

Influence of time delay on information exchanges between coupled linear stochastic systemsM. L. Rosinberg,^{1,*} G. Tarjus,¹ and T. Munakata²¹*Laboratoire de Physique Théorique de la Matière Condensée, Sorbonne Université, CNRS, 75005 Paris, France*²*Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan*

(Received 3 July 2018; published 24 September 2018)

Time lags are ubiquitous in biophysiological processes and more generally in real-world complex networks. It has been recently proposed to use information-theoretic tools such as transfer entropy to detect and estimate a possible delay in the couplings. In this work we focus on stationary linear stochastic processes in continuous time and compute the transfer entropy in the presence of delay and correlated noises, using an approximate but numerically effective solution to the relevant Wiener-Hopf factorization problem. Our results rectify and complete the recent study of Barnett and Seth [L. Barnett and A. K. Seth, *J. Neurosci. Methods* **275**, 93 (2017)].

DOI: [10.1103/PhysRevE.98.032130](https://doi.org/10.1103/PhysRevE.98.032130)**I. INTRODUCTION**

There is no need to overstate the prevalence of time lags in biological processes and more generally in complex networks, from neural and gene regulatory circuits to climate, traffic, communication, or computer systems, to name just a few (see, e.g., [1–3] and references therein). Delays, arising from finite propagation or processing times, also play a crucial role in sensor-actuator feedback applications [4,5]. Particularly significant is the interplay of delays and noise which is at the origin of the complex dynamical behavior observed in a host of experimental systems [6]. Although these issues are increasingly the focus of theoretical and experimental investigations, there are many examples in which very little is known about the magnitude of the time lags, or even their existence, and how they are distributed in the network. As a result, this may lead to a wrong identification of the causal relationships between the various physical or chemical processes occurring in the network.

A common method in biology or climate science for estimating delays and the direction of information transfer between coupled systems is to consider temporal cross-correlation functions extracted from time-series data (see, e.g., [7–10]). A peak in these functions is then interpreted as the time it takes for an upstream signal (e.g., a protein concentration) to influence the downstream one (e.g., a target gene). The reliability of this method is questionable, however, and the true physical meaning of the maxima in the cross correlations is often unclear [11]. As another option, it has been recently proposed to use information-theoretic measures built on the concept of Shannon entropy and mutual information, such as transfer entropy (TE). The idea is to identify a possible interaction delay by searching for a maximum in the TE (or some variant of it) as a function of an additional parameter, typically the prediction horizon [12–16]. Transfer entropy [17,18], which is essentially a generalization of Wiener-Granger causality principle [19,20], characterizes

the directional information flow between two interacting random processes and has become a popular tool for analyzing networks of interacting agents or processes, in particular in neuroscience [21–23]. Whether or not this method is effective is still debated [24], and before applying it to real data it is worth checking the results on systems where the dynamical equations are known and that can be fully analyzed numerically or even analytically.

With this perspective in mind, the present work is motivated by a recent study [25] that focuses on the determination of Granger causality (GC) from empirically sampled data produced by an underlying continuous-time process. As a working example, the case of a bivariate linear stochastic process with delayed interaction is investigated in detail. Our objective here is not to discuss the important issue of subsampling, which is meticulously treated in [25] (see also [26]), but merely to revisit the calculation of the continuous-time GC, which is equivalent to the TE in the case of multivariate Gaussian processes [27]. Indeed, it turns out that the analytical solution proposed in [25] for the process with delay is incorrect when the noises acting on each subsystem are correlated. This is not an academic issue because such correlations are often required when modeling real systems, for instance, cell metabolic networks [7,8,28–30]. It is thus important to have the correct expression of the continuous-time TE before investigating the issue of delay identification (and, in a second stage, the effects of subsampling). We also take this opportunity to rephrase the problem into a language that is perhaps more familiar to physicists, in particular those concerned with the use of information-theoretic concepts and tools in the field of stochastic and information thermodynamics [31–35].

The paper is organized as follows. In Sec. II we recall the definition of TE and its relationship with GC in the context of time-continuous stochastic processes. We also present the class of linear stochastic systems that will be considered. In Sec. III we then focus on a bivariate system with a time lag in one of the couplings and we describe the calculation of the TEs in both directions. As usual with delay systems, the complication arises from the fact that the state space is infinite dimensional. We then introduce an approximation scheme in

*mlr@lptmc.jussieu.fr

the frequency domain which allows us to solve the relevant Wiener-Hopf factorization problem. We also emphasize some points which in our opinion are not clearly stated in [25], in particular the condition for the spectral expression of the TE rate to be valid. The numerical calculations presented in Sec. IV show that our method leads to a rapidly convergent solution, and we then investigate the issue of delay identification. A summary of the results is provided in Sec. V. In addition, the analytical expression of an alternative, simplified version of TE is derived in the Appendixes.

