

Information Propagation in Multilayer Systems with Higher-Order Interactions across Timescales

Giorgio Nicoletti^{1,2,3,†} and Daniel Maria Busiello^{4,*}

¹*ECHO Laboratory, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland*

²*Laboratory of Interdisciplinary Physics, Department of Physics and Astronomy “Galileo Galilei”, University of Padova, Padova, Italy*

³*Department of Mathematics “Tullio Levi-Civita”, University of Padova, Padova, Italy*

⁴*Max Planck Institute for the Physics of Complex Systems, Dresden, Germany*



(Received 19 December 2023; revised 20 February 2024; accepted 8 March 2024; published 8 April 2024)

Complex systems are characterized by multiple spatial and temporal scales. A natural framework to capture their multiscale nature is that of multilayer networks, where different layers represent distinct physical processes that often regulate each other indirectly. We model these regulatory mechanisms through triadic higher-order interactions between nodes and edges. In this work, we focus on how the different timescales associated with each layer impact their reciprocal effective couplings. First, we rigorously derive a decomposition of the joint probability distribution of any dynamical process acting on such multilayer networks. By inspecting this probabilistic structure, we unravel the general principles governing how information propagates across timescales, elucidating the interplay between mutual information and causality in multiscale systems. In particular, we show that feedback interactions, i.e., those representing regulatory mechanisms from slow to fast variables, generate mutual information between layers. On the contrary, direct interactions, i.e., from fast to slow layers, can propagate this information only under certain conditions that depend solely on the structure of the underlying higher-order couplings. We introduce the mutual information matrix for multiscale observables to capture these emergent functional couplings. We apply our results to study archetypal examples of biological signaling networks and effective environmental dependencies in stochastic processes. Our framework generalizes to any dynamics on multilayer networks, paving the way for a deeper understanding of how the multiscale nature of real-world systems shapes their information content and complexity.

DOI: [10.1103/PhysRevX.14.021007](https://doi.org/10.1103/PhysRevX.14.021007)

Subject Areas: Complex Systems,
Interdisciplinary Physics,
Statistical Physics

I. INTRODUCTION

Real-world systems are often characterized by interconnected dynamical processes occurring at multiple scales, both spatial and temporal ones. This multiscale structure is a fundamental ingredient in shaping the properties of complex systems [1]. Further, such intrinsic temporal separation and the interplay between slow and fast processes have been known to be a crucial feature of biological, chemical, and ecological systems [2–10].

A natural framework to describe these interconnected multiscale systems is that of multilayer networks [11–15]. Nodes in a layer may represent either physical or biological units, habitats on complex landscapes, or states in a chemical system, with each layer embedding different processes—from random walk, agent-based, and master equation dynamics to spreading and stochastic processes [16–21]. Here, we consider the different layers as physically separated processes that take place at distinct timescales and regulate each other indirectly. In this scenario, as we will detail later on, interactions across timescales are not pairwise, but rather higher order in nature. Higher-order interactions [22,23] and their effects on a variety of phenomena have been extensively studied in recent years, from synchronization to epidemic spreading and the stability of ecological systems [24–31]. In particular, the higher-order interactions considered herein represent how a node evolving with a given timescale regulates the pairwise link between two nodes evolving with another timescale. These kinds of interactions are known as triadic

*Corresponding author: busiello@pks.mpg.de

†giorgio.nicoletti@epfl.ch

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.

interactions [32]. This setting is especially relevant in neuroscience, ecology, and climate science [24–26,33–35], but it has also been used to characterize the control of stochastic reaction networks [36,37], a general framework with applications in signaling and regulatory biochemical processes [38,39]. Furthermore, the presence of multiple timescales associated with distinct dynamical processes has been shown to be paramount in various contexts. Their interplay may fundamentally alter, for example, the onset of large-scale propagation of epidemics in temporal networks [40,41], the selection of chemical states in reaction-diffusion systems [42,43], and the quantification of dissipative properties at the nanoscale [44,45]. Moreover, recent studies investigated the interplay between timescales and network structure in general diffusive processes [46,47].

Here, we unravel the intimate relationship between these multiscale structures and the ability to propagate information in a wide class of complex systems. In Sec. II, we present such a link in a simple network as a proof of concept, while a general framework is introduced in Sec. III. By inspecting the dynamical properties of multi-layer networks with triadic interactions between layers defined by their timescales, we first derive a general decomposition for the joint probability distribution of any process taking place in these systems. To this aim, we generalize the timescale separation approach to deal with multiple interconnected timescales and higher-order interactions of any functional form. Such a decomposition allows us to compute the mutual information between layers exactly, hence providing the setting to unravel the interplay between the causal higher-order structure and the emergence of functional couplings between timescales. In Sec. IV, we uncover the principles governing how this information propagates across timescales. Section V presents an intriguing consequence of these principles, showing that disconnected components become effectively coupled due to shared slow modulations, an archetypal model for the interplay between environmental and internal dependencies in complex systems. To prove the versatility of our framework, we employ it to study biological signaling networks in Sec. VI. Section VII summarizes the main messages of our work, also offering hints on potential applications and useful implications.

II. INFORMATION ACROSS TIMESCALES

As a proof of concept, we first consider the simple case of a two-layer system. Each layer has two nodes, A_μ and B_μ , with $\mu = 1, 2$, and two directed links connecting them with weights $w_\mu^{A \rightarrow B}/\tau_\mu$ and $w_\mu^{B \rightarrow A}/\tau_\mu$ [see Fig. 1(a)]. Here, τ_μ is the characteristic timescale of the layer dynamics. We name x_μ the state of the layer μ , which can thus take two possible values, A_μ or B_μ . We can think of x_μ as describing a random walk in each layer, or equivalently a molecule switching between two possible configurations, a two-state chemical network, or any other master equation dynamics [48,49].

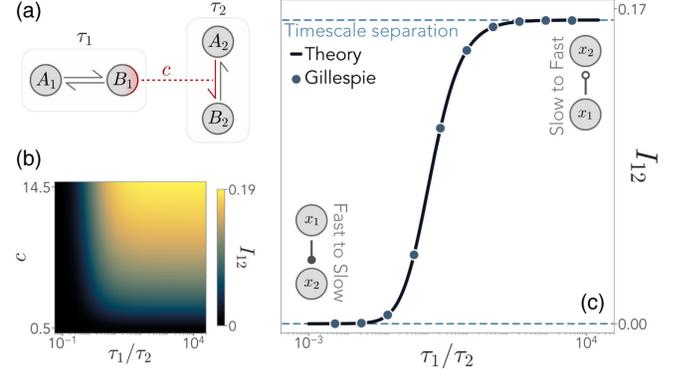


FIG. 1. Information between two layers with states x_1 and x_2 is influenced by their timescale ratio τ_1/τ_2 . (a) Sketch of the system with the node B_1 influencing the link $A_2 \rightarrow B_2$. (b) Mutual information I_{12} between x_1 and x_2 as a function of τ_1/τ_2 and the triadic interaction strength c . (c) At fixed $c = 10$, I_{12} vanishes when x_1 is faster than x_2 . When $\tau_1/\tau_2 \rightarrow \infty$, the regulating layer is slower, and I_{12} converges to a nonzero value. The master equation of the system can be solved exactly (solid black line), simulated via a Gillespie algorithm [50] (blue dots), and solved in a timescale separation regime (dashed blue line).

The two layers represent two stochastic systems physically separated from each other. In other words, a random walker moving in the first layer cannot jump on the second one, or a molecule cannot switch between states belonging to different layers. Since interactions across layers do not act through direct edges between nodes, the weighted adjacency matrix of the network \hat{W} is a block diagonal matrix. Rather, we causally connect layers via higher-order triadic interactions, so that a node in the first layer influences an edge in the second one [dashed line in Fig. 1(a)] [32]. To fix the ideas, let us consider that the node B_1 in the first layer interacts with the $A_2 \rightarrow B_2$ edge in the second layer. We write the modified edge weight as

$$\frac{(w_2^H)^{A \rightarrow B}}{\tau_2} = \frac{w_2^{A \rightarrow B} + c\delta(x_1, B_1)}{\tau_2}, \quad (1)$$

where $\delta(\cdot)$ is the Kronecker delta, and c represents the regulatory interaction that stimulates the transitions to B_2 when the first random walker is in the state B_1 .

To describe these kinds of systems, we introduce the joint probability of the two layers p_{12} , which quantifies the probability of finding the two random walkers in their respective node, i.e., of finding the system in one of the 2^2 possible configurations $\{(A_1, A_2), (B_1, A_2), (A_1, B_2), (B_1, B_2)\}$. In components, the master equation for p_{12} reads

$$\partial_t (p_{12})_n = \sum_{m=1}^{2^2} \left[\frac{(\hat{\Omega}_1)_{nm}}{\tau_1} + \frac{(\hat{\Omega}_2)_{nm}}{\tau_2} \right] (p_{12})_m, \quad (2)$$

where $\hat{\Omega}_1$ and $\hat{\Omega}_2$ are the transition matrices containing all the transitions involving the first and the second layer,

respectively. A detailed description of this system is given in Appendix A.

We can solve the master equation for the joint probability of the two layers exactly, and study the mutual information I_{12} between them (see Appendix B). In Figs. 1(b) and 1(c), we show the behavior of I_{12} as a function of the triadic interaction strength c and the relative timescale separation between the two layers, τ_1/τ_2 . As expected, I_{12} increases with c , since the two layers are decoupled when $c = 0$. Additionally, we find that no information is shared between the two layers when $\tau_1 \ll \tau_2$, i.e., when the regulating layer—the first one—evolves faster than the regulated layer—the second one. However, information is maximal and saturates when x_1 undergoes a much slower dynamics than x_2 , i.e., when $\tau_1 \gg \tau_2$, with a steep transition between these two regimes.

This simple example suggests that the relative characteristic timescales among physically separated systems shape the information content shared by different layers. Our goal is to unveil this connection on more general grounds and, therefore, the main principles governing information propagation in multilayer systems.

