Hidden Quantum Memory: Is Memory There When Somebody Looks?

Philip Taranto[1 2 3], Thomas J. Elliott[4 5], and Simon Milz[6 3 7]

[1]Department of Physics, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo City, Tokyo 113-0033, Japan

[2]Atominstitut, Technische Universität Wien, 1020 Vienna, Austria

[3]Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria

[4]Department of Physics & Astronomy, University of Manchester, Manchester M13 9PL, United Kingdom

[5]Department of Mathematics, University of Manchester, Manchester M13 9PL, United Kingdom

[6]School of Physics, Trinity College Dublin, Dublin 2, Ireland

[7]Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Vienna, Austria

April 26, 2023

In classical physics, memoryless dynamics and Markovian statistics are one and the same. This is not true for quantum dynamics, first and foremost because quantum measurements are invasive. Going beyond measurement invasiveness, here we derive a novel distinction between classical and quantum processes, namely the possibility of *hidden quantum memory*. While Markovian statistics of classical processes can always be reproduced by a memoryless dynamical model, our main result shows that this is not true in quantum mechanics: We first provide an example of quantum non-Markovianity whose manifestation depends on whether or not a previous measurement is performed—an impossible phenomenon for memoryless dynamics; we then strengthen this result by demonstrating statistics that are Markovian independent of how they are probed, but are nonetheless *still* incompatible with memoryless quantum dynamics. Thus, we establish the existence of Markovian statistics gathered by probing a quantum process that nevertheless *fundamentally* require memory for their creation.

Our ability to understand and control memory effects in the evolution of open quantum systems is becoming increasingly important as technology allows us to manipulate interactions with increasing levels of speed, precision and complexity [1, 2]. Control over memory can be advantageous in various tasks, such as creating, manipulating and preserving coherences and correlations [3, 4], simulating complex dynamics [5–18], implementing randomised benchmarking and error correction [19–21], performing optimal dynamical decoupling [22–24], designing quantum circuit architectures [25–31], and improving the efficiency of thermodynamic machines [32–35].

One has no choice but to account for complex noise and memory effects when modelling realistic dynamical systems, as no system is truly isolated; in general, the environment stores information about the past and propagates it in time, leading to memory effects that manifest themselves as complex multi-time correlations [36–40]. A special case of open dynamics are memoryless dynamics, for which the environment retains *no* memory of its previous interactions with the system. Such dynamics have been studied extensively due to their accuracy in many practically relevant situations and their exponentially reduced complexity from the general scenario. Both in the classical and quantum setting, such efficient descriptions arise by way of (time-local) master equations that efficiently simulate the system dynamics [41–43]; in practice, the assumption of memorylessness is often made for simplicity and describes many 'real-world' scenarios with a high degree of accuracy [44–47].

Philip Taranto: philipguy.taranto@phys.s.u-tokyo.ac.jp

However, experimentally determining that a quantum process is memoryless requires full process tomography, which necessitates a myriad of complex sequential measurements and has consequently only been done for low-dimensional/few timestep cases [28–30, 48]. A more tractable situation is the sequential probing of a fixed observable via sharp, projective measurements. In this case, memoryless quantum processes—like their classical counterparts—lead to Markovian statistics, i.e., statistics where the future is conditionally independent of the past. Thus, at first glance, memorylessness of the dynamics seems to manifest on the experimental level identically for classical and quantum processes. However, this is not the case; for one, quantum measurements of any observable are generally invasive, leading to inconsistent (sub-)statistics [49] and the violation of Leggett-Garg inequalities [50–52]. In contradistinction, measuring an observable in the classical world can be done non-invasively. Moreover—beyond measurement invasiveness—here we demonstrate that quantum processes can yield Markovian statistics *that fundamentally require memory* for their creation.

More concretely, in classical physics, *any* Markovian statistics can be described by a memoryless dynamical model (i.e., as emerging from a sequence of independent stochastic matrices). In the quantum case, measuring a fixed observable no longer constitutes a tomographically complete procedure; consequently, the existence of processes with memory that nonetheless lead to Markovian statistics when said observable is probed is not surprising *per se* and has been demonstrated [53–55]. This phenomenon notwithstanding, for any (quantum) experiment that yields Markovian statistics, it is reasonable to believe that there *always exists* some memoryless quantum dynamics that faithfully reproduces the observed statistics. Such a description is known as the *quantum regression formula* **(QRF)** [44, 45, 56] and is a widely used assumption that links operational quantities—namely, recorded statistics—to dynamical ones—namely, a model of the underlying dynamics.

Here, we ask the question: *Can Markovian statistics always be faithfully reproduced by a memoryless dynamical model?* In other words, can the QRF *always* be employed to describe Markovian statistics? Our main result, perhaps surprisingly, answers this in the negative. Since this contradicts the counterpart answer within classical physics (i.e., for sharp measurements of a given observable), we thus uncover a new type of genuinely quantum phenomenon: *Hidden quantum memory.* This observation makes quantum memory an emergent phenomenon: Observing Markovianity with respect to a fixed measurement basis is not sufficient to guarantee the existence of a memoryless dynamical descriptor. Such hidden quantum memory is similar in spirit to other quantum traits that require precisely the resource in their implementation that they ultimately hide, such as quantum channels that preserve all separable states but cannot be implemented via local operations and classical communication [57–59], non-signalling maps that require signalling [60], and maximally incoherent operations that necessitate coherent resources [61–63]. We begin by outlining the envisaged setup before detailing key properties of memoryless dynamics (both classical and quantum).

1 Framework

In any experimental scheme concerning temporal processes, an experimenter probes a system of interest at (any subset of) times $\mathcal{T}_n := \{t_1, \dots, t_n\}$ (with $t_n > \cdots > t_1$) and records the corresponding probability distributions $\{\mathbb{P}(\boldsymbol{x}_\Gamma)\}$, where $\Gamma \subseteq \mathcal{T}_n$ and $\boldsymbol{x}_\Gamma := \{x_j | t_j \in \Gamma\}$ (see Fig. 1). These capture, for instance, the probability that x_1 is observed at time t_1 *and* x_2 at t_2, and so on, with all possible combinations of measurement times. Note that the experimenter can also *not* make a measurement at any intermediate time, e.g., record $\mathbb{P}(x_3, x_1)$ without measuring at t_2.

Independent of the physical scenario—it could be classical, quantum, or even post-quantum—one can define the concept of *Markovianity* based on the observed statistics alone, as conditional independence of any current outcome from all but the most recent one. Concretely, we have the following working definition:

Definition 1. A *Markovian statistics* on a set of times $\mathcal{T}_n$ is a collection of conditional probability distributions $\{\mathbb{P}(x_j | x_{j-1}, \dots, x_1)\}_{t_j \in \mathcal{T}_n}$ for which

$$\mathbb{P}(x_j | x_{j-1}, \dots, x_1) = \mathbb{P}(x_j | x_{j-1}) \qquad (1)$$

for all $t_j \in \mathcal{T}_n$.

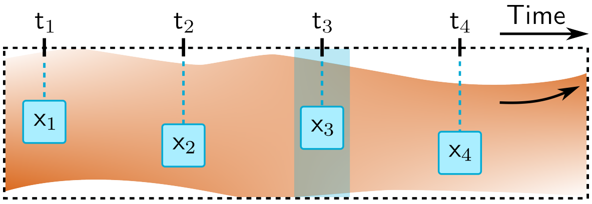

Figure 1: *Probing dynamics.* By probing a process—e.g., Brownian motion, or the evolution of a spin degree of freedom—sequentially (here, at times $\mathcal{T}_4 = \{t_1, t_2, t_3, t_4\}$), an experimenter can deduce the probability distribution $\mathbb{P}(x_4, x_3, x_2, x_1)$. In the classical case, this also includes all 'contained' distributions, e.g., $\mathbb{P}(x_4, x_2, x_1)$ via marginalisation [see Eq. (4)]. In the quantum case, due to invasiveness, deducing said distributions requires a new experiment where *no* measurement is performed at t_3 (depicted by the shaded box).

Defined as such, the question of Markovianity is, a priori, *theory independent* and concerns only the observed statistics. As we shall see, though, the concept of conditional probabilities is a subtle issue that depends on the envisaged physical scenario. Throughout this article, we distinguish Markovianity from the notion of *memoryless dynamics*, which corresponds to the memory properties of the *underlying* dynamics that engenders the observed statistics, thereby making the latter a *theory-dependent* concept.

Specifically, the question of memorylessness concerns whether, throughout the evolution of a system that is coupled to some inaccessible environment, said environment perpetuates past information about the system forward in time or irretrievably dissipates it.[1] The description of such open evolution differs across physical theories: In the classical setting, the most general state transformations are stochastic matrices, whereas in the quantum realm, these are quantum channels. Probability distributions arising from interrogating either classical or quantum processes therefore have different properties since they are calculated via different rules. Our main result shows that such a distinction holds for the relationship between Markovianity and memorylessness: Although equivalent in the classical case, in the quantum realm the observation of Markovian statistics does *not* guarantee even the existence of a memoryless dynamics that engenders them.

[1] We consider memory to be a property that is external to the system, i.e., stored in the environment, rather than information encoded in the system itself.

2 Classical Dynamics

We begin with a discussion of memoryless *classical* dynamics:

Definition 2. A *memoryless classical dynamics* on $\mathcal{T}_n$ is a set of mutually independent stochastic matrices $\{S_{j:j-1}\}_{j=2,\dots,n}$ and an initial state (i.e., probability vector) $\boldsymbol{p}_1$ such that the probability distribution over any sequence of outcomes $x_1, \dots, x_n$ is given by

$$\mathbb{P}(x_n, \dots, x_1) = \langle x_n | S_{n:n-1} | x_{n-1} \rangle\langle x_{n-1} | \dots | x_2 \rangle\langle x_2 | S_{2:1} | x_1 \rangle\langle x_1 | \boldsymbol{p}_1, \tag{2}$$

where $|x_j\rangle\langle x_j|$ are projectors corresponding to measurement outcomes x_j.

Although the environment plays a role in dictating the state transitions between any times t_{j-1} and t_j—namely via the stochastic matrices $S_{j:j-1}$, which are matrices with non-negative elements whose columns sum to unity—in *memoryless* processes, the environment does not propagate information, i.e., the stochastic matrices in Eq. (2) are mutually *independent*. On the other hand, Markovianity (see Def. 1) concerns only the observed statistics [l.h.s. of Eq. (2)]. In classical physics, we can make the following simple observation (see App. A):

Observation 1. *In the classical setting, memoryless dynamics are equivalent to Markovian statistics.*

Specifically, this equivalence is given by setting $\langle x_j | S_{j:j-1} | x_{j-1} \rangle = \mathbb{P}(x_j | x_{j-1})$, and it follows from Eq. (2) that for any Markovian statistics/memoryless classical dynamics we have

$$\mathbb{P}(x_n, \dots, x_1) = \mathbb{P}(x_n | x_{n-1}) \mathbb{P}(x_{n-1} | x_{n-2}) \dots \times \dots \mathbb{P}(x_2 | x_1)\, \mathbb{P}(x_1). \tag{3}$$

In one direction, Obs. 1 states that for any memoryless dynamics, the observed statistics are Markovian—this is also true in the quantum setting (see below). Conversely, if one records Markovian statistics by probing a classical process, then one can always construct a unique, memoryless dynamical model that faithfully reproduces them—as we will see, this is not true for statistics gathered from quantum processes.