II. SETUP

A. Transfer entropy in continuous time

Consider two subsystems X_1 and X_2 of a stochastic system \mathbf{X} . The corresponding random variables or states at time t are denoted by $X_1(t)$ and $X_2(t)$, respectively. As originally defined in [17] in the discrete-time framework, the transfer entropy from X_i (the source) to X_j (the target) quantifies the reduction of uncertainty in the value of $X_j(t)$ when learning the past of $X_i(t)$, if the past of $X_j(t)$ is already known. In continuous time, one must introduce infinitesimal increments, just as in the case of GC [36], which leads us to define TE as the rate

$$\mathcal{T}_{i \rightarrow j} = \lim_{h \rightarrow 0^+} \frac{1}{h} I[X_j(t+h) : X_i^-(t) | X_j^-(t)], \quad (1)$$

where $X_i^-(t) \equiv \{X_i(s) : s \leq t\}$ and I is the conditional mutual information [37]. In terms of probability distributions, this is rewritten as

$$\mathcal{T}_{i \rightarrow j} = \lim_{h \rightarrow 0^+} \frac{1}{h} \left\langle \ln \frac{p[X_j(t+h) | \mathbf{X}^-(t)]}{p[X_j(t+h) | X_j^-(t)]} \right\rangle, \quad (2)$$

where $\mathbf{X}^-(t) \equiv (X_1^-(t), X_2^-(t))$. Since we only focus on stationary processes, $\mathcal{T}_{i \rightarrow j}$ is here independent of t . Note that the *whole* past history of both the source and the target up to time t are taken into account in Eqs. (1) and (2). Conditioning the mutual information on $X_j^-(t)$ is natural because the marginal (or coarse-grained) processes X_1 and X_2 are generally non-Markovian even when the joint process \mathbf{X} is Markovian. However, it is also sensible to take into account the whole vector $X_i^-(t)$ and not only the latest state $X_i(t)$, as is done in another context [35,38]. This seems particularly justified for the class of non-Markovian processes that are studied in the following [cf. Eqs. (8)–(10)]. [In the original definition of TE in discrete time [17,18], the lengths of the two state vectors $X_1^-(t)$ and $X_2^-(t)$, i.e., the number of time bins in the past, are generally finite and possibly different. This definition can also be extended to continuous time [39].]

Although the rate $\mathcal{T}_{i \rightarrow j}$ is an interesting quantity *per se*, for instance, in the context of stochastic thermodynamics [35], it cannot be used to infer a possible time lag in the couplings. Just as in the discrete-time framework [16], this role is devoted to the finite-horizon TE

$$\begin{aligned} T_{i \rightarrow j}(h) &= I[X_j(t+h) : X_i^-(t) | X_j^-(t)] \\ &= \left\langle \ln \frac{p[X_j(t+h) | \mathbf{X}^-(t)]}{p[X_j(t+h) | X_j^-(t)]} \right\rangle, \end{aligned} \quad (3)$$

which in the terminology of forecasting [40] quantifies how much the prediction of $X_j(t+h)$ is improved by using both $X_i^-(t)$ and $X_j^-(t)$ rather than $X_j^-(t)$ alone. This is clearly quite similar to the definition of the continuous-time finite-horizon GC in [25], which itself extends the classic discrete-time definition [40,41]. [Note in passing that $T_{i \rightarrow j}(h)$, as a relative entropy, is a non-negative quantity and vanishes at $h=0$ by construction.] The main difference is that TE is model independent whereas GC is commonly defined in the context of vector autoregressive (VAR) processes. However, TE and GC become fully equivalent when the variables are Gaussian distributed, with a simple factor of $\frac{1}{2}$ relating the two quantities [27]. Indeed, since the entropy of Gaussian distributions is directly expressed in terms of their covariance matrix [37], Eq. (3) then yields

$$T_{i \rightarrow j}(h) = \frac{1}{2} \ln \frac{\sigma'_{jj}(h)}{\sigma_{jj}(h)}, \quad (4)$$

where

$$\sigma_{jj}(h) = \langle [X_j(t+h) - \langle X_j(t+h) | \mathbf{X}^-(t) \rangle]^2 \rangle \quad (5)$$

and

$$\sigma'_{jj}(h) = \langle [X_j(t+h) - \langle X_j(t+h) | X_j^-(t) \rangle]^2 \rangle \quad (6)$$

are the variances of $p[X_j(t+h) | \mathbf{X}^-(t)]$ and $p[X_j(t+h) | X_j^-(t)]$, respectively. Note that we have assumed that the two variables $X_1(t)$ and $X_2(t)$ are univariate, which will be the situation considered hereafter (see [27] for the generalization to multivariate variables). In the language of forecasting, $\langle X_j(t+h) | \mathbf{X}^-(t) \rangle$ and $\langle X_j(t+h) | X_j^-(t) \rangle$ are interpreted as the minimum mean-square error (MMSE) estimates of $X_j(t+h)$, and $\sigma_{jj}(h)$ and $\sigma'_{jj}(h)$ are the corresponding mean-square prediction errors. The present work is mainly concerned with the calculation of these quantities in the presence of delayed interactions.

