

## Emergence of Memory in Equilibrium versus Nonequilibrium Systems

Xizhu Zhao<sup>1,2</sup>, David Hartich,<sup>1</sup> and Aljaž Godec<sup>1,\*</sup>

<sup>1</sup>Mathematical bioPhysics Group, Max Planck Institute for Multidisciplinary Sciences, Am Faßberg 11, 37077 Göttingen

<sup>2</sup>Max Planck School Matter to Life, Jahnstraße 29, D-69120 Heidelberg, Germany

(Received 18 November 2023; accepted 1 February 2024; published 2 April 2024)

Experiments often probe observables that correspond to low-dimensional projections of high-dimensional dynamics. In such situations distinct microscopic configurations become lumped into the same observable state. It is well known that correlations between the observable and the hidden degrees of freedom give rise to memory effects. However, how and under which conditions these correlations emerge remain poorly understood. Here we shed light on two fundamentally different scenarios of the emergence of memory in minimal stationary systems, where observed and hidden degrees of freedom either evolve cooperatively or are coupled by a hidden nonequilibrium current. In the reversible setting the strongest memory manifests when the timescales of hidden and observed dynamics overlap, whereas, strikingly, in the driven setting maximal memory emerges under a clear timescale separation. Our results hint at the possibility of fundamental differences in the way memory emerges in equilibrium versus driven systems that may be utilized as a “diagnostic” of the underlying hidden transport mechanism.

DOI: 10.1103/PhysRevLett.132.147101

Observables coupled to hidden degrees of freedom that do not relax sufficiently fast [1] or selected reaction coordinates that do not locally equilibrate in mesostates [2] generically display memory. In fact, this holds for most high-dimensional dynamics probed on a coarse-grained level [3–14]. Tremendous progress has been made over the years in describing and understanding kinetic aspects of non-Markovian dynamics [1,12,15–37]. More recently, coarse-grained, partially observed dynamics have become of great interest from the point of view of thermodynamic inference [2,37–52]. Namely, while there are efficient methods to detect [53–55] and quantify [55] the existence of memory, it conversely turns out to be quite challenging to quantify [2,42,44,56] or even infer [37,50,51,57] irreversibility from lower-dimensional, projected dynamics. Thus, understanding potential differences in the emergence of memory in equilibrium and nonequilibrium systems is a difficult task.

When one considers, in particular, ergodic dynamics in the sense that the probability distribution to be found in a given microscopic state at long times relaxes to a unique stationary equilibrium or nonequilibrium steady state from any initial condition, the extent of memory is necessarily finite and is more prominent if the hidden degrees of freedom are slow [1,2,58]. Yet, even in this “well behaved,” thermodynamically consistent [59] setting quite little is

known about the possible ways in which the dynamics of observables can become correlated with that of hidden degrees of freedom on different timescales. A particularly intriguing question is whether there are any characteristic differences between how memory emerges in reversible versus irreversible driven systems when the observed dynamics is much faster than the hidden one.

Addressing this problem in full generality is a daunting task. Here we focus on the minimal “cooperative” setting [60–63], where the microscopic dynamics is a Markov process on a planar network and we observe only the vertical coordinate, whereas the horizontal transitions are hidden (see Fig. 1). This setting is important for understanding “active secondary transport”—transporter proteins exploiting the energy stored in the transmembrane gradient

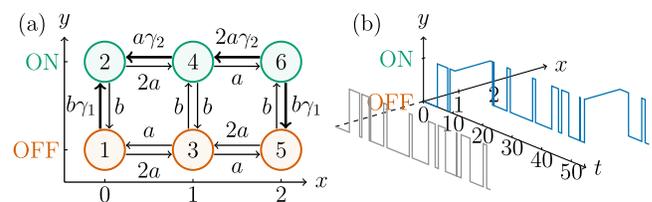


FIG. 1. (a) Schematics of the full Markov network for the respective models. The upper green states are lumped to the observed “ON” state and the lower orange ones to the observed “OFF” state. To separate the time scale of horizontal and vertical dynamics, we choose the rates (in arbitrary units)  $a = 0.5$  and  $b = 25$  to make the vertical process always much faster than the horizontal one. In the driven model (see Fig. 2) we set  $\gamma_1 > 1$  and  $\gamma_2 = 1$ , whereas in the reversible “allosteric” model (Fig. 3) we chose  $\gamma_1 = \gamma_2 \equiv \gamma \geq 1$ . (b) Trajectory of the full dynamics (blue) and its corresponding projection (gray).

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Open access publication funded by the Max Planck Society.

of one type of molecules to transport another type against their gradients [64–66]. Moreover, it is also relevant as the physical basis of the sensitivity of the flagellar motor in *E. coli* in sensing concentrations of its regulator [60–62,67], where two fundamentally different explanations were proposed to explain the “ultrasensitivity” of the motor’s response, an equilibrium allosteric and a dissipative non-equilibrium model [61,67,68].