A major distinction between classical and quantum processes (memoryless or not) is that in the classical realm, the single n-time probability distribution $\mathbb{P}(x_n, \dots, x_1)$ contains the entire set of statistics on all subsets of times $\Gamma \subseteq \mathcal{T}_n$. That is, the probability $\mathbb{P}(\boldsymbol{x}_\Gamma, \mathcal{I}_{\overline{\Gamma}})$ to observe a sequence of outcomes $\boldsymbol{x}_\Gamma$ when probing the process at times Γ and *not* measuring (denoted by the 'do-nothing' instrument $\mathcal{I}_{\overline{\Gamma}}$) at the remaining times $\overline{\Gamma} := \mathcal{T}_n \setminus \Gamma$ can be deduced via marginalisation

$$\mathbb{P}(\boldsymbol{x}_\Gamma, \mathcal{I}_{\overline{\Gamma}}) = \sum_{\boldsymbol{x}_{\overline{\Gamma}}} \mathbb{P}(x_n, \dots, x_1), \tag{4}$$

This *non-invasiveness* of measurements in classical physics underlies Obs. 1 and similarly fails to hold in quantum mechanics. As a direct consequence of measurement non-invasiveness, the properties of a memoryless classical process on $\mathcal{T}_n$ translate to all 'sub-processes' that are probed only at times $\Gamma \subset \mathcal{T}_n$ (see App. A):

Corollary 1. *All sub-statistics of a memoryless classical dynamics are Markovian and the corresponding conditional probabilities are compatible.*

By *compatible*, we mean that all conditional probabilities are independent of how they are obtained, i.e.,

$$\frac{\mathbb{P}(x_j, \boldsymbol{x}_{\Gamma^{(i)}})}{\mathbb{P}(\boldsymbol{x}_{\Gamma^{(i)}})} = \frac{\mathbb{P}(x_j, \boldsymbol{x}_{\Gamma^{(i)\prime}})}{\mathbb{P}(\boldsymbol{x}_{\Gamma^{(i)\prime}})} =: \mathbb{P}(x_j | x_i), \tag{5}$$

for all $t_j, t_i \in \mathcal{T}_n$ (with $t_j > t_i$) and all subsets $\Gamma^{(i)}, \Gamma^{(i)\prime} \subseteq \mathcal{T}_n$ that contain t_i as their largest time. For a classical memoryless dynamics, knowledge of any outcome x_i suffices to erase all historic information (including whether or not a previous measurement was made) and is therefore the only relevant parameter for predicting future outcomes. Such compatibility between Markovian sub-statistics of a memoryless quantum dynamics also holds (for sharp measurements of a fixed observable), although it is less obvious, and we will later employ the breakdown of compatibility as a witness for memory.

3 Quantum Dynamics

In contrast to classical physics, in quantum mechanics, measurements are generally *invasive* such that there is a difference between averaging over outcomes and not performing a measurement. In this article, we focus on the generally considered situation of sharp measurements of an observable, e.g., position in the classical case or spin in the quantum case. This allows us to fairly compare 'classical' and 'quantum' processes in time. Within this setting, the measurements themselves do not 'actively' change the state of the observed system (in the sense that no active interventions are performed), and thus measurement invasiveness only manifests itself in quantum mechanics (due to the loss of coherences in the observed state).

Subsequently, this makes (conditional) probabilities in the quantum realm protocol-dependent entities that require further specification; in what follows, whenever we consider a probability distribution $\mathbb{P}(\boldsymbol{x}_\Gamma)$, we mean the statistics obtained from *only* performing measurements at times in the set $\Gamma \subseteq \mathcal{T}_n$, and doing nothing (denoted by $\mathcal{I}_{\overline{\Gamma}}$) at the remaining times $\overline{\Gamma} = \mathcal{T}_n \setminus \Gamma$. Importantly, in quantum mechanics—in contrast to Eq. (4)—$\mathbb{P}(\boldsymbol{x}_\Gamma) := \mathbb{P}(\boldsymbol{x}_\Gamma, \mathcal{I}_{\overline{\Gamma}}) \neq \sum_{\boldsymbol{x}_{\overline{\Gamma}}} \mathbb{P}(x_n, \dots, x_1)$. Such measurement invasiveness is well-studied and has been used to witness the non-classicality of physical processes [49, 64–66]. Despite these added subtleties in the definition of (conditional) probabilities, memoryless quantum dynamics lead—just like in the classical case—to well-defined, compatible Markovian statistics and sub-statistics. To see this, we first generalise Def. 2 to the quantum case:

Definition 3. A *memoryless quantum dynamics* on $\mathcal{T}_n$ is a set of mutually independent completely positive and trace preserving **(CPTP)** maps $\{\Lambda_{j:j-1}\}_{j=2,\dots,n}$ and an initial state (density operator) ρ_1 such that the probability distribution over any sequence of outcomes $x_1, \dots, x_n$ is given by

$$\mathbb{P}(x_n, \dots, x_1) = \mathrm{tr}\left[\mathcal{P}_n^{(x_n)} \Lambda_{n:n-1} \dots \Lambda_{2:1} \mathcal{P}_1^{(x_1)}[\rho_1]\right], \tag{6}$$

where $\mathcal{P}_j^{(x_j)}[\,\bullet\,] := |x_j\rangle\langle x_j| \bullet |x_j\rangle\langle x_j|$ are maps corresponding to sharp (i.e., rank-1) projective measurements.

Analogous to the classical case, CPTP maps are the most general state transformations in the presence of environmental noise, and the absence of memory in the dynamics corresponds to the mutual independence of the maps $\Lambda_{j:j-1}$ in the definition. The above equation to compute probabilities is commonly known as the *quantum regression formula* **(QRF)** [44, 45, 56]. Importantly, it allows for the computation of sub-statistics on any $\Gamma \subseteq \mathcal{T}_n$, not via marginalisation, but by replacing the projection operators corresponding to probing times in $\overline{\Gamma}$ in Eq. (6) with identity maps. Of course, one need not perform projective measurements, and the above formula can be used to calculate the probability distribution over *any* sequence of outcomes for arbitrary instruments. However, in contrast to the non-invasive measurements typically considered in classical stochastic processes, such general quantum measurements do not necessarily reset the state of the system, which means that memoryless quantum dynamics can lead to non-Markovian statistics for general instruments [53, 54, 67, 68]. Nonetheless, when restricted to sharp, projective measurements of a given observable, then—just as in the classical setting—memorylessness in the quantum realm manifests itself on the observational level as Markovianity (see App. B):

Lemma 1. *Any memoryless quantum dynamics leads to Markovian statistics (for sharp, projective measurements).*

We saw earlier that memoryless classical processes also lead to (compatible) Markovian *sub*-statistics (see Cor. 1), where compatibility is given by Eq. (5). This is also true for memoryless quantum processes, with the important difference that sub-statistics are not obtained by marginalisation, but by 'doing nothing' at the excluded times, i.e., by explicitly performing the experiment in a different way. Probing sub-statistics in this manner yields meaningful conditional probabilities and we have the following (see App. B):

Lemma 2. *Any memoryless quantum dynamics leads to Markovian sub-statistics (for sharp, projective measurements) that are compatible.*

In both quantum mechanics and classical physics, memoryless dynamics—when probed sharply in a fixed basis—*always* lead to Markovian statistics and Markovian, compatible sub-statistics. In the classical setting, the converse is also true: From the observation of Markovian statistics one can always construct a (unique)

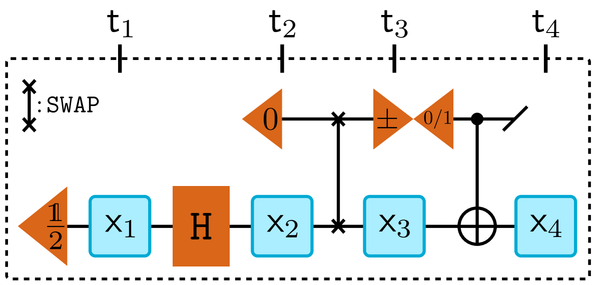

Figure 2: *Markovian statistics that require memory.* When the σ_z observable is measured (sharply) at *all* times, the circuit yields Markovian statistics. Memory becomes apparent in the joint statistics $\mathbb{P}(x_4, x_3, \mathcal{I}_2, x_1)$ when no measurement is performed at t_2, which is in contradiction with the possibility of a memoryless dynamical model (Thm. 1).

memoryless process describing the situation at hand. As discussed, measuring a fixed observable of a time-evolving quantum system cannot provide enough information to fully determine the underlying dynamics. Nonetheless, it is reasonable to assume that whenever one observes Markovian statistics, there should exist *some* memoryless description that correctly reproduces them (indeed, this is the assumption of employing the QRF to describe Markovian statistics). Thus, we now ask the question: *Given Markovian statistics (deduced via sharp, projective measurements), does there always exist a memoryless quantum dynamical model that faithfully reproduces them?*

4 Hidden Quantum Memory & Incompatability

We answer the above question in the negative, first by demonstrating a quantum process that leads to Markovian statistics with non-Markovian sub-statistics, and then by constructing a process with Markovian statistics and sub-statistics that are nonetheless incompatible.

Theorem 1. *Given Markovian statistics on $\mathcal{T}_n$ (deduced via sharp, projective measurements), there does* not *always exist a memoryless quantum dynamics that faithfully reproduces them.*

Proof. Our proof is by way of constructing an explicit example, depicted in Fig. 2. The dynamics is over four times and the experimenter always measures the σ_z observable. An initial state $\rho_1 = \frac{\mathbb{1}}{2}$ is sent to the experimenter, who measures it. The dynamics between times t_1 and t_2 is a Hadamard gate. Following the measurement at t_2, the system is swapped with a fiducial environment state $\tau = |0\rangle$, which is what the experimenter measures at time t_3. Meanwhile, the dynamics of the environment consists of a measurement in the σ_x-basis, followed by a preparation of $|0\rangle(|1\rangle)$ whenever $+(-)$ is recorded. Between times t_3 and t_4, the dynamics comprises a `CNOT` gate, controlled on the environment. Due to the gates that act on the system *and* environment, this circuit can, in principle, display memory effects for the system dynamics. In App. C, we calculate the full statistics $\mathbb{P}(x_4, x_3, x_2, x_1)$ and show them to be Markovian, i.e., $\mathbb{P}(x_4|x_3, x_2, x_1) = \mathbb{P}(x_4|x_3)$ and $\mathbb{P}(x_3|x_2, x_1) = \mathbb{P}(x_3|x_2)$. This is because the measurement of σ_z at t_2 yields an output state that is unbiased with respect to the σ_x-basis measurement on the environment and therefore all memory of x_1 is lost. However, by calculating the sub-statistics where the experimenter does *not* measure at time t_2, i.e., $\mathbb{P}(x_4, x_3, \mathcal{I}_2, x_1)$, we see that they are *non*-Markovian since information about x_1 is now *not* fully scrambled by the 'intervention' (or rather lack thereof) at t_2, and we have $\mathbb{P}(x_4|x_3, \mathcal{I}_2, x_1) \neq \mathbb{P}(x_4|x_3)$ with dependence on x_1. As we proved in Lem. 2, such behaviour cannot happen for *any* memoryless quantum dynamics. Thus, even though the statistics on $\mathcal{T}_n$ is Markovian, there is *no* memoryless quantum dynamics that faithfully reproduces the statistics on all four times, since the sub-statistics fail to be Markovian. □

Here, we have uncovered a new temporal quantum phenomenon: *Hidden quantum memory.* The fact that full statistics can be Markovian but sub-statistics can be non-Markovian for measurements of a given observable is impossible in the classical realm. Moreover, this property cannot occur for memoryless quantum dynamics either (whenever said observable is measured sharply). Thus, we have shown the existence of Markovian statistics that, not only potentially come from a quantum dynamics with memory (which can happen, as is well known, when measured in a fixed basis), but *fundamentally require* memory for their reproduction.