Notwithstanding the valuable arguments for conditioning $X_j(t+h)$ on the infinite past histories of $X_1(t)$ and $X_2(t)$, it is also useful from a practical viewpoint to consider a simplified version of $T_{i \rightarrow j}(h)$ that only involves the states at time t ,

$$\begin{aligned} \bar{T}_{i \rightarrow j}(h) &= I[X_j(t+h) : X_i(t) | X_j(t)] \\ &= \left\langle \ln \frac{p[X_j(t+h) | \mathbf{X}(t)]}{p[X_j(t+h) | X_j(t)]} \right\rangle. \end{aligned} \quad (7)$$

In the case of Gaussian distributed variables, $\bar{T}_{i \rightarrow j}(h)$ is given by an expression similar to Eq. (4) whose explicit calculation is presented in Appendix A. It is worth noticing that $\bar{T}_{i \rightarrow j}(h)$ is an upper bound on $T_{i \rightarrow j}(h)$ if the joint process \mathbf{X} is Markovian¹ (and in turn $\bar{T}_{i \rightarrow j}$, the slope at the origin, is an upper bound on $\mathcal{T}_{i \rightarrow j}$ [33,35]). However, this is no longer true in the general case.

¹If \mathbf{X} is a Markov process, one has $p[X_j(t+h) | \mathbf{X}^-(t)] = p[X_j(t+h) | \mathbf{X}(t)]$ and thus $\bar{T}_{i \rightarrow j}(h) - T_{i \rightarrow j}(h) = H[X_j(t) | X_j(t)] - H[X_j(t) | X_j^-(t)]$, where $H[\cdot]$ is the conditional Shannon entropy. This difference is always non-negative.

B. Class of models

In [25], the class of linear stochastic integro-differential equations

$$\dot{\mathbf{X}}(t) = - \int_0^\infty ds \mathbf{A}(s)\mathbf{X}(t - s) + \boldsymbol{\xi}(t) \quad (8)$$

was introduced, where \mathbf{X} is an n -dimensional vector process, $\mathbf{A}(s)$ is an $n \times n$ matrix of functions or generalized functions (distributions), and $\boldsymbol{\xi}(t)$ is an n -dimensional vector of (generally correlated) Gaussian white noises. Equation (8) is viewed as the continuous-time analog of a VAR representation (see [36] for mathematical details), and this type of equation, which can be obtained through the linearization of nonlinear problems, appears in various research fields where the history of the state variables must be taken into account, e.g., in econometry, biology, or control theory. Depending on the context, the time lags may then be either discrete or distributed according to some density function. This latter case often occurs in the modeling of biological processes [42–44]. In the following, we will instead focus on the case of discrete delays, so Eq. (8) takes the form of a linear stochastic differential delay equation

$$\dot{\mathbf{X}}(t) = - \sum_{\alpha=1}^N \mathbf{A}_\alpha \mathbf{X}(t - \tau_\alpha) + \boldsymbol{\xi}(t), \quad (9)$$

with possibly N distinct delays τ_α [45]. In recent years, such multivariate, multidelayed equations have been used to study synchronization problems in complex networks (see, e.g., [46]). However, for simplicity, and given the purpose of this work, we will only introduce a single delay τ in one of the couplings and consider a bivariate system, as already stated.

III. BIVARIATE LINEAR PROCESS WITH A TIME-DELAYED COUPLING

For definiteness, let us assume that the delay takes place in the feedback from X_2 to X_1 . Equation (9) then becomes

$$\dot{\mathbf{X}}(t) = - \begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix} \mathbf{X}(t) - \begin{pmatrix} 0 & a_{12} \\ 0 & 0 \end{pmatrix} \mathbf{X}(t - \tau) + \boldsymbol{\xi}(t), \quad (10)$$

where $\xi_1(t)$ and $\xi_2(t)$ are zero-mean Gaussian white noises with covariances $\langle \xi_i(t)\xi_j(t') \rangle = 2D_{ij}\delta(t - t')$. We stress that we do not assume independent noises as is usually done in the context of stochastic thermodynamics (the so-called bipartite assumption [31–35]). It is clear that the case of a time lag in the coupling from X_1 to X_2 follows by exchanging the labels 1 and 2. On the other hand, the two directions are not equivalent for a given model, and for the process described by Eq. (10) we will see that the computation of the TE in the direction $2 \rightarrow 1$ is significantly more difficult than in the direction $1 \rightarrow 2$. One should also keep in mind that time-delayed interactions typically lead to bifurcations and complicated dynamics [1,3]. This is an interesting issue in itself, but to simplify the forthcoming discussion we assume that the delay and the coupling parameters a_{ij} are such that a stable stationary solution exists (in other words, the spectral density matrix is bounded for all values of the frequency ω).

Moreover, in Sec. III B, to further simplify the model, we will completely suppress the possible occurrence of instabilities by setting $a_{21} = 0$, which corresponds to the model studied in Sec. 4 of [25].

