

ARTICLE OPEN



Non-Markovian memory strength bounds quantum process recoverability

Philip Taranto^{1,2}, Felix A. Pollock³ and Kavan Modi³

Generic non-Markovian quantum processes have infinitely long memory, implying an exact description that grows exponentially in complexity with observation time. Here, we present a finite memory ansatz that approximates (or recovers) the true process with errors bounded by the strength of the non-Markovian memory. The introduced memory strength is an operational quantity and depends on the way the process is probed. Remarkably, the recovery error is bounded by the smallest memory strength over all possible probing methods. This allows for an unambiguous and efficient description of non-Markovian phenomena, enabling compression and recovery techniques pivotal to near-term technologies. We highlight the implications of our results by analyzing an exactly solvable model to show that memory truncation is possible even in a highly non-Markovian regime.

npj Quantum Information (2021)7:149; <https://doi.org/10.1038/s41534-021-00481-4>

INTRODUCTION

Our ability to manipulate quantum systems underpins potential advantages over classical technologies¹. Their dynamics are often idealized as noiseless or, if unavoidable, noise is assumed to be uncorrelated. However, interactions with the environment generally perpetuate past information about the system to the future, thereby serving as a memory. Memory effects thus pervade physical evolutions, resulting in non-Markovian dynamics². The complexity of describing processes grows exponentially with memory length; therefore many simulation techniques invoke memory cutoffs³. Although several metrics have been proposed to quantify memory (and the consequences of neglecting it)⁴, most do not consider the influence of interventions, overlooking the operational reality of sequentially probed dynamics. Indeed, the impact of memory depends on how a system is controlled^{3–14}: generically, the system–environment state at any time is correlated, so an interrogation directly influences the system state and conditions the environment, both of which affect the future. Detected memory properties are thus naturally related to the interrogation method¹⁴. This operational perspective has important consequences for dynamical decoupling^{15–17}, erasure or transmission of information^{18,19}, correlated error correction and characterization^{20,21}, and (operational) quantum thermodynamics^{22–27}.

This can be illustrated with the shallow pocket model^{16,17,28,29}, comprising a qubit coupled to a continuous degree of freedom (see Fig. 1 and Supplementary Methods A). The joint dynamics induces pure-dephasing Lindblad evolution for the qubit, with exponentially decaying coherences. Non-classical correlations between any preparation and measurement similarly vanish, so the reduced dynamics forgets the initial state. However, the evolution following a σ_x unitary reverts the system to its original state; in this sense, the process displays infinitely long memory. This example highlights that although certain temporal correlations of the unperturbed system may decay rapidly, they do not account for the whole story; more generally, there exist detectable correlations between the history and future processes.

Disentangling the non-Markovian memory, carried by the environment, from the temporal correlations that result from probing the system is a challenge that has only recently been resolved through an operational framework for describing quantum stochastic processes^{10,11}. Subsequently, a notion of memory length—or quantum Markov order—was introduced, which reduces the complexity of describing processes with short-term memory by only retaining the minimal number of recent timesteps relevant to the dynamics^{12,13}. Higher-order Markov models capture all memory effects below the Markov order and are therefore more accurate than Markovian approximations. However, a missing element from this theory so far has been a method for quantifying the strength of non-Markovian memory and developing efficient approximations accordingly: truncating weak temporal-correlations will yield a more efficient description for a process without compromising the accuracy.

In this article, we construct an operational notion of memory strength for quantum processes. That is, we quantify the correlations between history and future processes with respect to an intermediate (multi-time) probing schema (see Fig. 2). Our main result links said memory strength with process recoverability: with respect to any interrogation sequence, one can approximate the process by discarding future-history correlations. If the memory strength is small for some instrument (family), then the approximate process accurately predicts expectation values for related observables, namely those in the linear span of the original instrument elements. This connection is akin to that between the conditional mutual information (CMI) and the fidelity of recovery for quantum states^{30,31}, and involves a generalization of the measured relative entropy³² to quantum stochastic processes. A corollary follows for the “do nothing” instrument—which many approximations assume—which bounds the accuracy of predicting future states. Moreover, the memory strength for an informationally complete instrument bounds the distinguishability between the actual and recovered process for any experimental protocol. Lastly, we demonstrate our results via a solvable non-Markovian model, highlighting the complex memory structures amenable to our framework.

¹Institute for Quantum Optics and Quantum Information—IQOQI Vienna, Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria. ²Atominstytut, Technische Universität Wien, 1020 Vienna, Austria. ³School of Physics & Astronomy, Monash University, Clayton, Vic 3800, Australia. ✉email: philip.taranto@oeaw.ac.at; kavan.modi@monash.edu

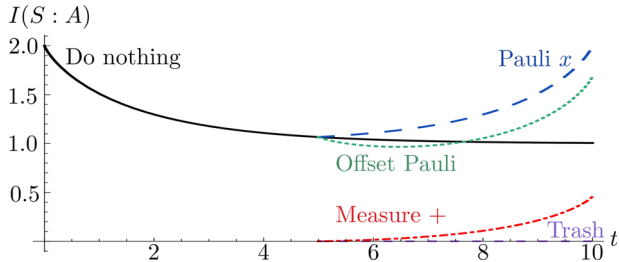


Fig. 1 Memory in shallow pocket dynamics. The mutual information $I(S:A)$ between a system and auxiliary initially in a Bell pair decays exponentially as the system undergoes shallow pocket evolution (black). However, this is not the case if an intervention is applied at t_1 ($=5$ above). We depict σ_x (blue); an offset rotation $\sqrt{0.95}\sigma_x + \sqrt{0.05}\sigma_z$ (green); measurement of $+$ in the x -basis (red); and trash-and-reprepare (purple).

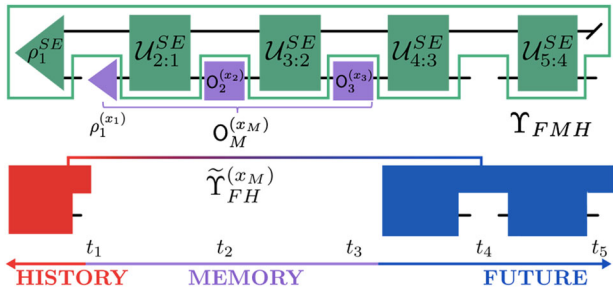


Fig. 2 Quantum stochastic process. An open quantum process with an initial system–environment state and unitary evolutions (green) can be represented as a process tensor Υ_{FMH} (outline). For each event of a memory instrument sequence, $\mathcal{J}_M = \{O_M^{(x_M)}\}$ (purple), a conditional future–history process $\tilde{\Upsilon}_{FH}^{(x_M)}$ (blue, red) results; future–history correlations evidence memory effects across $\ell := |M|$.

Classical stochastic processes

We begin by reviewing the pertinent ingredients from the theory of classical stochastic processes before turning our attention to quantum stochastic processes. A classical stochastic process over a discrete set of times $\mathcal{T} := \{t_1, \dots, t_n\}$ is described by an n -point joint probability distribution $\mathbb{P}(x_n, \dots, x_1)$. A process has finite-length memory whenever the probability of each event x_k at time $t_k \in \mathcal{T}$ only conditionally depends upon the past ℓ events

$$\mathbb{P}(x_k | x_{k-1}, \dots, x_1) = \mathbb{P}(x_k | x_{k-1}, \dots, x_{k-\ell}). \quad (1)$$

Here, ℓ , the minimum number for which Eq. (1) holds, denotes the Markov order; a Markovian process has $\ell \leq 1$. Markov order captures the complexity of characterizing a process, which grows exponentially in ℓ . Although ℓ may be large for many processes, their memory can be truncated, permitting efficient approximation. Grouping the times into three segments: the history $H = \{t_1, \dots, t_{k-\ell-1}\}$, memory $M = \{t_{k-\ell}, \dots, t_{k-1}\}$ and future $F = \{t_k, \dots, t_n\}$ (see Fig. 2), Markov order ℓ implies the conditional factorization

$$\mathbb{P}_{FH}(x_F, x_H | x_M) = \mathbb{P}_F(x_F | x_M) \mathbb{P}_H(x_H | x_M), \quad (2)$$

i.e., the future and history are conditionally independent given memory events. The Markov order condition is equivalently expressed by the vanishing classical CMI

$$I(F : H | M) := h_{FM} + h_{MH} - h_{FMH} - h_M, \quad (3)$$

where $h_x := -\sum_x \mathbb{P}_X(x) \log[\mathbb{P}_X(x)]$. The CMI is interpreted as the memory strength of the process. For processes with Markov order ℓ , $I(F : H | M) = 0$ by Eq. (2); however, in general \mathbb{P}_{FMH} does not conditionally factorize, and the CMI quantifies the correlations between F and H , given M .