We are not interested in the biophysical implications of the model. Instead, we use the above two distinct settings merely as a minimal model of non-Markovian two-state dynamics, where we can “turn on” memory in a controlled manner via an equilibrium versus nonequilibrium mechanism. We consider the scenario where the observed dynamics is faster than the hidden one [see Fig. 1(a)] but there is not necessarily a large timescale separation present. Interestingly, in the driven setting we only change

$$L = \begin{pmatrix} -b\gamma_1 - 2a & b & a & 0 & 0 & 0 \\ b\gamma_1 & -b - 2a & 0 & a\gamma_2 & 0 & 0 \\ 2a & 0 & -b - 2a & b & 2a & 0 \\ 0 & 2a & b & -b - a - a\gamma_2 & 0 & 2a\gamma_2 \\ 0 & 0 & a & 0 & -b - 2a & b\gamma_1 \\ 0 & 0 & 0 & a & b & -b\gamma_1 - 2a\gamma_2 \end{pmatrix}, \quad (1)$$

where we choose  $a = 1/2$  and  $b = 25$  such that in the “baseline” model horizontal transitions are slower than vertical ones; i.e., the hidden dynamics relaxes slower than the observable ON  $\rightleftharpoons$  OFF transitions. We adopt the Dirac bra-ket notation and denote the transition probability from microscopic state  $i$  to microscopic state  $j$  in time  $t$  as  $G(j, t|i) \equiv \langle j|e^{Lt}|i\rangle$  and the stationary probability of state  $j$  as  $P_s(j) = \lim_{t \rightarrow \infty} G(j, t|i)$ . In the baseline model with  $\gamma_1 = \gamma_2 = 1$  the dynamics is reversible and the ON-OFF states are equiprobable in the steady state. The parameters  $\gamma_1$  and  $\gamma_2$  are *acceleration factors* when they are larger than 1. In the cooperative allosteric regime with  $\gamma_1 = \gamma_2 \equiv \gamma > 1$  the dynamics obeys detailed balance, i.e.,  $P_s(i)L_{j,i} = P_s(j)L_{i,j}$ ,  $\forall i, j$ . Conversely, in the driven model we set  $\gamma_1 > 1, \gamma_2 = 1$ ; here detailed balance is violated, i.e.,  $\exists i, j$  for which  $P_s(i)L_{j,i} \neq P_s(j)L_{i,j}$  (that is, the model is irreversible yet thermodynamically consistent [59]).

It turns out that  $L$  defined this way is diagonalizable, i.e., we can find a biorthonormal basis  $\{|\psi_k^L\rangle, |\psi_k^R\rangle\}$  of left  $\langle \psi_k^L|$  and right  $|\psi_k^R\rangle$  eigenvectors with eigenvalue  $-\lambda_k$  and  $k = 0, \dots, 5$  and  $\langle \psi_k^L|\psi_l^R\rangle = \delta_{kl}$ . Thus, we have  $L = \sum_k -\lambda_k |\psi_k^R\rangle \langle \psi_k^L|$ , and in turn we can expand  $G(j, t|i) = \sum_{k=0}^5 \langle j|\psi_k^R\rangle \langle \psi_k^L|i\rangle e^{-\lambda_k t}$ . In this notation the steady-state probability of state  $i$  is given by  $P_s(i) = \langle i|\psi_0^R\rangle$ .

The eigenvalues of the baseline model are  $\lambda_0 = 0, \lambda_1 = 2a, \lambda_2 = 4a, \lambda_3 = 2b, \lambda_4 = 2(a+b)$ , and  $\lambda_5 = 2(2a+b)$ . We can also determine the eigenspectrum analytically for the

the fastest timescales, whereas in the reversible cooperative model we alter all timescales. In both cases we find a maximal capacity for memory; i.e., the maximal magnitude of memory saturates at a finite coupling and nonequilibrium driving, respectively. Interestingly, in the reversible setting the memory manifests strongest when the timescale separation between the hidden and observed transitions becomes partially lifted and the hidden and observable timescales overlap, whereas in the driven setting maximal memory occurs in the presence of a timescale separation. Our results provide deeper insight into the emergence of memory in the distinct situations when observed and hidden dynamics either evolve cooperatively or become coupled to a hidden nonequilibrium current.

*Setup.*—We consider a six-state continuous-time Markov process [see Fig. 1(a)] with generator  $L$ , whose elements  $L_{n,m}$  are transition rates between states  $m \rightarrow n$  given by

driven model, which has eigenvalues  $\lambda_0^{\gamma_1} = 0, \lambda_1^{\gamma_1} = 2a, \lambda_2^{\gamma_1} = 4a, \lambda_3^{\gamma_1} = \frac{1}{2}[4a + (\gamma_1 + 3)b - \sqrt{16a^2 + (\gamma_1 - 1)^2 b^2}], \lambda_4^{\gamma_1} = 2a + (\gamma_1 + 1)b$ , and  $\lambda_5^{\gamma_1} = \frac{1}{2}[4a + (\gamma_1 + 3)b + \sqrt{16a^2 + (\gamma_1 - 1)^2 b^2}]$ . The reversible allosteric model ( $\gamma > 1$ ) cannot be diagonalized analytically, and we therefore provide numerical results instead.