Since we only focus on the stationary regime, we can assume that the process started at $t_0 = -\infty$ and forget about the initial condition. The solution of Eq. (10) then reads

$$\mathbf{X}(t) = \int_{-\infty}^t ds \mathbf{H}(t - s)\boldsymbol{\xi}(s), \quad (11)$$

where $\mathbf{H}(t)$ is the response (or Green’s or transfer) functions matrix. Equivalently, in Fourier space or the frequency domain,

$$\mathbf{X}(\omega) = \mathbf{H}(\omega)\boldsymbol{\xi}(\omega), \quad (12)$$

where

$$\begin{aligned} \mathbf{H}(\omega) &\equiv \int_{-\infty}^{+\infty} dt e^{i\omega t} \mathbf{H}(t) \\ &= \frac{1}{(a_{11} - i\omega)(a_{22} - i\omega) - a_{12}a_{21}e^{i\omega\tau}} \\ &\quad \times \begin{pmatrix} a_{22} - i\omega & -a_{12}e^{i\omega\tau} \\ -a_{21} & a_{11} - i\omega \end{pmatrix}. \end{aligned} \quad (13)$$

The power-spectrum matrix whose elements are the Fourier transform of the stationary time-dependent correlation functions $\phi_{ij}(t) = \langle X_i(t')X_j(t' + t) \rangle$ is then given by

$$\mathbf{S}(\omega) = \mathbf{H}(\omega)(2\mathbf{D})\mathbf{H}^*(\omega), \quad (14)$$

where $2\mathbf{D}$ is the diffusion matrix with elements $2D_{ij}$ and the superscript $*$ denotes complex conjugate and matrix transpose.

A. Transfer entropy in the direction $1 \rightarrow 2$

1. Finite-horizon TE

We begin with the calculation of the TE in the direction $1 \rightarrow 2$, which is fairly straightforward. Although part of the material in this section may be viewed as a mere application to the bivariate case of the formalism presented in [25] (with Granger causality replaced by transfer entropy), it is included to keep the paper self-contained. This is also a useful preparation for the calculations of Sec. III B.

The starting point is Eq. (4) with $i = 1$ and $j = 2$, which requires one to compute $\sigma_{22}(h)$ and $\sigma'_{22}(h)$ and thus the associated MMSEs. The essential ingredient for computing these quantities is to have a one-to-one correspondence between the stationary process or subprocess under consideration and the corresponding forcing white noise(s). In other words, the process or subprocess must be invertible (or *minimum phase* in the language of control theory [4,5]): Fixing the trajectory of the process or subprocess up to time t must be equivalent to fixing the trajectory of the noise(s) and vice versa.

When the conditioning involves the past of the joint process \mathbf{X} [represented by either Eq. (10) or (11), which are the continuous-time analogs of the vector autoregressive and moving average representations [40]], the calculation of the

MMSE is immediate. Starting from

$$X_2(t+h) = \int_{-\infty}^{t+h} ds [H_{21}(t+h-s)\xi_1(s) + H_{22}(t+h-s)\xi_2(s)], \quad (15)$$

we readily obtain

$$\langle X_2(t+h) | \mathbf{X}^-(t) \rangle = \int_{-\infty}^t ds [H_{21}(t+h-s)\xi_1(s) + H_{22}(t+h-s)\xi_2(s)], \quad (16)$$

since the noises are fixed for $s \leq t$ by Eq. (10) and average to zero in the time interval $[t, t+h]$. Equation (5) then yields

$$\sigma_{22}(h) = 2 \int_0^h dt [D_{11}H_{21}^2(t) + D_{22}H_{22}^2(t) + 2D_{12}H_{22}(t)H_{21}(t)]. \quad (17)$$

The calculation of $\langle X_2(t+h) | X_2^-(t) \rangle$ is less straightforward because fixing the marginal process X_2 alone does not fix the noises ξ_1 and ξ_2 . Instead, one must find a coarse-grained representation of X_2 similar to Eq. (10),

$$\dot{X}_2(t) = - \int_0^\infty ds A'_{22}(s)X_2(t-s) + \xi'_2(t), \quad (18)$$

where $A'_{22}(s)$ is a kernel to be determined and $\xi'_2(t)$ is a Gaussian white noise, for instance, with the same variance $2D_{22}$ as $\xi_2(t)$. Then, starting from the equation

$$X_2(t) = \int_{-\infty}^t ds H'_{22}(t-s)\xi'_2(s), \quad (19)$$

where $H'_{22}(t)$ is the inverse of $A'_{22}(t)$ [see below Eq. (27)], and using the same reasoning as above, we obtain

$$\langle X_2(t+h) | X_2^-(t) \rangle = \int_{-\infty}^t ds H'_{22}(t+h-s)\xi'_2(s) \quad (20)$$