The significance of approximately-finite Markov order is best encapsulated through the *recovery map*, $\mathcal{R}_{M \rightarrow FM}$, that acts only on M to approximate the correct future statistics: $\mathbb{P}_{FMH}(x_F, x_M, x_H) \simeq \mathcal{R}_{M \rightarrow FM}[\mathbb{P}_{MH}(x_M, x_H)]$, with equality holding for processes with Markov order $\leq \ell$. Intuitively, the recovery map discards conditional future–history correlations and uses the r.h.s. of Eq. (2) as an approximate description. Importantly, the recovery map approximates the process with an error that is bounded by the CMI^{30,31}, with complexity reduced to the approximate Markov order. Thus, whenever the memory is weak, the recovery map provides an accurate and efficient approximation.

Quantum stochastic processes

We now move to the realm of quantum stochastic processes. To quantify the strength of memory in quantum processes, we begin by introducing the process tensor framework, detailed in “Methods”, which generalizes Eq. (1) to the quantum domain. Consider a joint system–environment SE , with (correlated) initial state ρ_1^{SE} . The system is interrogated at t_1 and a quantum event x_1 is observed, with an associated completely positive (CP) transformation $\mathcal{O}_1^{(x_1)}$ on S . The events are described by an *instrument* $\mathcal{J}_1 = \{\mathcal{O}_1^{(x_1)}\}$, which is trace preserving (TP), i.e., $\mathcal{O}_1^{\mathcal{J}_1} := \sum_{x_1} \mathcal{O}_1^{(x_1)}$ is CPTP³³. Following this intervention, SE evolves unitarily for time $t_2 - t_1$ according to the superoperator $U_{2:1}^{SE}$. Then, S is probed at t_2 , and so on, until t_n . The probability to observe $x_1, \dots, x_n =: x_{n:1}$ using instruments $\mathcal{J}_1, \dots, \mathcal{J}_n =: \mathcal{J}_{n:1}$ is

$$\mathbb{P}(x_{n:1} | \mathcal{J}_{n:1}) = \text{tr} \left[\mathcal{O}_n^{(x_n)} U_{n:n-1}^{SE} \dots U_{2:1}^{SE} \mathcal{O}_1^{(x_1)} \rho_1^{SE} \right]. \quad (4)$$

As per Fig. 2, abstracting everything outside of experimental control defines the process itself and yields a multi-linear map from instruments to probability distributions called the *process tensor*^{10,11}. On the other hand, the instruments on S are collected to define the generalized instrument $\mathcal{J}_{n:1} = \{\mathcal{O}_{n:1}^{(x_{n:1})}\}$. A generalized instrument can be temporally correlated, although we focus on uncorrelated instrument sequences for simplicity. Intuitively, the process tensor encapsulates the uncontrollable effect of the environment, i.e., the process per se, whereas the generalized instrument represents the controllable influence. Any open quantum dynamics probed at a (causally ordered) number of times can be described by a process tensor, and any probing sequence by a generalized instrument^{8,34}. (These objects are commonly known as quantum combs and have appeared elsewhere^{8,29,33–41}.)

Both the process tensor and the instrument elements are higher-order quantum maps^{8,34,36,42} that can respectively be represented as quantum states $\Upsilon_{n:1}$ and $\{\mathcal{O}_{n:1}^{(x_{n:1})}\}$ via the Choi–Jamiołkowski isomorphism. The joint probability of realizing any sequence of events is given by the generalized Born-rule⁴³ (see Methods)

$$\mathbb{P}_{n:1}(x_{n:1} | \mathcal{J}_{n:1}) = \text{tr} \left[\mathcal{O}_{n:1}^{(x_{n:1})T} \Upsilon_{n:1} \right] =: \langle \mathcal{O}_{n:1}^{(x_{n:1})} \rangle_{\Upsilon_{n:1}}. \quad (5)$$

The process tensor encodes all probabilities for any choice of instruments and thus characterizes the process. Grouping the times as before, the quantum generalization of the l.h.s. of Eq. (2) is obtained by projecting the memory of Υ_{FMH} onto the conditioning element of $\mathcal{J}_M = \{\mathcal{O}_M^{(x_M)}\}$. The result is the *conditional future–history process* (see Fig. 2)

$$\tilde{\Upsilon}_{FH}^{(x_M)} = \text{tr}_M \left[\mathcal{O}_M^{(x_M)T} \Upsilon_{FMH} \right]. \quad (6)$$

The tilde in Eq. (6) signifies that the conditional objects are not necessarily proper process tensors, since realizing memory events post-selects the history^{35,44}; nonetheless, summing these yields a

proper process tensor $\Upsilon_{FH}^{\mathcal{J}_M} := \sum_{x_M} \tilde{\Upsilon}_{FH}^{(x_M)}$. Mirroring the classical setting, if the conditional processes are uncorrelated, i.e., $\tilde{\Upsilon}_{FH}^{(x_M)} = \Upsilon_F^{(x_M)} \otimes \tilde{\Upsilon}_H^{(x_M)}$, then the process has Markov order ℓ with respect to \mathcal{J}_M ¹² [see r.h.s. of Eq. (2)]. The causality constraint on the process tensor ensures that $\Upsilon_F^{(x_M)}$ is causally ordered for each event. This operational notion means that by applying \mathcal{J}_M , no future-history correlations are possible for any history and future instruments, i.e., no memory lasting longer than ℓ is detectable. In general, the conditional processes are correlated. However, long memory does not imply strong memory; we now show how weak correlations can be truncated for accurate and efficient approximation.

RESULTS

The first ingredient to developing finite Markov order approximations is to quantify the memory strength over a block of length ℓ , which we provide in Eq. (7). We then construct an approximate process that neglects long-term memory. We first use the knowledge from the interrogation to (partially) tomographically reconstruct the process on the memory block, ensuring that the approximate process acts correctly on all instruments lying in the span of the original one. This reconstruction typically exhibits future-history correlations, which are subsequently discarded to yield the more efficient approximate process of Eq. (10). We then bound the error of expectation values computed with the approximate process in terms of the memory strength, thereby endowing it with operational meaning. By iterating our procedure over translations of the memory block, significant savings in the complexity of description are possible.

Memory strength and recoverability

Given a particular realization x_M of instrument \mathcal{J}_M , the future and history can be more or less correlated. Such correlations can be detected by applying any choice of instruments $\mathcal{J}_F, \mathcal{J}_H$ on the conditional future-history process. Taking the supremum over such instruments provides an operational definition of memory strength, quantifying the largest detectable conditional future-history correlations. We thus define

$$\Theta(\mathcal{J}_M) := \sup_{\mathcal{J}_F, \mathcal{J}_H \in \mathbb{J}} \sum_{x_M} p(x_M | \mathcal{J}_M) I_{\mathbb{J}}(F : H | x_M)_{\Upsilon}, \quad (7)$$

where $p(x_M | \mathcal{J}_M) := \text{tr}[\tilde{\Upsilon}_{FH}^{(x_M)}]$ is the probability of observing x_M and the supremum is taken over uncorrelated instruments $\mathcal{J}_Y := \{\mathcal{O}_Y^{(x_Y)}\}$ for $Y \in \{F, H\}$ belonging to a set \mathbb{J} . We have defined the *measured CMI*

$$I_{\mathbb{J}}(F : H | x_M)_{\Gamma} := I(F : H)_{\mathbb{P}_{\Gamma}(x_F, x_H | x_M)}, \quad \text{where} \quad (8)$$

$$\mathbb{P}_{\Gamma}(x_F, x_H | x_M) := \text{tr} \left[(\mathcal{O}_F^{(x_F)} \otimes \mathcal{O}_H^{(x_H)})^{\top} \tilde{\Upsilon}_{FH}^{(x_M)} \right] \quad (9)$$

is the conditional probability distribution for any process Γ and independent future and history instruments conditioned on a fixed memory event. Intuitively, the memory strength in Eq. (7) captures the largest detectable future-history correlation, conditioned on the outcomes recorded on the memory block, aggregated to the level of \mathcal{J}_M by averaging over x_M .