III. GENERAL MULTILAYER DYNAMICS

Consider a system of N layers. The μ th layer is a network with M_μ nodes, defined by a $M_\mu \times M_\mu$ adjacency matrix \hat{W}_μ that encodes the pairwise relations between nodes, e.g., transition rates or interactions. For easiness of notation and interpretation, here we focus on the case in which each layer supports a distinct random walk or master equation dynamics, so that we can identify the state of a layer x_μ with the occupied node, as in the previous example. However, our framework is valid for any stochastic dynamics on networks, from Ornstein-Uhlenbeck processes to models of neural dynamics, where the state of a layer is a vector of node states (see Appendix B) [51,52].

Causal interactions across layers are triadic, going from nodes of one layer to edges of another. As outlined above, this setting naturally stems from the requirement that layers represent physically separated systems. Such higher-order interactions are defined by a tensor $C_{\mu\nu}^{k,i \rightarrow j}$, which denotes how the node k in layer μ influences the $i \rightarrow j$ edge in layer ν . For the sake of brevity, we denote with $\hat{C}_{\mu\nu}$ the tensor that contains the set of all possible triadic interactions from μ to ν . We also introduce the modified adjacency matrix \hat{W}_μ^H of a layer, which contains both the intralayer pairwise edges, as well as the contribution of higher-order interactions coming from the nodes of other layers. To fix the ideas, when triadic interactions are additive with respect to pairwise couplings, we write its elements as

$$\hat{W}_\mu^H(\{x\}_{\rightarrow\mu})^{i \rightarrow j} = (\hat{W}_\mu)^{i \rightarrow j} + \sum_{\nu \neq \mu} \sum_{k=1}^{M_\nu} C_{\nu\mu}^{k,i \rightarrow j} \phi_\mu^k(x_\nu). \quad (3)$$

In Eq. (3), $\phi_\mu^k(x_\nu)$ is a generic nonlinear function of the state of a layer, and $\{x\}_{\rightarrow\mu}$ is a shorthand notation to denote the state of all layers connected to μ by higher-order interactions, i.e., all ν for which $\hat{C}_{\nu\mu}$ is nonzero. Let us stress that the additive structure in Eq. (3) is only exemplary and, in particular, allows for the presence of a baseline transition matrix \hat{W}_μ even in the absence of higher-order couplings. However, our results are independent of the specific form of \hat{W}_μ^H , where triadic interactions may enter in any functional shape.

We now order the layers by their timescales—that is, we take $\tau_1 < \tau_2 < \dots < \tau_N$, so that the first layer is the fastest and the last the slowest. We name triadic interactions going from a fast to a slow layer as “direct” interactions. Vice versa, when a slow node influences a fast link, we call the corresponding interaction a “feedback” one. Hence, the tensor containing all causal interactions $\hat{C}_{\mu\nu}$ can be written as the sum of two contributions,

$$\hat{C}_{\mu\nu} = \underbrace{\hat{D}_{\mu\nu}}_{\mu < \nu} + \underbrace{\hat{F}_{\mu\nu}}_{\mu > \nu}, \quad (4)$$

where $\hat{D}_{\mu\nu}$ is a lower-triangular tensor in the indices μ and ν and contains direct interactions, whereas $\hat{F}_{\mu\nu}$ is upper triangular and describes feedback interactions. Intuitively, these two kinds of interactions capture different physical mechanisms: direct couplings act as controls modulated by underlying fast processes, while feedback interactions might encapsulate regulatory schemes. As we will see, they play a very different role in determining how

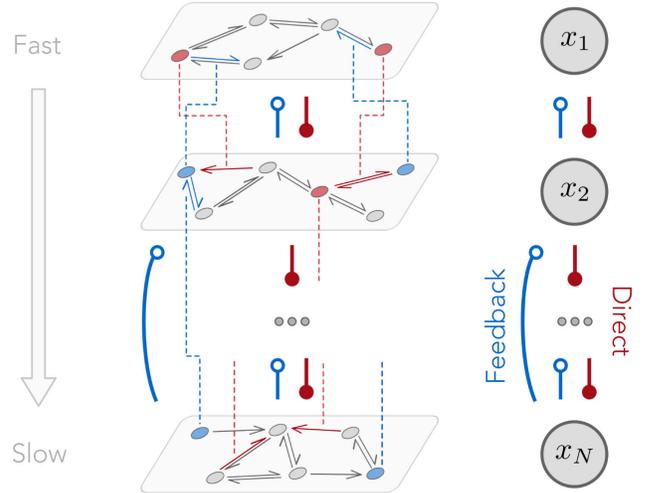


FIG. 2. Sketch of a multilayer system with intralayer pairwise interactions and triadic higher-order interactions across layers. We name interlayer interactions going from a faster layer to a slower one as “direct” interactions, as they follow the ordering induced by the timescales. On the contrary, we refer to those that go toward faster layers as “feedback” interactions, as they provide a slow modulation to fast processes.

information is created and propagated across layers. This general model is depicted in Fig. 2.

We denote the joint probability distribution of all layers at time t with $p_{1,\dots,N}(t)$. This is a probability distribution over all the $S = \prod_{\nu=1}^N M_{\nu}$ possible states of the distinct random walks (see Appendix B). The corresponding master equation reads

$$\partial_t p_{1,\dots,N}(t) = \sum_{\mu=1}^N \frac{\hat{\Omega}_{\mu}(\{x\}_{\rightsquigarrow\mu})}{\tau_{\mu}} p_{1,\dots,N}(t), \quad (5)$$

where $\hat{\Omega}_{\mu}$ is the $S \times S$ transition matrix containing all the transitions involving the μ th layer. Clearly, $\hat{\Omega}_{\mu}$ can be fully specified in terms of \hat{W}_{μ}^H , and vice versa. Inspired by the results of the previous section, we are particularly interested in the timescale separation regime $\tau_{\mu+1}/\tau_{\mu} \rightarrow 0$ [53]. The presence of direct and feedback interactions at once allows us to explore both the extreme behaviors—maximum and zero information—highlighted in the previous section. In this regime, we can solve Eq. (5) analytically (see Appendix C) and show that $p_{1,\dots,N}$ takes the following factorized form:

$$p_{1,\dots,N}(t) = p_N^{\text{eff}}(t) \prod_{\nu=1}^{N-1} p_{\nu|\rho(\nu)}^{\text{st,eff}}. \quad (6)$$

This result extends and generalizes the standard timescale separation approach to deal with systems with multiple interconnected timescales and higher-order interactions. Here, in the conditional probabilities in Eq. (6), we

denote by $\rho(\mu)$ the set of all layers interacting with μ via either a single feedback link or a minimal propagation path (mPP). Any propagation path (PP) to μ is a directional path coming from layers slower than μ and containing at least one direct link. An mPP from ν to μ is then a PP on the induced subgraph obtained by removing all layers slower than μ except ν (see Appendix C). Finally, $p_{\mu|\rho(\mu)}^{\text{st,eff}}$ is the stationary probability of the effective operator,

$$\hat{\Omega}_{\mu|\rho(\mu)}^{\text{eff}} = \sum_{x_1} \cdots \sum_{x_{\mu-1}} \hat{\Omega}_{\mu}(\{x\}_{\rightsquigarrow\mu}) \prod_{\nu=1}^{\mu-1} p_{\nu|\rho(\nu)}^{\text{st,eff}}, \quad (7)$$

that averages the effect of all the layers faster than μ and explicitly depends on those slower than μ belonging to $\rho(\mu)$. Clearly, $p^{\text{st,eff}} = p^{\text{st}}$ if there is no direct link toward the layer μ . Hence, the timescale separation defines a hierarchy of effective operators and their stationary probabilities, starting from the fastest layer. In terms of these operators, the multilayer joint probability can be written exactly as a product of suitable conditional probabilities, and the effective operator of the slowest layer determines its time evolution. From the stationary joint probability $p_{1,\dots,N}^{\text{st}}$, we are able to compute all marginal probabilities and thus the mutual information between the layers:

$$I_{\mu\nu} = \sum_{x_{\mu}, x_{\nu}} p_{\mu\nu} \log_2 \frac{p_{\mu\nu}}{p_{\mu} p_{\nu}}, \quad (8)$$

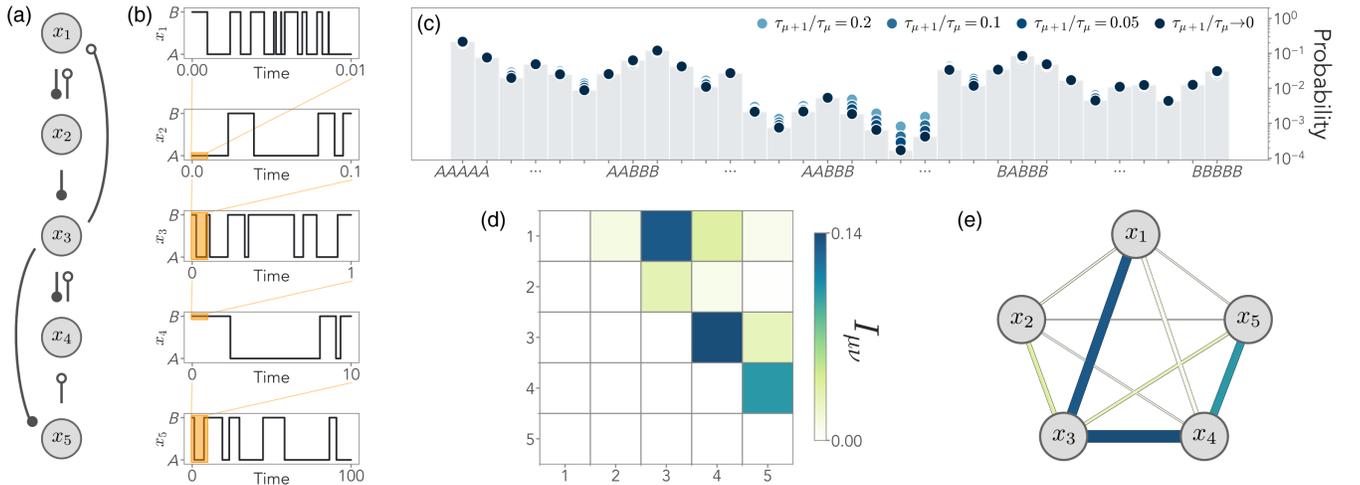


FIG. 3. Accuracy of the timescale separation solution and mutual information in a five-layer system with triadic interactions. (a) Scheme of the interactions, defined in Eq. (3). For simplicity, we set feedback interactions as $F_{\mu\nu}^{k,i \rightarrow j} = F_{\mu\nu}^{\text{eq}} \delta_{jk} + F_{\mu\nu}^{\text{cr}} (1 - \delta_{jk})$, and similarly for direct interactions $D_{\mu\nu}^{k,i \rightarrow j}$, with $F_{\mu\nu}^{\text{eq}} = 10$, $D_{\mu\nu}^{\text{eq}} = 0.1 = F_{\mu\nu}^{\text{cr}}$, and $D_{\mu\nu}^{\text{cr}} = 5$ (see Appendix D). (b) Simulated trajectories in each layer with $\tau_{\mu+1}/\tau_{\mu} = 0.1$, so that each layer is an order of magnitude faster than the previous one, using a Gillespie algorithm. (c) Comparison of the joint probability estimated from simulated trajectories at different separations between the timescales of each layer. Gray bars represent the analytical solution in Eq. (6), which is well approximated already for $\tau_{\mu+1}/\tau_{\mu} = 0.1$. (d) MIMMO associated with this system when $\tau_{\mu+1}/\tau_{\mu} \rightarrow 0$. (e) Mutual information between each pair of layers is shaped by the pathways among them defined by direct and feedback interactions.

defining a symmetric matrix that we name mutual information matrix for multiscale observables (MIMMO).

