$$2\text{Re}\{\overline{YZ^*}\} \leq \min \left\{ 2\sqrt{g_1 g_2 \|A_2\|_p^2 \|A_{2:1}\|_p^2 d_{\text{eff}}^{-1}[\rho]_{\min}}, 8s_1 s_2 \right\}, \quad (51)$$

in which case we arrive at the infinite time limit of Ref. [25]. We will omit this for clarity, as the bound proportional to the effective dimension is the most meaningful in the case of finite time intervals.

### 3.2.5 Extension to arbitrary $k$

The derivation of the  $k = 2$  result Eq. (49) dealt with the three types of terms that appear in the generalization of Eq. (35) for arbitrary  $k$ -time instruments. Therefore, the methods here can be directly extended to a similar derivation for a bound for any  $k$ . We give the result explicitly for  $k = 3$  in App. C (49 terms). The key factor is the inverse proportionality with respect to the effective dimension, which will typically be large compared to the other multiplicative factors, and compared to the total multiplicity.

Finally, using that  $s_\ell = \text{sinc}(\Delta E_{\max} T_\ell / 2) \leq 1$ ,  $g_\ell \geq 1$ , and that  $\|A\| \leq 1$ , we can generalize the above derivation to arrive at Eq. (24), together with a multiplicity argument which can be found in App. C. This completes the proof.

## 4 Equilibration of geometric measures

The multitime correlation equilibration results of Eqs. (19) and (24) are stronger than previous equilibration results for singletime observables [20–22, 46, 50]. To see this, in this section we show how process equilibration directly implies the equilibration of a number of geometric measures describing physical properties of a quantum process. We first provide a general theorem on the equilibration of geometric measures, and then specify some examples of physical measures. A geometric measure  $\mathcal{E}_{\mathcal{M}}$  is defined as the minimum distance to the closest process  $\Lambda \in \mathcal{K}$ , where  $\mathcal{K}$  is some restricted set of processes that defines the measure,

$$\mathcal{E}_{\mathcal{M}}(\Upsilon) := \min_{\Lambda \in \mathcal{K}} (D_{\mathcal{M}}(\Upsilon, \Lambda)). \quad (52)$$

Our chosen distance metric here is the operational diamond norm under the restricted set of instruments  $\mathcal{M}$ , as defined in Eq. (15). This allows us to derive results about the equilibration of various geometric measures.

**Result 2.** *For any geometric measure  $\mathcal{E}_{\mathcal{M}}$  of processes, in terms of the operational diamond norm distance restricted to the set of at most  $k$  time instruments  $\mathcal{M}$ ,*

$$\overline{|\mathcal{E}_{\mathcal{M}}(\Upsilon) - \mathcal{E}_{\mathcal{M}}(\Omega)|}^{\vec{T}} \leq \frac{\mathcal{S}_{\mathcal{M}} \sqrt{C_k(\epsilon, T_{\min})}}{2\sqrt{d_{\text{eff}}[\rho]_{\min}}}. \quad (53)$$

*Proof.* Consider without loss of generality that  $\mathcal{E}_{\mathcal{M}}(\Upsilon) \geq \mathcal{E}_{\mathcal{M}}(\Omega)$ ; an equivalent argument

applies in the complementary case. Then,

$$\begin{aligned}
 |\Delta \mathcal{E}_{\mathcal{M}}| &= |\mathcal{E}_{\mathcal{M}}(\Upsilon) - \mathcal{E}_{\mathcal{M}}(\Omega)| \\
 &= \mathcal{E}_{\mathcal{M}}(\Upsilon) - \mathcal{E}_{\mathcal{M}}(\Omega) \\
 &= \min_{\Lambda_{\Upsilon} \in \mathcal{K}} D_{\mathcal{M}}(\Upsilon, \Lambda_{\Upsilon}) - \mathcal{E}_{\mathcal{M}}(\Omega) \\
 &\leq D_{\mathcal{M}}(\Upsilon, \Lambda'_{\Omega}) - \mathcal{E}_{\mathcal{M}}(\Omega) \\
 &= D_{\mathcal{M}}(\Upsilon, \Lambda'_{\Omega}) - D_{\mathcal{M}}(\Omega, \Lambda'_{\Omega}),
 \end{aligned} \tag{54}$$

where  $\Lambda'_{\Omega}$  is chosen to be the process that minimizes  $D_{\mathcal{M}}(\Omega, \Lambda_{\Omega})$ , i.e. that in the definition of  $\mathcal{E}_{\mathcal{M}}(\Omega)$ . To arrive at the penultimate line, we have used that  $\mathcal{E}_{\mathcal{M}}(\Upsilon)$  is a minimum over all processes in the restricted set  $\mathcal{K}$ , and so  $\Lambda'_{\Omega}$  satisfies the inequality for the first term. Now we apply the triangle inequality to arrive at,

$$|\Delta \mathcal{E}_{\mathcal{M}}| \leq D_{\mathcal{M}}(\Upsilon, \Omega). \tag{55}$$

We can then take the multitime average and apply the bound Eq. (27). □

This allows us to prove that (geometric) time-dependent dynamical properties of a process equilibrate to a time-independent quantity in finite time intervals. All these bounds will result in meaningful equilibration under the same conditions as when the right hand side of Eq. (27) is small. That is, for a large effective dimension, a realistic number of total outcomes of measurements  $\mathcal{S}(\mathcal{M})$  that act at a not too large number of times  $k$ , and given that no energy gap is hugely degenerate, the quantity  $\mathcal{E}(\Upsilon)$  is approximately equal to the time-independent  $\mathcal{E}(\Omega)$ .

Examples of such geometric measures include the non-Markovianity [2, 3, 5, 60, 61], the entanglement in time (including genuinely multipartite entanglement) [7], and the classicality of a process [8, 9]. Given that the implications of process equilibration on non-Markovianity was investigated in Ref. [25], we will here focus on classicality.

A classical stochastic process is a joint probability distribution on a multi-time random variable,  $\mathbb{P}(x_k, \dots, x_1)$ . The process tensor is the quantum generalization of this, reducing to it in the correct limit [2, 8, 9], preserving the causal order of a process and satisfying a *generalized Kolmogorov extension theorem* (GET), in that one can marginalize over time steps through the insertion of the identity super-operator  $\mathcal{I}$  for an instrument, and so show the existence of an underlying process on all times [30].

A quantum process is deemed classical when it satisfies the Kolmogorov consistency condition inherent to classical stochastic processes [8, 9],

$$\mathbb{P}(x_1, \dots, \cancel{x_i}, \dots, x_k) := \sum_i \mathbb{P}(x_1, \dots, x_i, \dots, x_k), \tag{56}$$

for all  $i$ . This means that ignoring a step of the process is equivalent to summing over all outcomes. Recalling the definition of single time instruments as CP maps on some space  $S$  (see section 2.1), one can expand them in terms of projectors onto their eigenspace,  $\mathcal{A}_i(\cdot) \equiv \sum_{x_i} x_i \mathcal{P}_{x_i}(\cdot)$ , where  $x_i$  is an outcome of the measurement. Then, in terms of the process tensor, the classical condition Eq. (56) means that if

$$\langle \mathbb{P}_{x_k} \otimes \dots \otimes \Delta_i \dots \otimes \mathbb{P}_{x_1} \rangle_{\Upsilon} = \langle \mathbb{P}_{x_k} \otimes \dots \otimes \mathbb{1} \dots \otimes \mathbb{P}_{x_1} \rangle_{\Upsilon}, \tag{57}$$

then the quantum process  $\Upsilon$  is classical. Here, for the projector superoperator  $\mathcal{P}_{x_i}$ , we have called its Choi state  $\mathbb{P}_{x_i}$ , and defined the dephasing operation  $\Delta_i := \sum_{x_i} \mathcal{P}_{x_i}(\cdot)$ . That is,  $\Upsilon$  is classical when marginalizing in the quantum sense (insertion of  $\mathcal{I}$ ) is equivalent to marginalizing in the classical sense (tracing over all outcomes). To ground ourselves here among the

formalism, consider the following explicit example of an experimental realization of Eq. (57). Take an electron traversing a sequence of Stern-Gerlach apparatuses, within a noisy environment such as outside photons interacting with the electron. Then the electron corresponds to the system which we can measure, the noisy photons correspond to the environment, and the outcomes  $x_j$  are whether the electron is measured as spin up or spin down according to the  $j^{\text{th}}$  apparatus. Then Eq. (57) corresponds to whether the experimenter leaves the  $i^{\text{th}}$  device in the sequence, allowing both spin up and spin down to continue on (left hand side), compared to removing the  $i^{\text{th}}$  device altogether (right hand side). Then we call a process classical if this has no effects on the statistics of the measurements on devices  $1, \dots, i-1, i+1, \dots, k$ .