Unlike classical memory, quantum memory effects depend upon the choice of probing scheme, suggesting that a universal quantum recovery map may not exist. However, we now construct a quantum process recovery map that efficiently builds up longer processes from shorter ones whenever the process is stationary. We begin with an ansatz process with finite quantum

Markov order with respect to $\mathcal{J}_M = \{\mathcal{O}^{(x_M)}\}$:

$$\underline{\Delta}_{FMH}^{\mathcal{J}_M} := \sum_{x_M} \Upsilon_F^{(x_M)} \otimes D_M^{(x_M)} \otimes \tilde{\Upsilon}_H^{(x_M)}, \quad (10)$$

where the $D_M^{(x_M)}$ are dual operators satisfying $\text{tr}[D_M^{(x_M)\top} \mathcal{O}_M^{(x_M)}] = \delta_{x_M x_M'}$ ^{7,8}. This is the tomographic representation of the process via linear inversion of the instrument outcomes, with conditional FH correlations discarded. The recovered process can exhibit correlations between H and M , M and F (as well as within each block), but not between H and F . By construction, $\underline{\Delta}_{FMH}^{\mathcal{J}_M}$ is positive on its domain, which is the span of \mathcal{J}_M . When \mathcal{J}_M is not informationally complete, i.e., does not span the full space, then $\underline{\Delta}_{FMH}^{\mathcal{J}_M}$ only approximates the original process in said subspace, which we denote by the underline. Such processes are called restricted process tensors⁴⁵ and are commonly encountered in experiments^{46–49}. Nevertheless, its action on its domain is guaranteed to reproduce the correct statistics for any multi-time observable of the form $C = \sum_x c_x \mathcal{O}^{(x)}$ with $\mathcal{O}^{(x)} = \sum_{x_M} E_{FH}^{(x, x_M)} \otimes \mathcal{O}_M^{(x_M)}$, with arbitrary $E_{FH}^{(x, x_M)}$. This is a consequence of linearity, as the observable form ensures a linear decomposition in terms of $\mathcal{O}_M^{(x_M)}$, upon which the recovered process acts correctly due to $\text{tr}[D_M^{(x_M)\top} \mathcal{O}_M^{(x_M)}] = \delta_{x_M x_M'}$.

The above ansatz is the process analogue of a quantum Markov chain state^{30,31}, which is widely studied in the context of Petz's recovery map^{50,51}. Like a quantum Markov chain state, the process above is generally not separable, as F and H can share entanglement with M , as per the example in Supplementary Methods B. However, since the dual elements here can be non-positive and cannot necessarily be decomposed into orthogonal parts, care must be taken in defining the recovery map $\mathcal{R}_{M \rightarrow FM}^{\mathcal{J}_M} : \underline{\Delta}_{MH}^{\mathcal{J}_M} \rightarrow \underline{\Delta}_{FMH}^{\mathcal{J}_M}$ (see Supplementary Methods C). The concept of the quantum recovery map is analogous to the classical case [see below Eq. (3)]. The key advantage of the quantum process recovery map is that its repeated action on the ansatz propagates the process arbitrarily far into the future with fixed ℓ -dependent complexity. If the process Υ_{FMH} has weak memory $\Theta(\mathcal{J}_M)$, the expectation value of any valid observable calculated from the recovered process $\underline{\Delta}_{FMH}^{\mathcal{J}_M}$ accurately approximates that of the original:

Theorem 1. For any multi-time observable C with support on M within the span of the elements of \mathcal{J}_M ,

$$\left| \langle C \rangle_{\Upsilon_{FMH}} - \langle C \rangle_{\underline{\Delta}_{FMH}^{\mathcal{J}_M}} \right| \leq |C| \sqrt{2\Theta(\mathcal{J}_M)}, \quad (11)$$

with $|C| := \inf_{\mathcal{J}} \sqrt{\sum_x |c_x|^2}$.

This and the following statements are proven in the Methods section. Thm. 1 is fully general inasmuch as it holds without any assumptions on the dynamics or instruments employed. The r.h.s. involves a supremum over instruments on the future and history; as the memory strength takes the form of a generalized divergence, recent numerical techniques can be used for its estimation^{52–54}. In Supplementary Methods D, we provide an easier-to-compute (and looser) bound based on the relative entropy between the original and recovered process—which foregoes the requirement for optimization—by adapting results from refs. 30,32,51,55 to first bound a generalized measured relative entropy and then the l.h.s. of Eq. (11) via Pinsker's inequality. We also prove another bound, which is tighter in some cases, by restricting to unbiased instruments satisfying $\text{tr}_M[\mathcal{O}_{FMH}^{\mathcal{J}}] \propto \mathbb{1}_{FH}$. Such instruments have the unconditional action of a completely depolarising channel, e.g., a randomly sampled Clifford gate.

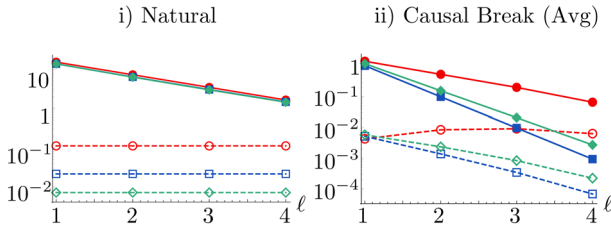


Fig. 3 Case study. We plot $|\langle C \rangle_{\Upsilon_{FMH}} - \langle C \rangle_{\Delta_{FMH}^{\mathcal{J}_M}}|$ (hollow, dashed) and the proxy memory strength based on the relative entropy between the Choi states of the true and recovered processes (see Supplementary Methods D) (solid) for (i) the identity map and (ii) a causal break. We construct a six-step process tensor in the strongly non-Markovian (red, circles), CP-divisible (blue, squares), and intermediate (green, diamonds) regimes, and consider the memory strength over ℓ applications of said instruments.

Deriving tighter bounds under various assumptions remains an open problem.

While Thm. 1 applies to multi-time observables, often one only requires the time-evolved density operator; a corollary bounds its prediction error:

Corollary 2. Let $\rho_j^{(x_M)}$ be the true density operator at any time $t_j \in F$ following outcome x_M of \mathcal{J}_M applied to the memory, and $\rho_j^{(x_M)}$ be the approximated one. Then:

$$\left\| \rho_j^{(x_M)} - \rho_j^{(x_M)} \right\|_1 \leq \sqrt{2\Theta(\mathcal{J}_M)} \quad \forall t_j \in F. \quad (12)$$

The future states result from applying identity maps at all history and future times except t_j and the instrument \mathcal{J}_M to the memory block to both the true and recovered process. See ref. 48 for a detailed analysis of using finite Markov order approximations to accurately prepare future states.

Whenever \mathcal{J}_M spans the full space, i.e., is informationally complete, then any multi-time expectation value can be accurately approximated. In this case the distinguishability, by any means, is bounded by the memory strength:

Theorem 3. For informationally complete \mathcal{J}_M , the recovered process $\Lambda_{FMH}^{\mathcal{J}_M}$ gives sensible predictions for any instrument on M and

$$\left\| \Upsilon_{FMH} - \Lambda_{FMH}^{\mathcal{J}_M} \right\| \leq \sqrt{2\Theta(\mathcal{J}_M)}, \quad (13)$$

where $\|X\| := \sup_{\mathcal{J}=\{O^{(x)}\}} \left\| \sum_x \text{tr}[O^{(x)} X \otimes \mathbb{1}] |x\rangle\langle x| \right\|_1$ generalizes the diamond norm⁵⁶ to quantum processes.

The memory strength thus provides an operationally-clear measure: if there exists some informationally complete instrument for which $\Theta(\mathcal{J}_M)$ is small, then Thm. 3 states that one can closely approximate the process for all instruments, even those for which the memory strength is large. If, additionally, $\Upsilon_{FH}^{(x_M)} \approx \Upsilon_{FH}^{(x'_M)} \forall x_M, x'_M$ in the informationally complete instrument, then the process has small memory strength for all instruments; such processes resemble similar properties to approximately finite-memory classical stochastic processes (where there is only one instrument).