Another way of viewing this result is that non-Markovian sub-statistics serves as a witness for the necessity of memory in the underlying quantum dynamics. This naturally begs the question: *If the full statistics and all sub-statistics are Markovian, does there always exist a memoryless quantum dynamical model that faithfully reproduces them*? In other words, is the ability to detect non-Markovian sub-statistics a requirement for ruling out a memoryless description of the quantum dynamics? Here, we also answer this in the negative, providing an even stronger result than above:

Theorem 2. *Given Markovian statistics and sub-statistics on $\mathcal{T}_n$ and all subsets thereof (deduced via sharp, projective measurements), there does not always exist a memoryless quantum process that faithfully reproduces them.*

Proof. The proof is again by constructing an explicit example, with the corresponding circuit depicted in Fig. 3. In App. C, we calculate the full statistics $\mathbb{P}(x_4, x_3, x_2, x_1)$ and all relevant sub-statistics [e.g., $\mathbb{P}(x_4, x_3, \mathcal{I}_2, x_1)$, etc.] and show them to be Markovian. This latter fact can easily be seen directly: Since the state of the system is discarded and reprepared in a fixed state $|0\rangle$ between times t_2 and t_3, the only way in which memory from t_1 and/or t_2 can influence the statistics observed at t_4—thus potentially rendering the conditional probabilities $\mathbb{P}(x_4|x_3, x_2, x_1)$, $\mathbb{P}(x_4|x_3, \mathcal{I}_2, x_1)$ and $\mathbb{P}(x_4|x_3, x_2, \mathcal{I}_1)$ non-Markovian—is via the state of the environment at time t_3. However, while this state explicitly depends on *whether* previous measurements were performed, it crucially does *not* depend on the respective measurement outcomes. For example, if measurements at t_1 *and* t_2 are performed, then the state of the environment at t_3 is proportional to the maximally mixed state, independent of the respective measurement outcomes. On the other hand, if only a measurement at t_1 is performed, then the state of the environment at t_3 is (proportional to) $|0\rangle\langle 0|$, again independent of the measurement outcome at t_1. The same independence of previous outcomes (but not of whether or not the respective measurements were performed) holds true for all other potential combinations of performed and unperformed measurements (as we show explicitly in App. C), such that *all* conditional probabilities observed at times t_4 and t_3 are indeed Markovian. However, as can already be seen

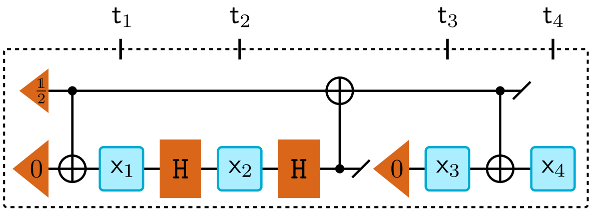

Figure 3: *Incompatible Markovian statistics.* When the σ_z observable is measured on the system at all subsets of times, the circuit yields Markovian statistics and sub-statistics. Despite this Markovianity, the respective conditional probabilities are incompatible, i.e., they depend on whether or not previous measurements were performed (Thm. 2). Such behaviour is *only* possible in the presence of memory.

from the above discussion, they are not compatible. Indeed, since the state of the environment at time t_3 depends upon whether or not a measurement was performed at t_2, the resulting conditional probabilities at t_4 differ depending upon whether or not previous measurements were performed, making them incompatible. □

5 Conclusions

In this article, we have presented the concept of hidden quantum memory, i.e., the existence of processes that yield Markovian statistics (for sharp measurements of an observable) that *cannot* be explained without underlying memory. In a similar vein to the violation of Leggett-Garg inequalities, this phenomenon can only occur when the performed measurements are invasive, since otherwise the observed statistics are classical and hence Markovianity and memorylessness coincide. However, hidden quantum memory is not merely a different manifestation of measurement invasiveness, but arises due to its interplay with memory effects; while memoryless quantum dynamics can violate Leggett-Garg inequalities, they cannot propagate information about whether or not measurements were performed at earlier times other than the most recent—i.e., they cannot exhibit hidden quantum memory.

It is important to stress the assumptions that underpin our observations, in particular with respect to the condition that the measurements of the observable are sharp. Indeed, allowing for classical stochastic processes to be probed via *active* interventions [69, 70] or noisy measurements [71–74] can also lead to seemingly 'non-classical' effects such as the violation of Leggett-Garg inequalities [50–52], breakdown of Kolmogorov consistency [49], or the device-dependence of memory [53, 54], and even the notion of Markovianity itself becomes either obfuscated or trivial [75]. In fact, with such interventions the same statistics as generated by the dynamics of Figs. (2) and (3) can be reproduced with classical dynamics (with memory). However, in the classical case, allowing for active manipulations of the state to permit such incompatible statistics is rather ad hoc—essentially requiring the observer to explicitly communicate to the environment whether or not a measurement is made—and not intrinsic to the fundamental properties of measurements themselves. This is in contrast to the situation in quantum mechanics, where measurements fundamentally disturb the system state in general. By only considering sharp measurements of an observable, we restrict our focus to the properties of sequential measurements per se and fairly compare the two theories at the expense of rendering our results device-dependent (which is necessary for any meaningful distinction between classical and quantum temporal effects [76]). Within this paradigm, then, activation of hidden memory can only be seen for quantum dynamics, making the phenomena we uncover genuinely quantum effects.

Importantly, our results differ from the (known) fact that probing a quantum processes with memory in a fixed basis—i.e., sharply measuring a fixed observable—can yield Markovian statistics. For the Markovian statistics that we reported, there exists *no* memoryless model that reproduces them, either because they become non-Markovian when measurements are not performed at some times, or because all observed statistics and sub-statistics are Markovian but incompatible. In turn, this implies that even if one observes Markovian statistics in a given basis, one cannot confidently employ a QRF to compute the statistics in said basis. As a consequence, even detecting the possibility of a memoryless description of a process is an experimentally complex undertaking that not only requires one to deduce joint probabilities on $\mathcal{T}_n$, but also on all subsets thereof. Naturally, one might expect that simultaneously demanding Markovianity *and* compatibility of all observed sub-statistics should suffice to guarantee a memoryless description. However, even under such strong requirements, the existence of a memoryless model is a priori not clear, and investigations into this question are subject to future work.

Together, our results expose a novel genuinely quantum effect in time and demonstrate the richness of effects that arise from the intricate interplay of measurement invasiveness, memory, and the freedom to choose different instruments that quantum mechanics affords.

Acknowledgments

We would like to thank Dario Egloff, Andrea Smirne, Kavan Modi, and Top Notoh for interesting discussions. P.T. acknowledges funding from the Japan Society for the Promotion of Science (JSPS) Postdoctoral Fellowships for Research in Japan (Short-term Program) and the KAKENHI grant No. 21H03394, Austrian Science Fund (FWF) project: Y879-N27 (START), and the European Research Council (Consolidator grant 'Cocoquest' 101043705). T.J.E. is supported by the University of Manchester Dame Kathleen Ollerenshaw Fellowship. S.M. acknowledges funding from the Austrian Science Fund (FWF): ZK3 (Zukunftkolleg) and Y879-N27 (START project), the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska Curie grant agreement No. 801110, and the Austrian Federal Ministry of Education, Science and Research (BMBWF). The opinions expressed in this publication are those of the authors, the EU Agency is not responsible for any use that may be made of the information it contains. This project/research was supported by grant number FQXi-RFP-IPW-1910 from the Foundational Questions Institute and Fetzer Franklin Fund, a donor advised fund of Silicon Valley Community Foundation.

References

[1] John Preskill, *Quantum Computing in the NISQ era and beyond,* Quantum **2**, 79 (2018), arXiv:1801.00862.

[2] Antonio Acín, Immanuel Bloch, Harry Buhrman, Tommaso Calarco, Christopher Eichler, Jens Eisert, Daniel Esteve, Nicolas Gisin, Steffen J. Glaser, Fedor Jelezko, Stefan Kuhr, Maciej Lewenstein, Max F. Riedel, Piet O. Schmidt, Rob Thew, Andreas Wallraff, Ian Walmsley, and Frank K. Wilhelm, *The quantum technologies roadmap: a European community view,* New J. Phys. **20**, 080201 (2018), arXiv:1712.03773.

[3] Konrad Banaszek, Andrzej Dragan, Wojciech Wasilewski, and Czesław Radzewicz, *Experimental Demonstration of Entanglement-Enhanced Classical Communication over a Quantum Channel with Correlated Noise,* Phys. Rev. Lett. **92**, 257901 (2004), arXiv:quant-ph/0403024.

[4] Julio T. Barreiro, Philipp Schindler, Otfried Gühne, Thomas Monz, Michael Chwalla, Christian F. Roos, Markus Hennrich, and Rainer Blatt, *Experimental multiparticle entanglement dynamics induced by decoherence,* Nat. Phys. **6**, 943 (2010), arXiv:1005.1965.

[5] Clément Sayrin, Igor Dotsenko, Xingxing Zhou, Bruno Peaudecerf, Théo Rybarczyk, Sébastien Gleyzes, Pierre Rouchon, Mazyar Mirrahimi, Hadis Amini, Michel Brune, Jean-Michel Raimond, and Serge Haroche, *Real-time quantum feedback prepares and stabilizes photon number states,* Nature **477**, 73 (2011), arXiv:1107.4027.

[6] Arne L. Grimsmo, *Time-Delayed Quantum Feedback Control,* Phys. Rev. Lett. **115**, 060402 (2015), arXiv:1502.06959.

[7] Jake Iles-Smith, Arend G. Dijkstra, Neill Lambert, and Ahsan Nazir, *Energy transfer in structured and unstructured environments: Master equations beyond the Born-Markov approximations,* J. Chem. Phys. **144**, 44110 (2016).

[8] Javier Cerrillo, Maximilian Buser, and Tobias Brandes, *Nonequilibrium quantum transport coefficients and transient dynamics of full counting statistics in the strong-coupling and non-Markovian regimes,* Phys. Rev. B **94**, 214308 (2016), arXiv:1606.05074.

[9] A. Metelmann and A. A. Clerk, *Nonreciprocal quantum interactions and devices via autonomous feedforward,* Phys. Rev. A **95**, 013837 (2017), arXiv:1610.06621.

[10] S. J. Whalen, A. L. Grimsmo, and H. J. Carmichael, *Open quantum systems with delayed coherent feedback,* Quantum Sci. Technol. **2**, 044008 (2017), arXiv:1702.05776.

[11] Daniel Basilewitsch, Rebecca Schmidt, Do-

minique Sugny, Sabrina Maniscalco, and Christiane P. Koch, *Beating the limits with initial correlations,* New J. Phys. **19**, 113042 (2017), arXiv:1703.04483.

[12] J. Fischer, D. Basilewitsch, C. P. Koch, and D. Sugny, *Time-optimal control of the purification of a qubit in contact with a structured environment,* Phys. Rev. A **99**, 033410 (2019), arXiv:1901.05756.

[13] I. A. Luchnikov, S. V. Vintskevich, H. Ouerdane, and S. N. Filippov, *Simulation Complexity of Open Quantum Dynamics: Connection with Tensor Networks,* Phys. Rev. Lett. **122**, 160401 (2019), arXiv:1812.00043.

[14] Mathias R. Jørgensen and Felix A. Pollock, *Exploiting the Causal Tensor Network Structure of Quantum Processes to Efficiently Simulate Non-Markovian Path Integrals,* Phys. Rev. Lett. **123**, 240602 (2019), arXiv:1902.00315.

[15] Qing Liu, Thomas J. Elliott, Felix C. Binder, Carlo Di Franco, and Mile Gu, *Optimal stochastic modeling with unitary quantum dynamics,* Phys. Rev. A **99**, 062110 (2019), arXiv:1810.09668.

[16] Mathias R. Jørgensen and Felix A. Pollock, *Discrete memory kernel for multitime correlations in non-Markovian quantum processes,* Phys. Rev. A **102**, 052206 (2020), arXiv:2007.03234.