We can now define a geometric measure of the *classicality*  $\mathcal{C}_{\mathcal{M}}$  of a process, as the distance to the closest classical process, given the restricted set  $\mathcal{M}$  with which the process can be probed. Therefore choosing  $\mathcal{E} \equiv \mathcal{C}$  in Eq. (53), we arrive at the equilibration of the classicality of a process to the time-independent equilibrium quantity  $\mathcal{C}_{\mathcal{M}}(\Omega)$ . This does not mean that the equilibrium process  $\Omega$  is necessarily classical. However, it does say that for coarse multi-time observables and large enough typical systems, how classical the statistics of your measurements look is overwhelmingly likely to be close to this constant. This means that, solely within a (generalized) Born rule quantum measurement framework, it is extremely likely that the classicality of your measurement statistics are close to some value  $\mathcal{C}_{\mathcal{M}}(\Omega)$ . Note that no semiclassical limit is taken here, and so this is a step towards a quantum process description of the emergence of classical stochasticity. If there is extra structure on the quantum process, such as an assumption of quantum Darwinism [62, 63], this could have profound implications for the emergence of macroscopic, objective determinism.

Our ultimate goal is to find the constraints that lead to nontrivial dynamical phenomena for a system, i.e., non-Markovian processes, from an underlying system-environment unitary quantum process. This has the potential to bridge the gaps between the quantum and classical theories.

## 5 Conclusions and discussion

In this work we have proven the conditions under which equilibration of quantum processes occurs in finite time intervals. This is a generalization of the infinite time process equilibration results of Ref. [25], and the extension of the finite time equilibration results of Refs. [46, 50] to multitime observables.

The time scales involved are generally much less than the recurrence times, and so this work offers a method on approximating equilibration times based on the properties of quantum processes. An example of a possible application of the bounds on geometric measures (Result 2) would be to enforce that  $\mathcal{E}_{\mathcal{M}}(\Omega) \stackrel{!}{=} 0$ , to determine bounds on different process equilibration times. This may motivate conditions on the generic Markovianization of process  $\Upsilon$ , or when a quantum process produces classical statistics, and what minimum times  $\vec{T}$  would result in this. This addresses the question: how long does a quantum process take to lose memory? How coarse in time do measurements need to be to arrive at classical observed phenomena? And what extra assumptions on the system and dynamics are needed for this to occur [26]?

In the recent, related work [59], it is shown that for a large class of translation invariant Hamiltonians which satisfy a weak version of the ETH, two point correlation functions computed over a thermal state factorize for large times, with small deviations from this on average. Their setup therefore goes beyond the minimal one proposed in this work and the infinite time result of Ref. [25]. Interestingly, their results imply a kind of weak Markovianization, in that the factorization of thermal, temporal correlations implies that no memory is carried between

observables at two different times. This shows the kind of extra physical assumptions that may lead to emergent, general Markovianization and thermalization. Indeed, we expect that results presented here may lead to a generalization of those presented in Ref. [59] to arbitrary multitime correlations and to finite time intervals, which in turn could stimulate new insight regarding the question of Markovianization. It should also be noted that there are known cases where quantum non-Markovianity is only seen for three or higher order correlations [64], and that non-Markovianity can be hidden for arbitrarily long times [65, 66].

One could similarly investigate the times needed for a process to look classical. Both of these cases are physically relevant, as macroscopically it is highly typical to observe Markovian [67, 68] and classical phenomena. This would be an interesting avenue for further investigation, and the methods used here may be used to address parallel questions to the contemporary issue of equilibration time scales [46, 50–55].

## Acknowledgments

ND is supported by an Australian Government Research Training Program Scholarship and the Monash Graduate Excellence Scholarship. PS acknowledges financial support from a fellowship from “la Caixa” Foundation (ID 100010434, fellowship code LCF/BQ/PR21/11840014) and from the Spanish Agencia Estatal de Investigación (project no. PID2019-107609GB-I00), the Spanish MINECO (FIS2016-80681-P, AEI/FEDER, UE), and the Generalitat de Catalunya (CIRIT 2017-SGR-1127). KM is supported through Australian Research Council Future Fellowship FT160100073, Discovery Project DP210100597, and the International Quantum U Tech Accelerator award by the US Air Force Research Laboratory.

## A Proof of Eq. (38)

This proof is essentially reproduced from the Appendix of Ref. [25].

To prove Eq. (38), we require to further specify the process  $\Upsilon$  to encompass a pure initial state, and only allow pure instruments (a pure instrument is one which preserves the purity of the input state). This allows us to argue that the evolution is according to an effective degenerate Hamiltonian. We will show that this generalizes to mixed states and instruments in the following.

Intuitively, more mixing means more statistical uncertainty, and so mixed instruments/states may only further equilibrate the system. Precisely,  $|\langle \mathbf{A} \rangle_{\Upsilon} - \langle \mathbf{A} \rangle_{\Omega}|^2$  is convex in mixtures of pure instruments/states, so any bound for pure instruments/states may be used in a straightforward manner to produce a bound for mixed instruments/states. Therefore, we consider only exclusively pure instruments/states. This allows us to specify that the evolution is equivalently according to a non-degenerate (rank 1) Hamiltonian,  $H' = \sum E_n |n\rangle \langle n|$ , where  $\{|n\rangle\}$  could be different for different ‘steps’ in the process. This is done by specifying a basis for each unitary evolution such that only one basis state  $|n\rangle$  of any degenerate subspace may overlap with the  $SE$  state, for each distinct energy.

First, consider Eq. (38) for a single sum over  $n_i \neq m_i$ . Expanding the instrument in its Kraus representation  $\mathcal{A}(\cdot) \equiv \sum_{\beta} K^{\beta}(\cdot)K^{\beta\dagger}$ , for any density operator  $\sigma$ ,

$$\begin{aligned}
 \sum_{n_i \neq m_i} |\text{tr}[\mathcal{A}\mathcal{P}_{n_i m_i}(\sigma)]|^2 &= \sum_{n_i \neq m_i} \left| \sum_{\beta} \text{tr}[K^{\beta}(|n_i\rangle \langle n_i| \sigma |m_i\rangle \langle m_i|)K^{\beta\dagger}] \right|^2 \\
 &= \sum_{n_i \neq m_i} |\text{tr}[A |n_i\rangle \langle n_i| \sigma |m_i\rangle \langle m_i|]|^2 \\
 &= \sum_{n_i \neq m_i} |\sigma_{n_i m_i}|^2 \langle m_i | A |n_i\rangle \langle n_i | A^{\dagger} |m_i\rangle \\
 &\leq \sum_{n_i \neq m_i} \sigma_{n_i n_i} \sigma_{m_i m_i} \langle m_i | A |n_i\rangle \langle n_i | A |m_i\rangle \tag{A.1} \\
 &\leq \sum_{n_i, m_i} \text{tr}[A \sigma_{n_i n_i} |n_i\rangle \langle n_i| A \sigma_{m_i m_i} |m_i\rangle \langle m_i|] \\
 &= \text{tr}[A \$_i(\sigma) A \$_i(\sigma)] \\
 &\leq \sqrt{\text{tr}[A \$_i(\sigma) \$_i(\sigma) A] \text{tr}[\$_i(\sigma) A A \$_i(\sigma)]} \\
 &= \sqrt{\text{tr}[A A (\$_i(\sigma))^2] \text{tr}[A A (\$_i(\sigma))^2]} \\
 &\leq \|A^2\|_p \text{tr}[(\$_i(\sigma))^2] = \|A\|_p^2 d_{\text{eff}}^{-1}[\sigma],
 \end{aligned}$$