We assume that the full system is prepared in a steady state  $P_s(j)$  and only vertical ON  $\rightleftharpoons$  OFF transitions are observed with observable sets ON = {2, 4, 6} and OFF = {1, 3, 5}. We determine the non-Markovian transition probability of the observed process  $\hat{k}_t, Q_{P_s}(\hat{n}, t|\hat{m})$  with  $\hat{m}, \hat{n} \in \{\text{ON}, \text{OFF}\}$  as [1]

$$Q_{P_s}(\hat{n}, t|\hat{m}) \equiv \frac{\sum_{j=1}^6 \mathbb{1}_{\hat{n}}[j] \sum_{i=1}^6 \mathbb{1}_{\hat{m}}[i] G(j, t|i) P_s(i)}{\sum_{i=1}^6 \mathbb{1}_{\hat{m}}[i] P_s(i)}, \quad (2)$$

where  $\mathbb{1}_{\Omega}$  is the indicator function of the set  $\Omega$ . The non-Markovian transition probability between two fixed observed states  $\hat{m} \rightarrow \hat{n}$  as well as the observable return probability  $\hat{m} \rightarrow \hat{m}$  depends on the preparation of the full system [69]. Moreover, in spite of the full system being prepared in the stationary state  $P_s$ , by specifying the initial observed state (here either ON or OFF) we “quench” the full system out of the steady state by conditioning on the state of the observable [1,69]. Without loss of generality we will focus on the scenario where the observable is initially in the ON state, i.e.,  $\hat{k}_0 = \text{ON}$ .

To quantify the magnitude of memory in the projected dynamics, we follow [55] and construct the auxiliary Chapman-Kolmogorov (CK) transition density

$$Q^{\text{CK}}(\hat{n}, t_1 + t_2 | \hat{m}) \equiv \sum_{\hat{k}} Q_{P_s}(\hat{n}, t_2 | \hat{k}) Q_{P_s}(\hat{k}, t_1 | \hat{m}). \quad (3)$$

Note that for a non-Markovian process  $Q^{\text{CK}}$  depends on both  $t_1$  and  $t_2$ . The Chapman-Kolmogorov construction  $Q^{\text{CK}}(\hat{n}, t_1 + t_2 | \hat{m})$  corresponds to a fictitious dynamics where we *force* at time  $t_1$  all hidden degrees of freedom to their stationary distribution and thereby erase all memory of their initial condition. When the observed ON  $\rightleftharpoons$  OFF dynamics is Markovian we have  $Q^{\text{CK}}(\hat{n}, t_1 + t_2 | \hat{m}) = Q_{P_s}(\hat{n}, t_1 + t_2 | \hat{m}) \forall t_1, t_2 \geq 0$ , but the converse is not true in general [1,55]. As soon as  $Q^{\text{CK}}(\hat{n}, t_1 + t_2 | \hat{m}) \neq Q_{P_s}(\hat{n}, t_1 + t_2 | \hat{m})$  for some  $\hat{n}$ , the observable at time  $t_1 + t_2$  “remembers” the state of hidden degrees of freedom at time  $t_1$ .

We use the Kullback-Leibler divergence  $D_k[p||q] \equiv \sum_k p(k) \ln[p(k)/q(k)] \geq 0$  to quantify the difference between  $Q_{P_s}$  and  $Q^{\text{CK}}$  [55]

$$D_{\hat{m}}^{\text{CK}}(t_1, t_2) \equiv D_{\hat{n}}[Q_{P_s}(\hat{n}, t_1 + t_2 | \hat{m}) || Q^{\text{CK}}(\hat{n}, t_1 + t_2 | \hat{m})], \quad (4)$$

where the superscript  $k$  in  $D_k[p||q]$  denotes the independent dummy variable of the measures  $p$  and  $q$ . In the absence of memory  $D_{\hat{m}}^{\text{CK}}(t_1, t_2) = 0 \forall t_1, t_2$ . Conversely, as we are interested in ergodic dynamics prepared in a steady state, we have  $D_{\hat{m}}^{\text{CK}} \rightarrow 0$  whenever  $t_1 + t_2 \rightarrow 0$  or  $t_1 + t_2 \rightarrow \infty$  [55]. Therefore, by the positivity of  $D_{\hat{m}}^{\text{CK}}$  we will have at least one maximum in the half-space to  $t_1, t_2 > 0$ . We quantify the magnitude of memory in terms of the global maximum on  $t_1, t_2 > 0$  as

$$D_{\text{max}}^{\text{CK}}(\hat{m}) \equiv \sup_{t_1, t_2 > 0} D_{\hat{m}}^{\text{CK}}(t_1, t_2). \quad (5)$$

In the baseline setting ( $\gamma_1 = \gamma_2 = 1$ ) the observed and hidden dynamics are decoupled (i.e., all microscopic pathways are equivalent). As a result, the observed dynamics is Markovian and  $D_{\text{max}}^{\text{CK}}(\hat{m}) = 0$ . We are interested in the dependence of  $D_{\text{max}}^{\text{CK}}(\hat{m})$  as we couple the vertical and horizontal dynamics cooperatively or by a dissipative current, that is, on  $\gamma_1$  and  $\gamma$  in the driven and cooperative models, respectively.

*Driven setting.*—We first consider the driven scenario with  $\gamma_1 > 1$  and  $\gamma_2 = 1$ . Instead of  $\gamma_1$ , we use the steady-state entropy production rate of the microscopic dynamics  $\dot{S}(\gamma_1) = \sum_{i,j} P_s(j) L_{i,j} \ln[P_s(j) L_{i,j} / (P_s(i) L_{j,i})]$  to indicate how far the system is driven out of equilibrium, and we change  $\gamma_1$  in equidistant units of the chemical potential  $\ln(\gamma_1/\gamma_2) = \ln \gamma_1$  that drives the system out of equilibrium, i.e.,  $\gamma_1$  increases exponentially. We first provide some intuition about the microscopic dynamics.