and in turn

$$\sigma'_{22}(h) = 2D_{22} \int_0^h dt H_{22}^2(t). \quad (21)$$

The response function $H'_{22}(t)$ must be causal and is easily found by going to Fourier space. Indeed, Eq. (19) implies that the power spectral density (PSD) $S_{22}(\omega) = \langle X_2(\omega)X_2(-\omega) \rangle$ is given by

$$S_{22}(\omega) = 2D_{22}|H_{22}(\omega)|^2. \quad (22)$$

On the other hand, Eq. (14) tells us that

$$S_{22}(\omega) = 2D_{22}|H_{22}(\omega)|^2 + 2D_{11}|H_{21}(\omega)|^2 + 2D_{12}[H_{22}(\omega)H_{21}(-\omega) + H_{22}(-\omega)H_{21}(\omega)], \quad (23)$$

which is conveniently rewritten as

$$S_{22}(\omega) = 2D_{22}|H_{22}(\omega)|^2 \frac{\omega^2 + r_2^2}{\omega^2 + a_{11}^2}, \quad (24)$$

where

$$r_2 = \sqrt{a_{11}^2 + \frac{D_{11}}{D_{22}}a_{21}^2 - 2\frac{D_{12}}{D_{22}}a_{11}a_{21}}. \quad (25)$$

Since $H_{22}(\omega)$ is the Fourier transform of a causal function, the Wiener-Hopf factorization of $S_{22}(\omega)$ is simple and gives

$$H'_{22}(\omega) = H_{22}(\omega) \frac{r_2 - i\omega}{a_{11} - i\omega} = \frac{r_2 - i\omega}{(a_{11} - i\omega)(a_{22} - i\omega) - a_{12}a_{21}e^{i\omega\tau}}. \quad (26)$$

(In turn, one can readily check that the noise defined by Eq. (19) and given in Fourier space by $\xi'(\omega) = [-a_{21}\xi_1(\omega) + (a_{11} - i\omega)\xi_2(\omega)]/(r_2 - i\omega)$ is indeed white.) By construction, $H'_{22}(\omega)$ has no poles in the upper half of the complex plane and since we have chosen $r_2 > 0$ in Eq. (25), it is also zero-free in this region. The minimum-phase condition, a prerequisite for Eq. (20), is thus satisfied.

For a given choice of the model parameters, the response functions $H_{ij}(t)$ and $H'_{22}(t)$ can be computed numerically by taking the corresponding inverse Fourier transforms, and $\sigma_{22}(h)$ and $\sigma'_{22}(h)$ are then obtained from Eqs. (17) and (21). For brevity, we do not present a numerical study here. On the other hand, it is instructive to look at the explicit representation of the marginal process X_2 provided by Eq. (18). By construction, the Fourier transform of the kernel $A'_{22}(t)$ is obtained as

$$A'_{22}(\omega) \equiv \frac{1}{H'_{22}(\omega)} + i\omega, \quad (27)$$

which yields

$$A'_{22}(\omega) = \frac{(a_{11} - i\omega)(a_{22} - i\omega)}{r_2 - i\omega} - a_{12}a_{21} \frac{e^{i\omega\tau}}{r_2 - i\omega} = a_{11} + a_{22} - r_2 + \frac{(r_2 - a_{11})(r_2 - a_{22})}{r_2 - i\omega} - a_{12}a_{21} \frac{e^{i\omega\tau}}{r_2 - i\omega}. \quad (28)$$

As a result,

$$A'_{22}(t) = (a_{11} + a_{22} - r_2)\delta(t) + (r_2 - a_{11})(r_2 - a_{22}) \times e^{-r_2 t} \Theta(t) - a_{12}a_{21}e^{-r_2(t-\tau)} \Theta(t-\tau) \quad (29)$$

and Eq. (18) reads

$$\dot{X}_2(t) = -(a_{11} + a_{22} - r_2)X_2(t) - (r_2 - a_{11})(r_2 - a_{22}) \times \int_{-\infty}^t ds e^{-r_2(t-s)} X_2(s) + a_{12}a_{21} \int_{-\infty}^t ds e^{-r_2(t-s)} X_2(s-\tau) + \xi'_2(t). \quad (30)$$

Finally, by splitting the integrals into two parts $\int_{-\infty}^t ds = \int_{-\infty}^{t-\tau} ds + \int_{t-\tau}^t ds$ and performing some simple manipulations, we can transform the equation into

$$\dot{X}_2(t) = -(a_{11} + a_{22} - r_2)X_2(t) - [(r_2 - a_{11})(r_2 - a_{22}) - a_{12}a_{21}e^{r_2\tau}] \int_{-\infty}^{t-\tau} ds e^{-r_2(t-s)} X_2(s) - (r_2 - a_{11})(r_2 - a_{22}) \int_{t-\tau}^t ds e^{-r_2(t-s)} X_2(s) + \xi'_2(t). \quad (31)$$

This is an interesting representation of the coarse-grained dynamics of X_2 because it shows that a significant simplification occurs if the delay τ satisfies the condition

$$a_{12}a_{21}e^{r_2\tau} = (r_2 - a_{11})(r_2 - a_{22}). \quad (32)$$

The second term on the right-hand side of Eq. (31) then vanishes, and although the dynamics is still non-Markovian,

the dependence on the past is now limited to a finite-time interval of duration τ .