where we have defined the POVM element  $A := \sum_{\beta} K^{\beta\dagger} K^{\beta} = A^{\dagger}$ , and the energy eigenstate decomposition  $\sigma := \sum_{n_i, m_i} \sigma_{n_i m_i} |n_i\rangle \langle m_i|$ . In the fifth line we have used the Cauchy-Schwarz inequality  $|\sigma_{nm}|^2 \leq \sigma_{nn} \sigma_{mm}$ , valid for any positive hermitian operator  $\sigma$  (equality for pure states). In the sixth line we have added the (non-negative) terms where  $m_i = n_i$  to the sum. In the penultimate line we have again used Cauchy-Schwarz, but for the Hilbert-Schmidt inner product,  $\text{tr}[A^{\dagger} B] \leq \|A\|_{\text{HS}} \|B\|_{\text{HS}}$  with  $\|A\|_{\text{HS}} := \sqrt{\text{tr}[A^{\dagger} A]}$  and noting that  $A = A^{\dagger}$  and  $\sigma = \sigma^{\dagger}$ . Finally, we have used the identity  $\text{tr}(XY) \leq \|X\|_p \text{tr}(Y)$  for positive operators  $X$  and  $Y$ , and that operator norms satisfy  $\|X^{\dagger} X\|_p = \|X\|_p^2$ .

Next, proving Eq. (38) for two sums over  $n_i \neq m_i$ ,

$$\begin{aligned}
 \sum_{\substack{n_1 \neq m_1 \\ n_2 \neq m_2}} |\text{tr}[\mathcal{A}_2 \mathcal{P}_{n_2 m_2} \mathcal{A}_1 \mathcal{P}_{n_1 m_1}(\rho)]|^2 &= \sum_{\substack{n_1 \neq m_1 \\ n_2 \neq m_2}} \left| \sum_{\alpha} \text{tr}[\mathcal{A}_2 |n_2\rangle \langle n_2| K_1^{\alpha} |n_1\rangle \langle n_1| \rho |m_1\rangle \langle m_1| K_1^{\alpha\dagger} |m_2\rangle \langle m_2|] \right|^2 \\
 &= \sum_{\substack{n_1 \neq m_1 \\ n_2 \neq m_2}} \sum_{\alpha, \beta} |\rho_{n_1 m_1}|^2 \langle m_2 | \mathcal{A}_2 |n_2\rangle \langle n_2 | K_1^{\alpha} |n_1\rangle \langle m_1 | K_1^{\alpha\dagger} |m_2\rangle \langle m_2 | K_1^{\beta} |m_1\rangle \langle n_1 | K_1^{\beta\dagger} |n_2\rangle \langle n_2 | \mathcal{A}_2 |m_2\rangle \\
 &\leq \sum_{\substack{n_1, m_1 \\ n_2, m_2}} \sum_{\alpha} \rho_{n_1 n_1} \rho_{m_1 m_1} |\langle m_2 | \mathcal{A}_2 |n_2\rangle|^2 \langle n_2 | K_1^{\alpha} (|n_1\rangle \langle n_1|) K_1^{\alpha\dagger} |n_2\rangle \langle m_2 | K_1^{\alpha} (|m_1\rangle \langle m_1|) K_1^{\alpha\dagger} |m_2\rangle \\
 &\quad + \sum_{\substack{n_1, m_1 \\ n_2, m_2}} \sum_{\alpha \neq \beta} \rho_{n_1 n_1} \rho_{m_1 m_1} |\langle m_2 | \mathcal{A}_2 |n_2\rangle|^2 \langle n_2 | K_1^{\alpha} |n_1\rangle \langle m_1 | K_1^{\alpha\dagger} |m_2\rangle \langle m_2 | K_1^{\beta} |m_1\rangle \langle n_1 | K_1^{\beta\dagger} |n_2\rangle \\
 &\leq \sum_{\alpha} \sum_{n_2, m_2} |\langle m_2 | \mathcal{A}_2 |n_2\rangle|^2 \langle n_2 | K_1^{\alpha}(\omega_1) K_1^{\alpha\dagger} |n_2\rangle \langle m_2 | K_1^{\alpha}(\omega_1) K_1^{\alpha\dagger} |m_2\rangle \\
 &\quad + \max_{i,j} |(A_2)_{ij}|^2 \sum_{\alpha \neq \beta} \sum_{\substack{n_1, m_1 \\ n_2, m_2}} \langle m_2 | K_1^{\beta} (|m_1\rangle \langle m_1|) K_1^{\alpha\dagger} |m_2\rangle \langle n_2 | K_1^{\alpha} (|n_1\rangle \langle n_1|) K_1^{\beta\dagger} |n_2\rangle \\
 &= \sum_{\alpha} \sum_{n_2, m_2} (K_1^{\alpha}(\omega_1) K_1^{\alpha\dagger})_{n_2 n_2} (K_1^{\alpha}(\omega_1) K_1^{\alpha\dagger})_{m_2 m_2} \text{tr}[\mathcal{A}_2 |n_2\rangle \langle n_2| \mathcal{A}_2 |m_2\rangle \langle m_2|] \\
 &\quad + \max_{i,j} |(A_2)_{ij}|^2 \sum_{\alpha \neq \beta} \text{tr}[K_1^{\beta} K_1^{\alpha\dagger}] \text{tr}[K_1^{\alpha} K_1^{\beta\dagger}] \\
 &= \sum_{\alpha} \text{tr}[\mathcal{A}_2 \mathcal{S}_2(K_1^{\alpha}(\omega_1) K_1^{\alpha\dagger}) \mathcal{A}_2 \mathcal{S}_2(K_1^{\alpha}(\omega_1) K_1^{\alpha\dagger})] \\
 &\leq \sum_{\alpha} \|\mathcal{A}_2\|_p^2 \text{tr}[(\mathcal{S}_2(K_1^{\alpha}(\omega_1) K_1^{\alpha\dagger}))^2] \\
 &\leq \|\mathcal{A}_2\|_p^2 \text{tr}\left[\left(\sum_{\alpha} \mathcal{S}_2(K_1^{\alpha}(\omega_1) K_1^{\alpha\dagger})\right)^2\right] = \|\mathcal{A}_2\|_p^2 d_{\text{eff}}^{-1}[\mathcal{A}_1(\omega_1)],
 \end{aligned}$$

(A.2)

where in the third line we have again used that  $|\rho_{n_1 m_1}|^2 \leq \rho_{n_1 n_1} \rho_{m_1 m_1}$ , and also split the sums  $\sum_{\alpha, \beta} = \sum_{\alpha \neq \beta} + \sum_{\alpha} \delta_{\alpha\beta}$ , adding the (non-negative) extra terms  $n_1 = m_1$  and  $n_2 = m_2$ s. In the antepenultimate line we have chosen an orthogonal (canonical) Kraus representation for  $\mathcal{A}_1$ , a minimal representation such that  $\text{tr}[K_1^{\alpha\dagger} K_1^{\beta}] \propto \delta^{\alpha\beta}$  [45]. At that point we have a term of the form of the seventh line of Eq. (A.1), and so we use that result to arrive at the next inequality. In the final line we bring the sum inside by the linearity of the trace, and as  $\sum |x_i|^2 \leq |\sum x_i|^2$  for positive  $x_i$ .