Case study

Consider a qubit S coupled to another qubit E , which is cooled by an external bath. The joint evolution follows:

$$\frac{\partial \rho_t^{SE}}{\partial t} = -i\xi[\sigma_x^S \otimes \sigma_x^E, \rho_t^{SE}] + \kappa \mathcal{L}[\sigma_-^E](\rho_t^{SE}), \quad (14)$$

where the dissipator acts on E : $\mathcal{L}[\sigma_-^E](\rho_t^{SE}) := \sigma_-^E \rho_t^{SE} \sigma_+^E - \frac{1}{2} \{\sigma_+^E \sigma_-^E, \rho_t^{SE}\}$, with $\sigma_\pm^E := \sigma_x^E \pm i\sigma_y^E$. In ref. 57, it was shown that

for $\kappa^2 \geq 64\xi^2$, the process is CP-divisible, which is a common proxy for quantum Markovianity^{58,59}; however, CP-divisibility only implies an absence of some kinds of memory^{60,61}. Non-Markovianity ‘measures’ built upon two-time considerations—many of which are contradictory⁴—overlook multi-time effects. Indeed, this model contains higher-order correlations; by constructing the process tensor, we quantify the (non-vanishing) non-Markovianity for all $(\xi, \kappa) \in [0, 2] \times [0, 10]$ (see Supplementary Methods E). We then examine the memory strength in various regimes by constructing three 6-time process tensors $\Upsilon_{6;1}(\xi, \kappa)$: one CP-divisible, one strongly non-Markovian, and one intermediate, and let M range from t_2 to t_5 . We consider: (i) the identity map, which captures the *natural memory strength*; (ii) the “causal break” instrument, where the system is measured and independently reprepared in an informationally complete manner, breaking information flow through the system; and (iii) the completely noisy instrument, which replaces the state with white noise, quantifying *noise-resistant memory*¹⁴.

All processes have vanishing memory strength for the completely-noisy instrument, which can be implemented by applying random unitaries sampled from a set whose average is the depolarizing channel, providing a convenient way to bound memory⁴⁶. For cases (i) and (ii), in Fig. 3 we plot the error in the multi-time expectation value (i.e., l.h.s. of Thm. 1) and a memory strength proxy based on the relative entropy between the Choi states of the true and recovered processes (see Supplementary Methods D) which upper bounds the r.h.s. of Thm. 1. The observable C is chosen as an initial preparation, followed by doing nothing for four steps, before a final measurement. Each process displays significant memory strength for the identity instrument, indicating that unperturbed memory does not decay rapidly. In contrast, the effects of interventions are seen for the causal break here, all memory effects detected result from environmental interactions since the causal break ensures no temporal correlations can be transmitted through the system (cf. the identity instrument). The CP-divisible process exhibits negligible memory strength, the intermediate process some, and the strongly non-Markovian one stronger still. We emphasize that the unperturbed evolution is better approximated by the process recovered from the informationally complete recovery scheme than that from the identity instrument, demonstrating Thm. 3.

DISCUSSION

We have introduced the concept of memory strength for quantum stochastic processes, which is shown to bound process recoverability. Its applicability is exemplified by the case study, where we are able to accurately and efficiently reconstruct dynamics with a memory cutoff, even in a highly non-Markovian regime. We expect these tools to be broadly applicable to modern techniques for efficient simulation, where operationally motivated memory approximations with quantitative error bounds are desired, such as transfer tensor^{62–66} and machine-learning methods that either attempt to learn non-Markovian features^{67,68} or compress the memory to low-dimensional effective environments^{69,70}. Our notion of memory strength and the associated concept of recoverability will play an important role in the characterization and mitigation of noise in quantum experiments where multi-time memory effects are present^{46–49}. Of particular relevance in this direction, in ref. 46, the authors recovered the restricted process tensors on four IBM quantum computers and reported a reconstruction fidelity of order 10^{-3} . In addition, in refs. 48,49, the authors directly applied our tools to drastically reduce the number of conditional circuits required to be estimated to characterize a multi-time process on the IBM quantum computer. In particular, they demonstrate that a finite Markov order model of length $\ell = 2, 3$ suffices to prepare future target states on a five step non-Markovian process with approximately 88% and 93%

fidelity, respectively, in the presence of correlated noise. This is a huge improvement over previous ‘gold standard’ techniques using gate-set tomography to mitigate state-preparation-and-measurement errors, which gives fidelity values of around 75% in the same circumstances. Our present work lays the conceptual foundations for approximating processes with finite memory with requirements only as large as the complexity of the memory and these examples highlight the efficacy of our framework in realistic settings. Further developments in this direction will bridge the gap between efficient characterization and simulation of quantum processes with memory^{48,49,71,72}.

METHODS

Introduction to process tensor

Here we provide an introduction to the process tensor formalism; for details, see, e.g., refs. ^{8,11,34,73}.

A discrete-time classical stochastic process is characterized by the joint probability distribution \mathbb{P} over all sequences of events, $\mathbb{P}(x_n, \dots, x_1)$, where we drop the explicit time labels with the understanding that x_j represents an event at time t_j . In multi-time quantum processes, it is important to not only capture the outcome of a measurement, but also the transformation induced on the state, which together constitute an event. Thus, an interrogation of a quantum stochastic process at time t_j is described by an instrument $\mathcal{J}_j = \{\mathcal{O}_j^{(x_j)}\}$, which is a collection of completely positive (CP) maps that sum to a completely positive and trace preserving (CPTP) map. Instruments represent general quantum operations, including projective measurements, unitary transformations, and anything in between. Each CP map corresponds to a particular event realized and the fact that the maps sum to a CPTP one encodes the assumption that some event is observed. A discrete-time quantum stochastic process is uniquely described once the probability $\mathbb{P}(x_n, \dots, x_1 | \mathcal{J}_n, \dots, \mathcal{J}_1)$ for all possible events $\{x_1, \dots, x_n\}$ for all possible instruments $\{\mathcal{J}_1, \dots, \mathcal{J}_n\}$ are known. As a consequence of the linearity of mixing principle⁴³, there exists a multi-linear functional $\mathcal{T}_{n:1}$ that takes any sequence of CP maps to the correct probability distribution via $\mathbb{P}(x_n, \dots, x_1 | \mathcal{J}_n, \dots, \mathcal{J}_1) = \mathcal{T}_{n:1}[\mathcal{O}_n^{(x_n)}, \dots, \mathcal{O}_1^{(x_1)}]$, known as the process tensor¹¹. The process tensor generalizes classical stochastic processes⁹ and reproduces classical properties appropriately^{74,75}. As it encodes all detectable memory effects, it has been used to develop operationally meaningful notions of quantum Markovianity^{10,11} and memory length^{12–14}.

Since all of the CP maps constituting the instruments, as well as the process tensor itself, are linear maps, they can be represented as matrices through the Choi–Jamiołkowski isomorphism (CJI)⁸. Any map $\mathcal{O} : \mathcal{B}(\mathcal{H}_i) \rightarrow \mathcal{B}(\mathcal{H}_o)$, where $\mathcal{B}(\mathcal{X})$ denotes the space of bounded linear operators on \mathcal{X} , can be mapped isomorphically to a matrix $O \in \mathcal{B}(\mathcal{H}_o \otimes \mathcal{H}_i)$ through its action on half of an (unnormalized) maximally entangled state $\Psi^+ := \sum_{i,j} |ii\rangle\langle jj| \in \mathcal{B}(\mathcal{H}_i \otimes \mathcal{H}_i)$, i.e., $O := (\mathcal{O} \otimes \mathcal{I})[\Psi^+]$. Note that the time of the event is associated to both an input and output Hilbert space, and the Choi matrix is a supernormalized bipartite state. In this representation, the properties of CP and TP for the maps respectively translate to $O \geq 0$ and $\text{tr}_o[O] = \mathbb{1}_i$. To aid intuition, note that the output state $\rho' := \mathcal{O}[\rho]$ of the map \mathcal{O} acting on an arbitrary input state ρ is computed in the Choi picture via $\rho' = \text{tr}_i[(\mathbb{1}_o \otimes \rho_i^T)O]$. If the initial state is not subject to a (CPTP) quantum channel but instead a particular event is observed, associated to a CP map of an instrument $\mathcal{J} = \{\mathcal{O}^{(x)}\}$, then the corresponding probability is computed via

$$\mathbb{P}(x | \mathcal{J}) = \text{tr}[(\mathbb{1}_o \otimes \rho_i^T)O_{oi}^{(x)}]. \quad (15)$$

Similarly, the action of a process tensor map $\mathcal{T}_{n:1}$ on a sequence of instrument elements $\{\mathcal{O}_1^{(x_1)}, \dots, \mathcal{O}_n^{(x_n)}\}$ can be expressed in terms of a multiplication of their Choi matrices and a trace as follows⁴³

$$\mathbb{P}(x_n, \dots, x_1 | \mathcal{J}_n, \dots, \mathcal{J}_1) = \text{tr}[(O_n^{(x_n)} \otimes \dots \otimes O_1^{(x_1)})^T \Upsilon_{n:1}], \quad (16)$$

where $\Upsilon_{n:1} \in \mathcal{B}(\mathcal{H}_{n'} \otimes \mathcal{H}_{n-1'} \otimes \dots \otimes \mathcal{H}_{1'})$ is the $(2n-1)$ -partite Choi matrix of the process tensor map $\mathcal{T}_{n:1}$ and each $O_j^{(x_j)}$ is the Choi matrix of $\mathcal{O}_j^{(x_j)}$. The Choi state of the process $\Upsilon_{n:1}$ plays the role of a quantum state over time, inasmuch as it encodes all observable probability distributions

for all possible instrument sequences (just as a quantum state encodes all observable probability distributions for any choice of POVM). In the (one-time) spatial setting, $\Upsilon_1 = \mathbb{1}_o \otimes \rho_i^T$ and Eq. (16) reduces to Eq. (15). Note that we label the Hilbert spaces logically from the perspective of the experimenter (i.e., the experimenter receives a state from the process that is ‘‘input’’ into their instrument of choice, transforming it into an ‘‘output’’ state that is fed back into the process); hence, i denotes outputs of the process and o denotes inputs to the process. Whenever a process tensor acts on an instrument sequence, the degrees of freedom with the same labels (timestep and input/output) are contracted over.