2. TE rate

By definition, the TE rate $\mathcal{T}_{1 \rightarrow 2}$ is the slope of $T_{1 \rightarrow 2}(h)$ at $h = 0^+$. After expanding $\sigma_{22}(h)$ and $\sigma'_{22}(h)$ in powers of h and using $H_{21}(0^+) = 0$ and $H_{22}(0^+) = H'_{22}(0^+) = 1$, we obtain

$$T_{1 \rightarrow 2}(h) = \frac{1}{2} \ln \frac{2D_{22}[h + \dot{H}'_{22}(0^+)h^2 + O(h^3)]}{2D_{22}h + 2[D_{22}\dot{H}_{22}(0^+) + D_{12}\dot{H}_{21}(0^+)]h^2 + O(h^3)} \quad (33)$$

and then

$$\mathcal{T}_{1 \rightarrow 2} = \frac{1}{2} \left[\dot{H}'_{22}(0^+) - \dot{H}_{22}(0^+) - \frac{D_{12}}{D_{22}} \dot{H}_{21}(0^+) \right], \quad (34)$$

which is the two-dimensional version of Eq. (61) in [25] (with the usual multiplicative factor $\frac{1}{2}$ coming from the replacement of GC by the corresponding TE). In order to obtain the explicit expressions of $\dot{H}_{21}(0^+)$ and $\dot{H}_{22}(0^+)$, we then use the equation

$$\dot{\mathbf{H}}(t) = - \int_0^t ds \mathbf{A}(s) \mathbf{H}(t-s), \quad t \geq 0, \quad (35)$$

which is obtained by differentiating Eq. (11) with respect to t and identifying with Eq. (8) (see, e.g., Appendix F in [25]). Specifically,

$$\dot{\mathbf{H}}(t) = - \begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix} \mathbf{H}(t) - \begin{pmatrix} 0 & a_{12} \\ 0 & 0 \end{pmatrix} \mathbf{H}(t-\tau) \Theta(t-\tau). \quad (36)$$

Together with the condition $\mathbf{H}(0^+) = \mathbf{I}$, where \mathbf{I} is the unity matrix, this readily yields $\dot{H}_{21}(0) = -a_{21}$ and $\dot{H}_{22}(0) = -a_{22}$. Likewise, $\dot{H}'_{22}(0^+)$ is obtained from the equation

$$\dot{H}'_{22}(t) = - \int_0^t ds A'_{22}(s) H'_{22}(t-s), \quad (37)$$

with $A'_{22}(t)$ given by Eq. (29). Explicitly,

$$\begin{aligned} \dot{H}'_{22}(t) &= -(a_{11} + a_{22} - r_2) H'_{22}(t) - (r_2 - a_{11})(r_2 - a_{22}) \\ &\times \int_0^t ds e^{-r_2(t-s)} H'_{22}(s) \\ &+ a_{12}a_{21}e^{r_2\tau} \Theta(t-\tau) \int_0^{t-\tau} ds e^{-r_2(t-s)} H'_{22}(s), \end{aligned} \quad (38)$$

from which we find that $\dot{H}'_{22}(0^+) = r_2 - a_{11} - a_{22}$. Inserting these expressions of $\dot{H}_{21}(0^+)$, $\dot{H}_{22}(0^+)$, and $\dot{H}'_{22}(0^+)$ into Eq. (34), we finally obtain

$$\mathcal{T}_{1 \rightarrow 2} = \frac{1}{2} \left[r_2 - a_{11} + \frac{D_{12}}{D_{22}} a_{21} \right]. \quad (39)$$

Therefore, the TE rate in the direction $1 \rightarrow 2$ does not depend on τ , a result that was not obvious from the outset because of the bidirectional character of the coupling between the two subprocesses. As a matter of fact, $\overline{\mathcal{T}}_{1 \rightarrow 2}$, the simplified version of the TE rate that only takes into account the information provided by the states at time t and whose expression is given by Eq. (A14) in Appendix A, does depend on τ .

3. Spectral expression of the TE rate

As originally introduced in the context of VAR processes [47], there is a spectral version of GC that is used, especially in neuroscience [48], to analyze causal relationships in the frequency domain. The continuous-time version is briefly presented in [25], but the conditions for the validity of this spectral representation are not discussed. This will play an important role in Sec. III B, and for completeness we revisit the derivation, focusing again on TE instead of GC.