[17] Lorenzo Magrini, Philipp Rosenzweig, Constanze Bach, Andreas Deutschmann-Olek, Sebastian G. Hofer, Sungkun Hong, Nikolai Kiesel, Andreas Kugi, and Markus Aspelmeyer, *Real-time optimal quantum control of mechanical motion at room temperature,* Nature **595**, 373 (2021), arXiv:2012.15188.

[18] Thomas J. Elliott, Mile Gu, Andrew J. P. Garner, and Jayne Thompson, *Quantum Adaptive Agents with Efficient Long-Term Memories,* Phys. Rev. X **12**, 011007 (2022), arXiv:2108.10876.

[19] Harrison Ball, Thomas M. Stace, Steven T. Flammia, and Michael J. Biercuk, *Effect of noise correlations on randomized benchmarking,* Phys. Rev. A **93**, 022303 (2016), arXiv:1504.05307.

[20] Pedro Figueroa-Romero, Kavan Modi, Robert J. Harris, Thomas M. Stace, and Min-Hsiu Hsieh, *Randomized Benchmarking for Non-Markovian Noise,* PRX Quantum **2**, 040351 (2021), arXiv:2107.05403.

[21] Pedro Figueroa-Romero, Kavan Modi, and Min-Hsiu Hsieh, *Towards a general framework of Randomized Benchmarking incorporating non-Markovian Noise,* Quantum **6**, 868 (2022), arXiv:2202.11338.

[22] Lorenza Viola, Emanuel Knill, and Seth Lloyd, *Dynamical Decoupling of Open Quantum Systems,* Phys. Rev. Lett. **82**, 2417 (1999), arXiv:quant-ph/9809071.

[23] Michael J. Biercuk, Hermann Uys, Aaron P. VanDevender, Nobuyasu Shiga, Wayne M. Itano, and John J. Bollinger, *Optimized dynamical decoupling in a model quantum memory,* Nature **458**, 996 (2009), arXiv:0812.5095.

[24] Carole Addis, Francesco Ciccarello, Michele Cascio, G. Massimo Palma, and Sabrina Maniscalco, *Dynamical decoupling efficiency versus quantum non-Markovianity,* New J. Phys. **17**, 123004 (2015), arXiv:1502.02528.

[25] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Quantum Circuit Architecture,* Phys. Rev. Lett. **101**, 060401 (2008), arXiv:0712.1325.

[26] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti, *Theoretical framework for quantum networks,* Phys. Rev. A **80**, 022339 (2009), arXiv:0904.4483.

[27] S. Mavadia, C. L. Edmunds, C. Hempel, H. Ball, F. Roy, T. M. Stace, and M. J. Biercuk, *Experimental quantum verification in the presence of temporally correlated noise,* npj Quantum Inf. **4**, 7 (2018), arXiv:1706.03787.

[28] 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), arXiv:2004.14018.

[29] G. A. L. White, F. A. Pollock, L. C. L. Hollenberg, K. Modi, and C. D. Hill, *Non-Markovian Quantum Process Tomography,* PRX Quantum **3**, 020344 (2022), arXiv:2106.11722.

[30] Yu Guo, Philip Taranto, Bi-Heng Liu, Xiao-Min Hu, Yun-Feng Huang, Chuan-Feng Li, and Guang-Can Guo, *Experimental Demonstration of Instrument-Specific Quan-*

tum Memory Effects and Non-Markovian Process Recovery for Common-Cause Processes, Phys. Rev. Lett. **126**, 230401 (2021), arXiv:2003.14045.

[31] Gregory A. L. White, Felix A. Pollock, Lloyd C. L. Hollenberg, Charles D. Hill, and Kavan Modi, *From many-body to many-time physics,* arXiv:2107.13934 (2021).

[32] Bogna Bylicka, Mikko Tukiainen, Dariusz Chruściński, Jyrki Piilo, and Sabrina Maniscalco, *Thermodynamic power of non-Markovianity,* Sci. Rep. **6**, 27989 (2016), arXiv:1504.06533.

[33] Akihito Kato and Yoshitaka Tanimura, *Quantum heat current under non-perturbative and non-Markovian conditions: Applications to heat machines,* J. Chem. Phys. **145**, 224105 (2016), arXiv:1609.08783.

[34] Philip Taranto, Faraj Bakhshinezhad, Philipp Schüttelkopf, Fabien Clivaz, and Marcus Huber, *Exponential Improvement for Quantum Cooling through Finite-Memory Effects,* Phys. Rev. Appl. **14**, 054005 (2020), arXiv:2004.00323.

[35] Philip Taranto, Faraj Bakhshinezhad, Andreas Bluhm, Ralph Silva, Nicolai Friis, Maximilian P.E. Lock, Giuseppe Vitagliano, Felix C. Binder, Tiago Debarba, Emanuel Schwarzhans, Fabien Clivaz, and Marcus Huber, *Landauer Versus Nernst: What is the True Cost of Cooling a Quantum System?* PRX Quantum **4**, 010332 (2023), arXiv:2106.05151.

[36] Ángel Rivas, Susana F. Huelga, and Martin B. Plenio, *Quantum non-Markovianity: Characterization, quantification and detection,* Rep. Prog. Phys. **77**, 094001 (2014), arXiv:1405.0303.

[37] Heinz-Peter Breuer, Elsi-Mari Laine, Jyrki Piilo, and Bassano Vacchini, *Colloquium: Non-Markovian dynamics in open quantum systems,* Rev. Mod. Phys. **88**, 021002 (2016), arXiv:1505.01385.

[38] Inés de Vega and Daniel Alonso, *Dynamics of non-Markovian open quantum systems,* Rev. Mod. Phys. **89**, 015001 (2017), arXiv:1511.06994.

[39] Li Li, Michael J. W. Hall, and Howard M. Wiseman, *Concepts of quantum non-Markovianity: A hierarchy,* Phys. Rep. **759**, 1 (2018), arXiv:1712.08879.

[40] Philip Taranto, *Memory Effects in Quantum Processes,* Int. J. Quantum Inf. **18**, 1941002 (2020), arXiv:1909.05245.

[41] G. Lindblad, *On the generators of quantum dynamical semigroups,* Commun. Math. Phys. **48**, 119 (1976).

[42] Vittorio Gorini, Andrzej Kossakowski, and E. C. G. Sudarshan, *Completely positive semigroups of N-level systems,* J. Math. Phys. **17**, 821 (1976).

[43] Daniel Manzano, *A short introduction to the Lindblad master equation,* AIP Adv. **10**, 025106 (2020), arXiv:1906.04478.

[44] Howard Carmichael, *An Open Systems Approach to Quantum Optics* (Springer-Verlag, Berlin, 1993).

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

[46] N. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, New York, 2011).

[47] Horacio Wio, Roberto Deza, and Juan López, *An Introduction to Stochastic Processes and Nonequilibrium Statistical Physics (Rev. Ed.)* (World Scientific, Singapore, 2012).

[48] M. Ringbauer, C. J. Wood, K. Modi, A. Gilchrist, A. G. White, and A. Fedrizzi, *Characterizing Quantum Dynamics with Initial System-Environment Correlations,* Phys. Rev. Lett. **114**, 090402 (2015), arXiv:1410.5826.

[49] Simon Milz, Fattah Sakuldee, Felix A. Pollock, and Kavan Modi, *Kolmogorov extension theorem for (quantum) causal modelling and general probabilistic theories,* Quantum **4**, 255 (2020), arXiv:1712.02589.

[50] A. J. Leggett and Anupam Garg, *Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks?* Phys. Rev. Lett. **54**, 857 (1985).

[51] A. J. Leggett, *Realism and the physical world,* Rep. Prog. Phys. **71**, 022001 (2008).

[52] Clive Emary, Neill Lambert, and Franco Nori, *Leggett–Garg inequalities,* Rep. Prog. Phys. **77**, 016001 (2014), arXiv:1304.5133.

[53] Philip Taranto, Felix A. Pollock, Simon Milz, Marco Tomamichel, and Kavan Modi, *Quantum Markov Order,* Phys. Rev. Lett. **122**, 140401 (2019), arXiv:1805.11341.

[54] Philip Taranto, Simon Milz, Felix A. Pollock, and Kavan Modi, *Structure of quantum stochastic processes with finite Markov order,* Phys. Rev. A **99**, 042108 (2019), arXiv:1810.10809.

[55] Philip Taranto, Felix A. Pollock, and Kavan Modi, *Non-Markovian memory strength bounds quantum process recoverability,* npj Quantum Inf. **7**, 149 (2021), arXiv:1907.12583.

[56] Melvin Lax, *Formal Theory of Quantum Fluctuations from a Driven State,* Phys. Rev. **129**, 2342 (1963).

[57] Martin B. Plenio and Shashank Virmani, *An Introduction to Entanglement Measures,* Quantum Info. Comput. **7**, 1 (2007), arXiv:quant-ph/0504163.

[58] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki, *Quantum entanglement,* Rev. Mod. Phys. **81**, 865 (2009), arXiv:quant-ph/0702225.

[59] Eric Chitambar, Julio I. de Vicente, Mark W. Girard, and Gilad Gour, *Entanglement manipulation beyond local operations and classical communication,* J. Math. Phys. **61**, 042201 (2020), arXiv:1711.03835.

[60] David Beckman, Daniel Gottesman, M. A. Nielsen, and John Preskill, *Causal and localizable quantum operations,* Phys. Rev. A **64**, 052309 (2001), arXiv:quant-ph/0102043.

[61] Eric Chitambar and Gilad Gour, *Critical Examination of Incoherent Operations and a Physically Consistent Resource Theory of Quantum Coherence,* Phys. Rev. Lett. **117**, 030401 (2016), arXiv:1602.06969.

[62] Eric Chitambar and Gilad Gour, *Comparison of incoherent operations and measures of coherence,* Phys. Rev. A **94**, 052336 (2016), arXiv:1602.06969.

[63] Iman Marvian and Robert W. Spekkens, *How to quantify coherence: Distinguishing speakable and unspeakable notions,* Phys. Rev. A **94**, 052324 (2016), arXiv:1602.08049.

[64] A. Smirne, D. Egloff, M. G. Díaz, M. B. Plenio, and S. F. Huelga, *Coherence and non-classicality of quantum Markov processes,* Quantum Sci. Technol. **4**, 01LT01 (2019), arXiv:1709.05267.

[65] Philipp Strasberg and María García Díaz, *Classical quantum stochastic processes,* Phys. Rev. A **100**, 022120 (2019), arXiv:1905.03018.

[66] Simon Milz, Dario Egloff, Philip Taranto, Thomas Theurer, Martin B. Plenio, Andrea Smirne, and Susana F. Huelga, *When Is a Non-Markovian Quantum Process Classical?* Phys. Rev. X **10**, 041049 (2020), arXiv:1907.05807.

[67] Felix A. Pollock, César Rodríguez-Rosario, Thomas Frauenheim, Mauro Paternostro, and Kavan Modi, *Operational Markov Condition for Quantum Processes,* Phys. Rev. Lett. **120**, 040405 (2018), arXiv:1801.09811.

[68] Felix A. Pollock, César Rodríguez-Rosario, Thomas Frauenheim, Mauro Paternostro, and Kavan Modi, *Non-Markovian quantum processes: Complete framework and efficient characterization,* Phys. Rev. A **97**, 012127 (2018), arXiv:1512.00589.

[69] J. Pearl, *Causality* (Oxford Univ. Press, 2000).

[70] Fabio Costa and Sally Shrapnel, *Quantum causal modelling,* New J. Phys. **18**, 063032 (2016), arXiv:1512.07106.

[71] M. Siefert, A. Kittel, R. Friedrich, and J. Peinke, *On a quantitative method to analyze dynamical and measurement noise,* EPL **61**, 466 (2003), arXiv:physics/0108034.