Despite its intricate structure, the proposed framework provides us with an instructive and intuitive way to study information-theoretic properties of multilayer systems operating at different timescales. To show this beyond the general theory, we study the example of the five-layer system in Fig. 3(a). Each layer has two nodes $\mathcal{N} = \{A, B\}$. Triadic interactions between layers are additive, as in Eq. (3), with $\phi_\mu^k(x_\nu) = \delta(x_\nu, k)$ for all layers μ , with $k \in \mathcal{N}$. In this way, they are switched on and off depending on the state of the regulating layer (see Appendix D for details about the choice of triadic interactions). Figure 3(b) shows the typical trajectories obtained through an exact Gillespie algorithm implemented by simulating the dynamics of the full transition matrix $\sum_\mu \hat{\Omega}_\mu / \tau_\mu$ [50]. From Eq. (6), the joint probability distribution reads: $p_{1,2,3,4,5}(t) = p_{1|2,3}^{\text{st}} p_2^{\text{st,eff}} p_{3|4}^{\text{st,eff}} p_{4|5}^{\text{st,eff}} p_5^{\text{eff}}(t)$. We can compare this analytical solution with simulations at different values of $\tau_{\mu+1}/\tau_\mu$. As we see in Fig. 3(c), the timescale separation assumption provides an excellent approximation for the system even when the timescales differ by just one order of magnitude, i.e., $\tau_{\mu+1}/\tau_\mu = 0.1$. Finally, we show the upper triangular section of the mutual information matrix in Fig. 3(d). We remark that, even for $\tau_{\mu+1}/\tau_\mu = 0.1$,

the MIMMO obtained directly from trajectories is qualitatively identical to the analytical one. Importantly, the MIMMO helps to quantify the dependence between the layers induced by the interlayer interactions. We can assign an undirected link between any two layers μ and ν with a weight equal to $I_{\mu\nu}$, resulting in a fully connected network that represents how much information each pair of layers share [Fig. 3(e)].

Already at the level of this paradigmatic example, it is evident that direct and feedback interactions both play a role in determining how information spreads among layers due to their causal relationships, i.e., how much layers are coupled to one another. To unravel the differences between these two classes of interactions, we now exploit the conditional structure of the complete solution.

IV. INFORMATION PROPAGATION AND EMERGENT DIRECTIONALITY

Let us start again from basic examples and consider the scheme in Fig. 4(a). This represents a three-layer system in which the structure of each layer is the same as in Fig. 3(a), for simplicity. Without loss of generality, triadic interactions have been set following Appendix D. The joint probability distribution is $p_{1,2,3}^{\text{st}} = p_{1|3}^{\text{st}} p_2^{\text{st}} p_3^{\text{st,eff}}$, and thus the only nonzero element of the MIMMO is I_{13}

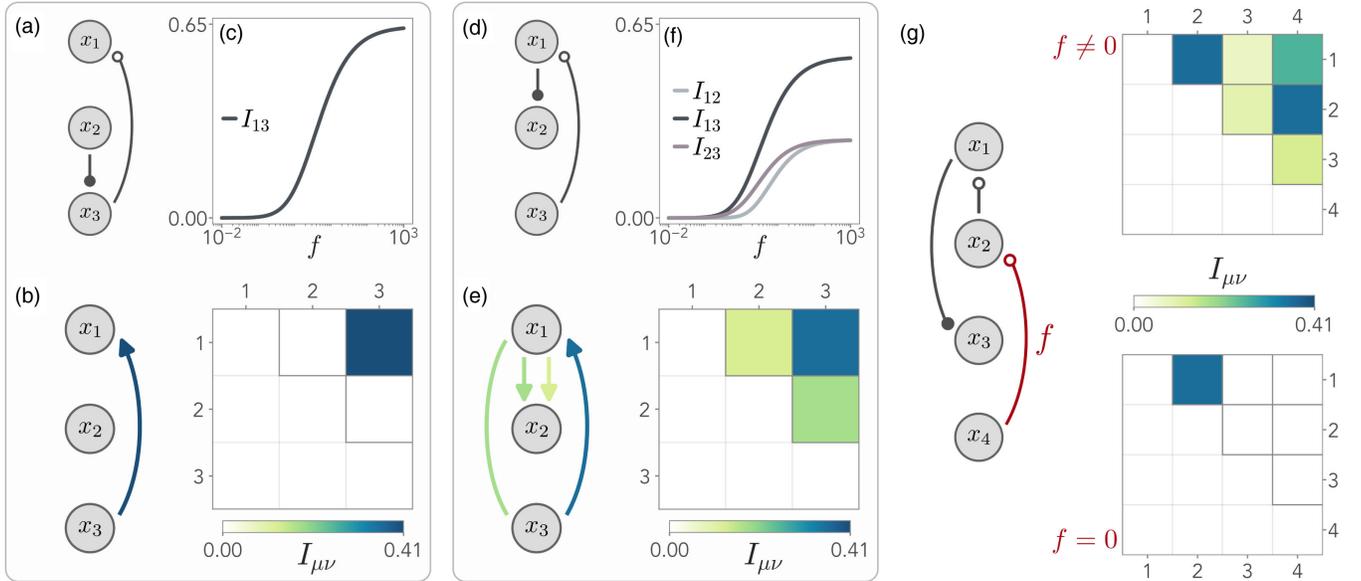


FIG. 4. The principles of information propagation in multilayer networks. (a) A three-layer system with a higher-order feedback interaction from layer 3 to 1 and a direct link from 2 to 3. Each layer has two nodes, as in the previous examples. (b) The slow layer acts as an information source for the fast one so that $I_{13} \neq 0$ in the MIMMO, while the direct link does not generate information. (c) A stronger feedback interaction f increases the mutual information between the two layers. (d) The same system as in panel (a), with the direct link from 1 to 2. (e) All the elements of the MIMMO are nonzero, since it exists a minimal propagation path $3 \rightarrow 1 \rightarrow 2$. The direct link propagates the information created in 1. (f) All the entries of the MIMMO depend on f , and they vanish as f goes to zero, i.e., no information is created. (g) A four-layer system with feedback interactions from 4 to 2 and 2 to 1, and a direct link from 1 to 3. While for $f \neq 0$ all the entries of the MIMMO are nonzero, when $f = 0$, the information created in 1 by 2 cannot be propagated to 3, since the dynamics of x_3 is faster than x_2 . As a consequence, only $I_{12} \neq 0$.

[see Figs. 4(b) and 4(c)]. This information is due to the explicit dependence on x_3 in $p_{1|3}^{\text{st}}$ induced and modulated by the feedback interaction of strength f . This asymmetric effect provides us with an emergent directionality of information propagation, ascribing to the feedback link the role of generating information from slow to fast layers. Instead, the isolated direct link does not contribute to the information content.

We now slightly modify the system, considering the one in Fig. 4(d). Again, the feedback interaction generates information from x_3 and x_1 . At variance with the previous case, the direct link can now propagate the information generated in the first layer on x_2 , leading to mutual information among all layers [see Figs. 4(e) and 4(f)]. This effect manifests into the joint probability distribution $p_{1,2,3}^{\text{st}} = p_{1|3}^{\text{st}} p_{2|3}^{\text{st,eff}} p_3^{\text{st}}$. The dependence on x_3 appears in all layers, as expected, but $p_{1|3}^{\text{st}}$ and $p_{2|3}^{\text{st,eff}}$ are conditionally independent. Indeed, the mutual information among them is not directly created by a feedback, but it stems solely from the third layer and is then propagated; i.e., the layers belong to an mPP.

Finally, we move to the four-layer system in Fig. 4(g). When $f \neq 0$, its joint probability distribution is $p_{1,2,3,4}^{\text{st}} = p_{1|2}^{\text{st}} p_{2|4}^{\text{st}} p_{3|4}^{\text{st,eff}} p_4^{\text{st}}$. Here, information is generated from x_2 into x_1 and from x_4 into x_2 , and this is reflected in the explicit dependencies in the joint probability distribution. Again, the fact that 2 and 3 are conditionally independent signals that information is propagated only through direct links. The same happens between x_1 and x_3 . Nevertheless, when $f = 0$, the only nonzero element of the MIMMO is I_{12} . Indeed, although x_2 creates information in x_1 , it cannot propagate to x_3 , since its dynamics is slower than the regulating layer.

Hence, the topological structure of higher-order interactions fully determines the propagation of information. It is clear that our results depend only on the general structure of the probability in Eq. (6), not on the dynamics of the layer nor the form of the higher-order interactions. As a consequence, the main result of the proposed framework resides in the identification of three basic physical principles of information propagation across timescales: (i) feedback interactions generate information from slow to fast layers; (ii) direct interactions alone do not generate information; (iii) information generated through feedback by a slow layer may be propagated through direct interactions to any faster layer. The first two principles highlight the different roles of feedback and direct interactions in affecting the information content of the system. The third one leads to a natural definition of directionality in the information propagation, despite the symmetry of the MIMMO. At the same time, it constrains the accessibility of any control mechanism in stochastic networks in terms of timescales, unveiling that they create

mutual information only when acting as regulatory processes, i.e., to faster layers.