The natural generalisation $\Upsilon_{n:1}$ of the CJI applied to the multilinear map $\mathcal{T}_{n:1}$ is constructed by feeding one half of an (unnormalized) maximally entangled state into the dynamics at each time⁸. More precisely, begin with the system–environment dilated dynamics shown in Fig. 2 (green), and denote the initial system–environment state by ρ and the unitary maps describing the joint evolution between times t_{j-1} and t_j by $\mathcal{U}_{j,j-1}$. Now consider $n-1$ additional maximally entangled pairs, $\Psi_{j,j'}^+$ associated to auxiliary systems $A_j \simeq S$, collectively described as $\Psi_{n-1}^+ := \bigotimes_{j=1}^{n-1} \Psi_{j,j'}^+$. Letting the unitary maps between each timestep act on the environment and one half of the appropriate auxiliary systems, i.e., $\mathcal{U}_{j,j-1} : \mathcal{B}(\mathcal{H}_{j-1'} \otimes \mathcal{H}_E) \rightarrow \mathcal{B}(\mathcal{H}_j \otimes \mathcal{H}_E)$ yields the Choi state of the process tensor

$$\Upsilon_{n:1} = \text{tr}_E[\mathcal{U}_{n,n-1} \dots \mathcal{U}_{2,1}(\rho \otimes \Psi_{n-1}^+)]. \quad (17)$$

It is straightforward (albeit arduous) to verify the correctness of Eq. (16) via direct insertion of Eq. (17). Natural generalizations of complete positivity and trace preservation to multi-time processes translate respectively to $\Upsilon_{n:1} \geq 0$ and the following hierarchy of trace conditions

$$\text{tr}_j[\Upsilon_{j:1}] = \mathbb{1}_{j-1'} \otimes \Upsilon_{j-1:1}, \quad \forall j. \quad (18)$$

Conversely, any operator satisfying the above represents some (causally-ordered) quantum dynamics inasmuch as it corresponds to a fixed underlying system–environment circuit^{11,34}.

Eq. (16) constitutes a special case of how (higher-order) quantum maps act on each other: here, we are contracting all open slots of the process tensor with an operation associated to each time in order to yield a probability distribution. More generally, it is possible to consider applying instruments to only a subset of times, yielding a conditional process defined upon the remaining times, which describes the correct behaviour of the concatenated dynamics. In other words, it contains all of the information required to compute the correct probability distribution for any instruments applied to the remaining times. To compute such an object in the Choi representation, one uses the *link product* defined in ref. ³⁶. Essentially, this amounts to restricting both the trace and the transposition in Eq. (16) to only the common Hilbert spaces associated to the relevant subset of times where the instrument is being applied.

For instance, grouping the times into history $H = \{t_1, \dots, t_k\}$, memory $M = \{t_{k+1}, \dots, t_{k+\ell}\}$ and future $F = \{t_{k+\ell+1}, \dots, t_n\}$ and choosing an instrument $\mathcal{J}_M = \{\mathcal{O}_M^{(x_M)}\}$ on the memory block, the conditional future–history process that occurs given any particular event sequence $O_M^{(x_M)}$ on M alone is

$$\tilde{\Upsilon}_{FH}^{(x_M)} = \text{tr}_M[\mathcal{O}_M^{(x_M)T} \Upsilon_{FMH}]. \quad (19)$$

Such a conditional process is generically correlated across F and H ; however, if it is of tensor product form $\tilde{\Upsilon}_{FH}^{(x_M)} = \Upsilon_F^{(x_M)} \otimes \tilde{\Upsilon}_H^{(x_M)}$ for each event x_M of the instrument \mathcal{J}_M , the process has Markov order $\ell := |M|$ with respect to said instrument¹². (In general, each $O_M^{(x_M)}$ may act on only a subspace of M , with the history and future retaining the rest, to yield $\tilde{\Upsilon}_{FM_H M_H H}^{(x_M)}$, where M_F and M_H can depend on x_M (see Supplementary Methods B); for brevity, we absorb these into F and H .)

Lastly, a Markovian process corresponds to one for which the process tensor has the specific tensor product structure of an uncorrelated sequence of CPTP maps $\{\Lambda_{j,j-1'}\}$ connecting adjacent timesteps, and an initial quantum state $\rho_{1'}$ ¹⁰

$$\Upsilon_{n:1}^{\text{Markov}} = \Lambda_{n':n-1'} \otimes \Lambda_{n-1':n-2'} \otimes \Lambda_{2':1'} \otimes \rho_{1'}. \quad (20)$$

Preliminaries

To begin with, we introduce the following definition:

Definition 4. (Instrument relative entropy). For any family of instruments \mathbb{J} and process tensors Υ, Γ ,

$$S_{\mathbb{J}}(\Upsilon \parallel \Gamma) := \sup_{\mathcal{J} \in \mathbb{J}} S(\mathcal{P}_{\mathcal{J}}[\Upsilon] \parallel \mathcal{P}_{\mathcal{J}}[\Gamma]), \quad (21)$$

where $S(A \parallel B) := \text{tr}[A(\log A - \log B)]$ is the quantum relative entropy and $\mathcal{P}_{\mathcal{J}}$ is a CP map from process tensors to classical pointer states, whose elements form a probability distribution over outcomes of the instrument $\mathcal{J} = \{O^{(x)}\}$

$$\mathcal{P}_{\mathcal{J}}[\Upsilon] := \sum_x \text{tr}[O^{(x)\top} \Upsilon] |x\rangle \langle x|. \quad (22)$$

Proposition 5. For any Υ_{FMH} , \mathcal{J}_M and $\underline{\Delta}_{FMH}^{\mathcal{J}_M}$ as defined in Eq. (10),

$$\Theta(\mathcal{J}_M) = S_{\mathbb{J} \cap \text{span}(\mathcal{J}_M)}(\Upsilon_{FMH} \parallel \underline{\Delta}_{FMH}^{\mathcal{J}_M}), \quad (23)$$

with $\Theta(\mathcal{J}_M)$ taken to be the measured CMI, $\mathbb{J} \cap \text{span}(\mathcal{J}_M)$ a family of instruments whose elements have support on M only in the linear span of the elements of \mathcal{J}_M .

Proof. Consider first the measured conditional probability distributions for a fixed memory instrument \mathcal{J}_M and arbitrary (independent) $\mathcal{J}_F, \mathcal{J}_H$ arising from the process tensor Υ_{FMH} and the recovered process $\underline{\Delta}_{FMH}^{\mathcal{J}_M}$, which are respectively given by:

$$\mathbb{P}_{\Upsilon_{FMH}}(x_F, x_H | x_M) = \text{tr}[(O_F^{(x_F)} \otimes O_H^{(x_H)})^\top \tilde{\Upsilon}_{FH}^{(x_M)}] \quad (24)$$

and

$$\mathbb{P}_{\underline{\Delta}_{FMH}^{\mathcal{J}_M}}(x_F, x_H | x_M) = \text{tr}[(O_F^{(x_F)} \otimes O_H^{(x_H)})^\top \tilde{\Delta}_{FH}^{(x_M)}]. \quad (25)$$

Here, we have defined $\tilde{\Delta}_{FH}^{(x_M)} := \text{tr}_M[O_M^{(x_M)\top} \underline{\Delta}_{FMH}^{\mathcal{J}_M}]$, which, by construction, factorizes as

$$\begin{aligned} \mathbb{P}_{\underline{\Delta}_{FMH}^{\mathcal{J}_M}}(x_F, x_H | x_M) &= \text{tr}[(O_F^{(x_F)} \otimes O_H^{(x_H)})^\top (\Upsilon_F^{(x_F)} \otimes \tilde{\Upsilon}_H^{(x_H)})] \\ &= \mathbb{P}_{\Upsilon_{FMH}}(x_F | x_M) \mathbb{P}_{\Upsilon_{FMH}}(x_H | x_M), \end{aligned} \quad (26)$$

since $\tilde{\Delta}_{FH}^{(x_M)} = \Upsilon_F^{(x_F)} \otimes \tilde{\Upsilon}_H^{(x_H)}$.