Since  $\lambda_{1,2}$  are independent of  $\gamma_1$  and because  $\lambda_3 \approx 2(a + b)$  for  $b \gg a$  and  $\gamma_1 \geq 2$ , by increasing  $\gamma_1$  we only alter  $\lambda_{4,5}$  [see Fig. 2(a) and Appendix B]. That is, we are essentially

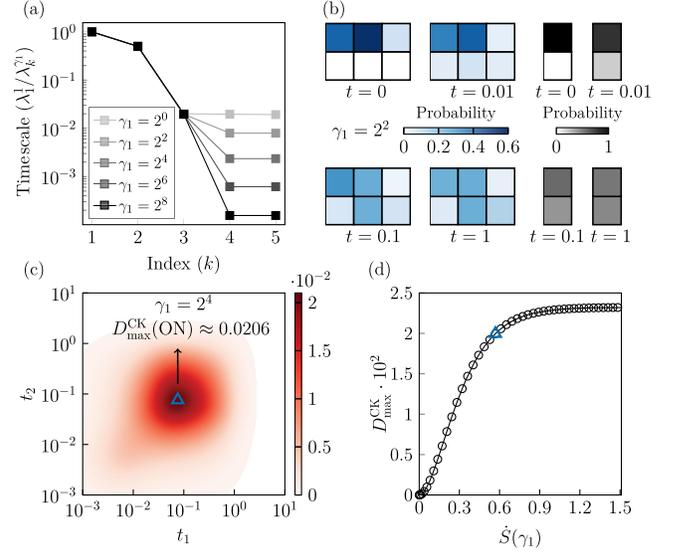


FIG. 2. Driven setting ( $\gamma_1 > 1, \gamma_2 = 1$ ). (a) Characteristic timescales  $1/\lambda_i^i$ ,  $i \geq 1$  as a function of  $\gamma_1$  relative to the baseline relaxation time  $1/\lambda_1^1$ . (b) Left sketch: microscopic transition probability  $G(j, t | \text{ON})$  for  $\gamma_1 = 2^2$  at four different times. Right sketch: corresponding observed transition probability  $Q_{P_s}(\hat{n}, t | \text{ON})$ . (c) Relative entropy  $D_{\text{ON}}^{\text{CK}}(t_1, t_2)$  in Eq. (4) for  $\gamma_1 = 2^4$ . The triangle depicts  $D_{\text{max}}^{\text{CK}}(\text{ON})$ . (d) Magnitude of memory  $D_{\text{max}}^{\text{CK}}(\text{ON})$  in Eq. (5) as a function of driving  $\dot{S}(\gamma_1)$ . The blue triangle denotes the position of the maximum in (c).

only tuning the fastest timescales, whereas the slow timescales remain unaffected by the driving. We also alter the stationary distribution  $P_s(\hat{m})$ . Because  $b \gg a$ , the system (even without driving) tends to first explore vertical paths to reach a “quasisteady state” between observed ON-OFF states within a timescale of approximately  $\sim 1/b = 0.04$  [see Fig. 2(b)]. Afterward, the probability redistributes horizontally within the observable states. However, because the transition rates from states 1 to 2 and from 6 to 5 are accelerated by a factor of  $\gamma_1$ , transitions  $2 \rightarrow 1$  and  $6 \rightarrow 5$  will be instantly followed by the reverse transitions  $1 \rightarrow 2$  and  $5 \rightarrow 6$ . Thus, the probability distribution dominantly redistributes along the microscopic path  $2 \leftrightarrow 4 \leftrightarrow 3 \leftrightarrow 5$  and finally reaches a steady state “skewed” in the hidden direction [see Fig. 2(b)].

We now address the magnitude of the emerging memory via  $D_{\text{max}}^{\text{CK}}$  in Eq. (5). We find that  $D_{\text{max}}^{\text{CK}}$  monotonically increases with  $\gamma_1$  (note that  $\dot{S}$  is a monotonically increasing function of  $\gamma_1$ ), and eventually at  $\approx 0.02324$ , where the location of the supremum approaches  $t_1 = t_2 \approx 0.08$  as  $\gamma_1 \rightarrow \infty$ . Note that the maximal memory is attained on a timescale that is longer than the local vertical equilibration time  $\approx 1/b = 0.04$ . The saturation may be explained by noticing that as  $\gamma_1 \rightarrow \infty$ , only states 2–5 have a nonzero probability and accelerated paths are almost never traversed. As a result,  $D_{\text{max}}^{\text{CK}}$  no longer changes with  $\gamma_1$ .

Note that in this driven setting vertical transitions are always much faster than horizontal ones, which maintains a separation of timescales between the observed and hidden