V. EFFECTIVE INFORMATION BETWEEN DISCONNECTED LAYERS

We now employ the principles outlined in the previous sections to study what happens when information is generated from a slow layer into either disconnected layers or a single layer with disjoint network components. This scenario is particularly relevant when interpreting this scheme as a model for two distinct degrees of freedom, evolving with two timescales τ_1 and τ_2 , under the influence of a shared environment that exhibits an independent dynamics on a different timescale, τ_E . Without loss of generality, we resort to the archetypal system in Fig. 5(a). Each disjoint component has two nodes and a random walk dynamics evolving on a timescale $\tau_1 = \tau_2 = \tau$. Triadic interactions stem from another layer with two nodes, E_{off} and E_{on} , and a random walk dynamics with timescale τ_E . In particular, c enhances both the transitions $A_1 \rightarrow B_1$ and $A_2 \rightarrow B_2$ (see Appendix G). Processes with this kind of structure have been extensively studied in different contexts, from neuroscience to ecology and chemical reaction networks [42,52,54–59].

In the limit $\tau_E \ll \tau$, the shared triadic interaction becomes a direct link and, as such, does not generate mutual information in the system. Conversely, when

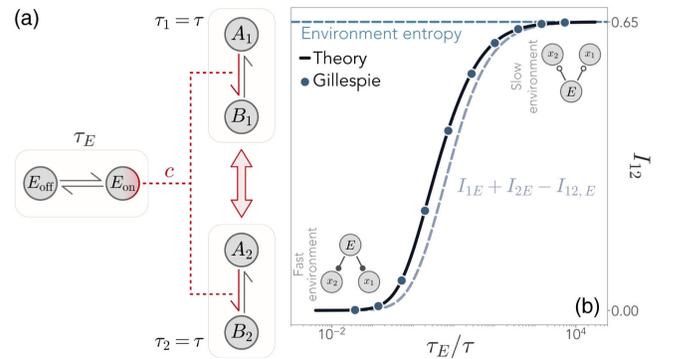


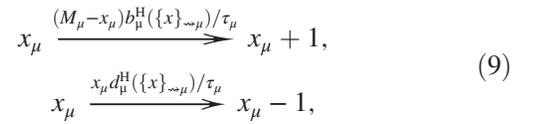
FIG. 5. Shared higher-order interactions generate information among random walks in disjoint network components. (a) An archetypal system where one layer has two disjoint components (A_1, B_1) and (A_2, B_2), with two random walks evolving with timescales $\tau_1 = \tau_2 = \tau$. The links of both these components are influenced by another layer E through a triadic interaction c that models an environmental-like dynamics with timescale τ_E influencing two separate degrees of freedom x_1 and x_2 . (b) In the limit $\tau_E \gg \tau$, the shared environment induces a nonzero mutual information between the two disjoint random walks I_{12} . The black curve shows I_{12} in the limit $c \rightarrow \infty$, while blue dots show the corresponding Gillespie simulation. The two random walks are conditionally independent in this limit, so their mutual information can be decomposed exactly (gray dashed line) and converges to the Shannon entropy of the environment E (blue dashed line).

$\tau_E \ll \tau$, there is a feedback link from E to the two disjoint components at once, so that I_{1E} and I_{2E} are both positive. As a nontrivial consequence, this feedback interaction induces a nonzero mutual information between the two disjoint components, x_1 and x_2 , as well [see Fig. 5(b)]. This phenomenon of dependencies induced by a switching environment was first highlighted in Ref. [56], but our framework allows us to immediately pinpoint it. Indeed, in the limit of a slow environment, $p_{1,2}(t) = \sum_E p_{1|E}^{\text{st}} p_{2|E}^{\text{st}} p_E(t)$. Thus, the two random walks are conditionally independent and their effective coupling stems solely from the shared environmental dynamics. In particular, I_{12} converges, in this limit, to the Shannon entropy of E , i.e., the maximal information content in the environment, and we can find an exact decomposition as in Fig. 5(b). Let us stress that, even if the two disjoint components evolved on two different timescales τ_1 and τ_2 , both faster than τ_E , they would share mutual information as a consequence of the principle (i) highlighted in Sec. IV. This observation emphasizes that, when faster but disconnected layers are modulated by a shared (slow) degree of freedom, their relative timescales are formally irrelevant to quantifying their information content. However, the crucial distinction arises in the presence of interactions between layer 1 and 2, where the ratio between their timescales determines whether their interaction is of direct or feedback type, and thus shapes the features of their joint probability distribution. In this scenario, when $\tau_1 = \tau_2 = \tau$, insightful

decompositions of the mutual information emerge, as in Refs. [56,57].

VI. MULTISCALE SIGNALING NETWORKS

As a biophysically relevant application, we focus on biochemical signaling networks, that have been extensively studied in different biological and living systems [60–63] and constitute an example of controlled stochastic reaction networks [36,37]. The formalism of the previous sections can be immediately translated to study these cases. Each layer now is associated with a biochemical species evolving on its intrinsic timescale. Nodes in a layer represent the number of molecules of the corresponding species; i.e., the state is given by $x_\mu = 1, \dots, M_\mu$ for layer μ . Thus, the number of nodes M_μ quantifies the maximum number of molecules of the μ th species. The random walk in each layer now describes a birth-and-death process with reactions:



where b_μ^H and d_μ^H are, respectively, the birth and death rates of species μ modified by higher-order interactions, and so they may depend on the state of other layers. We model inhibitory (excitatory) triadic interactions as increasing

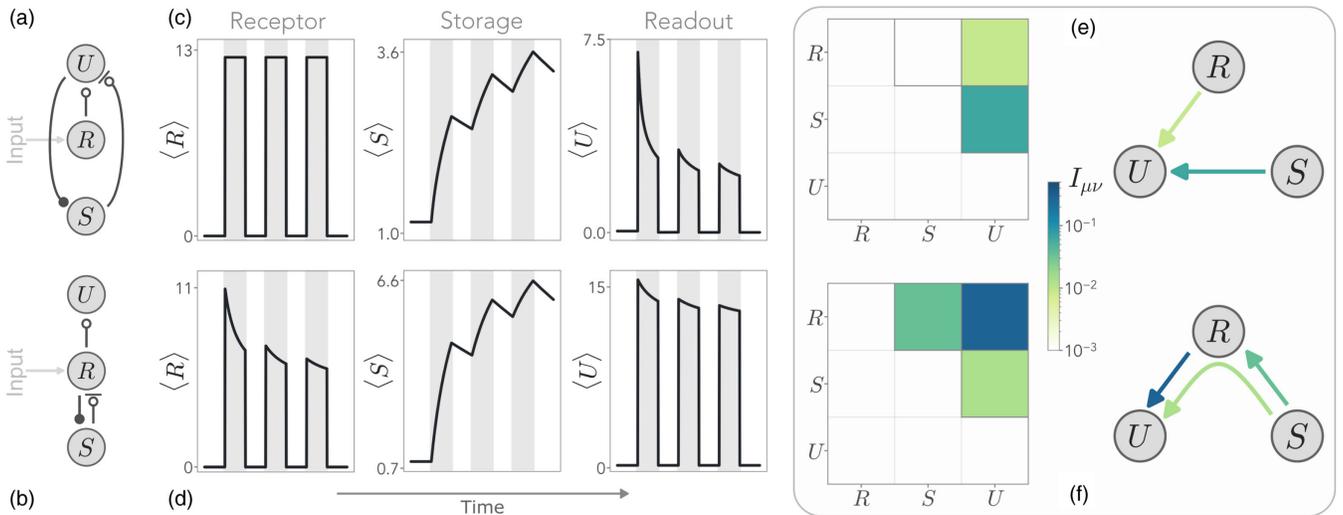


FIG. 6. Information in birth-and-death processes modeling multiscale signaling networks. (a),(b) Two biochemical architectures known to implement biological adaptation and habituation, encompassing a receptor R , a storage S , and a readout population U . Inhibitory links, which increase the death rate of a population, are indicated with a bar at the end. (c),(d) Both architectures display similar dynamical behavior in the presence of a repeated input to the receptor, with the readout average $\langle U \rangle$ decreasing and $\langle S \rangle$ increasing at each signal (gray shaded area). (e) The two architectures display important differences at the information level across layers. In the first, the receptor and the storage share no mutual information, and the regulation of S acts independently of the signal. (f) In the second, the storage is also informationally coupled to the receptor; hence the regulatory mechanism is tangled with the external signal and all entries of the MIMMO are nonzero. In this figure, we set the maximum population of all species, i.e., the nodes of each layer, to $M_R = 15$, $M_S = 10$, and $M_U = 20$.

the baseline death (birth) rate. The details and the full time-dependent solution of these systems are given in Appendix E. Note that, in principle, we may have layers with an infinite number of nodes, if there is no upper bound to the number of molecules of a species. The probabilistic structure and, in turn, the information propagation rules are determined solely by higher-order couplings across timescales, as shown in previous sections.

In Figs. 6(a) and 6(b), we show two representative three-layer signaling architectures that are well known to provide minimal models for biological adaptation, habituation, and adaptive responses [38,64]. A receptor R receives a time-varying input $h(t)$ (here a periodically switching signal) and stimulates in turn a fast response described by a readout population U . Such response is regulated by a slow storage population S , inhibiting either U [Fig. 6(a)] or R [Fig. 6(b)] and acting as an effective memory [65–67]. The specific forms of b_{μ}^H and d_{μ}^H for these two architectures are in Appendix F. The maximum numbers of molecules for each species, i.e., the nodes of each layer, are M_U , M_R , and M_S .

From the time-dependent solution of the master equation of this system, we find that, as expected, both architectures display qualitatively similar dynamical behaviors, with the only difference being whether the receptor population decreases its activity [Fig. 6(d)] or not [Fig. 6(c)]. In particular, the average system response $\langle U \rangle$ decreases in time upon repeated stimulation, as the storage increases [Figs. 6(c) and 6(d)].