With this, we can express the mutual information in the correlated distribution $I(F : H)_{\mathbb{P}_{\Upsilon_{FMH}}(x_F, x_H | x_M)} = I_{\mathbb{J}}(F : H)_{\Upsilon}$ [see Eq. (8)] in terms of the relative entropy between said distribution and the uncorrelated one arising from measurements on the recovered process, i.e.,

$$I_{\mathbb{J}}(F : H)_{\Upsilon} = S[\mathbb{P}_{\Upsilon_{FMH}}(x_F, x_H | x_M) \parallel \mathbb{P}_{\underline{\Delta}_{FMH}^{\mathcal{J}_M}}(x_F, x_H | x_M)]. \quad (27)$$

Thus, beginning with the definition of Eq. (7), we have

$$\Theta(\mathcal{J}_M) = \sup_{\mathcal{J}_F, \mathcal{J}_H \in \mathbb{J}} \left(\sum_{x_M} p(x_M | \mathcal{J}_M) I_{\mathbb{J}}(F : H | x_M)_{\Upsilon} \right) = S_{\mathbb{J}_{FMH}}(\Upsilon_{FMH}^{\mathcal{J}_M} \parallel \underline{\Delta}_{FMH}^{\mathcal{J}_M}). \quad (28)$$

Here \mathbb{J}_{FMH} is the original set of uncorrelated instruments \mathbb{J} on FH , combined with a POVM on the pointer space m , i.e., the supremum in the measured relative entropy is taken over $\mathcal{J} = \{O_{FMH}^{(x)}\} \in \mathbb{J}_{FMH}$ of the form

$$O_{FMH}^{(x)} = \sum_{x_M} E_{FH}^{(x, x_M)} \otimes |x_M\rangle \langle x_M|_m, \quad (29)$$

where $E_{FH}^{(x, x_M)}$ can be any operator, $O_{FMH}^{\mathcal{J}} = \sum_x O_{FMH}^{(x)}$ satisfies the relevant trace conditions on the FH part and $\text{tr}_{FH}[O_{FMH}^{\mathcal{J}}] = D_{FH}^j \mathbb{1}_m$. Since, for $\mathcal{J}_M = \{O_M^{(x_M)}\}$,

$$\text{tr}_m[|x_M\rangle \langle x_M|_m \Gamma_{FMH}^{\mathcal{J}_M}] = \text{tr}_M \left[O_M^{(x_M)} \sum_{y_M} \Gamma_{FH}^{(y_M)} \otimes D_M^{(y_M)} \right],$$

with $\Gamma \in \{\Upsilon, \Lambda\}$ and $\{D_M^{(x_M)}\}$ the dual set to \mathcal{J}_M , we have

$$S_{\mathbb{J}_{FMH}}(\Upsilon_{FMH}^{\mathcal{J}_M} \parallel \underline{\Delta}_{FMH}^{\mathcal{J}_M}) = S_{\mathbb{J} \cap \text{span}(\mathcal{J}_M)}(\Upsilon_{FMH} \parallel \underline{\Delta}_{FMH}^{\mathcal{J}_M}),$$

with $\underline{\Delta}_{FMH}^{\mathcal{J}_M} = \sum_{x_M} \Upsilon_F^{(x_M)} \otimes D_M^{(x_M)} \otimes \tilde{\Upsilon}_H^{(x_M)}$. Hence, the claim is asserted. \square
We are now in a position to prove our main results.

Proof of Thm. 1

First, we note that, for any set of instruments \mathbb{J} and process tensors Υ and Γ ,

$$S_{\mathbb{J}}(\Upsilon \parallel \Gamma) = \sup_{\mathcal{J} \in \mathbb{J}} \sum_x p_x^{\mathcal{J}} (\log p_x^{\mathcal{J}} - \log q_x^{\mathcal{J}}), \quad (30)$$

with $p_x^{\mathcal{J}} := \text{tr}[O^{(x)\top} \Upsilon_{FMH}]$ and $q_x^{\mathcal{J}} := \text{tr}[O^{(x)\top} \underline{\Delta}_{FMH}^{\mathcal{J}_M}]$ the probabilities associated with the instrument $\mathcal{J} = \{O^{(x)}\}$. We can then use Pinsker's inequality to write

$$\begin{aligned} S_{\mathbb{J}}(\Upsilon \parallel \Gamma) &\geq \frac{1}{2} \sup_{\mathcal{J} \in \mathbb{J}} (\sum_x |p_x^{\mathcal{J}} - q_x^{\mathcal{J}}|)^2 \\ &= \frac{1}{2} \left(\sup_{\mathcal{J} \in \mathbb{J}} \sum_x |\text{tr}[O^{(x)\top} (\Upsilon - \Gamma)]| \right)^2. \end{aligned} \quad (31)$$

Any multi-time operator C can be decomposed in terms of the elements of a single instrument $\mathcal{J} = \{O^{(x)}\} \in \mathbb{J}$, as long as those elements span a sufficiently large space. That is, $C = \sum_x c_x^{\mathcal{J}} O^{(x)}$ with $c_x^{\mathcal{J}} \in \mathbb{C}$; in general, the norm $|C|_{\mathcal{J}} := \sqrt{\sum_x |c_x^{\mathcal{J}}|^2}$ will vary with the instrument involved in the decomposition. We therefore have, using the Cauchy–Schwarz inequality

$$\begin{aligned} |\text{tr}[C \Xi]| &\leq \inf_{\mathcal{J} \in \mathbb{J}} \left(\sqrt{\sum_x |c_x^{\mathcal{J}}|^2} \sqrt{\sum_x |\text{tr}[O^{(x)\top} \Xi]|^2} \right) \\ &\leq \inf_{\mathcal{J} \in \mathbb{J}} (|C|_{\mathcal{J}} \sum_x |\text{tr}[O^{(x)\top} \Xi]|) \\ &\leq \inf_{\mathcal{J} \in \mathbb{J}} (|C|_{\mathcal{J}} \sup_{\mathcal{J} \in \mathbb{J}} (\sum_x |\text{tr}[O^{(x)\top} \Xi]|)), \end{aligned} \quad (32)$$

for any operator Ξ . Choosing $\Xi = \Upsilon_{FMH} - \underline{\Delta}_{FMH}^{\mathcal{J}_M}$ and restricting the set \mathbb{J} to $\mathbb{J} \cap \text{span}(\mathcal{J}_M)$, such that $\text{tr}[C \Xi] = \langle C \rangle_{\Upsilon_{FMH}} - \langle C \rangle_{\underline{\Delta}_{FMH}^{\mathcal{J}_M}}$ and combining Eqs. (31) and (32) leads to the bound

$$|\langle C \rangle_{\Upsilon_{FMH}} - \langle C \rangle_{\underline{\Delta}_{FMH}^{\mathcal{J}_M}}| \leq |C| \sqrt{2 S_{\mathbb{J} \cap \text{span}(\mathcal{J}_M)}(\Upsilon_{FMH} \parallel \underline{\Delta}_{FMH}^{\mathcal{J}_M})}, \quad (33)$$

where $|C| = \min_{\mathcal{J} \in \mathbb{J} \cap \text{span}(\mathcal{J}_M)} |C|_{\mathcal{J}}$ (equivalent to the definition given in Thm. 1).

Invoking Prop. 5 proves Thm. 1. \square

Cor. 2 follows directly as shown below.