The combination of Eqs. (A.1) and (A.2) generalize directly to arrive at Eq. (38) for arbitrarily many sums over  $n_i \neq m_i$ .

## B Proof of Eq. (40)

This proof is reproduced from Ref. [50].

Consider the matrix  $\bar{\mathcal{G}}^{(\ell)}$  as defined in Eq. (33). Then

$$\|\bar{\mathcal{G}}^{(\ell)}\| \leq \max_{n'_\ell m'_\ell} \sum_{n, m} |\bar{\mathcal{G}}_{n'_\ell m'_\ell n m}^{(\ell)}|, \tag{B.1}$$

noting that

$$\begin{aligned} \overline{\mathcal{G}}_{n_\ell m_\ell n'_\ell m'_\ell}^{(\ell)} &= \overline{\exp[-i\Delta t(E_{n_\ell} - E_{n'_\ell} - E_{m_\ell} + E_{m'_\ell})]}^{T_\ell} \\ &= \begin{cases} 1, & \text{if } (E_{n_\ell} - E_{n'_\ell} - E_{m_\ell} + E_{m'_\ell}) = 0, \\ \frac{\exp[i(E_{n_\ell} - E_{n'_\ell} - E_{m_\ell} + E_{m'_\ell})T] - 1}{i(E_{n_\ell} - E_{n'_\ell} - E_{m_\ell} + E_{m'_\ell})T}, & \text{otherwise.} \end{cases} \end{aligned} \quad (\text{B.2})$$

Now, the sum in Eq. (B.1) can be split into intervals of width  $\epsilon$ , such that there are at most  $N(\epsilon)$  energy gaps  $(n, m)$  satisfying,

$$(k + 1/2)\epsilon > E_{n_\ell} - E_{n'_\ell} - E_{m_\ell} + E_{m'_\ell} \geq (k - 1/2)\epsilon, \quad (\text{B.3})$$

for each  $k \in \mathbb{Z}$ . For  $k = 0$ , we just take  $|\overline{\mathcal{G}}_{n_\ell m_\ell n'_\ell m'_\ell}^{(\ell)}| \leq 1$ , giving the first term of the bound Eq. (40). If  $k$  is non-zero, then  $|E_{n_\ell} - E_{n'_\ell} - E_{m_\ell} + E_{m'_\ell}| \geq (|k| - 1/2)\epsilon$ , and so considering that there are  $d_H(d_H - 1)$  terms in the sum, we can use Eq. (B.2) to arrive at,

$$\sum_{n,m} |\overline{\mathcal{G}}_{n_\ell m_\ell n'_\ell m'_\ell}^{(\ell)}| \leq N(\epsilon) \left( 1 + 2 \sum_{k=1}^{d_H(d_H-1)/2} \frac{2}{(k - 1/2)\epsilon T} \right). \quad (\text{B.4})$$

Now, due to convexity,

$$\sum_{n=2}^{d_H(d_H-1)/2} \frac{1}{n - 1/2} \leq \int_1^{d_H(d_H-1)/2} \frac{1}{x} dx = \ln \left( \frac{d_H(d_H - 1)}{2} \right), \quad (\text{B.5})$$

and so

$$\sum_{n=1}^{d_H(d_H-1)/2} \frac{1}{n - 1/2} \leq 2 + \ln \left( \frac{d_H(d_H - 1)}{2} \right) \leq 2 \log_2(d_H). \quad (\text{B.6})$$

The final inequality can be checked explicitly for  $d_H = 2$  and  $d_H = 3$ , and confirmed for higher  $d_H$  by comparing the derivatives of both sides. Using this in Eq. (B.4), we arrive at the bound Eq. (40).

### C Eq. (49) extended to $k = 3$ time instruments

The proof for higher  $k$  proceeds in the same way as way described in the main body for  $k = 2$ , with the three types of terms addressed individually in the Sections 3.2.1, 3.2.2, and 3.2.3 equivalently appearing in the expansion for  $k > 2$ . The only non-trivial aspect is the counting of the multiplicity of particular terms. However, the total multiplicity is  $(2^k - 1)^2$ , which is what we take to be relevant for our main results for arbitrary  $k$  (found explicitly in the definition of

the constant  $C_k$  in Eq. (24)). Expanding Eq. (31) for  $k = 3$  we have,

$$\begin{aligned}
 \langle \mathbf{A}_k \rangle_{\Gamma-\Omega} & \stackrel{(k=3)}{=} \overbrace{\sum_{\substack{n_1 \neq m_1 \\ n_2 \neq m_2 \\ n_3 \neq m_3}} \text{tr} [\mathcal{A}_3 \mathcal{P}_{n_3 m_3} \mathcal{A}_2 \mathcal{P}_{n_2 m_2} \mathcal{A}_1 \mathcal{P}_{n_1 m_1}(\rho)] \left[ e^{-i\Delta t_3(E_{n_3}-E_{m_3})} e^{-i\Delta t_2(E_{n_2}-E_{m_2})} e^{-i\Delta t_1(E_{n_1}-E_{m_1})} \right]}^{=:X} \\
 & + \overbrace{\sum_{\substack{n_1 \neq m_1 \\ n_2 \neq m_2}} \text{tr} [\mathcal{A}_3 \mathcal{A}_2 \mathcal{P}_{n_2 m_2} \mathcal{A}_1 \mathcal{P}_{n_1 m_1}(\rho)] \left[ e^{-i\Delta t_2(E_{n_2}-E_{m_2})} e^{-i\Delta t_1(E_{n_1}-E_{m_1})} \right]}^{=:Y_1} \\
 & + \overbrace{\sum_{\substack{n_1 \neq m_1 \\ n_3 \neq m_3}} \text{tr} [\mathcal{A}_3 \mathcal{P}_{n_3 m_3} \mathcal{A}_2 \mathcal{A}_1 \mathcal{P}_{n_1 m_1}(\rho)] \left[ e^{-i\Delta t_3(E_{n_3}-E_{m_3})} e^{-i\Delta t_1(E_{n_1}-E_{m_1})} \right]}^{=:Y_2} \\
 & + \overbrace{\sum_{\substack{n_2 \neq m_2 \\ n_3 \neq m_3}} \text{tr} [\mathcal{A}_3 \mathcal{P}_{n_3 m_3} \mathcal{A}_2 \mathcal{P}_{n_2 m_2} \mathcal{A}_1(\omega_1)] \left[ e^{-i\Delta t_3(E_{n_3}-E_{m_3})} e^{-i\Delta t_2(E_{n_2}-E_{m_2})} \right]}^{=:Y_3} \\
 & + \underbrace{\sum_{n_1 \neq m_1} \text{tr} [\mathcal{A}_3 \mathcal{A}_2 \mathcal{A}_1 \mathcal{P}_{n_1 m_1}(\rho)] e^{-i\Delta t_1(E_{n_1}-E_{m_1})}}_{=:Z_1} \\
 & + \underbrace{\sum_{n_2 \neq m_2} \text{tr} [\mathcal{A}_3 \mathcal{A}_2 \mathcal{P}_{n_2 m_2} \mathcal{A}_1(\omega_1)] e^{-i\Delta t_2(E_{n_2}-E_{m_2})}}_{=:Z_2} \\
 & + \underbrace{\sum_{n_3 \neq m_3} \text{tr} [\mathcal{A}_2 \mathcal{P}_{n_3 m_3} \mathcal{A}_2(\omega_2)] e^{-i\Delta t_3(E_{n_3}-E_{m_3})}}_{=:Z_3}.
 \end{aligned} \tag{C.1}$$