However, the information propagation framework allows us to discriminate between the two architectures through their different physical mechanisms that underlie adaptive and habituated behaviors. In the first case [Fig. 6(e)], there is no mutual information between receptor and storage; $I_{RS} = 0$ since information is generated solely in the readout U . Thus, the regulatory mechanism implemented by S is independent of the receptor activity at the information level. On the contrary, in the second architecture [Fig. 6(f)], all the entries in the MIMMO are nonzero, since all layers are sequentially connected via feedback interactions. The information generated in R by the storage is then propagated to the readout U . Thus, the regulatory function of S is dependent on the receptor activity at the information level, a completely different biochemical mechanism with respect to the previous case.

VII. DISCUSSIONS

In this work, we characterized how information propagates and couples the layers of multilayer networks. We introduced a general framework that extends the timescale separation approach to multiscale systems and higher-order interactions. In particular, the addition of multiple couplings from slow to fast timescales induces nontrivial properties in the joint probability distribution. By employing this framework, we find that the probabilistic structure of any multilayer network depends solely on the topology

of higher-order couplings between layers, being independent of their specific functional form. Ultimately, we identify three physical principles governing the propagation of information across timescales. They ascribe the functional role of creating information to feedback interactions representing regulatory and controlling mechanisms, while elucidating that direct interactions (from fast to slow processes) may propagate such information across layers. Importantly, triadic and higher-order interactions play a prime role in several biological and non-biological systems. A pivotal example is that of complex chemical networks under stochastic control [36,37] or exhibiting catalytic reaction schemes [68]. Their structure is formally identical to the one implemented to describe signaling networks in Sec. VI. This similarity allows for a direct evaluation of the MIMMO in these contexts and a clear identification of the role of the intrinsic reaction timescales to drive pattern-formation mechanisms [69] and selection phenomena [42,70].

Although we represented each layer as a physically separated process acting on different timescales and affecting each other through triadic higher-order interactions, our results can be applied to many other dynamical processes as well. For example, a relevant application would be that of multivariate (discrete or continuous) dynamical processes on multilayer networks with pairwise edges only, where once more nodes in a layer evolve with the same timescale. In fact, multiple subparts of a large number of biological and artificial systems are characterized by different timescales and thus can be seen as layers of a multilayer network [11–13,15]. By clustering the degrees of freedom on a given layer together, one could define the joint probability distribution and the mutual information between layers as in Appendix B. The factorization of the joint probability in Eq. (6) would still hold, as long as pairwise interactions between layers act on one of the layers' timescales.

Because of such generality, our work provides a quantitative framework with applications in a large number of fields. Examples can be found in epidemic spreading, by elucidating the functional interplay between the timescales of the disease and the evolution of the topology supporting its spreading [40,41]; in ecological population dynamics, by revealing how the timescales of environmental changes affect species strategies and survival [21,25,54]; in brain activity, where intrinsic neuronal timescales are known to be one of the prime mediators of behavior, processing, and cognition [71–75]. Moreover, experimental developments have enabled substantial progress in understanding complex systems operating at different timescales through the simultaneous recording of their internal variables. In principle, the MIMMO can be estimated directly from recorded time series, and it may provide a valuable tool to shed light on the information structure of the system and how regulatory mechanisms are implemented [36–39].

As a caveat, the precision of this procedure depends on the accuracy of temporal data and relies on the possibility of correctly identifying the timescales associated with the degrees of freedom at play.

Ultimately, our framework suggests an additional outlook. Ranking the degrees of freedom of a system by the similarity of their timescales allows for the construction of a coarse-grained multilayer model [76]. With this perspective, our framework might unveil the intrinsic information structure of the underlying dynamics and highlight how information is shared among coarse-grained degrees of freedom. The resulting metric will capture how dynamical processes propagate information across timescales and provide fundamental insights into the physics of regulatory and controlling mechanisms, both internal and environmentally driven.

ACKNOWLEDGMENTS

G.N. acknowledges funding provided by the Swiss National Science Foundation through its Grant No. CRSII5_186422. We thank Amos Maritan and Giacomo Barzon for critical discussions and suggestions. The authors thank the Max Planck Institute for the Physics of Complex Systems for hosting G.N. during the final phase of this work.

APPENDIX A: DETAILED DESCRIPTION OF THE SYSTEM IN SEC. II

In this appendix, we present a detailed solution for the system presented in Sec. II. We consider a simple two-layer network, where each layer has two nodes $\{A_\mu, B_\mu\}$ with $\mu = 1, 2$. The adjacency matrix \hat{W} of the network is a block-diagonal matrix, so that the first and the second layer are not connected by any edge:

$$\hat{W} = \begin{pmatrix} 0 & w_1^{A \rightarrow B} & 0 & 0 \\ w_1^{B \rightarrow A} & 0 & 0 & 0 \\ 0 & 0 & 0 & w_2^{A \rightarrow B} \\ 0 & 0 & w_2^{B \rightarrow A} & 0 \end{pmatrix}.$$

When we take into account triadic interactions, as in Eq. (1), the adjacency matrix becomes

$$\hat{W}^H = \begin{pmatrix} 0 & w_1^{A \rightarrow B} & 0 & 0 \\ w_1^{B \rightarrow A} & 0 & 0 & 0 \\ 0 & 0 & 0 & w_2^{A \rightarrow B} + c\delta(x_1, B_1) \\ 0 & 0 & w_2^{B \rightarrow A} & 0 \end{pmatrix},$$

where $x_1 \in \{A_1, B_1\}$ is the state of the first layer, i.e., the node occupied by the first random walker.

The dynamics of the two random walks is associated with a transition matrix $\hat{\Omega}_1/\tau_1 + \hat{\Omega}_2/\tau_2$, where

$$\hat{\Omega}_1 = \begin{pmatrix} -w_1^{A \rightarrow B} & w_1^{B \rightarrow A} & 0 & 0 \\ w_1^{A \rightarrow B} & -w_1^{B \rightarrow A} & 0 & 0 \\ 0 & 0 & -w_1^{A \rightarrow B} & w_1^{B \rightarrow A} \\ 0 & 0 & w_1^{A \rightarrow B} & -w_1^{B \rightarrow A} \end{pmatrix} \quad (\text{A1})$$

contains the transitions taking place in the first layer, and

$$\hat{\Omega}_2 = \begin{pmatrix} -w_2^{A \rightarrow B} & 0 & w_2^{B \rightarrow A} & 0 \\ 0 & -w_2^{A \rightarrow B} - c & 0 & w_2^{B \rightarrow A} \\ w_2^{A \rightarrow B} & 0 & -w_2^{B \rightarrow A} & 0 \\ 0 & w_2^{A \rightarrow B} + c & 0 & -w_2^{B \rightarrow A} \end{pmatrix} \quad (\text{A2})$$

the transitions in the second layer. Note that the joint states of the system are assumed to be ordered as $\{(A_1, A_2), (B_1, A_2), (A_1, B_2), (B_1, B_2)\}$, and the two transition matrices can be fully specified from the diagonal blocks of \hat{W}^H . Since both Eqs. (A1) and (A2) contain the exit rates along the diagonals, the master equation reads as in Eq. (2) in the main text.

APPENDIX B: JOINT PROBABILITIES AND MUTUAL INFORMATION BETWEEN LAYERS

In the case studied in the main text, where each layer represents a distinct random walk, the joint probability distribution $p_{1, \dots, N}$ between N layers is a probability distribution on the tensor product of the state spaces of the layers, i.e., on the space of the joint state of all random walkers. As such, it can be in general represented by a $M_1 \times \dots \times M_N$ tensor, where M_μ is the number of nodes in the μ th layer. From $p_{1, \dots, N}$, we can obtain the joint probabilities between any two layers μ and ν , which we denote by $p_{\mu\nu}$, and the marginal probability of every single layer p_μ . While the former is a $M_\mu \times M_\nu$ matrix, the latter is a M_μ -dimensional vector. To fix the ideas, in the example system presented in Appendix A, we have that the marginal distribution $p_1 = (p_1(x_1 = A_1), p_1(x_1 = B_1))$ is defined as

$$p_1 = \begin{pmatrix} \sum_{x'_2 \in \{A_2, B_2\}} p_{12}(x_1 = A_1, x_2 = x'_2) \\ \sum_{x'_2 \in \{A_2, B_2\}} p_{12}(x_1 = B_1, x_2 = x'_2) \end{pmatrix},$$

and similarly for p_2 .

With these definitions in mind, the mutual information between any two layers is defined as

$$I_{\mu\nu} = \sum_{x_\mu \in \mathcal{N}_\mu} \sum_{x_\nu \in \mathcal{N}_\nu} p_{\mu\nu}(x_\mu, x_\nu) \log_2 \frac{p_{\mu\nu}(x_\mu, x_\nu)}{p_\mu(x_\mu)p_\nu(x_\nu)}, \quad (\text{B1})$$

where we explicitly denoted the dependence of the probabilities on the state for clarity. In general, $I_{\mu\nu}$ will depend on time, but we focus on its stationary limit.

Note that Eq. (B1) can be readily applied to any stochastic process on networks. For instance, in the case of a multivariate continuous stochastic process, we need to replace sums with integrals and the layer state with a state vector $\vec{x}_\mu = (x_\mu^{(1)}, \dots, x_\mu^{(M_\mu)})$. Importantly, in this case, the joint probability distribution $p_{\mu\nu}$ is defined on an $(M_\mu + M_\nu)$ -dimensional space, since each node now represents a degree of freedom with a continuous state—e.g., in a multivariate Ornstein-Uhlenbeck process on a given multilayer network. In this case,

$$I_{\mu\nu} = \int d\vec{x}_\mu d\vec{x}_\nu p_{\mu\nu}(\vec{x}_\mu, \vec{x}_\nu) \log_2 \frac{p_{\mu\nu}(\vec{x}_\mu, \vec{x}_\nu)}{p_\mu(\vec{x}_\mu) p_\nu(\vec{x}_\nu)}, \quad (\text{B2})$$

where the probabilities are now probability density functions. The corresponding extension to the case in which each node can assume a finite number of states, such as the case of spreading or epidemic processes, is also immediate.