It is instructive to first consider the case $D_{12} = 0$ (i.e., the joint process \mathbf{X} is bipartite). The PSD $S_{22}(\omega)$ then reduces to two terms

$$S_{22}(\omega) = 2D_{22}|H_{22}(\omega)|^2 + 2D_{11}|H_{21}(\omega)|^2. \quad (40)$$

The first one can be viewed as the intrinsic contribution of the subprocess X_2 to its (auto)spectrum, whereas the second one can be viewed as the causal part due to X_1 . Following [47], this suggests that we adopt the quantity

$$t_{1 \rightarrow 2}(\omega) \equiv \frac{1}{2} \ln \frac{S_{22}(\omega)}{2D_{22}|H_{22}(\omega)|^2} \quad (41)$$

as a measure of the transfer entropy from X_1 to X_2 in the frequency domain [49]. However, two requirements must be fulfilled: (i) $t_{1 \rightarrow 2}(\omega)$ must be a non-negative quantity and (ii) the TE rate in the time domain must be the average of the spectral TE over all frequencies, i.e.,

$$\mathcal{T}_{1 \rightarrow 2} = \frac{1}{2} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \ln \frac{S_{22}(\omega)}{2D_{22}|H_{22}(\omega)|^2}. \quad (42)$$

The first condition is obviously fulfilled, and to check the second one we replace $H_{22}(\omega)$ and $S_{22}(\omega)$ by their expressions (13) and (24), respectively, and integrate over ω . This gives

$$\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} t_{1 \rightarrow 2}(\omega) = \frac{1}{2} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \ln \frac{r_2^2 + \omega^2}{a_{11}^2 + \omega^2} = \frac{1}{2} (r_2 - |a_{11}|), \quad (43)$$

which indeed coincides with Eq. (39) when $D_{12} = 0$, but under the condition that $a_{11} > 0$.²

²An expression similar to Eq. (42) is derived in [34] in the simpler case of a linear feedback system with no time delay. The spectral expression is obtained by discretizing the time evolution and integrating Eq. (2) over time so as to express the TE rate as a log-ratio of Gaussian path probabilities. This leads to the correct result for the TE

Following again [47] and the literature on GC [50,51], Eq. (42) can be generalized to the case of correlated noises ($D_{12} \neq 0$). This is done by performing a linear transformation $\tilde{\xi}(t) = \mathbf{U}\xi(t)$ that makes the covariance matrix of the transformed noises diagonal. Specifically, by choosing

$$\mathbf{U} = \begin{pmatrix} 1 & -\frac{D_{12}}{D_{22}} \\ 0 & 1 \end{pmatrix}, \quad (44)$$

$$\tilde{\mathbf{H}}(\omega) = \mathbf{U}\mathbf{H}(\omega)\mathbf{U}^{-1} = \begin{pmatrix} H_{11} - \frac{D_{12}}{D_{22}}H_{21} & H_{12} + \frac{D_{12}}{D_{22}}(H_{11} - H_{22}) - (\frac{D_{12}}{D_{22}})^2H_{21} \\ H_{21} & H_{22} + \frac{D_{12}}{D_{22}}H_{21} \end{pmatrix} \quad (46)$$

and

$$\tilde{\mathbf{S}}(\omega) = \mathbf{U}\mathbf{S}(\omega)\mathbf{U}^T = \begin{pmatrix} S_{11} - 2\frac{D_{12}}{D_{22}}S_{12} + (\frac{D_{12}}{D_{22}})^2S_{22} & S_{12} - \frac{D_{12}}{D_{22}}S_{22} \\ S_{12} - \frac{D_{12}}{D_{22}}S_{22} & S_{22} \end{pmatrix}, \quad (47)$$

where the dependence of the functions H_{ij} and S_{ij} on ω is dropped for brevity. The crucial feature is that the TE rate $\mathcal{T}_{1 \rightarrow 2}$ is invariant under the linear transformation defined by the matrix \mathbf{U} . Indeed, since $\tilde{D}_{12} = 0$, we have from Eq. (34)

$$\begin{aligned} \tilde{\mathcal{T}}_{1 \rightarrow 2} &= \frac{1}{2}[\dot{\tilde{H}}_2'(0^+) - \dot{\tilde{H}}_{22}(0^+)] \\ &= \frac{1}{2}\left[\dot{H}'_{22}(0^+) - \dot{H}_{22}(0^+) - \frac{D_{12}}{D_{22}}\dot{H}_{21}(0^+)\right] \\ &= \mathcal{T}_{1 \rightarrow 2}, \end{aligned} \quad (48)$$

where we have used the fact that $\tilde{S}_{22}(\omega) = S_{22}(\omega)$ implies $\dot{\tilde{H}}_2'(\omega) = \dot{H}'_{22}(\omega)$. Accordingly, by applying the spectral decomposition (42) to the transformed variables \tilde{X}_1 and \tilde{X}_2 and going back to the original variables, we obtain

$$\begin{aligned} \mathcal{T}_{1 \rightarrow 2} &= \frac{1}{2} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \ln \frac{\tilde{S}_{22}(\omega)}{2\tilde{D}_{22}|\tilde{H}_{22}(\omega)|^2} \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \ln \frac{S_{22}(\omega)}{2D_{22}|H_{22}(\omega) + (D_{12}/D_{22})H_{21}(\omega)|^2}. \end{aligned} \quad (49)$$