Then,

$$\begin{aligned}
 \overline{|\langle \mathbf{A}_k \rangle_{\Gamma-\Omega}|^2} & \stackrel{(k=3)}{=} \overline{X^2} + \overline{Y_1^2} + \overline{Y_2^2} + \overline{Y_3^2} + \overline{Z_1^2} + \overline{Z_2^2} + \overline{Z_3^2} \\
 & + 2\text{Re} \left\{ (\overline{Y_1 Z_1^*} + \overline{Y_2 Z_1^*}) + (\overline{Y_1 Z_2^*} + \overline{Y_3 Z_2^*}) \right. \\
 & \quad + (\overline{Y_2 Z_3^*} + \overline{Y_3 Z_3^*}) + \overline{X Y_1^*} \\
 & \quad + \overline{X Y_2^*} + \overline{X Y_3^*} \\
 & \quad + \overline{Z_1 Z_2^*} + \overline{Z_1 Z_3^*} + \overline{Z_2 Z_3^*} \\
 & \quad + (\overline{X Z_1^*} + \overline{Y_1 Y_2^*}) + (\overline{X Z_2^*} + \overline{Y_1 Y_3^*}) + (\overline{X Z_3^*} + \overline{Y_2 Y_3^*}) \\
 & \quad \left. + (\overline{Y_1 Z_3^*} + \overline{Y_2 Z_2^*} + \overline{Y_3 Z_1^*}) \right\}.
 \end{aligned} \tag{C.2}$$

Here we have grouped terms that will have similar bounds, such that the order of these groups here will agree with the equations below. Note that in the last two lines, we have grouped terms that are equal (thus getting an additional multiplicity). The key thing to note is that after carefully examination, one can see that the bounds of the individual terms in the proof for  $k = 2$  in Section 3.2 can be directly generalized to higher  $k > 2$ . We need only determine what type of terms appear in Eq. (C.2); “diagonal”, “cross” or “off-diagonal”. In particular, the terms in the first line of Eq. (C.2) are diagonal and so can be bound directly using the method

of Section 3.2.1; the second, third, fourth and sixth lines can all be bound using Section 3.2.2 as they are cross terms; and finally the fifth and seventh lines are off-diagonal and so can be bound using the methods of Section 3.2.3. Looking at a representative example of each type of term, the diagonal terms can be bounded as

$$\begin{aligned} \overline{X^2} &= \sum_{\substack{n_1^{(\prime)} \neq m_1^{(\prime)} \\ n_2^{(\prime)} \neq m_2^{(\prime)} \\ n_3^{(\prime)} \neq m_3^{(\prime)}}} \overline{\mathcal{G}}_{n_1 m_1 n_1' m_1'}^{(1)} \overline{\mathcal{G}}_{n_2 m_2 n_2' m_2'}^{(2)} \overline{\mathcal{G}}_{n_3 m_3 n_3' m_3'}^{(3)} f(\mathcal{P}_{n_3 m_3}, \mathcal{P}_{n_2 m_2}, \mathcal{P}_{n_1 m_1}) f(\mathcal{P}_{n_3' m_3'}, \mathcal{P}_{n_2' m_2'}, \mathcal{P}_{n_1' m_1'})^* \\ &\leq \|\overline{\mathcal{G}}^{(1)}\| \|\overline{\mathcal{G}}^{(2)}\| \|\overline{\mathcal{G}}^{(3)}\| \sum_{\substack{n_1 \neq m_1 \\ n_2 \neq m_2 \\ n_3 \neq m_3}} |f(\mathcal{P}_{n_2 m_2}, \mathcal{P}_{n_1 m_1})|^2 \\ &\leq g_1 g_2 g_3 \|A_3\|_p^2 d_{\text{eff}}^{-1}[\mathcal{A}_2(\omega_2)]. \end{aligned} \tag{C.3}$$

In the second line we have used the steps outlined below Eq. (36), and then finally used Eq. (38) together with (45) and (40). Note that we define the generalization of the shorthand function  $f$  (Eq. (34)) as

$$f(\mathcal{X}, \mathcal{Y}, \mathcal{Z}) := \text{tr}[\mathcal{A}_3 \mathcal{X} \mathcal{A}_2 \mathcal{Y} \mathcal{A}_1 \mathcal{Z}(\rho)], \tag{C.4}$$

with  $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$  being either a projector, dephasing, or the identity map. An example of how we can bound a cross term is

$$\begin{aligned} 2\text{Re}\{\overline{XY_1^*}\} &= 2\text{Re}\left\{ \sum_{\substack{n_1^{(\prime)} \neq m_1^{(\prime)} \\ n_2^{(\prime)} \neq m_2^{(\prime)} \\ n_3 \neq m_3}} \overline{\mathcal{G}}_{n_1 m_1 n_1' m_1'}^{(1)} \overline{\mathcal{G}}_{n_2 m_2 n_2' m_2'}^{(2)} \overline{\mathcal{G}}_{n_3 m_3}^{(3)} f(\mathcal{P}_{n_3 m_3}, \mathcal{P}_{n_2 m_2}, \mathcal{P}_{n_1 m_1}) f(\$, \mathcal{P}_{n_2' m_2'}, \mathcal{P}_{n_1' m_1'})^* \right\} \\ &\leq 2\overline{\mathcal{G}}_{\text{max}}^{(2)} \|\overline{\mathcal{G}}^{(1)}\| \sqrt{\sum_{\substack{n_1' \neq m_1' \\ n_2' \neq m_2'}} |f(\$, \mathcal{P}_{n_2' m_2'}, \mathcal{P}_{n_1' m_1'})|^2} \left( \sqrt{\sum_{\substack{n_1 \neq m_1 \\ n_2 \neq m_2}} |f(\mathcal{I}, \mathcal{P}_{n_2 m_2}, \mathcal{P}_{n_1 m_1})|^2} \right. \\ &\quad \left. + \sqrt{\sum_{\substack{n_1 \neq m_1 \\ n_2 \neq m_2}} |f(\$, \mathcal{P}_{n_2 m_2}, \mathcal{P}_{n_1 m_1})|^2} \right) \\ &\leq g_1 g_2 (2s_1) \|A_{3:2}\|_p^2 d_{\text{eff}}^{-1}[\rho]_{\text{min}}. \end{aligned} \tag{C.5}$$

Here, in the second line we used the triangle inequality argument as described above Eq. (43), and in the final line we have applied the bound Eq. (38) together with the bounds of the prefactors: (45) and (40). Finally, an example of the method to bound an off-diagonal term is

$$\begin{aligned} 2\text{Re}\{\overline{Y_1 Z_3^*}\} &\leq 2 |f(\$, \$, \mathcal{U}_1 - \$)|^{T_1} |f(\$, \mathcal{U}_2 - \$, \$)|^{T_2} |f(\mathcal{U}_3 - \$, \$, \$)|^{T_3} \\ &\leq \frac{\sqrt{g_1 \|A_{3:1}\|_p^2 g_2 \|A_{3:2}\|_p^2 g_3 \|A_3\|_p^2}}{d_{\text{eff}}[\rho]_{\text{min}}}, \end{aligned} \tag{C.6}$$

using the single time equilibration result of Ref. [50], analogous to the explanation of Eq. (48).