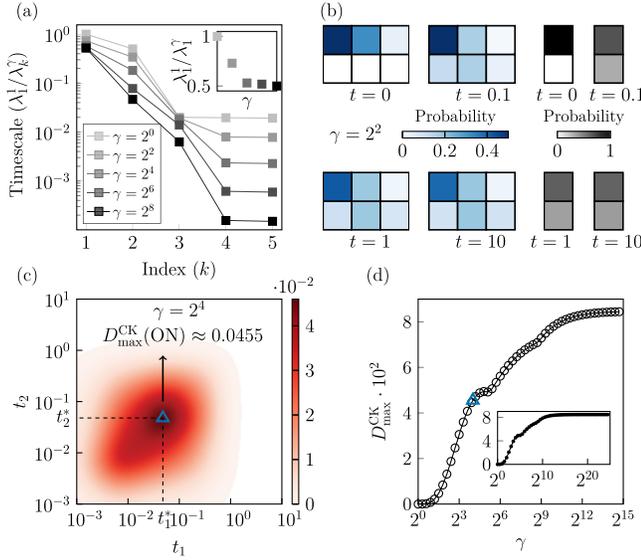


FIG. 3. Reversible setting ( $\gamma_1 = \gamma_2 = \gamma$ ). (a) Characteristic timescales  $1/\lambda_i^\gamma$ ,  $i \geq 1$  as a function of  $\gamma_1 = \gamma_2 = \gamma$  relative to the baseline relaxation time  $1/\lambda_1^1$ . Inset:  $\lambda_1^1/\lambda_i^\gamma$  on a linear scale. (b) Left sketch: microscopic transition probability  $G(j, t|\text{ON})$  for  $\gamma = 2^2$  at four different times. Right sketch: corresponding observed transition probability  $Q_p(\hat{n}, t|\text{ON})$ . (c) Relative entropy  $D_{\text{ON}}^{\text{CK}}(t_1, t_2)$  in Eq. (4) for  $\gamma = 2^4$  as a function of  $t_1, t_2$ . The triangle denotes  $D_{\text{max}}^{\text{CK}}(\text{ON})$ . (d) Magnitude of memory  $D_{\text{max}}^{\text{CK}}(\text{ON})$  in Eq. (5) as a function of  $\gamma$ .

dynamics. The memory that we observe is thus “only” a manifestation of the relaxation of hidden degrees of freedom.

*Reversible cooperative setting.*—We now inspect the reversible scenario where  $\gamma_1 = \gamma_2 \equiv \gamma \geq 1$ . As before, we first give some insight into the microscopic dynamics. As time evolves, in the first stage the system initially populates microscopic states 2 and 4, where  $\gamma$  is large, especially state 2. The accelerated transition paths do not instantly play a role. Thus, as in the driven model, the system tends to first explore vertical paths (paths  $2 \rightarrow 1$  and  $4 \rightarrow 3$ ) on a timescale of  $1/b \sim 0.04$ . In the second stage, also as in the driven model, transitions  $2 \rightarrow 1$  will instantly go back to state 2. During this stage, the probability redistributes in the horizontal direction. However, since the transition rates inside the ON state ( $L_{2,4}, L_{4,6}$ ) are also accelerated by a factor of  $\gamma$ , the timescale of the horizontal redistribution of probability is not necessarily larger than that of vertical dynamics. Here the two main frequently visited vertical paths are  $2 \rightarrow 1$  and  $4 \rightarrow 3$ .

Note that when we increase  $\gamma$  in this reversible setting such that  $a\gamma > b$ , some of the horizontal transition rates exceed the vertical ones, i.e., the timescales of the observable dynamics and the hidden dynamics overlap and there is no timescale separation. This overlapping (and “mixing”) of the hidden and observable timescales may contribute to the appearance of two shoulders in Fig. 3(d) at  $\gamma \approx 38$  and  $\gamma \approx 443$ . The shoulders are a result of the shift in position of the peak of  $D_{\hat{m}}^{\text{CK}}(t_1, t_2)$  (for details see Fig. 4 in Appendix A). As  $\gamma$  tends to become very large,  $D_{\text{max}}^{\text{CK}}$

saturates to  $\approx 0.08471$ , and the location of peak approaches  $t_1 = t_2 \approx 0.13$ .

*Conclusion.*—In this Letter we addressed the emergence of memory in a minimal setting, where the microscopic dynamics corresponds to a Markov process on a planar network and we observed only the vertical ON  $\rightleftharpoons$  OFF dynamics, whereby the horizontal dynamics are hidden. Our aim was to gain insight into how correlations between the hidden and observed dynamics emerge, in particular, if and how the nature of these correlations depends on whether the microscopic dynamics is reversible (i.e., whether it obeys detailed balance) or instead is driven. In the former scenario, the observed and hidden degrees of freedom are coupled cooperatively, whereas in the latter scenario the coupling emerges due to a nonequilibrium current. We focused on quantifying the magnitude of memory while tuning cooperativity or irreversible driving. Many features were found to be similar in the two settings. However, in the reversible setting the strongest memory was found in the expected situation, when the timescales of the hidden and observed dynamics overlap. Conversely, in the driven setting maximal memory is reached under a clear timescale separation. Our Letter therefore unraveled qualitative differences in the way that memory can emerge in equilibrium versus driven systems. While we focused on a simple model, our findings pave the way for more systematic studies. From a practical, “diagnostic” perspective, our results imply the possibility of gaining insight into the dynamic coupling underlying active secondary transport [64–66] from observations of memory in the transmembrane transport of either species.

Financial support from the Max Planck School Matter to Life, supported by the German Federal Ministry of Education and Research (BMBF) in collaboration with the Max Planck Society (fellowship for X. Z.), and from the European Research Council (ERC) under the European Union’s Horizon Europe Research and Innovation Program (Grant Agreement No. 101086182 for A. G.) is gratefully acknowledged.