[72] Frank Böttcher, Joachim Peinke, David Kleinhans, Rudolf Friedrich, Pedro G. Lind, and Maria Haase, *Reconstruction of Complex Dynamical Systems Affected by Strong Measurement Noise,* Phys. Rev. Lett. **97**, 090603 (2006), arXiv:nlin/0607002.

[73] David Kleinhans, Rudolf Friedrich, Matthias Wächter, and Joachim Peinke, *Markov properties in presence of measurement noise,* Phys. Rev. E **76**, 041109 (2007), arXiv:0705.1222.

[74] B. Lehle, *Analysis of stochastic time series in the presence of strong measurement noise,* Phys. Rev. E **83**, 021113 (2011), arXiv:1010.5641.

[75] Matheus Capela, Lucas C. Céleri, Rafael Chaves, and Kavan Modi, *Quantum Markov monogamy inequalities,* Phys. Rev. A **106**, 022218 (2022), arXiv:2108.11533.

[76] Costantino Budroni, Gabriel Fagundes, and Matthias Kleinmann, *Memory cost of temporal correlations,* New J. Phys. **21**, 093018 (2019), arXiv:1902.06517.

A Classical Dynamics

A.1 Memoryless Classical Dynamics and Markovian Statistics

Here, we prove Obs. 1 of the main text:

Observation 1. *In the classical setting, memoryless dynamics are equivalent to Markovian statistics.*

Naturally, this equivalence is well-known, but its explicit discussion exposes many of the subtleties with respect to marginalisation that play a crucial role in the quantum case. For the proof, in the forwards direction, beginning with Eq. (2), we have

$$\mathbb{P}(x_n, \dots, x_1) = \langle x_n | S_{n:n-1} | x_{n-1} \rangle\!\langle x_{n-1} | \dots | x_2 \rangle\!\langle x_2 | S_{2:1} | x_1 \rangle\!\langle x_1 | \boldsymbol{p}_1 \,, \tag{7}$$

where $\{S_{j:j-1}\}$ and $\boldsymbol{p}_1$ are, respectively, the stochastic matrices and the initial probability vector that define the memoryless classical dynamics. Computing the conditional probability for an arbitrary time t_j given the entire sequence of historic outcomes up until that time explicitly gives

$$\begin{aligned}\mathbb{P}(x_j | x_{j-1}, \dots, x_1) &= \frac{\langle x_j | S_{j:j-1} | x_{j-1} \rangle\!\langle x_{j-1} | S_{j-1:j-2} | x_{j-2} \rangle\!\langle x_{j-2} | \dots | x_2 \rangle\!\langle x_2 | S_{2:1} | x_1 \rangle\!\langle x_1 | \boldsymbol{p_1}}{\langle x_{j-1} | S_{j-1:j-2} | x_{j-2} \rangle\!\langle x_{j-2} | \dots | x_2 \rangle\!\langle x_2 | S_{2:1} | x_1 \rangle\!\langle x_1 | \boldsymbol{p_1}} \\ &= \langle x_j | S_{j:j-1} | x_{j-1} \rangle \qquad \forall\, x_{j-2}, \dots, x_1 .\end{aligned} \tag{8}$$

This expression is independent of all $x_1, \dots, x_{j-2}$ and it is indeed equivalent to the the conditional probability $\mathbb{P}(x_j | x_{j-1})$ an experimenter would observe when only making measurements at t_{j-1} and t_j, i.e., they do not measure (which we denote below by $\mathcal{I}_{j-2:1}$). Unlike in quantum mechanics, this conditional probability can equivalently be expressed by marginalising the full joint probability distribution [see Eq. (4)] as follows

$$\begin{aligned}\mathbb{P}(x_j | x_{j-1}) &= \frac{\mathbb{P}(x_j, x_{j-1}, \mathcal{I}_{j-2:1})}{\mathbb{P}(x_{j-1}, \mathcal{I}_{j-2:1})} \\ &= \frac{\sum_{x_{j-2}, \dots, x_1} \mathbb{P}(x_j, x_{j-1}, x_{j-2}, \dots, x_1)}{\sum_{x_{j-2}, \dots, x_1} \mathbb{P}(x_{j-1}, x_{j-2}, \dots, x_1)} \\ &= \frac{\sum_{x_{j-2}, \dots, x_1} \langle x_j | S_{j:j-1} | x_{j-1} \rangle\!\langle x_{j-1} | S_{j-1:j-2} | x_{j-2} \rangle\!\langle x_{j-2} | \dots | x_2 \rangle\!\langle x_2 | S_{2:1} | x_1 \rangle\!\langle x_1 | p_1 | x_0 \rangle}{\sum_{x_{j-2}, \dots, x_1} \langle x_{j-1} | S_{j-1:j-2} | x_{j-2} \rangle\!\langle x_{j-2} | \dots | x_2 \rangle\!\langle x_2 | S_{2:1} | x_1 \rangle\!\langle x_1 | p_1 | x_0 \rangle} \\ &= \langle x_j | S_{j:j-1} | x_{j-1} \rangle .\end{aligned} \tag{9}$$

Thus we have that for a memoryless classical dynamics, the conditional probabilities $\mathbb{P}(x_j | x_{j-1}, \dots, x_1)$ and $\mathbb{P}(x_j | x_{j-1})$ coincide (and both amount to $\langle x_j | S_{j:j-1} | x_{j-1} \rangle$), leading to Markovianity of the statistics and consequently the decomposition of the joint probability distribution expressed in Eq. (3).

Conversely, any Markovian statistics can be faithfully reproduced via a memoryless classical model: Given a joint probability distribution over measurement outcomes for a classical stochastic process, one can always write

$$\mathbb{P}(x_n, \dots, x_1) = \mathbb{P}(x_n | x_{n-1}, \dots, x_1) \mathbb{P}(x_{n-1} | x_{n-2}, \dots, x_1) \, \dots \, \mathbb{P}(x_2 | x_1) \, \mathbb{P}(x_1). \tag{10}$$

Equation (10) holds true for any probability distribution (by definition of conditional probabilities), with the decomposition on the r.h.s. encoding potential memory effects. For Markovian statistics, the above expression simplifies to Eq. (3). Then, one can simply define a set of matrices $\{S_{j:j-1}\}$ via

$$\langle x_j | S_{j:j-1} | x_{j-1} \rangle := \mathbb{P}(x_j | x_{j-1}). \tag{11}$$

These matrices are stochastic (as they contain only non-negative entries and each of the columns to unity since $\sum_{x_j} \mathbb{P}(x_j | x_{j-1}) = 1 \, \forall \, x_{j-1}$). One can also define the initial state via $\langle x_1 | \boldsymbol{p}_1 := \mathbb{P}(x_1)$. From these objects, one can reproduce the joint statistics faithfully via the memoryless dynamical model expressed in Eq. (2).

A.2 Sub-Statistics of Memoryless Classical Dynamics

Here, we prove Cor. 1 of the main text:

Corollary 1. *All sub-statistics of a memoryless classical dynamics are Markovian and the corresponding conditional probabilities are compatible.*

For the proof, consider a memoryless dynamics on $\mathcal{T}_n = \{t_1, \ldots, t_n\}$ and an arbitrary sub-statistics where the experimenter measures at time t_j and any subset of earlier times $\Gamma^{(i)}$ (with corresponding sequence of outcomes $\boldsymbol{x}_{\Gamma^{(i)}}$), where $t_j > t_i = \max(\Gamma^{(i)})$. Below, we assume both that $\Gamma^{(i)}$ does not 'skip times' (e.g., it can be of the form $\{t_3, t_2, t_1\}$, but not $\{t_3, t_1\}$) and that $\min(\Gamma^{(i)}) = t_1$. These assumptions are not crucial and do not affect the generality of the results, but significantly simplify notation. We denote by M the set of all times between t_i and t_j and by F that of all times after t_j, with corresponding outcome sequences $\boldsymbol{x}_M$ and $\boldsymbol{x}_F$ and do-nothing operations $\mathcal{I}_M$ and $\mathcal{I}_F$, respectively. For convenience, we also introduce the do-nothing operation $\mathcal{I}_{FjM}$ for all times after t_i. With this, we explicitly calculate the probability of observing x_j conditioned on the sequence of previous outcomes $\boldsymbol{x}_{\Gamma^{(i)}}$ as

$$\begin{aligned}
\mathbb{P}(x_j|\boldsymbol{x}_{\Gamma^{(i)}}) &= \frac{\mathbb{P}(\mathcal{I}_F, x_j, \mathcal{I}_M, x_i, \ldots, x_1)}{\mathbb{P}(\mathcal{I}_{FjM}, x_i, \ldots, x_1)} \\
&= \frac{\sum_{\boldsymbol{x}_F \boldsymbol{x}_M} \mathbb{P}(x_n, \ldots, x_1)}{\sum_{\boldsymbol{x}_F x_j \boldsymbol{x}_M} \mathbb{P}(x_n, \ldots, x_1)} \\
&= \frac{\sum_{\boldsymbol{x}_F \boldsymbol{x}_M} \mathbb{P}(x_n|x_{n-1}) \ldots \mathbb{P}(x_2|x_1)}{\sum_{\boldsymbol{x}_F x_j \boldsymbol{x}_M} \mathbb{P}(x_n|x_{n-1}) \ldots \mathbb{P}(x_2|x_1)} \\
&= \frac{\left[\sum_{\boldsymbol{x}_F} \mathbb{P}(x_n|x_{n-1}) \ldots \mathbb{P}(x_{j+1}|x_j)\right] \left[\sum_{\boldsymbol{x}_M} \mathbb{P}(x_j|x_{j-1}) \ldots \mathbb{P}(x_{i+1}|x_i)\right] \{\mathbb{P}(x_i|x_{i-1}) \ldots \mathbb{P}(x_2|x_1)\}}{\left[\sum_{\boldsymbol{x}_F x_j \boldsymbol{x}_M} \mathbb{P}(x_n|x_{n-1}) \ldots \mathbb{P}(x_{i+1}|x_i)\right] \{\mathbb{P}(x_i|x_{i-1}) \ldots \mathbb{P}(x_2|x_1)\}} \\
&= \sum_{\boldsymbol{x}_M} \mathbb{P}(x_j|x_{j-1}) \ldots \mathbb{P}(x_{i+1}|x_i) =: \mathbb{P}(x_j|x_i),
\end{aligned} \tag{12}$$

where in the second line we employed the marginalisation rule to compute the sub-statistics from the full process on $\mathcal{T}_n$, in the third line we invoked the Markovianity condition (on the full statistics), in the fourth line we split the sums into independent parts, in the fifth line we used the fact that the first sum in the numerator and the sum in the denominator both evaluate to unity, and the final line only depends on x_j and x_i and satisfies the properties of a conditional probability distribution. Thus we see that any sub-statistics of a memoryless classical dynamics are also Markovian, i.e., $\mathbb{P}(x_j|\boldsymbol{x}_{\Gamma^{(i)}}) = \mathbb{P}(x_j|x_i)$ for all $t_j > t_i$. As mentioned, this reasoning also holds for more 'complicated' sets $\Gamma^{(i)}$, albeit with a slightly more cumbersome notation than used in the proof above.