Therefore, using these methods to bound all terms in Eq. (C.2), Eq. (49) generalizes to  $(2^3 - 1)^2 = 49$  terms, with an extra multiplicity of 2 for each  $s_\ell$  in the cross terms

$$\begin{aligned} \overline{|\langle \mathbf{A}_k \rangle_\Upsilon - \langle \mathbf{A}_k \rangle_\Omega|^2}^{\bar{T}} &\stackrel{(k=3)}{\leq} \left\{ (1 + g_1 + g_2 + g_1 g_2) g_3 \|\mathbf{A}_3\|_p^2 + (g_1 + 1) g_2 \|\mathbf{A}_{3:2}\|_p^2 + g_1 \|\mathbf{A}_{3:1}\|_p^2 \right. \\ &+ 2 \left( (2s_2) + (2s_3) \right) g_1 \|\mathbf{A}_{3:1}\|_p^2 + \left( (2s_1) + (2s_3) \right) g_2 \|\mathbf{A}_{3:2}\|_p^2 \\ &+ \left( (2s_1) + (2s_2) \right) g_3 \|\mathbf{A}_3\|_p^2 + g_1 g_2 (2s_3) \|\mathbf{A}_{3:2}\|_p^2 \\ &+ g_1 g_3 (2s_2) \|\mathbf{A}_3\|_p^2 + g_2 g_3 (2s_1) \|\mathbf{A}_3\|_p^2 \\ &+ \left( \sqrt{g_1 g_2 \|\mathbf{A}_{3:1}\|_p^2 \|\mathbf{A}_{3:2}\|_p^2} + \sqrt{g_1 g_3 \|\mathbf{A}_{3:1}\|_p^2 \|\mathbf{A}_3\|_p^2} + \sqrt{g_2 g_3 \|\mathbf{A}_{3:2}\|_p^2 \|\mathbf{A}_3\|_p^2} \right) \\ &+ 2g_1 (2s_2) (2s_3) \|\mathbf{A}_{3:1}\|_p^2 + 2g_2 (2s_1) (2s_3) \|\mathbf{A}_{3:2}\|_p^2 + 2g_3 (2s_1) (2s_2) \|\mathbf{A}_3\|_p^2 \\ &\left. + 3 \sqrt{g_1 g_2 g_3 \|\mathbf{A}_{3:1}\|_p^2 \|\mathbf{A}_{3:2}\|_p^2 \|\mathbf{A}_3\|_p^2} \right\} d_{\text{eff}}^{-1}[\rho]_{\min}. \end{aligned} \tag{C.7}$$

To determine the total multiplicity, given that there are  $(2^k - 1)^2$  total terms in the  $k$ -time generalization of Eq. (35), there is also an additional factor of 2 for every  $s_\ell$  that appears due to the triangle inequality used in the proof. Taking the largest power out the front by introducing an additional inequality, we arrive at a total multiplicity bounded above by  $4^k 2^{k-1} = 2^{3k-1}$ . Recalling the definition (25) and that  $g \geq 1$ , we again introduce an additional inequality with  $g^k$  as a common factor, and arrive at the definition of  $C_k$ .

## D Proof of result 1

Consider that each  $\mathbf{A}_w \in \mathcal{M}_k$  has outcomes  $\{\vec{x}\} = \{(x_1, x_2, \dots, x_w)\}$  corresponding to the instrument  $\mathbf{A}_{\vec{x}}$ , where  $w \leq k$ . We can then bound the time averaged operational diamond norm, as defined in Eq. (15),

$$\begin{aligned} \overline{D_{\mathcal{M}}(\Upsilon, \Omega)}^{\bar{T}} &= \frac{1}{2} \overline{\max_{\mathbf{A}_w \in \mathcal{M}_k} \sum_{\vec{x}} |\text{tr}[\mathbf{A}_{\vec{x}}(\Upsilon_w - \Omega_w)]|}^{\bar{T}} \\ &\leq \frac{1}{2} \sum_{\mathbf{A}_w \in \mathcal{M}_k} \sum_{\vec{x}} \overline{|\langle \mathbf{A}_{\vec{x}} \rangle_\Upsilon - \langle \mathbf{A}_{\vec{x}} \rangle_\Omega|}^{\bar{T}} \\ &\leq \frac{1}{2} \sum_{\mathbf{A}_w \in \mathcal{M}_k} \sum_{\vec{x}} \sqrt{\overline{|\langle \mathbf{A}_{\vec{x}} \rangle_\Upsilon - \langle \mathbf{A}_{\vec{x}} \rangle_\Omega|^2}^{\bar{T}}} \\ &\leq \frac{1}{2} \sum_{\mathbf{A}_w \in \mathcal{M}_k} \sum_{\vec{x}} \sqrt{C_w(\epsilon, T_{\min}) d_{\text{eff}}^{-1}[\rho]_{\min}} \\ &\leq \frac{1}{2} \sum_{\mathbf{A}_w \in \mathcal{M}_k} \sum_{\vec{x}} \sqrt{C_k(\epsilon, T_{\min}) d_{\text{eff}}^{-1}[\rho]_{\min}} \\ &= \frac{S_{\mathcal{M}} \sqrt{C_k(\epsilon, T_{\min})}}{2 \sqrt{d_{\text{eff}}[\rho]_{\min}}}, \end{aligned} \tag{D.1}$$

where in the fourth line we have used Eq. (24), and in the fifth that  $w \leq k$ . Note that this proof follows closely to one given in Ref. [22], and is effectively reproduced from Ref. [25].

## References

- [1] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro and K. Modi, *Non-Markovian quantum processes: Complete framework and efficient characterization*, Phys. Rev. A **97**, 012127 (2018), doi:[10.1103/PhysRevA.97.012127](https://doi.org/10.1103/PhysRevA.97.012127).
- [2] S. Milz and K. Modi, *Quantum stochastic processes and quantum non-Markovian phenomena*, PRX Quantum **2**, 030201 (2021), doi:[10.1103/PRXQuantum.2.030201](https://doi.org/10.1103/PRXQuantum.2.030201).
- [3] F. Costa and S. Shrapnel, *Quantum causal modelling*, New J. Phys. **18**, 063032 (2016), doi:[10.1088/1367-2630/18/6/063032](https://doi.org/10.1088/1367-2630/18/6/063032).
- [4] G. Chiribella, G. M. D'Ariano and P. Perinotti, *Theoretical framework for quantum networks*, Phys. Rev. A **80**, 022339 (2009), doi:[10.1103/PhysRevA.80.022339](https://doi.org/10.1103/PhysRevA.80.022339).
- [5] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro and K. Modi, *Operational Markov condition for quantum processes*, Phys. Rev. Lett. **120**, 040405 (2018), doi:[10.1103/PhysRevLett.120.040405](https://doi.org/10.1103/PhysRevLett.120.040405).
- [6] G. A. L. White, C. D. Hill, F. A. Pollock, L. C. L. Hollenberg and K. Modi, *Demonstration of non-Markovian process characterisation and control on a quantum processor*, Nat. Commun. **11**, 6301 (2020), doi:[10.1038/s41467-020-20113-3](https://doi.org/10.1038/s41467-020-20113-3).
- [7] S. Milz, C. Spee, Z.-P. Xu, F. Pollock, K. Modi and O. Gühne, *Genuine multipartite entanglement in time*, SciPost Phys. **10**, 141 (2021), doi:[10.21468/SciPostPhys.10.6.141](https://doi.org/10.21468/SciPostPhys.10.6.141).
- [8] P. Strasberg and M. G. Díaz, *Classical quantum stochastic processes*, Phys. Rev. A **100**, 022120 (2019), doi:[10.1103/PhysRevA.100.022120](https://doi.org/10.1103/PhysRevA.100.022120).
- [9] S. Milz et al., *When is a non-Markovian quantum process classical?*, Phys. Rev. X **10**, 041049 (2020), doi:[10.1103/PhysRevX.10.041049](https://doi.org/10.1103/PhysRevX.10.041049).
- [10] J. M. Deutsch, *Quantum statistical mechanics in a closed system*, Phys. Rev. A **43**, 2046 (1991), doi:[10.1103/PhysRevA.43.2046](https://doi.org/10.1103/PhysRevA.43.2046).
- [11] M. Srednicki, *Chaos and quantum thermalization*, Phys. Rev. E **50**, 888 (1994), doi:[10.1103/PhysRevE.50.888](https://doi.org/10.1103/PhysRevE.50.888).
- [12] M. Srednicki, *The approach to thermal equilibrium in quantized chaotic systems*, J. Phys. A: Math. Gen. **32**, 1163 (1999), doi:[10.1088/0305-4470/32/7/007](https://doi.org/10.1088/0305-4470/32/7/007).
- [13] M. Rigol, V. Dunjko, V. Yurovsky and M. Olshanii, *Relaxation in a completely integrable many-body quantum system: An ab initio study of the dynamics of the highly excited states of 1D lattice hard-core bosons*, Phys. Rev. Lett. **98**, 050405 (2007), doi:[10.1103/PhysRevLett.98.050405](https://doi.org/10.1103/PhysRevLett.98.050405).
- [14] M. Rigol, V. Dunjko and M. Olshanii, *Thermalization and its mechanism for generic isolated quantum systems*, Nature **452**, 854 (2008), doi:[10.1038/nature06838](https://doi.org/10.1038/nature06838).
- [15] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn and Z. Papić, *Weak ergodicity breaking from quantum many-body scars*, Nat. Phys. **14**, 745 (2018), doi:[10.1038/s41567-018-0137-5](https://doi.org/10.1038/s41567-018-0137-5).
- [16] J. M. Deutsch, *Eigenstate thermalization hypothesis*, Rep. Prog. Phys. **81**, 082001 (2018), doi:[10.1088/1361-6633/aac9f1](https://doi.org/10.1088/1361-6633/aac9f1).