*Appendix A: Peak position of  $D_{\hat{m}}^{\text{CK}}$  in the reversible setting.*—Let  $(t_1^*, t_2^*) \equiv \arg \max_{t_1, t_2} D_{\hat{m}}^{\text{CK}}(t_1, t_2)$  be the values of  $t_1$  and  $t_2$  when  $D_{\text{ON}}^{\text{CK}}$  reaches its maximum at a

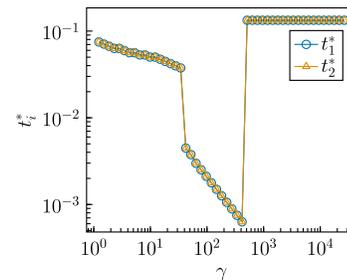


FIG. 4. Dependence of the peak position of  $D_{\text{ON}}^{\text{CK}}(t_1^*, t_2^*)$  on the cooperativity parameter  $\gamma$  in the reversible model. Note the two discontinuities.

TABLE I. Characteristic timescales relative to the baseline relaxation time in driven ( $\lambda_1^1/\lambda_k^1$ ) and reversible ( $\lambda_1^1/\lambda_k^1$ ) settings.

Index $k$	Baseline	Driven ( $\gamma_1 > 1, \gamma_2 = 1$ )		Reversible ( $\gamma_1 = \gamma_2 = \gamma$ )	
	$\gamma_1 = \gamma_2 = 1$	$\gamma_1 = 2^4$	$\gamma_1 = 2^8$	$\gamma = 2^4$	$\gamma = 2^8$
1	1	1	1	0.578	0.519
2	0.5	0.5	0.5	0.181	0.0465
3	0.02	0.0196	0.0196	0.0182	$6.29 \times 10^{-3}$
4	0.0196	$2.35 \times 10^{-3}$	$1.56 \times 10^{-4}$	$2.35 \times 10^{-3}$	$1.56 \times 10^{-4}$
5	0.0192	$2.35 \times 10^{-3}$	$1.56 \times 10^{-4}$	$2.27 \times 10^{-3}$	$1.50 \times 10^{-4}$

given  $\gamma$  in the reversible cooperative scenario [see Fig. 3(c)]. As stated in the main text, the two shoulders in Fig. 3(d) are the result of discontinuities in the shift of peak position. To visualize this, we show in Fig. 4 the dependence of the peak position on  $\gamma$ . Note that  $G^{\text{CK}}$  in Eq. (3) is a symmetric function of  $t_1$  and  $t_2$ , so the peak of  $D_{\text{ON}}^{\text{CK}}$  always occurs at  $t_1^* = t_2^*$ .

*Appendix B: Table of characteristic timescales.*—Table I lists some of the characteristic timescales under the driven and reversible settings shown in Figs. 2(a) and 3(a), respectively.

\*agodec@mpinat.mpg.de

- [1] A. Lapolla and A. Godec, Manifestations of projection-induced memory: General theory and the tilted single file, *Front. Phys.* **7**, 182 (2019).
- [2] D. Hartich and A. Godec, Emergent memory and kinetic hysteresis in strongly driven networks, *Phys. Rev. X* **11**, 041047 (2021).
- [3] S. S. Plotkin and P. G. Wolynes, Non-Markovian configurational diffusion and reaction coordinates for protein folding, *Phys. Rev. Lett.* **80**, 5015 (1998).
- [4] W. Min, G. Luo, B. J. Cherayil, S. C. Kou, and X. S. Xie, Observation of a power-law memory kernel for fluctuations within a single protein molecule, *Phys. Rev. Lett.* **94**, 198302 (2005).
- [5] T. Neusius, I. Daidone, I. M. Sokolov, and J. C. Smith, Subdiffusion in peptides originates from the fractal-like structure of configuration space, *Phys. Rev. Lett.* **100**, 188103 (2008).
- [6] A. K. Sangha and T. Keyes, Proteins fold by subdiffusion of the order parameter, *J. Phys. Chem. B* **113**, 15886 (2009).
- [7] S. M. Avdoshenko, A. Das, R. Satija, G. A. Papoian, and D. E. Makarov, Theoretical and computational validation of the Kuhn barrier friction mechanism in unfolded proteins, *Sci. Rep.* **7**, 269 (2017).
- [8] D. E. Makarov, Interplay of non-Markov and internal friction effects in the barrier crossing kinetics of biopolymers: Insights from an analytically solvable model, *J. Chem. Phys.* **138**, 014102 (2013).
- [9] M. Ozmaian and D. E. Makarov, Transition path dynamics in the binding of intrinsically disordered proteins: A simulation study, *J. Chem. Phys.* **151**, 235101 (2019).
- [10] H. Meyer, P. Pelagejcev, and T. Schilling, Non-Markovian out-of-equilibrium dynamics: A general numerical procedure to construct time-dependent memory kernels for coarse-grained observables, *Europhys. Lett.* **128**, 40001 (2020).
- [11] A. G. T. Pyo and M. T. Woodside, Memory effects in single-molecule force spectroscopy measurements of biomolecular folding, *Phys. Chem. Chem. Phys.* **21**, 24527 (2019).
- [12] C. Ayaz, L. Scalfi, B. A. Dalton, and R. R. Netz, Generalized Langevin equation with a nonlinear potential of mean force and nonlinear memory friction from a hybrid projection scheme, *Phys. Rev. E* **105**, 054138 (2022).
- [13] F. Ginot, J. Caspers, M. Krüger, and C. Bechinger, Barrier crossing in a viscoelastic bath, *Phys. Rev. Lett.* **128**, 028001 (2022).
- [14] N. Narinder, C. Bechinger, and J. R. Gomez-Solano, Memory-induced transition from a persistent random walk to circular motion for achiral microswimmers, *Phys. Rev. Lett.* **121**, 078003 (2018).
- [15] H. Haken, Cooperative phenomena in systems far from thermal equilibrium and in nonphysical systems, *Rev. Mod. Phys.* **47**, 67 (1975).
- [16] H. Grabert, P. Talkner, and P. Hänggi, Microdynamics and time-evolution of macroscopic non-Markovian systems, *Z. Phys. B* **26**, 389 (1977).
- [17] H. Grabert, P. Talkner, P. Hänggi, and H. Thomas, Microdynamics and time-evolution of macroscopic non-Markovian systems. II, *Z. Phys. B* **29**, 273 (1978).
- [18] J. Casademunt, R. Mannella, P. V. E. McClintock, F. E. Moss, and J. M. Sancho, Relaxation times of non-Markovian processes, *Phys. Rev. A* **35**, 5183 (1987).
- [19] M. Ferrario and P. Grigolini, The non-Markovian relaxation process as a “contraction” of a multidimensional one of Markovian type, *J. Math. Phys. (N.Y.)* **20**, 2567 (2008).
- [20] I. Goychuk, Viscoelastic subdiffusion: Generalized Langevin equation approach, in *Advances in Chemical Physics*, Vol. 150, edited by S. A. Rice and A. R. Dinner (Wiley, New York, 2012).
- [21] I. M. Sokolov, Solutions of a class of non-Markovian Fokker-Planck equations, *Phys. Rev. E* **66**, 041101 (2002).
- [22] I. M. Sokolov, Cyclization of a polymer: First-passage problem for a non-Markovian process, *Phys. Rev. Lett.* **90**, 080601 (2003).
- [23] B. Dybiec, I. M. Sokolov, and A. V. Chechkin, Relaxation to stationary states for anomalous diffusion, *Commun. Nonlinear Sci. Numer. Simul.* **16**, 4549 (2011).