Proof of Cor. 2

Choose $C = P_j \otimes \Psi^+ \otimes j^{k-2} \otimes O_M^{(x_M)} \otimes \Psi^+ \otimes j^{k-\ell-2}$, with $\Psi^+ := \sum_{\alpha\beta} |\alpha\rangle \langle \beta|$ the Choi state of the identity map and P_j a projector. Then $|C| = 1$, since C is an element of the instrument where the system is left to freely evolve on H , \mathcal{J}_M is applied, it again freely evolves to time t_j and then the POVM $\{P_j, 1 - P_j\}$ is applied. It follows that $\langle C \rangle_{\Upsilon} = \text{tr}[P_j \rho_j^{(x_M)}]$, with $\rho_j^{(x_M)} := \text{tr}_{FMH \setminus j} \left[(\Psi_H^+ \otimes O_M^{(x_M)} \otimes \Psi_{FV}^+) \Upsilon_{FMH} \right]$ the state, at time t_j , of the system undergoing the process specified by Υ_{FMH} , acted on by \mathcal{J}_M with outcome x_M occurring, and with no other active interventions. Similarly, the predicted state is $\rho_j^{(x_M)} := \text{tr}_{FMH \setminus j} \left[(\Psi_H^+ \otimes O_M^{(x_M)} \otimes \Psi_{FV}^+) \underline{\Delta}_{FMH}^{\mathcal{J}_M} \right]$. Here, Ψ^+ denotes the Choi state of the identity map and Ψ_X^+ is shorthand for a sequence of identity maps applied at all times in the block X , and j corresponds to a subspace of the future Hilbert space associated to time t_j . The l.h.s. of Eq. (11) of the main text then reduces to $|\text{tr}[P_j (\rho_j^{(x_M)} - \rho_j^{(x_M)})]|$; since the bound must be true for any P_j , it must be true for the one for which the l.h.s. is largest; i.e., $\sup_{P_j} |\text{tr}[P_j (\rho_j^{(x_M)} - \rho_j^{(x_M)})]| = \|\rho_j^{(x_M)} - \rho_j^{(x_M)}\|_1$ is bounded. \square

For informationally complete instruments, a combination of the results derived above leads to Thm. 3.

Proof of Thm. 3

When \mathcal{J}_M is informationally complete, $\underline{\Delta}_{FMH}^{\mathcal{J}_M}$ is a full process tensor and any instrument can be applied to it, since $\mathbb{J} \cap \text{span}(\mathcal{J}_M) = \mathbb{J}$ by definition.

Therefore, we can use Eq. (31), along with Prop. 5, to write:

$$\left(\sup_{\mathcal{J} \in \mathcal{J}} \sum_x \left| \text{tr} \left[O^{(x)} (\Upsilon_{FMH} - \Delta_{FMH}^{\mathcal{J}_M}) \right] \right| \right)^2 \leq 2\Theta(\mathcal{J}_M) \quad (34)$$

The square root of the l.h.s. of this equation is the generalized diamond distance $\|\Upsilon_{FMH} - \Delta_{FMH}^{\mathcal{J}_M}\|$ with $\|X\|_{\diamond}$ defined in Thm. 3. Equation (13) and hence Thm. 3 follows. \square

DATA AVAILABILITY

The data that support the findings of this study are available from P.T. upon reasonable request.

Received: 16 March 2021; Accepted: 2 September 2021;

Published online: 12 October 2021

REFERENCES

- Nielsen, M. and Chuang, I. *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
- Van Kampen, N. *Stochastic Processes in Physics and Chemistry* (Elsevier, New York, 2011).
- de Vega, I. & Alonso, D. Dynamics of non-Markovian open quantum systems. *Rev. Mod. Phys.* **89**, 015001 (2017).
- Li, L., Hall, M. J. W. & Wiseman, H. M. Concepts of quantum non-Markovianity: a hierarchy. *Phys. Rep.* **759**, 1–51 (2018).
- Modi, K. Preparation of states in open quantum mechanics. *Open Syst. Inf. Dyn.* **18**, 253–260 (2011).
- Modi, K. Operational approach to open dynamics and quantifying initial correlations. *Sci. Rep.* **2**, 581 (2012).
- Modi, K., Rodríguez-Rosario, C. A. & Aspuru-Guzik, A. Positivity in the presence of initial system-environment correlation. *Phys. Rev. A* **86**, 064102 (2012).
- Milz, S., Pollock, F. A. & Modi, K. An introduction to operational quantum dynamics. *Open Syst. Inf. Dyn.* **24**, 1740016 (2017).
- Milz, S., Sakuldee, F., Pollock, F. A. & Modi, K. Kolmogorov extension theorem for (quantum) causal modelling and general probabilistic theories. *Quantum* **4**, 255 (2020).
- Pollock, F. A., Rodríguez-Rosario, C., Frauenheim, T., Paternostro, M. & Modi, K. Operational Markov condition for quantum processes. *Phys. Rev. Lett.* **120**, 040405 (2018a).
- Pollock, F. A., Rodríguez-Rosario, C., Frauenheim, T., Paternostro, M. & Modi, K. Non-Markovian quantum processes: complete framework and efficient characterization. *Phys. Rev. A* **97**, 012127 (2018b).
- Taranto, P., Pollock, F. A., Milz, S., Tomamichel, M. & Modi, K. Quantum Markov order. *Phys. Rev. Lett.* **122**, 140401 (2019a).
- Taranto, P., Milz, S., Pollock, F. A. & Modi, K. Structure of quantum stochastic processes with finite Markov order. *Phys. Rev. A* **99**, 042108 (2019b).
- Taranto, P. Memory effects in quantum processes. *Int. J. Quantum Inf.* **18**, 1941002 (2020).
- Viola, L., Knill, E. & Lloyd, S. Dynamical decoupling of open quantum systems. *Phys. Rev. Lett.* **82**, 2417 (1999).
- Arenz, C., Hillier, R., Fraas, M. & Burgarth, D. Distinguishing decoherence from alternative quantum theories by dynamical decoupling. *Phys. Rev. A* **92**, 022102 (2015).
- Arenz, C., Burgarth, D., Facchi, P. & Hillier, R. Dynamical decoupling of unbounded Hamiltonians. *J. Math. Phys.* **59**, 032203 (2018).
- Kretschmann, D. & Werner, R. F. Quantum channels with memory. *Phys. Rev. A* **72**, 062323 (2005).
- Caruso, F., Giovannetti, V., Lupo, C. & Mancini, S. Quantum channels and memory effects. *Rev. Mod. Phys.* **86**, 1203 (2014).
- Morris, J., Pollock, F. A. & Modi, K. Non-Markovian memory in IBMQX4. Preprint at [arXiv https://arxiv.org/abs/1902.07980](https://arxiv.org/abs/1902.07980) (2019).
- Figueroa-Romero, P., Modi, K., Stace, T. M. & Hsieh, M.-H. Randomised benchmarking for non-Markovian noise. Preprint at [arXiv https://arxiv.org/abs/2107.05403](https://arxiv.org/abs/2107.05403) (2021a).
- Figueroa-Romero, P., Modi, K. & Pollock, F. A. Almost Markovian processes from closed dynamics. *Quantum* **3**, 136 (2019).
- Strasberg, P. Operational approach to quantum stochastic thermodynamics. *Phys. Rev. E* **100**, 022127 (2019a).
- Figueroa-Romero, P., Modi, K. & Pollock, F. A. Equilibration on average in quantum processes with finite temporal resolution. *Phys. Rev. E* **102**, 032144 (2020).