- [17] F. G. S. L. Brandão, E. Crosson, M. B. Şahinoğlu and J. Bowen, *Quantum error correcting codes in eigenstates of translation-invariant spin chains*, Phys. Rev. Lett. **123**, 110502 (2019), doi:[10.1103/PhysRevLett.123.110502](https://doi.org/10.1103/PhysRevLett.123.110502).
- [18] J. Richter, J. Gemmer and R. Steinigeweg, *Impact of eigenstate thermalization on the route to equilibrium*, Phys. Rev. E **99**, 050104 (2019), doi:[10.1103/PhysRevE.99.050104](https://doi.org/10.1103/PhysRevE.99.050104).
- [19] H. Tasaki, *From quantum dynamics to the canonical distribution: General picture and a rigorous example*, Phys. Rev. Lett. **80**, 1373 (1998), doi:[10.1103/PhysRevLett.80.1373](https://doi.org/10.1103/PhysRevLett.80.1373).
- [20] P. Reimann, *Foundation of statistical mechanics under experimentally realistic conditions*, Phys. Rev. Lett. **101**, 190403 (2008), doi:[10.1103/PhysRevLett.101.190403](https://doi.org/10.1103/PhysRevLett.101.190403).
- [21] N. Linden, S. Popescu, A. J. Short and A. Winter, *Quantum mechanical evolution towards thermal equilibrium*, Phys. Rev. E **79**, 061103 (2009), doi:[10.1103/PhysRevE.79.061103](https://doi.org/10.1103/PhysRevE.79.061103).
- [22] A. J. Short, *Equilibration of quantum systems and subsystems*, New J. Phys. **13**, 053009 (2011), doi:[10.1088/1367-2630/13/5/053009](https://doi.org/10.1088/1367-2630/13/5/053009).
- [23] L. D'Alessio, Y. Kafri, A. Polkovnikov and M. Rigol, *From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics*, Adv. Phys. **65**, 239 (2016), doi:[10.1080/00018732.2016.1198134](https://doi.org/10.1080/00018732.2016.1198134).
- [24] C. Gogolin and J. Eisert, *Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems*, Rep. Prog. Phys. **79**, 056001 (2016), doi:[10.1088/0034-4885/79/5/056001](https://doi.org/10.1088/0034-4885/79/5/056001).
- [25] N. Dowling, P. Figueroa-Romero, F. A. Pollock, P. Strasberg and K. Modi, *Relaxation of multitime statistics in quantum systems*, Quantum **7**, 1027 (2023), doi:[10.22331/q-2023-06-01-1027](https://doi.org/10.22331/q-2023-06-01-1027).
- [26] P. Strasberg, A. Winter, J. Gemmer and J. Wang, *Classicality, Markovianity and local detailed balance from pure state dynamics*, (arXiv preprint) doi:[10.48550/arXiv.2209.07977](https://doi.org/10.48550/arXiv.2209.07977).
- [27] S. Milz, F. A. Pollock and K. Modi, *An introduction to operational quantum dynamics*, Open Syst. Inf. Dyn. **24**, 1740016 (2017), doi:[10.1142/S1230161217400169](https://doi.org/10.1142/S1230161217400169).
- [28] J. Watrous, *The Theory of quantum information*, Cambridge University Press, Cambridge, UK, ISBN 9781316848142 (2018), doi:[10.1017/9781316848142](https://doi.org/10.1017/9781316848142).
- [29] G. Chiribella, G. M. D'Ariano and P. Perinotti, *Transforming quantum operations: Quantum supermaps*, Europhys. Lett. **83**, 30004 (2008), doi:[10.1209/0295-5075/83/30004](https://doi.org/10.1209/0295-5075/83/30004).
- [30] S. Milz, F. Sakuldee, F. A. Pollock and K. Modi, *Kolmogorov extension theorem for (quantum) causal modelling and general probabilistic theories*, Quantum **4**, 255 (2020), doi:[10.22331/q-2020-04-20-255](https://doi.org/10.22331/q-2020-04-20-255).
- [31] O. Oreshkov, F. Costa and Č. Brukner, *Quantum correlations with no causal order*, Nat. Commun. **3**, 1092 (2012), doi:[10.1038/ncomms2076](https://doi.org/10.1038/ncomms2076).
- [32] S. Shrapnel, F. Costa and G. Milburn, *Updating the Born rule*, New J. Phys. **20**, 053010 (2018), doi:[10.1088/1367-2630/aabe12](https://doi.org/10.1088/1367-2630/aabe12).
- [33] E. B. Davies and J. T. Lewis, *An operational approach to quantum probability*, Commun. Math. Phys. **17**, 239 (1970), doi:[10.1007/BF01647093](https://doi.org/10.1007/BF01647093).