Moreover, the Markovian sub-statistics are compatible in the sense that it does not matter what occurred at any time prior to that of the most recent conditioning argument, i.e., t_i. For instance, if one computes $\mathbb{P}(x_j|x_i, \mathcal{I}_{i-1:\ell+1}, x_\ell, \mathcal{I}_{\ell-1:1})$, this should also be independent of x_ℓ (i.e., Markovian sub-statistics) *and* equal to $\mathbb{P}(x_j|\boldsymbol{x}_{\Gamma^{(i)}}) = \mathbb{P}(x_j|x_i)$ computed above (i.e., compatible). This can be seen by noting that for any historic sequence (of either measuring or not at any times $t_1, \ldots, t_{i-1}$, which we denote with $(x \cup \mathcal{I})_{i-1:1}$), the logic of Eq. (12) holds, since the only changes would appear in the terms in curly parentheses in the fourth line, which always cancel. Hence, we have the compatibility $\mathbb{P}(x_j|x_i, (x \cup \mathcal{I})_{i-1:1}) = \mathbb{P}(x_j|x_i)$ for all possible combinations of measuring or not in the history leading up to time t_i. Again, this argument can be run in exactly the same vein for any two subsets of times $\Gamma^{(i)}$ and $\Gamma^{(i)\prime}$ satisfying $\max(\Gamma^{(i)}) = \max(\Gamma^{(i)\prime}) = t_i$, with the result that $\mathbb{P}(x_j|\boldsymbol{x}_{\Gamma^{(i)}}) = \mathbb{P}(x_j|\boldsymbol{x}_{\Gamma^{(i)\prime}}) = \mathbb{P}(x_j|x_i)$ for all $t_j > t_i$.

B Quantum Dynamics

B.1 Memoryless Quantum Dynamics and Markovian Statistics

Here, we prove Lem. 1 of the main text:

Lemma 1. *Any memoryless quantum dynamics leads to Markovian statistics (for sharp, projective measurements).*

Beginning with Eq. (6), we have that for any memoryless quantum dynamics

$$\mathbb{P}(x_n,\dots,x_1) = \mathrm{tr}\left[\mathcal{P}_n^{(x_n)}\Lambda_{n:n-1}\dots\Lambda_{2:1}\mathcal{P}_1^{(x_1)}\rho_1\right], \tag{13}$$

where $\{\Lambda_{j:j-1}\}$ are mutually independent CPTP maps, ρ_1 is an initial quantum state, and $\mathcal{P}_j^{(x_j)}[\bullet] = |x_j\rangle\langle x_j|\bullet|x_j\rangle\langle x_j|$. The statistics up to any time t_j is given by $\mathbb{P}(\mathcal{I}_{n:j+1}, x_j, \dots, x_1) =: \mathbb{P}(x_j,\dots,x_1) = \mathrm{tr}\left[\mathcal{P}_j^{(x_j)}\Lambda_{j:j-1}\dots\Lambda_{2:1}\mathcal{P}_1^{(x_1)}\rho_1\right]$ (this can be seen either by direct computation or by invoking causality), where $\mathcal{I}_{n:j+1}$ denotes 'do-nothing' operations from t_j to t_n. With this, computing the conditional probability for an arbitrary time t_j given the entire sequence of historic outcomes up until that time explicitly gives

$$\begin{aligned}
\mathbb{P}(x_j|x_{j-1},\dots,x_1) &= \frac{\mathrm{tr}\left[\mathcal{P}_j^{(x_j)}\Lambda_{j:j-1}\dots\Lambda_{2:1}\mathcal{P}_1^{(x_1)}\rho_1\right]}{\mathrm{tr}\left[\mathcal{P}_{j-1}^{(x_{j-1})}\Lambda_{j-1:j-2}\dots\Lambda_{2:1}\mathcal{P}_1^{(x_1)}\rho_1\right]}\\
&= \frac{\mathrm{tr}\left[\sum_\ell |x_j\rangle\langle x_j|L^\ell_{j:j-1}|x_{j-1}\rangle\langle x_{j-1}|\left(\Lambda_{j-1:j-2}\dots\Lambda_{2:1}\mathcal{P}_1^{(x_1)}\rho_1\right)|x_{j-1}\rangle\langle x_{j-1}|L^{\ell\dagger}_{j:j-1}|x_j\rangle\langle x_j|\right]}{\mathrm{tr}\left[|x_{j-1}\rangle\langle x_{j-1}|\left(\Lambda_{j-1:j-2}\dots\Lambda_{2:1}\mathcal{P}_1^{(x_1)}\rho_1\right)|x_{j-1}\rangle\langle x_{j-1}|\right]}\\
&= \frac{\sum_\ell\langle x_j|L^\ell_{j:j-1}|x_{j-1}\rangle\langle x_{j-1}|L^{\ell\dagger}_{j:j-1}|x_j\rangle\langle x_{j-1}|\left(\Lambda_{j-1:j-2}\dots\Lambda_{2:1}\mathcal{P}_1^{(x_1)}\rho_1\right)|x_{j-1}\rangle}{\langle x_{j-1}|\left(\Lambda_{j-1:j-2}\dots\Lambda_{2:1}\mathcal{P}_1^{(x_1)}\rho_1\right)|x_{j-1}\rangle}\\
&= \sum_\ell\langle x_j|L^\ell_{j:j-1}|x_{j-1}\rangle\langle x_{j-1}|L^{\ell\dagger}_{j:j-1}|x_j\rangle\\
&= \langle x_j|\Lambda_{j:j-1}[|x_{j-1}\rangle\langle x_{j-1}|]|x_j\rangle,
\end{aligned} \tag{14}$$

where we wrote $\Lambda_{j:j-1}[\bullet] := \sum_\ell L^\ell_{j:j-1}\bullet L^{\ell\dagger}_{j:j-1}$ in Kraus operator form in the second line, and then made use of the cyclicity of the trace and the fact that the measurements are sharp (rank-1) projectors in the third line (importantly, if the projectors are not rank-1, corresponding, e.g., to the measurement of an observable with degeneracies, then memoryless processes do *not* necessarily lead to Markovian statistics [40, 54], since in this case the state after the measurement is not fully determined by the outcome; this fact is also true in the classical setting). This expression is independent of all $x_1,\dots,x_{j-2}$ and therefore the conditional probabilities are Markovian. We now show that it is indeed equivalent to the conditional probability $\mathbb{P}(x_j|x_{j-1},\mathcal{I}_{j-2:1}) =: \mathbb{P}(x_j|x_{j-1})$, where the experimenter does *not* measure at all on times $t_1,\dots,t_{j-2}$. Explicitly, we have

$$\begin{aligned}
\mathbb{P}(x_j|x_{j-1}) &= \frac{\mathrm{tr}\left[\mathcal{P}_j^{(x_j)}\Lambda_{j:j-1}\mathcal{P}_{j-1}^{(x_{j-1})}\Lambda_{j-1:j-2}\mathcal{I}_{j-2}\dots\Lambda_{2:1}\mathcal{I}_1\rho_1\right]}{\mathrm{tr}\left[\mathcal{P}_{j-1}^{(x_{j-1})}\Lambda_{j-1:j-2}\mathcal{I}_{j-2}\dots\Lambda_{2:1}\mathcal{I}_1\rho_1\right]}\\
&= \frac{\mathrm{tr}\left[\sum_\ell |x_j\rangle\langle x_j|L^\ell_{j:j-1}|x_{j-1}\rangle\langle x_{j-1}|\left(\Lambda_{j-1:j-2}\mathcal{I}_{j-2}\dots\Lambda_{2:1}\mathcal{I}_1\rho_1\right)|x_{j-1}\rangle\langle x_{j-1}|L^{\ell\dagger}_{j:j-1}|x_j\rangle\langle x_j|\right]}{\mathrm{tr}\left[|x_{j-1}\rangle\langle x_{j-1}|\left(\Lambda_{j-1:j-2}\mathcal{I}_{j-2}\dots\Lambda_{2:1}\mathcal{I}_1\rho_1\right)|x_{j-1}\rangle\langle x_{j-1}|\right]}\\
&= \frac{\sum_\ell\langle x_j|L^\ell_{j:j-1}|x_{j-1}\rangle\langle x_{j-1}|L^{\ell\dagger}_{j:j-1}|x_j\rangle\langle x_{j-1}|\left(\Lambda_{j-1:j-2}\mathcal{I}_{j-2}\dots\Lambda_{2:1}\mathcal{I}_1\rho_1\right)|x_{j-1}\rangle}{\langle x_{j-1}|\left(\Lambda_{j-1:j-2}\mathcal{I}_{j-2}\dots\Lambda_{2:1}\mathcal{I}_1\rho_1\right)|x_{j-1}\rangle}
\end{aligned}$$

$$= \sum_\ell \langle x_j | L^\ell_{j:j-1} | x_{j-1} \rangle\langle x_{j-1} | L^{\ell\dagger}_{j:j-1} | x_j \rangle$$
$$= \langle x_j | \Lambda_{j:j-1}[|x_{j-1}\rangle\langle x_{j-1}|] | x_j \rangle. \tag{15}$$

Thus, we see that the conditional statistics in both situations above coincide and are indeed Markovian $\mathbb{P}(x_j|x_{j-1},\dots,x_1) = \mathbb{P}(x_j|x_{j-1},\mathcal{I}_{j-2:1}) = \mathbb{P}(x_j|x_{j-1}) = \langle x_j|\Lambda_{j:j-1}[|x_{j-1}\rangle\langle x_{j-1}|]|x_j\rangle$.

B.2 Sub-Statistics of Memoryless Quantum Dynamics

Here, we prove Lem. 2 from the main text:

Lemma 2. *Any memoryless quantum dynamics leads to Markovian sub-statistics (for sharp, projective measurements) that are compatible.*

Similar to App. A.2, we will restrict the discussion again to subsets of $\mathcal{T}_n$ of the form $\Gamma^{(i)} = \{t_1,\dots,t_i\}$ and show that $\mathbb{P}(x_j|\boldsymbol{x}_{\Gamma^{(i)}}) = \mathbb{P}(x_j|x_i)$ holds for all $t_j > t_i$ and $t_i = \max(\Gamma^{(i)})$. Let $\mathcal{I}_{j-1:i+1}$ denote the 'do-nothing' operation at all times between t_i and t_j. With this, we obtain

$$\begin{aligned}
\mathbb{P}(x_j|\mathcal{I}_{j-1:i+1},\boldsymbol{x}_{\Gamma^{(i)}}) &= \frac{\mathrm{tr}\left[\mathcal{P}_j^{(x_j)}\Lambda_{j:j-1}\mathcal{I}_{j-1}\Lambda_{j-1:j-2}\dots\mathcal{I}_{i+1}\Lambda_{i+1:i}\mathcal{P}_i^{(x_i)}\Lambda_{i:i-1}\mathcal{P}_{i-1}^{(x_{i-1})}\dots\mathcal{P}_1^{(x_1)}\rho_1\right]}{\mathrm{tr}\left[\mathcal{P}_i^{(x_i)}\Lambda_{i:i-1}\mathcal{P}_{i-1}^{(x_{i-1})}\dots\mathcal{P}_1^{(x_1)}\rho_1\right]} \\
&= \frac{\langle x_j|\Lambda_{j:j-1}\mathcal{I}_{j-1}\Lambda_{j-1:j-2}\dots\mathcal{I}_{i+1}\Lambda_{i+1:i}[|x_i\rangle\langle x_i|]|x_j\rangle\langle x_i|\Lambda_{i:i-1}\mathcal{P}_{i-1}^{(x_{i-1})}\dots\mathcal{P}_1^{(x_1)}\rho_1|x_i\rangle}{\langle x_i|\Lambda_{i:i-1}\mathcal{P}_{i-1}^{(x_{i-1})}\dots\mathcal{P}_1^{(x_1)}\rho_1|x_i\rangle} \\
&= \langle x_j|\Lambda_{j:j-1}\mathcal{I}_{j-1}\Lambda_{j-1:j-2}\dots\mathcal{I}_{i+1}\Lambda_{i+1:i}[|x_i\rangle\langle x_i|]|x_j\rangle \quad \forall\, x_{i-1},\dots,x_1.
\end{aligned} \tag{16}$$