APPENDIX C: EXACT SOLUTION FOR THE MULTILAYER PROBABILITY AND MINIMAL PROPAGATION PATHS

We start from Eq. (5) and seek a solution of the form

$$p_{1,\dots,N} = p_{1,\dots,N}^{(1,0)} + \sum_{\mu=2}^{N-2} \left(\prod_{\nu=1}^{\mu-1} \epsilon_\nu \right) p_{1,\dots,N}^{(\mu,0)} + \prod_{\mu=1}^N \epsilon_\mu p_{1,\dots,N}^{(N,1)} + \mathcal{O}(\epsilon^2), \quad (\text{C1})$$

where $\epsilon_\mu = \tau_\mu / \tau_N$, and $\mathcal{O}(\epsilon^2)$ denotes second-order contributions for any $\mu = 1, \dots, N$. In Eq. (C1), $p_{1,\dots,N}^{(\mu,0)}$ denotes the zeroth-order contribution in ϵ_μ , and $p_{1,\dots,N}^{(N,1)}$ is instead the first-order contribution in the slowest timescale, $\epsilon_N = 1$. This expansion can be readily obtained by subsequent expansions in $\epsilon_1, \dots, \epsilon_N$ of the joint probability. In the limit $\epsilon_\mu \rightarrow 0$, the solution at the leading order $p_{1,\dots,N}^{(1,0)}$ obeys

$$\partial_t p_{1,\dots,N}^{(1,0)}(t) = \left[\hat{\Omega}_N + \sum_{\mu=1}^{N-1} \frac{\hat{\Omega}_\mu}{\epsilon_\mu} \right] p_{1,\dots,N}^{(1,0)} + \hat{\Omega}_1 p_{1,\dots,N}^{(2,0)}, \quad (\text{C2})$$

where we rescaled the time by the slowest timescale. Note that the solution does not depend on how interlayer interactions enter in \hat{W}_μ^H and thus in $\hat{\Omega}_\mu$, so our results generalize immediately to the case of pairwise interactions across layers.

For what follows, it is useful to introduce the notion of minimal propagation paths (mPPs) for a multilayer

network. Consider the directed graph \mathcal{G} with N nodes associated with a multilayer network, where each node represents a layer, and edges are given by interactions between layers. We define a propagation path (PP) on \mathcal{G} as a path from μ to ν with $\mu > \nu$ and such that it contains at least one direct interaction, i.e., one edge $\alpha \rightarrow \beta$ with $\alpha < \beta$. In Figs. 4(a) and 4(d), both $2 \rightarrow 3 \rightarrow 1$ and $3 \rightarrow 1 \rightarrow 2$ are PPs, respectively.

We then define the graph $\mathcal{G}^{(\nu)}(\nu^*)$ as the induced subgraph obtained by removing all nodes $\alpha > \nu$ —i.e., all layers slower than ν —except ν^* . A propagation path from μ to ν is a minimal propagation path if it is a PP both in \mathcal{G} and in the induced subgraph $\mathcal{G}^{(\nu)}(\mu)$. In Figs. 4(a) and 4(d), $2 \rightarrow 3 \rightarrow 1$ is not an mPP, but $3 \rightarrow 1 \rightarrow 2$ is.

With these ideas in mind, we solve Eq. (C2) order by order, with each order corresponding to a given layer. The first order is immediately solved by

$$0 = \hat{\Omega}_1(\{x\}_{\rightarrow 1}) p_{1|\rho(1)}^{\text{st}}, \quad (\text{C3})$$

where $\rho(1)$ is the set of all slower layers that connect to the first through a feedback link, $\{x\}_{\rightarrow 1}$ is the set of layers directly connected to the first, and we have that $p_{1,\dots,N}^{(1,0)} = p_{1|\rho(1)}^{\text{st}} p_{2,\dots,N}^{(1,0)}$. Note that for the first layer, by definition, $\{x\}_{\rightarrow 1} = \rho(1)$.

The second order, after a summation over x_1 , reads

$$0 = \left[\sum_{x_1} \hat{\Omega}_2(\{x\}_{\rightarrow 2}) p_{1|\rho(1)}^{\text{st}} \right] p_{2|\rho(2)}^{\text{st,eff}}, \quad (\text{C4})$$

where the operator in the bracket is the effective operator $\hat{\Omega}_{2|\rho(2)}^{\text{eff}}$. Its stationary probability $p_{2|\rho(2)}^{\text{st,eff}}$ inherits the dependence on slow layers through both the set $\{\rightarrow x_2\}$ and the dependencies of faster layers—in this case, $\rho(1)$ appearing in $p_{1|\rho(1)}^{\text{st}}$. Therefore, the conditional dependencies contained in $\rho(2)$ now include all slower layers that are connected to 2 through higher-order interactions, as well as all slower layers directly connected to 1 if there is a direct interaction from 1 to 2. Otherwise, in the absence of such directed connection, $\{x\}_{\rightarrow 2}$ does not include x_1 , and thus the effective operator coincides with $\hat{\Omega}_2$.

By recursively solving each order and marginalizing over the slower layers, we obtain Eq. (6) and the effective operators in Eq. (7). In particular, the set

$$\rho(\mu) = \{\nu > \mu : \exists \text{ mPP } \nu \rightarrow \mu \text{ or } \hat{C}_{\nu\mu} \neq 0\} \quad (\text{C5})$$

is the set of all layers connected to μ whether via an mPP or a single feedback link, and it determines the conditional structure of each term of the joint probability $p_{1,\dots,N}(t)$. Let us stress that, by construction, the only time dependence arises at order $\mathcal{O}(1)$, and thus it is ascribed to the effective operator of the slowest layer $\hat{\Omega}_N^{\text{eff}}$. Note that, although $\rho(N)$

is an empty set by definition, such an effective operator depends on all the previous layers by which it is influenced through the marginalization in Eq. (7).

APPENDIX D: BINARY TRIADIC INTERACTIONS

In Figs. 3 and 4 we considered, without loss of generality, systems where each layer is a simple two-node network. Denoting such nodes with $\mathcal{N}_\mu = \{A_\mu, B_\mu\}$ for all layers, we take triadic interactions to be

$$\hat{W}_\mu^H = (\hat{W}_\mu)^{i \rightarrow j} + \sum_{\nu \neq \mu} \sum_{k \in \mathcal{N}_\nu} C_{\nu\mu}^{k,i \rightarrow j} \delta(x_\nu, k), \quad (\text{D1})$$

so that the transition rates of a layer—i.e., its adjacency matrix—change in a binary fashion depending on the state of another layer, and acting additively with respect to pairwise interactions. For simplicity, we set $C_{\mu\nu}^{k,i \rightarrow j} = C_{\mu\nu}^{\text{eq}} \delta_{jk} + C_{\mu\nu}^{\text{cr}} (1 - \delta_{jk})$, where $C_{\mu\nu}^{\text{eq}}$ represents the triadic interactions of nodes influencing transitions to the same nodes—e.g., A_μ favoring the transition to A_ν in another layer—and $C_{\mu\nu}^{\text{cr}}$ interactions with the opposite effect—e.g., A_μ favoring the transition to B_ν . We stress that this specific choice does not affect the structure of information, which is fully determined by the conditional structure of Eq. (6) and thus is independent of the details of the dynamics.

APPENDIX E: MULTILAYER SOLUTION FOR BIRTH-AND-DEATH PROCESSES

A birth-and-death process of a species X may be interpreted as a random walk on a network, with each node representing a given number of particles x . In particular, the case of a finite reservoir with a maximum number of particles M_X , birth rate b , and death rate d is represented by a network with a fixed number of M_X nodes and adjacency matrix with elements:

$$W_{x \rightarrow j} = \begin{cases} (M_X - x)b & \text{if } j = x + 1 \\ xd & \text{if } j = x - 1 \\ 0 & \text{otherwise,} \end{cases} \quad (\text{E1})$$

so that each node is associated with the corresponding number of particles $x = 1, \dots, M_X$. In this representation, the node x is connected only to the neighbors $x \pm 1$, with the weight of the edge $x \rightarrow x + 1$ being $(M_X - x)b$, and the weight of $x \rightarrow x - 1$ being xd . This formally corresponds to the microscopic reactions,



where \emptyset_X represents the finite reservoir.

Then, the formalism presented in the main text can be straightforwardly applied by assuming that a given layer μ

is a network characterized by a $M_\mu \times M_\mu$ adjacency matrix \hat{W}_μ as in Eq. (E1). In this way, it represents a birth-and-death process for a given species μ with birth rate b_μ , death rate d_μ , and a maximum number of particles M_μ . In particular, a physically and biologically meaningful implementation of triadic interactions in this scenario is to modify b_μ and d_μ of a given layer so that they depend on the concentration of the number of particles in another layer. That is,

$$b_\mu^H = b_\mu + \sum_{\nu \neq \mu} C_{\nu\mu}^{\text{ex}} \frac{x_\nu}{M_\nu}, \quad d_\mu^H = d_\mu + \sum_{\nu \neq \mu} C_{\nu\mu}^{\text{in}} \frac{x_\nu}{M_\nu}, \quad (\text{E2})$$

where $C_{\mu\nu}^{\text{ex}}$ and $C_{\mu\nu}^{\text{in}}$ are the strength of triadic interactions. Note that, in general, higher-order birth and death rates can be time dependent through the states of the layers. From these rates, it is possible to define a matrix \hat{W}^H that takes the same form as in Eq. (E1), i.e., $(\hat{W}_\mu^H)_{i \rightarrow j} = \delta_{i,x_\mu} [(M_\mu - x_\mu) b_\mu^H \delta_{j,x_\mu+1} + x_\mu d_\mu^H \delta_{j,x_\mu-1}]$. Thus, in this case, triadic interactions affect all edges of another layer at once, solely distinguishing between birth and death transition rates. In particular, we consider inhibitory interactions $C_{\nu\mu}^{\text{in}}$ those that increase the death rate, and excitatory interactions $C_{\nu\mu}^{\text{ex}}$ those that increase the birth rate. Since the resulting interactions are linear, the effective operators in Eq. (7) depend only on the average of the probability distributions $p_{\mu|\rho(\mu)}^{\text{st,eff}}$.