- [34] D. Kretschmann and R. F. Werner, *Quantum channels with memory*, Phys. Rev. A **72**, 062323 (2005), doi:[10.1103/PhysRevA.72.062323](https://doi.org/10.1103/PhysRevA.72.062323).
- [35] F. Caruso, V. Giovannetti, C. Lupo and S. Mancini, *Quantum channels and memory effects*, Rev. Mod. Phys. **86**, 1203 (2014), doi:[10.1103/RevModPhys.86.1203](https://doi.org/10.1103/RevModPhys.86.1203).
- [36] C. Portmann, C. Matt, U. Maurer, R. Renner and B. Tackmann, *Causal boxes: Quantum information-processing systems closed under composition*, IEEE Trans. Inf. Theory **63**, 3277 (2017), doi:[10.1109/TIT.2017.2676805](https://doi.org/10.1109/TIT.2017.2676805).
- [37] S. Khatri and M. M. Wilde, *Principles of quantum communication theory: A modern approach*, (arXiv preprint) doi:[10.48550/arXiv.2011.04672](https://doi.org/10.48550/arXiv.2011.04672).
- [38] L. Hardy, *Towards quantum gravity: A framework for probabilistic theories with non-fixed causal structure*, J. Phys. A: Math. Theor. **40**, 3081 (2007), doi:[10.1088/1751-8113/40/12/s12](https://doi.org/10.1088/1751-8113/40/12/s12).
- [39] L. Hardy, *The operator tensor formulation of quantum theory*, Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. **370**, 3385 (2012), doi:[10.1098/rsta.2011.0326](https://doi.org/10.1098/rsta.2011.0326).
- [40] L. Hardy, *Operational general relativity: Possibilistic, probabilistic, and quantum*, (arXiv preprint) doi:[10.48550/arXiv.1608.06940](https://doi.org/10.48550/arXiv.1608.06940).
- [41] J. Cotler, C.-M. Jian, X.-L. Qi and F. Wilczek, *Superdensity operators for spacetime quantum mechanics*, J. High Energy Phys. **09**, 093 (2018), doi:[10.1007/JHEP09\(2018\)093](https://doi.org/10.1007/JHEP09(2018)093).
- [42] P. Strasberg, *Operational approach to quantum stochastic thermodynamics*, Phys. Rev. E **100**, 022127 (2019), doi:[10.1103/PhysRevE.100.022127](https://doi.org/10.1103/PhysRevE.100.022127).
- [43] P. Strasberg, *Quantum stochastic thermodynamics: Foundations and selected applications*, Oxford University Press, Oxford, UK, ISBN 9780192895585 (2022), doi:[10.1093/oso/9780192895585.001.0001](https://doi.org/10.1093/oso/9780192895585.001.0001).
- [44] P. Taranto, F. A. Pollock and K. Modi, *Non-Markovian memory strength bounds quantum process recoverability*, npj Quantum Inf. **7**, 149 (2021), doi:[10.1038/s41534-021-00481-4](https://doi.org/10.1038/s41534-021-00481-4).
- [45] M. M. Wilde, *Quantum information theory*, Cambridge University Press, Cambridge, UK, ISBN 9781107034259 (2013), doi:[10.1017/CBO9781139525343](https://doi.org/10.1017/CBO9781139525343).
- [46] P. Reimann and M. Kastner, *Equilibration of isolated macroscopic quantum systems*, New J. Phys. **14**, 043020 (2012), doi:[10.1088/1367-2630/14/4/043020](https://doi.org/10.1088/1367-2630/14/4/043020).
- [47] P. Bocchieri and A. Loinger, *Quantum recurrence theorem*, Phys. Rev. **107**, 337 (1957), doi:[10.1103/PhysRev.107.337](https://doi.org/10.1103/PhysRev.107.337).
- [48] A. Peres, *Recurrence phenomena in quantum dynamics*, Phys. Rev. Lett. **49**, 1118 (1982), doi:[10.1103/PhysRevLett.49.1118](https://doi.org/10.1103/PhysRevLett.49.1118).
- [49] L. C. Venuti, *The recurrence time in quantum mechanics*, (arXiv preprint) doi:[10.48550/arXiv.1509.04352](https://doi.org/10.48550/arXiv.1509.04352).
- [50] A. J. Short and T. C. Farrelly, *Quantum equilibration in finite time*, New J. Phys. **14**, 013063 (2012), doi:[10.1088/1367-2630/14/1/013063](https://doi.org/10.1088/1367-2630/14/1/013063).
- [51] L. Masanes, A. J. Roncaglia and A. Acín, *Complexity of energy eigenstates as a mechanism for equilibration*, Phys. Rev. E **87**, 032137 (2013), doi:[10.1103/PhysRevE.87.032137](https://doi.org/10.1103/PhysRevE.87.032137).

- [52] T. Monnai, *Generic evaluation of relaxation time for quantum many-body systems: Analysis of the system size dependence*, J. Phys. Soc. Jpn. **82**, 044006 (2013), doi:[10.7566/JPSJ.82.044006](https://doi.org/10.7566/JPSJ.82.044006).
- [53] A. S. L. Malabarba, L. P. García-Pintos, N. Linden, T. C. Farrelly and A. J. Short, *Quantum systems equilibrate rapidly for most observables*, Phys. Rev. E **90**, 012121 (2014), doi:[10.1103/PhysRevE.90.012121](https://doi.org/10.1103/PhysRevE.90.012121).
- [54] L. P. García-Pintos, N. Linden, A. S. L. Malabarba, A. J. Short and A. Winter, *Equilibration time scales of physically relevant observables*, Phys. Rev. X **7**, 031027 (2017), doi:[10.1103/PhysRevX.7.031027](https://doi.org/10.1103/PhysRevX.7.031027).
- [55] T. R. de Oliveira, C. Charalambous, D. Jonathan, M. Lewenstein and A. Riera, *Equilibration time scales in closed many-body quantum systems*, New J. Phys. **20**, 033032 (2018), doi:[10.1088/1367-2630/aab03b](https://doi.org/10.1088/1367-2630/aab03b).
- [56] M. Rigol, A. Muramatsu and M. Olshanii, *Hard-core bosons on optical superlattices: Dynamics and relaxation in the superfluid and insulating regimes*, Phys. Rev. A **74**, 053616 (2006), doi:[10.1103/PhysRevA.74.053616](https://doi.org/10.1103/PhysRevA.74.053616).
- [57] E. J. Torres-Herrera and L. F. Santos, *Quench dynamics of isolated many-body quantum systems*, Phys. Rev. A **89**, 043620 (2014), doi:[10.1103/PhysRevA.89.043620](https://doi.org/10.1103/PhysRevA.89.043620).
- [58] M. Schiulaz, A. Silva and M. Müller, *Dynamics in many-body localized quantum systems without disorder*, Phys. Rev. B **91**, 184202 (2015), doi:[10.1103/PhysRevB.91.184202](https://doi.org/10.1103/PhysRevB.91.184202).
- [59] Á. M. Alhambra, J. Riddell and L. P. García-Pintos, *Time evolution of correlation functions in quantum many-body systems*, Phys. Rev. Lett. **124**, 110605 (2020), doi:[10.1103/PhysRevLett.124.110605](https://doi.org/10.1103/PhysRevLett.124.110605).
- [60] P. Taranto, F. A. Pollock, S. Milz, M. Tomamichel and K. Modi, *Quantum Markov order*, Phys. Rev. Lett. **122**, 140401 (2019), doi:[10.1103/PhysRevLett.122.140401](https://doi.org/10.1103/PhysRevLett.122.140401).
- [61] P. Taranto, S. Milz, F. A. Pollock and K. Modi, *Structure of quantum stochastic processes with finite Markov order*, Phys. Rev. A **99**, 042108 (2019), doi:[10.1103/PhysRevA.99.042108](https://doi.org/10.1103/PhysRevA.99.042108).
- [62] W. H. Zurek, *Decoherence, einselection, and the quantum origins of the classical*, Rev. Mod. Phys. **75**, 715 (2003), doi:[10.1103/RevModPhys.75.715](https://doi.org/10.1103/RevModPhys.75.715).
- [63] W. H. Zurek, *Quantum Darwinism*, Nat. Phys. **5**, 181 (2009), doi:[10.1038/nphys1202](https://doi.org/10.1038/nphys1202).
- [64] S. Milz, M. S. Kim, F. A. Pollock and K. Modi, *Completely positive divisibility does not mean Markovianity*, Phys. Rev. Lett. **123**, 040401 (2019), doi:[10.1103/PhysRevLett.123.040401](https://doi.org/10.1103/PhysRevLett.123.040401).
- [65] D. Burgarth, P. Facchi, M. Ligabò and D. Lonigro, *Hidden non-Markovianity in open quantum systems*, Phys. Rev. A **103**, 012203 (2021), doi:[10.1103/PhysRevA.103.012203](https://doi.org/10.1103/PhysRevA.103.012203).
- [66] D. Burgarth, P. Facchi, D. Lonigro and K. Modi, *Quantum non-Markovianity elusive to interventions*, Phys. Rev. A **104**, L050404 (2021), doi:[10.1103/PhysRevA.104.L050404](https://doi.org/10.1103/PhysRevA.104.L050404).
- [67] P. Figueroa-Romero, K. Modi and F. A. Pollock, *Almost Markovian processes from closed dynamics*, Quantum **3**, 136 (2019), doi:[10.22331/q-2019-04-30-136](https://doi.org/10.22331/q-2019-04-30-136).
- [68] P. Figueroa-Romero, F. A. Pollock and K. Modi, *Markovianization with approximate unitary designs*, Commun. Phys. **4**, 127 (2021), doi:[10.1038/s42005-021-00629-w](https://doi.org/10.1038/s42005-021-00629-w).