In the case where *no* measurements are made until time t_i, we similarly have

$$\begin{aligned}
\mathbb{P}(x_j|\mathcal{I}_{j-1:i+1},x_i,\mathcal{I}_{i-1:1}) &= \frac{\mathrm{tr}\left[\mathcal{P}_j^{(x_j)}\Lambda_{j:j-1}\mathcal{I}_{j-1}\Lambda_{j-1:j-2}\dots\mathcal{I}_{i+1}\Lambda_{i+1:i}\mathcal{P}_i^{(x_i)}\Lambda_{i:i-1}\mathcal{I}_{i-1}\dots\mathcal{I}_1\rho_1\right]}{\mathrm{tr}\left[\mathcal{P}_k^{(x_i)}\Lambda_{i:i-1}\mathcal{I}_{i-1}\dots\mathcal{I}_1\rho_1\right]} \\
&= \frac{\langle x_j|\Lambda_{j:j-1}\mathcal{I}_{j-1}\Lambda_{j-1:j-2}\dots\mathcal{I}_{i+1}\Lambda_{i+1:i}[|x_i\rangle\langle x_i|]|x_j\rangle\langle x_i|\Lambda_{i:i-1}\mathcal{I}_{i-1}\dots\mathcal{I}_1\rho_1|x_i\rangle}{\langle x_i|\Lambda_{i:i-1}\mathcal{I}_{i-1}\dots\mathcal{I}_1\rho_1|x_i\rangle} \\
&= \langle x_j|\Lambda_{j:j-1}\mathcal{I}_{j-1}\Lambda_{j-1:j-2}\dots\mathcal{I}_{i+1}\Lambda_{i+1:i}[|x_i\rangle\langle x_i|]|x_j\rangle.
\end{aligned} \tag{17}$$

Thus, we see that both conditional probabilities are equal and independent of all measurement outcomes prior to t_i i.e., we have $\mathbb{P}(x_j|\mathcal{I}_{j-1:i+1},x_i,x_{i-1},\dots,x_1) = \mathbb{P}(x_j|\mathcal{I}_{j-1:i+1},x_i,\mathcal{I}_{i-1:1}) =: \mathbb{P}(x_j|x_i)$ and the sub-statistics are indeed Markovian. Regarding compatibility, note that for any combination of measuring or not in the times prior to t_i, the only changes to the above expressions occur in the numerator term that always cancels with the corresponding part in the denominator, and so compatibility also holds true. In other words, we have $\mathbb{P}(x_j|\mathcal{I}_{j-1:i+1},x_i,(x\cup\mathcal{I})_{i-1:1}) = \mathbb{P}(x_j|x_i)$ for all possible choices of $(x\cup\mathcal{I})_{i-1:1}$, i.e., all possible choices of measuring or not at times $\{t_1,\dots,t_{i-1}\}$ in the history. As for the classical case we demonstrated in App. A.2, the argument above can be run in exactly the same way for more 'complicated' subsets $\Gamma^{(i)} \subset \mathcal{T}_n$, with the only difference being that the notation becomes slightly more cumbersome.

C Hidden Quantum Memory and Incompatibility

C.1 Hidden Quantum Memory

Here we explicitly calculate all sub-statistics of the example used regarding Thm. 1 and show that, while the full statistics is Markovian, there are non-Markovian sub-statistics, i.e., we uncover hidden quantum memory. This phenomenon acts as a witness to the impossibility of a memoryless quantum dynamical

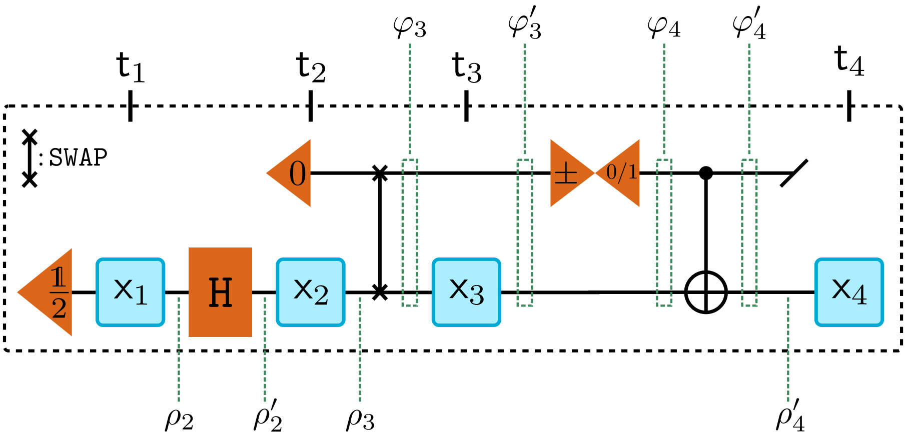

Figure 4: *Circuit with hidden quantum memory.* For convenience, we reproduce the circuit provided in Fig. 2 in the main text. Additionally, to better facilitate orientation, the states that are explicitly mentioned throughout the proof are marked in green, i.e., the points in the circuit where the states $\rho_2, \rho_2', \rho_3, \varphi_3, \dots$ occur.

model by way of contradiction with (the first part of) Lem. 2, which states that any memoryless dynamics leads to Markovian sub-statistics. The circuit corresponding to the process we discuss is shown in Fig. 4, where, for convenience, the states we explicitly calculate in the discussion below are annotated.

We begin with the full statistics. The probability over measurement outcomes at time t_1 are set by the initial state of the process, i.e.,

$$\mathbb{P}(x_1) = \operatorname{tr}\left[|x_1\rangle\langle x_1|\rho_1\right], \tag{18}$$

with the post-measurement (sub-normalised) state given by $\rho_2(x_1) = \mathbb{P}(x_1)|x_1\rangle\langle x_1|$. Without loss of generality, we choose $\rho_1 = \frac{\mathbb{1}}{2}$ and thereby set $\mathbb{P}(x_1 = 0) = \mathbb{P}(x_1 = 1) = \frac{1}{2}$. The process then consists of a Hadamard gate, which rotates said post-measurement state (which is diagonal in the σ_z-basis) to the σ_x-basis, and we have

$$\rho_2'(x_1) := H\rho_2(x_1)H = \begin{cases} \frac{1}{2}|+\rangle\langle+| & \text{for } x_1 = 0 \\ \frac{1}{2}|-\rangle\langle-| & \text{for } x_1 = 1. \end{cases} \tag{19}$$

The experimenter then measures the σ_z observable again, yielding the joint two-time statistics

$$\mathbb{P}(x_2, x_1) = \frac{1}{4} \quad \forall\; x_2, x_1. \tag{20}$$

The state after the second measurement is independent of x_1 and given by

$$\rho_3(x_2, x_1) = \frac{1}{4}|x_2\rangle\langle x_2| \quad \forall\; x_2, x_1. \tag{21}$$

This state is then swapped with the environment, which is prepared in an arbitrary fiducial state τ, which we set as the blank state $|0\rangle$. The joint system-environment state φ_3 immediately prior to the measurement at t_3 is given by

$$\varphi_3(x_2, x_1) = \texttt{SWAP}[\rho_3(x_2, x_1) \otimes \tau] = \tau \otimes \rho_3(x_2, x_1) = \frac{1}{4}|0\rangle\langle 0| \otimes |x_2\rangle\langle x_2|. \tag{22}$$

The experimenter then measures the system at time t_3, recording the probabilities

$$\mathbb{P}(x_3, x_2, x_1) = \begin{cases} \frac{1}{4} & \text{for } x_3 = 0 \\ 0 & \text{for } x_3 = 1 \end{cases} \quad \forall\; x_2, x_1. \tag{23}$$

This distribution is Markovian, as we have the conditional probabilities

$$\mathbb{P}(x_3|x_2,x_1)=\begin{cases}1 & \text{for } x_3=0\\ 0 & \text{for } x_3=1\end{cases} \quad \forall\ x_2,x_1, \tag{24}$$

which are independent of x_1 [in fact, the statistics are 'super'-Markovian as the conditional probabilities are even independent of x_2, so we have $\mathbb{P}(x_3|x_2,x_1)=\mathbb{P}(x_3|x_2)=\mathbb{P}(x_3)$]. The system-environment state $\varphi_3'(x_3,x_2,x_1)$ following the measurement at t_3 is

$$\varphi_3'(x_3,x_2,x_1)=(\mathcal{P}_3^{(x_3)}\otimes\mathcal{I})[\varphi_3(x_2,x_1)]=\frac{1}{4}|0\rangle\langle 0|\otimes|x_2\rangle\langle x_2|, \tag{25}$$

i.e., the measurement at t_3 is non-invasive (note that the outcome $x_3=1$ cannot occur). Subsequently, a channel occurs that measures the environment in the σ_x-basis and feeds forward $|0\rangle$ ($|1\rangle$) whenever the measurement outcome is $+$ ($-$). The corresponding CPTP map is given by $\Upsilon[\bullet]=\sum_k Y^k\bullet Y^{k\dagger}$ with Kraus operators $Y^0=|0\rangle\langle +|$ and $Y^1=|1\rangle\langle -|$. Since $\langle\pm|x_2\rangle\langle x_2|\pm\rangle=\frac{1}{2}\,\forall\, x_2$, this yields the system-environment state

$$\varphi_4(x_3,x_2,x_1)=(\mathcal{I}\otimes\Upsilon)[\varphi_3'(x_3,x_2,x_1)]=\frac{1}{8}|0\rangle\langle 0|\otimes\mathbb{1}. \tag{26}$$

After this, a `CNOT` gate on system and environment occurs (with the environment qubit acting as control), leading to

$$\varphi_4'(x_3,x_2,x_1)=\texttt{CNOT}[\varphi_4(x_3,x_2,x_1)]=\frac{1}{8}(|00\rangle\langle 00|+|11\rangle\langle 11|). \tag{27}$$

The experimenter performs the final measurement at t_4, recording the probabilities

$$\mathbb{P}(x_4,x_3,x_2,x_1)=\begin{cases}\frac{1}{8} & \text{for } x_3=0\\ 0 & \text{for } x_3=1\end{cases} \quad \forall\ x_4,x_2,x_1. \tag{28}$$

This distribution is indeed Markovian, as the conditional probabilities are

$$\mathbb{P}(x_4|x_3,x_2,x_1)=\begin{cases}\frac{1}{2} & \text{for } x_3=0\\ 0 & \text{for } x_3=1\end{cases} \quad \forall\ x_4,x_2,x_1, \tag{29}$$

[where we take the convention that conditioning on an event that cannot occur (i.e., $x_3=1$) gives conditional probability 0]. Thus, the full joint statistics $\mathbb{P}(x_4,x_3,x_2,x_1)$ is Markovian.