Although the solution to a standard birth-and-death process with a finite reservoir and linear rates can be found analytically and is a binomial distribution (or a Poisson distribution if there is no maximum number of particles, so that we formally have an infinite number of nodes) [49], the effective operators in Eq. (7) may not admit a general closed-form solution. Thus, we introduce an efficient numerical scheme to obtain the solution to the master equation at all times. We write the solution in the recursive form,

$$p_{1,\dots,N}(t + \Delta t) = \prod_{\nu=1}^{N-1} p_{\mu|\rho(\mu)}^{\text{st}(\rightsquigarrow\mu)} \sum_{\tilde{x}_N} P_N(x_N, t + \Delta t; \tilde{x}_N, t) p_N^{\text{st}(\rightsquigarrow N)}(\tilde{x}_N, t), \quad (\text{E3})$$

where Δt is the time step and $P_N(x_N, t + \Delta t; \tilde{x}_N, t)$ is the propagator associated to the effective operator $\hat{\Omega}_N^{\text{eff}}$, which we can write as

$$P_N(x_N, t + \Delta t; \tilde{x}_N, t) = p_N^{\text{st}(\rightsquigarrow N)} + \sum_{i=2}^N \vec{\omega}_i a^{(i)} e^{\lambda_i \Delta t}, \quad (\text{E4})$$

where $p_N^{\text{st}(\rightsquigarrow N)}$ is the stationary probability of the N layer, and $\vec{\omega}_i$ and λ_i are, respectively, the eigenvectors and

eigenvalues of $\hat{\Omega}_N^{\text{eff}}$, ordered in such a way that $\lambda_0 = 0$. This recursive form is by definition exact when $\Delta t \ll 1$, and it is particularly useful when the input to the system $h(t)$ varies with time, since all effective operators would depend explicitly on time as well.

APPENDIX F: SIGNALING ARCHITECTURES

The architecture depicted in Fig. 6(a) corresponds to the following birth and death rates. The receptor receives a time-varying input $h(t)$, so that

$$b_R^H(t) = b_R + h(t), \quad d_R^H = d_R, \quad (\text{F1})$$

with b_R and d_R the baseline birth and death rates. We typically set $b_R = 0$, so the receptor is entirely input driven and time dependent. The readout population is stimulated by the receptor and inhibited by the storage, so that

$$b_U^H(t) = b_U + C_{RU}^{\text{ex}} \frac{r(t)}{N_R}, \quad d_U^H(t) = d_U + C_{SU}^{\text{in}} \frac{s(t)}{N_S}, \quad (\text{F2})$$

with $r(t)$ and $s(t)$ the number of receptor and storage particles at time t , i.e., the node in the corresponding layer. Finally, the storage follows

$$b_S^H(t) = b_S + C_{US}^{\text{ex}} \frac{u(t)}{N_U}, \quad d_S^H = d_S, \quad (\text{F3})$$

so that, at a given time, it is excited by the number of readout units $u(t)$.

The architecture in Fig. 6(b) is instead specified by

$$b_R^H(t) = b_R + h(t), \quad d_R^H(t) = d_R + C_{RS}^{\text{in}} \frac{s(t)}{N_S} \quad (\text{F4})$$

for the receptor, which is inhibited by the storage population. The readout layer is simply given by

$$b_U^H(t) = b_U + C_{RU}^{\text{ex}} \frac{r(t)}{N_R}, \quad d_U^H = d_U, \quad (\text{F5})$$

and, finally,

$$b_S^H(t) = b_S + C_{RS}^{\text{ex}} \frac{r(t)}{N_R}, \quad d_S^H = d_S, \quad (\text{F6})$$

so that the storage is excited directly by the receptor.

APPENDIX G: PROCESSES ON DISJOINTED COMPONENTS

We consider a two-layer network, where one layer has two disjointed components with independent states x_1 and x_2 , each with two nodes (A, B). The other layer, with state E and nodes ($E_{\text{off}}, E_{\text{on}}$), influences both components at the same time. The rates are

$\omega_1^{A \rightarrow B} = \omega_2^{A \rightarrow B} = [w_B + c\delta(E, E_{\text{on}})]/\tau$, $\omega_1^{B \rightarrow A} = \omega_2^{B \rightarrow A} = w_A$, $\omega_E^{\text{on} \rightarrow \text{off}} = w_0/\tau_E$, $\omega_E^{\text{off} \rightarrow \text{on}} = w_e/\tau_E$. For Fig. 5, we set $w_B = 0$, $w_A = 1$, $w_0 = 5$, $w_e = 1$. In the numerical simulation, to recover the $c \rightarrow \infty$ limit obtained analytically, we set $c = 10^3$.

When $\tau_E \gg \tau$, the interlayer interaction is of the feedback type, so that $p_{1,2,E}(t) = p_{1|E}^{\text{st}} p_{2|E}^{\text{st}} p_E(t)$. In this case, we can write the sum of the mutual information:

$$\begin{aligned} I_{12} + I_{12,E} &= \sum_{1,2,E} p_{1,2,E} \left[\log_2 \frac{p_{1,2}}{p_1 p_2} + \log_2 \frac{p_{1|E} p_{2|E} p_E}{p_{1,2} p_E} \right] \\ &= I_{1,E} + I_{2,E}, \end{aligned} \quad (\text{G1})$$

which shows that in this limit the mutual information $I_{1,2}$ induced by the shared influence of E is equal to the sum of the information generated by E in 1 and 2 minus the information with E that both components have. Note that this result, as well as the factorization of the probability, does not change if we introduce two different timescales for the two independent processes, τ_1 and τ_2 , as long as $\tau_E \gg \tau_1, \tau_2$. In this case, rather than having a single layer with timescale τ and two disjointed components, we would have two disconnected layers.

-
- [1] M. De Domenico, *More is different in real-world multilayer networks*, *Nat. Phys.* **19**, 1247 (2023).
 - [2] L. V. Schaffer and T. Ideker, *Mapping the multiscale structure of biological systems*, *Cell Syst.* **12**, 622 (2021).
 - [3] A. Hastings, *Timescales, dynamics, and ecological understanding*, *Ecology* **91**, 3471 (2010).
 - [4] O. Radulescu, A. N. Gorban, A. Zinovyev, and A. Lilienbaum, *Robust simplifications of multiscale biochemical networks*, *BMC Syst. Biol.* **2**, 1 (2008).
 - [5] T. Poisot, D. B. Stouffer, and D. Gravel, *Beyond species: Why ecological interaction networks vary through space and time*, *Oikos* **124**, 243 (2015).
 - [6] N. Timme, S. Ito, M. Myroshnychenko, F.-C. Yeh, E. Hsiolki, P. Hottowy, and J. M. Beggs, *Multiplex networks of cortical and hippocampal neurons revealed at different timescales*, *PLoS One* **9**, e115764 (2014).
 - [7] S. E. Cavanagh, L. T. Hunt, and S. W. Kennerley, *A diversity of intrinsic timescales underlie neural computations*, *Front. Neural Circuits* **14**, 615626 (2020).
 - [8] C. J. Honey, E. L. Newman, and A. C. Schapiro, *Switching between internal and external modes: A multiscale learning principle*, *Netw. Neurosci.* **1**, 339 (2017).
 - [9] B. Mélykúti, J. P. Hespanha, and M. Khammash, *Equilibrium distributions of simple biochemical reaction systems for time-scale separation in stochastic reaction networks*, *J. R. Soc. Interface* **11**, 20140054 (2014).
 - [10] G. Nicoletti, L. Saravia, F. Momo, A. Maritan, and S. Suweis, *The emergence of scale-free fires in Australia*, *Iscience* **26**, 106181 (2023).
 - [11] G. Bianconi, *Multilayer Networks: Structure and Function* (Oxford University Press, New York, 2018).