- Strasberg, P. & Winter, A. Stochastic thermodynamics with arbitrary interventions. *Phys. Rev. E* **100**, 022135 (2019).
- Strasberg, P. Repeated interactions and quantum stochastic thermodynamics at strong coupling. *Phys. Rev. Lett.* **123**, 180604 (2019b).
- Figueroa-Romero, P., Pollock, F. A. & Modi, K. Markovianization with approximate unitary designs. *Commun. Phys.* **4**, 127 (2021b).
- Lindblad, G. *Response of Markovian and non-Markovian Quantum Stochastic Systems to Time-Dependent Forces*. (unpublished) (Stockholm, 1980).
- Accardi, L., Frigerio, A. & Lewis, J. T. Quantum stochastic processes. *Publ. Res. Inst. Math. Sci.* **18**, 97–133 (1982).
- Fawzi, O. & Renner, R. Quantum conditional mutual information and approximate Markov chains. *Commun. Math. Phys.* **340**, 575–611 (2015).
- Sutter, D., Fawzi, O. & Renner, R. Universal recovery map for approximate Markov chains. *Proc. R. Soc. A* **472**, 20150623 (2016).
- Piani, M. Relative entropy of entanglement and restricted measurements. *Phys. Rev. Lett.* **103**, 160504 (2009).
- Lindblad, G. Non-Markovian quantum stochastic processes and their entropy. *Commun. Math. Phys.* **65**, 281–294 (1979).
- Chiribella, G., D'Ariano, G. M. & Perinotti, P. Theoretical framework for quantum networks. *Phys. Rev. A* **80**, 022339 (2009).
- Chiribella, G., D'Ariano, G. M. & Perinotti, P. Transforming quantum operations: quantum supermaps. *EPL* **83**, 30004 (2008a).
- Chiribella, G., D'Ariano, G. M. & Perinotti, P. Quantum circuit architecture. *Phys. Rev. Lett.* **101**, 060401 (2008b).
- Hardy, L. The operator tensor formulation of quantum theory. *Philos. Trans. R. Soc. A* **370**, 3385–3417 (2012).
- Oreshkov, O., Costa, F. & Brukner, Č. Quantum correlations with no causal order. *Nat. Commun.* **3**, 1092 (2012).
- Costa, F. & Shrapnel, S. Quantum causal modelling. *New J. Phys.* **18**, 063032 (2016).
- Hardy, L. Operational general relativity: possibilistic, probabilistic, and quantum. Preprint at [arXiv https://arxiv.org/abs/1608.06940](https://arxiv.org/abs/1608.06940) (2016).
- Coecke, B. & Kissinger, A. *Picturing Quantum Processes: A First Course in Quantum Theory and Diagrammatic Reasoning* (Cambridge University Press, 2017) <https://doi.org/10.1017/9781316219317>.
- Bisio, A. & Perinotti, P. Theoretical framework for higher-order quantum theory. *Proc. R. Soc. A* **475**, 20180706 (2019).
- Shrapnel, S., Costa, F. & Milburn, G. Updating the Born rule. *New J. Phys.* **20**, 053010 (2018a).
- Milz, S., Pollock, F. A., Le, T. P., Chiribella, G. & Modi, K. Entanglement, non-Markovianity, and causal non-separability. *New J. Phys.* **20**, 033033 (2018a).
- Milz, S., Pollock, F. A. & Modi, K. Reconstructing non-Markovian quantum dynamics with limited control. *Phys. Rev. A* **98**, 012108 (2018b).
- White, G. A. L., Hill, C. D., Pollock, F. A., Hollenberg, L. C. L. & Modi, K. Demonstration of non-Markovian process characterisation and control on a quantum processor. *Nat. Commun.* **11**, 6301 (2020).
- Guo, Y. et al. Experimental demonstration of instrument-specific quantum memory effects and non-markovian process recovery for common-cause processes. *Phys. Rev. Lett.* **126**, 230401 (2021).
- White, G.A. L., Pollock, F. A., Hollenberg, L.C. L., Modi, K. & Hill, C.D. Non-Markovian quantum process tomography. Preprint at [arXiv https://arxiv.org/abs/2106.11722](https://arxiv.org/abs/2106.11722) (2021a).
- White, G.A. L., Pollock, F. A., Hollenberg, L.C. L., Hill, C.D. & Modi, K. Diagnosing temporal quantum correlations: compressed non-Markovian calipers. Preprint at [arXiv https://arxiv.org/abs/2107.13934](https://arxiv.org/abs/2107.13934) (2021b).
- Petz, D. Sufficient subalgebras and the relative entropy of states of a von Neumann algebra. *Commun. Math. Phys.* **105**, 123–131 (1986).
- Petz, D. Monotonicity of quantum relative entropy revisited. *Rev. Math. Phys.* **15**, 79–91 (2003).
- Fawzi, H. & Fawzi, O. Efficient optimization of the quantum relative entropy. *J. Phys. A* **51**, 154003 (2018).
- Fawzi, H., Saunderson, J. & Parrilo, P. A. Semidefinite approximations of the matrix logarithm. *Found. Comput. Math.* **19**, 259–296 (2019).
- Fang, K. & Fawzi, H. Geometric Rényi divergence and its applications in quantum channel capacities. *Commun. Math. Phys.* **384**, 1615–1677 (2021).
- Paulsen, V. *Completely Bounded Maps and Operator Algebras* (Cambridge University Press, 2002).
- Gilchrist, A., Langford, N. K. & Nielsen, M. A. Distance measures to compare real and ideal quantum processes. *Phys. Rev. A* **71**, 062310 (2005).
- Pang, S., Brun, T. A. & Jordan, A. N. Abrupt transitions between Markovian and non-Markovian dynamics in open quantum systems. Preprint at [arXiv https://arxiv.org/abs/1712.10109](https://arxiv.org/abs/1712.10109) (2017).
- Rivas, A., Huelga, S. F. & Plenio, M. B. Quantum non-Markovianity: characterization, quantification and detection. *Rep. Prog. Phys.* **77**, 094001 (2014).

59. Breuer, H.-P., Laine, E.-M., Piilo, J. & Vacchini, B. Colloquium: non-Markovian dynamics in open quantum systems. *Rev. Mod. Phys.* **88**, 021002 (2016).
60. Milz, S., Kim, M. S., Pollock, F. A. & Modi, K. Completely positive divisibility does not mean Markovianity. *Phys. Rev. Lett.* **123**, 040401 (2019).
61. Hsieh, Y.-Y., Su, Z.-Y. & Goan, H.-S. Non-Markovianity, information backflow, and system-environment correlation for open-quantum-system processes. *Phys. Rev. A* **100**, 012120 (2019).
62. Cerrillo, J. & Cao, J. Non-Markovian dynamical maps: numerical processing of open quantum trajectories. *Phys. Rev. Lett.* **112**, 110401 (2014).
63. Rosenbach, R., Cerrillo, J., Huelga, S. F., Cao, J. & Plenio, M. B. Efficient simulation of non-Markovian system-environment interaction. *New J. Phys.* **18**, 023035 (2016).
64. Kananenka, A. A., Hsieh, C.-Y., Cao, J. & Geva, E. Accurate long-time mixed quantum-classical liouville dynamics via the transfer tensor method. *J. Phys. Chem. Lett.* **7**, 4809–4814 (2016).
65. Pollock, F. A. & Modi, K. Tomographically reconstructed master equations for any open quantum dynamics. *Quantum* **2**, 76 (2018).
66. Jørgensen, M. R. & Pollock, F. A. Exploiting the causal tensor network structure of quantum processes to efficiently simulate non-Markovian path integrals. *Phys. Rev. Lett.* **123**, 240602 (2019).
67. Banchi, L., Grant, E., Rocchetto, A. & Severini, S. Modelling non-Markovian quantum processes with recurrent neural networks. *New J. Phys.* **20**, 123030 (2018).
68. Shrapnel, S., Costa, F. & Milburn, G. Quantum Markovianity as a supervised learning task. *Int. J. Quantum Inf.* **16**, 1840010 (2018b).
69. Luchnikov, I. A., Vintskevich, S. V., Grigoriev, D. A. & Filippov, S. N. Machine learning non-Markovian quantum dynamics. *Phys. Rev. Lett.* **124**, 140502 (2020).
70. Guo, C., Modi, K. & Poletti, D. Tensor-network-based machine learning of non-Markovian quantum processes. *Phys. Rev. A* **102**, 062414 (2020).
71. Luchnikov, I. A., Vintskevich, S. V. & Filippov, S. N. Dimension truncation for open quantum systems in terms of tensor networks. Preprint at *arXiv* <https://arxiv.org/abs/1801.07418> (2018).
72. Luchnikov, I. A., Vintskevich, S. V., Ouerdane, H. & Filippov, S. N. Simulation complexity of open quantum dynamics: connection with tensor networks. *Phys. Rev. Lett.* **122**, 160401 (2019).
73. Milz, S. & Modi, K. Quantum stochastic processes and quantum non-Markovian phenomena. *PRX Quantum* **2**, 030201 (2021).
74. Strasberg, P. & Diaz, M. G. Classical quantum stochastic processes. *Phys. Rev. A* **100**, 022120 (2019).
75. Milz, S. et al. When is a non-Markovian quantum process classical? *Phys. Rev. X* **10**, 041049 (2020).

ACKNOWLEDGEMENTS

We thank Top Notoh, Joshua Morris, and Simon Milz for discussions. P.T. was supported by the Australian Government Research Training Program Scholarship, the J.L. William Scholarship, and the Austrian Science Fund (FWF): Y879-N27

(START project). K.M. is supported through the Australian Research Council Future Fellowship FT160100073. This research was supported in part by the National Science Foundation under Grant no. NSF PHY-1748958.

AUTHOR CONTRIBUTIONS

All authors contributed to conception of the project, proved and interpreted the results, and discussed and edited the paper. P.T. wrote the text and performed all numerical simulations for the case study.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41534-021-00481-4>.

Correspondence and requests for materials should be addressed to Philip Taranto or Kavan Modi.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021