On the other hand, consider the situation in which the experimenter does not measure at time t_2, i.e., observes the sub-statistics $\mathbb{P}(x_4,x_3,\mathcal{I}_2,x_1)$. Everything until Eq. (19) remains the same, but without measurement at t_2 we have the state

$$\rho_3(\mathcal{I}_2,x_1)=\mathcal{I}_2[\rho_2'(x_1)]=\frac{1}{2}\begin{cases}|+\rangle\langle +| & \text{for } x_1=0\\ |-\rangle\langle -| & \text{for } x_1=1.\end{cases} \tag{30}$$

The system is then swapped with the environment, yielding the joint state

$$\varphi_3(\mathcal{I}_2,x_1)=\texttt{SWAP}[\rho_3(\mathcal{I}_2,x_1)\otimes\tau]=\tau\otimes\rho_3(\mathcal{I}_2,x_1)=\frac{1}{2}\begin{cases}|0\rangle\langle 0|\otimes|+\rangle\langle +| & \text{for } x_1=0\\ |0\rangle\langle 0|\otimes|-\rangle\langle -| & \text{for } x_1=1.\end{cases} \tag{31}$$

Measurement of the system at t_3 leads to the joint statistics

$$\mathbb{P}(x_3,\mathcal{I}_2,x_1)=\begin{cases}\frac{1}{2} & \text{for } x_3=0\\ 0 & \text{for } x_3=1\end{cases} \quad \forall\ x_1. \tag{32}$$

Thus, we have the conditional probabilities

$$\mathbb{P}(x_3|\mathcal{I}_2, x_1) = \begin{cases} 1 & \text{for } x_3 = 0 \\ 0 & \text{for } x_3 = 1 \end{cases} \quad \forall\, x_1. \tag{33}$$

The system-environment state $\varphi_3'(x_3, \mathcal{I}_2, x_1)$ following the measurement at t_3 is

$$\varphi_3'(x_3, \mathcal{I}_2, x_1) = (\mathcal{P}_3^{(x_3)} \otimes \mathcal{I})[\varphi_3(\mathcal{I}_2, x_1)] = \frac{1}{2} \begin{cases} |0\rangle\langle 0| \otimes |+\rangle\langle +| & \text{for } x_1 = 0 \\ |0\rangle\langle 0| \otimes |-\rangle\langle -| & \text{for } x_1 = 1. \end{cases} \tag{34}$$

After the map Υ on the environment, we have the system-environment state

$$\begin{aligned} \varphi_4(x_3, \mathcal{I}_2, x_1) = (\mathcal{I} \otimes \Upsilon)[\varphi_3'(x_3, \mathcal{I}_2, x_1)] &= \frac{1}{2} \begin{cases} |0\rangle\langle 0| \otimes |0\rangle\langle 0| & \text{for } x_1 = 0 \\ |0\rangle\langle 0| \otimes |1\rangle\langle 1| & \text{for } x_1 = 1 \end{cases} \\ &= \frac{1}{2}|0\rangle\langle 0| \otimes |x_1\rangle\langle x_1|. \end{aligned} \tag{35}$$

Upon application of the `CNOT` gate, the system-environment state is

$$\varphi_4'(x_3, \mathcal{I}_2, x_1) = \texttt{CNOT}[\varphi_4(x_3, \mathcal{I}_2, x_1)] = \frac{1}{2}|x_1\rangle\langle x_1| \otimes |x_1\rangle\langle x_1|. \tag{36}$$

The experimenter finally performs the measurement at t_4 on the reduced state of the system $\rho_4'(x_3, \mathcal{I}_2, x_1) = \frac{1}{2}|x_1\rangle\langle x_1|$ (with $x_3 = 0$ being the only possibility), recording the statistics

$$\mathbb{P}(x_4, x_3, \mathcal{I}_2, x_1) = \begin{cases} \frac{1}{2}\delta_{x_4 x_1} & \text{for } x_3 = 0 \\ 0 & \text{for } x_3 = 1. \end{cases} \tag{37}$$

This sub-statistics is, however, non-Markovian, since the conditional probability at time t_4 depends on x_1. Explicitly, we have

$$\mathbb{P}(x_4|x_3, \mathcal{I}_2, x_1) = \frac{\mathbb{P}(x_4, x_3, \mathcal{I}_2, x_1)}{\mathbb{P}(x_3, \mathcal{I}_2, x_1)} = \begin{cases} \delta_{x_4 x_1} & \text{for } x_3 = 0 \\ 0 & \text{for } x_3 = 1 \end{cases} \neq \mathbb{P}(x_4|x_3). \tag{38}$$

As we have discussed in the main text, such non-Markovian sub-statistics cannot arise for a memoryless quantum dynamics probed by sharp, projective measurements (as is the case in this example), and therefore we conclude that the statistics observed—although Markovian on the whole—cannot be faithfully reproduced by a memoryless quantum dynamical model.

C.2 Incompatibility

Here we explicitly calculate all sub-statistics of the example used regarding Thm. 2 and show that although they are all Markovian (i.e., unlike in Thm. 1, there is no explicit activation of hidden quantum memory witnessed via non-Markovian sub-statistics), they are nonetheless incompatible and therefore serve to witness the impossibility of a memoryless quantum dynamical description by way of contradiction with (the second part of) Lem. 2.

Intuitively, the corresponding circuit is such that the state of the environment at time t_3, where system and environment are in a product state, does *not* depend upon any previously observed measurement outcomes at times t_1 and t_2 (thus yielding Markovian (sub-)statistics for all conceivable subsets $\Gamma \subseteq \{t_1, t_2, t_3, t_4\}$ at which measurements can be performed), but rather only on *whether or not* any such prior measurements were performed (thus leading to Markovian but incompatible (sub-)statistics. To see this explicitly, we now calculate the intermediate system-environment states at all relevant points throughout the circuit (see Fig. 5 for better orientation).

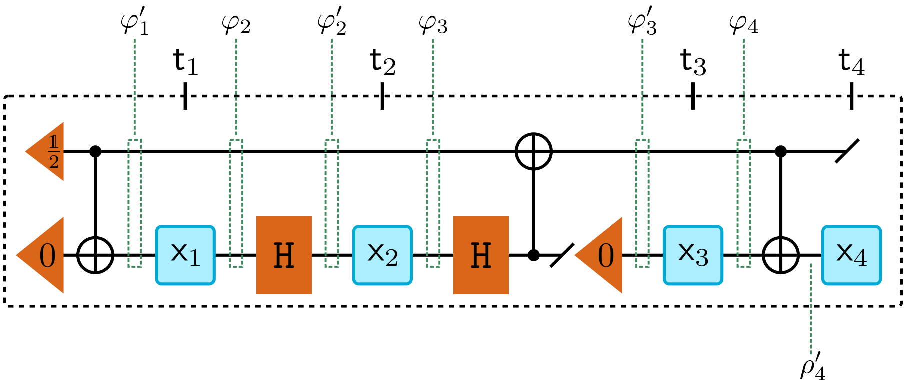

Figure 5: *Circuit with incompatible Markovian (sub-)statistics.* For convenience, we reproduce the circuit provided in Fig. 3 in the main text. Additionally, to better facilitate orientation, the states that are explicitly mentioned throughout the proof are marked in green, i.e., the points in the circuit where the states $\varphi_1', \varphi_2, \dots$ occur.

Firstly, we have

$$\varphi_1' = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|)\,. \tag{39}$$

After time t_1, this state is transformed to either

$$\varphi_2(x_1) = \frac{1}{2}|x_1x_1\rangle\langle x_1x_1| \quad \text{or} \quad \varphi_2(\mathcal{I}_1) = \varphi_1' \tag{40}$$

depending on whether a measurement was performed or not. Following the first Hadamard gate applied to the system, we then have either

$$\varphi_2'(x_1) = \frac{1}{2}|\pm\rangle\langle\pm|^{x_1} \otimes |x_1\rangle\langle x_1| \quad \text{or} \quad \varphi_2'(\mathcal{I}_1) = \frac{1}{2}(|{+0}\rangle\langle{+0}| + |{-1}\rangle\langle{-1}|). \tag{41}$$

Here, $|\pm\rangle\langle\pm|^{x_1}$ is used to denote the state $|+\rangle\langle+|$ for $x_1 = 0$ and $|-\rangle\langle-|$ for $x_1 = 1$. Analogously, after time t_2, we have the system-environment states

$$\varphi_3(x_2, x_1) = \frac{1}{4}|x_2x_1\rangle\langle x_2x_1|, \quad \varphi_3(\mathcal{I}_2, x_1) = \frac{1}{2}|\pm\rangle\langle\pm|^{x_1} \otimes |x_1\rangle\langle x_1|, \quad \text{or} \quad \varphi_3(x_2, \mathcal{I}_1) = \frac{1}{4}|x_2\rangle\langle x_2| \otimes \mathbb{1}. \tag{42}$$

Between times t_2 and t_3, the system and environment undergo a Hadamard gate on the system, followed by a `CNOT` gate (with system acting as control)), after which the system is discarded and re-prepared in the $|0\rangle$ state. Consequently, the system-environment state immediately prior to the measurement at t_3 is one of

$$\varphi_3'(x_2, x_1) = \frac{1}{8}|0\rangle\langle 0| \otimes \mathbb{1}, \quad \varphi_3'(\mathcal{I}_2, x_1) = \frac{1}{2}|00\rangle\langle 00|, \quad \text{or} \quad \varphi_3'(x_2, \mathcal{I}_1) = \frac{1}{4}|0\rangle\langle 0| \otimes \mathbb{1}\,. \tag{43}$$

Consequently, the state of the environment depends on whether or not measurements were performed at t_1 and t_2 (e.g., it is proportional to $\mathbb{1}$ if both measurements were performed, and proportional to $|0\rangle\langle 0|$ if only the measurement at t_1 was performed), while the reduced state of the system is always equal to $|0\rangle\langle 0|$ for any combination of measurements and outcomes. With this, we see directly that

$$\mathbb{P}(x_3|x_2, x_1) = \mathbb{P}(x_3|\mathcal{I}_2, x_1) = \mathbb{P}(x_3|x_2, \mathcal{I}_1) = \delta_{x_3 0}, \tag{44}$$

and therefore the three-point statistics ending at time t_3 are Markovian (and compatible), since the corresponding conditional probabilities do not depend on *any* previous outcomes. Analogously, since the system-environment state φ_4 immediately before the final `CNOT` gate depends (at most) on the outcome x_3, but not on outcomes x_2 or x_1, *all* statistics ending at time t_4 are also Markovian (note that for a four-step process, the statistics ending at times t_3 and t_4 are the only ones that have to be checked with respect to Markovianity).

However, the resulting four-point conditional probabilities are not compatible. For example, we have

$$\varphi_4(x_3, x_2, x_1) = \frac{1}{8}\delta_{x_3 0}|x_3\rangle\langle x_3| \otimes \mathbb{1} \quad \text{and} \quad \varphi_4(x_3, \mathcal{I}_2, x_1) = \frac{1}{2}\delta_{x_3 0}|x_3 0\rangle\langle x_3 0|\,, \tag{45}$$

such that following the final `CNOT`, the reduced system state is either

$$\rho_4'(x_3, x_2, x_1) = \frac{1}{8}\delta_{x_3 0}\mathbb{1} \quad \text{or} \quad \rho_4'(x_3, \mathcal{I}_2, x_1) = \frac{1}{2}\delta_{x_3 0}|x_3\rangle\langle x_3|, \tag{46}$$

depending on whether or not a measurement was made at time t_2. Consequently, we can compute the corresponding statistics in either scenario:

$$\mathbb{P}(x_4, x_3, x_2, x_1) = \frac{1}{8}\delta_{x_3 0} \quad \text{and} \quad \mathbb{P}(x_4, x_3, \mathcal{I}_2, x_1) = \frac{1}{2}\delta_{x_3 0}\delta_{x_4 x_3}. \tag{47}$$

In a similar vein, from Eq. (45), we can directly compute the relevant statistics for the processes ending at time t_3 to be

$$\mathbb{P}(x_3, x_2, x_1) = \frac{1}{4}\delta_{x_3 0} \quad \text{and} \quad \mathbb{P}(x_3, \mathcal{I}_2, x_1) = \frac{1}{2}\delta_{x_3 0}. \tag{48}$$

Finally, combining the above two equations, we obtain

$$\mathbb{P}(x_4|x_3, x_2, x_1) = \frac{1}{2}\delta_{x_3 0} \quad \text{and} \quad \mathbb{P}(x_4|x_3, \mathcal{I}_2, x_1) = \delta_{x_3 0}\delta_{x_4 x_3}. \tag{49}$$

Although both of these conditional probability distributions are Markovian (i.e., show no dependence on outcomes prior to x_3), they nonetheless differ depending upon whether some intervention was made overall in the past, i.e., they are incompatible. Since these conditional probabilities are incompatible, there cannot be a memoryless quantum dynamics that faithfully reproduces them (as demonstrated in Lem. 2), proving the claim.