- [24] R. Metzler and J. Klafter, The random walk's guide to anomalous diffusion: A fractional dynamics approach, *Phys. Rep.* **339**, 1 (2000).
- [25] R. Metzler and J. Klafter, The restaurant at the end of the random walk: Recent developments in the description of anomalous transport by fractional dynamics, *J. Phys. A* **37**, R161 (2004).
- [26] E. Barkai, R. Metzler, and J. Klafter, From continuous time random walks to the fractional Fokker-Planck equation, *Phys. Rev. E* **61**, 132 (2000).
- [27] J. H. P. Schulz, E. Barkai, and R. Metzler, Aging renewal theory and application to random walks, *Phys. Rev. X* **4**, 011028 (2014).
- [28] S. Carmi, L. Turgeman, and E. Barkai, On distributions of functionals of anomalous diffusion paths, *J. Stat. Phys.* **141**, 1071 (2010).
- [29] E. Barkai and I. M. Sokolov, Multi-point distribution function for the continuous time random walk, *J. Stat. Mech.* (2007) P08001.
- [30] A. J. Bray, S. N. Majumdar, and G. Schehr, Persistence and first-passage properties in nonequilibrium systems, *Adv. Phys.* **62**, 225 (2013).
- [31] S. N. Majumdar and A. J. Bray, Persistence with partial survival, *Phys. Rev. Lett.* **81**, 2626 (1998).
- [32] S. N. Majumdar, A. J. Bray, S. J. Cornell, and C. Sire, Global persistence exponent for nonequilibrium critical dynamics, *Phys. Rev. Lett.* **77**, 3704 (1996).
- [33] A. Lapolla and A. Godec, Single-file diffusion in a bi-stable potential: Signatures of memory in the barrier-crossing of a tagged-particle, *J. Chem. Phys.* **153**, 194104 (2020).
- [34] P. Talkner and P. Hänggi, Colloquium: Statistical mechanics and thermodynamics at strong coupling: Quantum and classical, *Rev. Mod. Phys.* **92**, 041002 (2020).
- [35] I. Di Terlizzi, F. Ritort, and M. Baiesi, Explicit solution of the generalised Langevin equation, *J. Stat. Phys.* **181**, 1609 (2020).
- [36] R. R. Netz, Derivation of the nonequilibrium generalized Langevin equation from a generic time-dependent Hamiltonian, [arXiv:2310.00748](https://arxiv.org/abs/2310.00748).
- [37] R. R. Netz, Multi-point distribution for Gaussian nonequilibrium non-Markovian observables, [arXiv:2310.08886](https://arxiv.org/abs/2310.08886).
- [38] I. D. Terlizzi and M. Baiesi, A thermodynamic uncertainty relation for a system with memory, *J. Phys. A* **53**, 474002 (2020).
- [39] J. Mehl, B. Lander, C. Bechinger, V. Blickle, and U. Seifert, Role of hidden slow degrees of freedom in the fluctuation theorem, *Phys. Rev. Lett.* **108**, 220601 (2012).
- [40] M. Esposito, Stochastic thermodynamics under coarse graining, *Phys. Rev. E* **85**, 041125 (2012).
- [41] J. van der Meer, B. Ertel, and U. Seifert, Thermodynamic inference in partially accessible Markov networks: A unifying perspective from transition-based waiting time distributions, *Phys. Rev. X* **12**, 031025 (2022).
- [42] P. E. Harunari, A. Dutta, M. Polettini, and E. Roldán, What to learn from a few visible transitions' statistics?, *Phys. Rev. X* **12**, 041026 (2022).
- [43] C. Dieball and A. Godec, Mathematical, thermodynamical, and experimental necessity for coarse graining empirical densities and currents in continuous space, *Phys. Rev. Lett.* **129**, 140601 (2022).
- [44] J. van der Meer, J. Degünther, and U. Seifert, Time-resolved statistics of snippets as general framework for model-free entropy estimators, *Phys. Rev. Lett.* **130**, 257101 (2023).
- [45] A. Godec and D. E. Makarov, Challenges in inferring the directionality of active molecular processes from single-molecule fluorescence resonance energy transfer trajectories, *J. Phys. Chem. Lett.* **14**, 49 (2023).
- [46] D. Andrieux, Bounding the coarse graining error in hidden Markov dynamics, *Appl. Math. Lett.* **25**, 1734 (2012).
- [47] A. Gomez-Marin, J. M. R. Parrondo, and C. Van den Broeck, Lower bounds on dissipation upon coarse graining, *Phys. Rev. E* **78**, 011107 (2008).
- [48] S. Rahav and C. Jarzynski, Fluctuation relations and coarse-graining, *J. Stat. Mech.* (2007) P09012.
- [49] G. Teza and A. L. Stella, Exact coarse graining preserves entropy production out of equilibrium, *Phys. Rev. Lett.* **125**, 110601 (2020).
- [50] E. Roldán and J. M. R. Parrondo, Estimating dissipation from single stationary trajectories, *Phys. Rev. Lett.* **105**, 150607 (2010).
- [51] E. Roldán and J. M. R. Parrondo, Entropy production and Kullback-Leibler divergence between stationary trajectories of discrete systems, *Phys. Rev. E* **85**, 031129 (2012).
- [52] M. Baiesi, G. Falasco, and T. Nishiyama, Effective estimation of entropy production with lacking data, [arXiv:2305.04657](https://arxiv.org/abs/2305.04657).
- [53] A. M. Berezhkovskii and D. E. Makarov, Single-molecule test for Markovianity of the dynamics along a reaction coordinate, *J. Phys. Chem. Lett.* **9**, 2190 (2018).
- [54] K. Engbring, D. Boriskovsky, Y. Roichman, and B. Lindner, A nonlinear fluctuation-dissipation test for Markovian systems, *Phys. Rev. X* **13**, 021034 (2023).
- [55] A. Lapolla and A. Godec, Toolbox for quantifying memory in dynamics along reaction coordinates, *Phys. Rev. Res.* **3**, L022018 (2021).
- [56] D. Hartich and A. Godec, Violation of local detailed balance upon lumping despite a clear timescale separation, *Phys. Rev. Res.* **5**, L032017 (2023).
- [57] A. Godec and D. E. Makarov, Challenges in inferring the directionality of active molecular processes from single-molecule fluorescence resonance energy transfer trajectories, *J. Phys. Chem. Lett.* **14**, 49 (2023).
- [58] D. Hartich and A. Godec, Thermodynamic uncertainty relation bounds the extent of anomalous diffusion, *Phys. Rev. Lett.* **127**, 080601 (2021).
- [59] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, *Rep. Prog. Phys.* **75**, 126001 (2012).
- [60] J. Monod, J. Wyman, and J.-P. Changeux, On the nature of allosteric transitions: A plausible model, *J. Mol. Biol.* **12**, 88 (1965).
- [61] Y. Tu, The nonequilibrium mechanism for ultrasensitivity in a biological switch: Sensing by Maxwell's demons, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 11737 (2008).
- [62] E. A. Korobkova, T. Emonet, H. Park, and P. Cluzel, Hidden stochastic nature of a single bacterial motor, *Phys. Rev. Lett.* **96**, 058105 (2006).
- [63] S. Marzen, H. G. Garcia, and R. Phillips, Statistical mechanics of Monod-Wyman-Changeux (MWC) models, *J. Mol. Biol.* **425**, 1433 (2013).