- [12] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas, *Mathematical formulation of multilayer networks*, *Phys. Rev. X* **3**, 041022 (2013).
- [13] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardenes, M. Romance, I. Sendina-Nadal, Z. Wang, and M. Zanin, *The structure and dynamics of multilayer networks*, *Phys. Rep.* **544**, 1 (2014).
- [14] O. Artime, B. Benigni, G. Bertagnolli, V. d'Andrea, R. Gallotti, A. Ghavasieh, S. Raimondo, and M. De Domenico, *Multilayer Network Science: From Cells to Societies* (Cambridge University Press, Cambridge, England, 2022).
- [15] A. Aleta and Y. Moreno, *Multilayer networks in a nutshell*, *Annu. Rev. Condens. Matter Phys.* **10**, 45 (2019).
- [16] M. Salehi, R. Sharma, M. Marzolla, M. Magnani, P. Siyari, and D. Montesi, *Spreading processes in multilayer networks*, *IEEE Trans. Netw. Sci. Eng.* **2**, 65 (2015).
- [17] M. De Domenico, C. Granell, M. A. Porter, and A. Arenas, *The physics of spreading processes in multilayer networks*, *Nat. Phys.* **12**, 901 (2016).
- [18] G. F. de Arruda, F. A. Rodrigues, and Y. Moreno, *Fundamentals of spreading processes in single and multilayer complex networks*, *Phys. Rep.* **756**, 1 (2018).
- [19] A. Ghavasieh, C. Nicolini, and M. De Domenico, *Statistical physics of complex information dynamics*, *Phys. Rev. E* **102**, 052304 (2020).
- [20] S. Flatt, D. M. Busiello, S. Zamuner, and P. De Los Rios, *ABC transporters are billion-year-old Maxwell demons*, *Commun. Phys.* **6**, 205 (2023).
- [21] G. Nicoletti, P. Padmanabha, S. Azaele, S. Suweis, A. Rinaldo, and A. Maritan, *Emergent encoding of dispersal network topologies in spatial metapopulation models*, *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2311548120 (2023).
- [22] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, *Networks beyond pairwise interactions: Structure and dynamics*, *Phys. Rep.* **874**, 1 (2020).
- [23] F. Battiston, E. Amico, A. Barrat, G. Bianconi, G. Ferraz de Arruda, B. Franceschiello, I. Iacopini, S. Kéfi, V. Latora, Y. Moreno *et al.*, *The physics of higher-order interactions in complex systems*, *Nat. Phys.* **17**, 1093 (2021).
- [24] E. Bairey, E. D. Kelsic, and R. Kishony, *High-order species interactions shape ecosystem diversity*, *Nat. Commun.* **7**, 12285 (2016).
- [25] J. Grilli, G. Barabás, M. J. Michalska-Smith, and S. Allesina, *Higher-order interactions stabilize dynamics in competitive network models*, *Nature (London)* **548**, 210 (2017).
- [26] A. D. Letten and D. B. Stouffer, *The mechanistic basis for higher-order interactions and non-additivity in competitive communities*, *Ecol. Lett.* **22**, 423 (2019).
- [27] T. Carletti, F. Battiston, G. Cencetti, and D. Fanelli, *Random walks on hypergraphs*, *Phys. Rev. E* **101**, 022308 (2020).
- [28] A. P. Millán, J. J. Torres, and G. Bianconi, *Explosive higher-order Kuramoto dynamics on simplicial complexes*, *Phys. Rev. Lett.* **124**, 218301 (2020).
- [29] G. St-Onge, H. Sun, A. Allard, L. Hébert-Dufresne, and G. Bianconi, *Universal nonlinear infection kernel from heterogeneous exposure on higher-order networks*, *Phys. Rev. Lett.* **127**, 158301 (2021).
- [30] S. Majhi, M. Perc, and D. Ghosh, *Dynamics on higher-order networks: A review*, *J. R. Soc. Interface* **19**, 20220043 (2022).
- [31] T. Gibbs, S. A. Levin, and J. M. Levine, *Coexistence in diverse communities with higher-order interactions*, *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2205063119 (2022).
- [32] H. Sun, F. Radicchi, J. Kurths, and G. Bianconi, *The dynamic nature of percolation on networks with triadic interactions*, *Nat. Commun.* **14**, 1308 (2023).
- [33] W.-H. Cho, E. Barcelon, and S. J. Lee, *Optogenetic glia manipulation: Possibilities and future prospects*, *Exp. Neurobiol.* **25**, 197 (2016).
- [34] N. Boers, B. Goswami, A. Rheinwalt, B. Bookhagen, B. Hoskins, and J. Kurths, *Complex networks reveal global pattern of extreme-rainfall teleconnections*, *Nature (London)* **566**, 373 (2019).
- [35] J. Faskowitz, R. F. Betzel, and O. Sporns, *Edges in brain networks: Contributions to models of structure and function*, *Netw. Neurosci.* **6**, 1 (2022).
- [36] A. Hilfinger, T. M. Norman, G. Vinnicombe, and J. Paulsson, *Constraints on fluctuations in sparsely characterized biological systems*, *Phys. Rev. Lett.* **116**, 058101 (2016).
- [37] J. Yan, A. Hilfinger, G. Vinnicombe, J. Paulsson, *Kinetic uncertainty relations for the control of stochastic reaction networks*, *Phys. Rev. Lett.* **123**, 108101 (2019).
- [38] W. Ma, A. Trusina, H. El-Samad, W. A. Lim, and C. Tang, *Defining network topologies that can achieve biochemical adaptation*, *Cell* **138**, 760 (2009).
- [39] S. J. Rahi, J. Larsch, K. Pecani, A. Y. Katsov, N. Mansouri, K. Tsaneva-Atanasova, E. D. Sontag, and F. R. Cross, *Oscillatory stimuli differentiate adapting circuit topologies*, *Nat. Methods* **14**, 1010 (2017).
- [40] E. Valdano, L. Ferreri, C. Poletto, and V. Colizza, *Analytical computation of the epidemic threshold on temporal networks*, *Phys. Rev. X* **5**, 021005 (2015).
- [41] A. Koher, H. H. K. Lentz, J. P. Gleeson, and P. Hövel, *Contact-based model for epidemic spreading on temporal networks*, *Phys. Rev. X* **9**, 031017 (2019).
- [42] D. M. Busiello, S. Liang, F. Piazza, and P. De Los Rios, *Dissipation-driven selection of states in non-equilibrium chemical networks*, *Commun. Chem.* **4**, 16 (2021).
- [43] A. V. Dass, T. Georgelin, F. Westall, F. Foucher, P. De Los Rios, D. M. Busiello, S. Liang, and F. Piazza, *Equilibrium and non-equilibrium furanose selection in the ribose isomerisation network*, *Nat. Commun.* **12**, 2749 (2021).
- [44] D. M. Busiello, D. Gupta, and A. Maritan, *Coarse-grained entropy production with multiple reservoirs: Unraveling the role of time scales and detailed balance in biology-inspired systems*, *Phys. Rev. Res.* **2**, 043257 (2020).
- [45] A. Celani, S. Bo, R. Eichhorn, and E. Aurell, *Anomalous thermodynamics at the microscale*, *Phys. Rev. Lett.* **109**, 260603 (2012).
- [46] M. De Domenico and J. Biamonte, *Spectral entropies as information-theoretic tools for complex network comparison*, *Phys. Rev. X* **6**, 041062 (2016).
- [47] P. Villegas, T. Gili, G. Caldarelli, and A. Gabrielli, *Laplacian renormalization group for heterogeneous networks*, *Nat. Phys.* **19**, 445 (2023).

- [48] J. Schnakenberg, *Network theory of microscopic and macroscopic behavior of master equation systems*, *Rev. Mod. Phys.* **48**, 571 (1976).
- [49] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences* Springer Series in Synergetics, 3rd ed. (Springer-Verlag, Berlin, 2004), Chap. 7.
- [50] D. T. Gillespie, *Exact stochastic simulation of coupled chemical reactions*, *J. Phys. Chem.* **81**, 2340 (1977).
- [51] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences* (Ref. [49]), Chap. 3.
- [52] B. Mariani, G. Nicoletti, M. Bisio, M. Maschietto, S. Vassanelli, and S. Suweis, *Disentangling the critical signatures of neural activity*, *Sci. Rep.* **12**, 10770 (2022).
- [53] S. Bo and A. Celani, *Multiple-scale stochastic processes: Decimation, averaging and beyond*, *Phys. Rep.* **670**, 1 (2017).
- [54] P. Villa Martín, M. A. Muñoz, and S. Pigolotti, *Bet-hedging strategies in expanding populations*, *PLoS Comput. Biol.* **15**, e1006529 (2019).
- [55] C. Berton, D. M. Busiello, S. Zamuner, E. Solari, R. Scopelliti, F. Fadaei-Tirani, K. Severin, and C. Pezzato, *Thermodynamics and kinetics of protonated merocyanine photoacids in water*, *Chem. Sci.* **11**, 8457 (2020).
- [56] G. Nicoletti and D. M. Busiello, *Mutual information disentangles interactions from changing environments*, *Phys. Rev. Lett.* **127**, 228301 (2021).
- [57] G. Nicoletti and D. M. Busiello, *Mutual information in changing environments: non-linear interactions, out-of-equilibrium systems, and continuously-varying diffusivities*, *Phys. Rev. E* **106**, 014153 (2022).
- [58] G. Nicoletti, A. Maritan, and D. M. Busiello, *Information-driven transitions in projections of underdamped dynamics*, *Phys. Rev. E* **106**, 014118 (2022).
- [59] A. V. Chechkin, F. Seno, R. Metzler, and I. M. Sokolov, *Brownian yet non-Gaussian diffusion: From superstatistics to subordination of diffusing diffusivities*, *Phys. Rev. X* **7**, 021002 (2017).
- [60] Y. Tu and W.-J. Rappel, *Adaptation of living systems*, *Annu. Rev. Condens. Matter Phys.* **9**, 183 (2018).
- [61] G. Lan, P. Sartori, S. Neumann, V. Sourjik, and Y. Tu, *The energy–speed–accuracy trade-off in sensory adaptation*, *Nat. Phys.* **8**, 422 (2012).
- [62] H. Wajant, K. Pfizenmaier, and P. Scheurich, *Tumor necrosis factor signaling*, *Cell Death Differ.* **10**, 45 (2003).
- [63] J. Selimkhanov, B. Taylor, J. Yao, A. Pilko, J. Albeck, A. Hoffmann, L. Tsimring, and R. Wollman, *Accurate information transmission through dynamic biochemical signaling networks*, *Science* **346**, 1370 (2014).
- [64] J. Benda, *Neural adaptation*, *Curr. Biol.* **31**, R110 (2021).
- [65] A. Celani, T. S. Shimizu, and M. Vergassola, *Molecular and functional aspects of bacterial chemotaxis*, *J. Stat. Phys.* **144**, 219 (2011).
- [66] G. Lan and Y. Tu, *Information processing in bacteria: memory, computation, and statistical physics: A key issues review*, *Rep. Prog. Phys.* **79**, 052601 (2016).
- [67] G. Nicoletti, M. Bruzzone, S. Suweis, M. Dal Maschio, and D. M. Busiello, *Adaptation maximizes information and minimizes dissipation across biological scales*, [arXiv:2301.12812](https://arxiv.org/abs/2301.12812).
- [68] F. Avanzini, N. Freitas, and M. Esposito, *Circuit theory for chemical reaction networks*, *Phys. Rev. X* **13**, 021041 (2023).
- [69] F. Brauns, J. Halatek, and E. Frey, *Phase-space geometry of mass-conserving reaction-diffusion dynamics*, *Phys. Rev. X* **10**, 041036 (2020).
- [70] S. Liang, D. M. Busiello, and P. De Los Rios, *Emergent thermophoretic behavior in chemical reaction systems*, *New J. Phys.* **24**, 123006 (2022).
- [71] S. J. Kiebel, J. Daunizeau, and K. J. Friston, *A hierarchy of time-scales and the brain*, *PLoS Comput. Biol.* **4**, e1000209 (2008).
- [72] M. Golesorkhi, J. Gomez-Pilar, F. Zilio, N. Berberian, A. Wolff, M. C. Yagoub, and G. Northoff, *The brain and its time: Intrinsic neural timescales are key for input processing*, *Commun. Biol.* **4**, 970 (2021).
- [73] R. Rossi-Pool, A. Zainos, M. Alvarez, S. Parra, J. Zizumbo, and R. Romo, *Invariant timescale hierarchy across the cortical somatosensory network*, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2021843118 (2021).
- [74] A. Wolff, N. Berberian, M. Golesorkhi, J. Gomez-Pilar, F. Zilio, and G. Northoff, *Intrinsic neural timescales: Temporal integration and segregation*, *Trends Cognit. Sci.* **26**, 159 (2022).
- [75] R. Zeraati, Y.-L. Shi, N. A. Steinmetz, M. A. Gieselmann, A. Thiele, T. Moore, A. Levina, and T. A. Engel, *Intrinsic timescales in the visual cortex change with selective attention and reflect spatial connectivity*, *Nat. Commun.* **14**, 1858 (2023).
- [76] M. De Domenico, S. Sasai, and A. Arenas, *Mapping multiplex hubs in human functional brain networks*, *Front. Neurosci.* **10**, 326 (2016).