- [64] A. Berlaga and A. B. Kolomeisky, Molecular mechanisms of active transport in antiporters: Kinetic constraints and efficiency, *J. Phys. Chem. Lett.* **12**, 9588 (2021).
- [65] A. Berlaga and A. B. Kolomeisky, Understanding mechanisms of secondary active transport by analyzing the effects of mutations and stoichiometry, *J. Phys. Chem. Lett.* **13**, 5405 (2022).
- [66] A. Berlaga and A. B. Kolomeisky, Theoretical study of active secondary transport: Unexpected differences in molecular mechanisms for antiporters and symporters, *J. Chem. Phys.* **156**, 085102 (2022).
- [67] D. Hartich, A. C. Barato, and U. Seifert, Nonequilibrium sensing and its analogy to kinetic proofreading, *New J. Phys.* **17**, 055026 (2015).
- [68] J. A. Owen, T. R. Gingrich, and J. M. Horowitz, Universal thermodynamic bounds on nonequilibrium response with biochemical applications, *Phys. Rev. X* **10**, 011066 (2020).
- [69] A. Lapolla, J. C. Smith, and A. Godec, Ubiquitous dynamical time asymmetry in measurements on materials and biological systems, [arXiv:2102.01666](https://arxiv.org/abs/2102.01666).