This expression (with the labels 1 and 2 exchanged) will play an important role in Sec. III B as it will give a closed-form expression of the TE rate $\mathcal{T}_{2 \rightarrow 1}$. However, there is a serious caveat. Replacing $H_{22}(\omega)$, $H_{21}(\omega)$, and $S_{22}(\omega)$ by their expressions and integrating over ω , we find that Eq. (39) is recovered *only if the condition* $\tilde{a}_{11} = a_{11} - (D_{12}/D_{22})a_{21} > 0$ *is satisfied*. This of course generalizes the condition $a_{11} > 0$ that was found in the case $D_{12} = 0$. Otherwise, Eq. (49) underestimates the actual value of the TE rate in the time domain, as was already pointed out in [47] for the discrete-time GC (see also footnote 6 in [25]).

rate because the condition $a_{11} > 0$ is naturally satisfied in the model. The spectral expression also includes an additional term coming from the so-called sensitivity function of the feedback system, but this term does not contribute to the TE rate after integration over frequency due to Bode's integral formula [4].

we get

$$2\tilde{\mathbf{D}} = 2\mathbf{U}\mathbf{D}\mathbf{U}^T = 2 \begin{pmatrix} D_{11} - \frac{D_{12}^2}{D_{22}} & 0 \\ 0 & D_{22} \end{pmatrix}, \quad (45)$$

and the dynamics of the transformed vector $\tilde{\mathbf{X}}(t) = \mathbf{U}\mathbf{x}(t)$ is now governed by the equation $\dot{\tilde{\mathbf{X}}}(t) = -\int_0^\infty \tilde{\mathbf{A}}(s)\tilde{\mathbf{X}}(t-s) + \tilde{\xi}(t)$ with $\tilde{\mathbf{A}} = \mathbf{U}\mathbf{A}\mathbf{U}^{-1}$. Likewise,

What is the rationale for the condition $\tilde{a}_{11} > 0$? Since $\tilde{H}_{22}(\omega) = (\tilde{a}_{11} - i\omega)/[(\tilde{a}_{11} - i\omega)(\tilde{a}_{22} - i\omega) - \tilde{a}_{12}\tilde{a}_{21}e^{i\omega\tau}]$ [cf. Eq. (13) with a_{ij} replaced by \tilde{a}_{ij}], this condition guarantees that $\tilde{H}_{22}(\omega)$ has no zeros in the upper half of the complex ω plane [it has no poles in this region since $\tilde{H}_{22}(t)$ is causal]. To summarize, the condition for the spectral expression to be valid is that the stationary process $\tilde{X}_2(\omega) = \tilde{H}_{22}(\omega)\tilde{\xi}_2(\omega)$ is minimum phase. There is no reason for this condition to be always satisfied for a time-delayed process governed by Eq. (10), not to mention the more general Eq. (8), and it must be carefully checked on a case-by-case basis.

B. Transfer entropy in the direction $2 \rightarrow 1$

We now turn to the calculation of the TE in the direction $2 \rightarrow 1$ and to simplify the forthcoming analysis we set the parameter a_{21} to zero from the outset. This makes the coupling unidirectional, and $X_2(t)$ becomes a simple Ornstein-Uhlenbeck process that drives $X_1(t)$ at a fixed delay τ . This is the model introduced in Sec. 4 of Ref. [25], which is regarded as the minimal continuous-time version of a VAR process. Interestingly, this also corresponds to the model of a cellular signaling pathway considered in [52], in which $X_1(t)$ and $X_2(t)$ represent the deviations from the mean of active kinase populations (these quantities can be treated as continuous variables by assuming a chemical Langevin description).³ Of course, the fact that X_2 is now an autonomous process implies that $\mathcal{T}_{1 \rightarrow 2}(h)$ and thus $\mathcal{T}_{1 \rightarrow 2}$ vanish identically. Moreover, the stationary state is stable for all values of τ .

Following [25], we set $a_{11} = a > 0$, $a_{22} = b > 0$, and $a_{12} = -c$ and we assume that the two noises ξ_1 and ξ_2 have the same variance $2D_{11} = 2D_{22} = 1$ to further restrict the parameter space. The parameter $\rho = 2D_{12}$ (with $-1 \leq \rho \leq 1$)

³The continuous population approximation is generally valid in signaling cascades where molecular populations are large. On the other hand, the linearization of the Langevin equations is an approximation whose validity depends on the details of the underlying biochemical reaction network and which must be checked on a case-by-case basis.

quantifies the correlation between the noises. The response functions for $t \geq 0$ now have very simple expressions in the time domain,

$$\begin{aligned} H_{11}(t) &= e^{-at}, \\ H_{12}(t) &= c \frac{e^{-a(t-\tau)} - e^{-b(t-\tau)}}{b-a} \Theta(t-\tau), \\ H_{21}(t) &= 0, \\ H_{22}(t) &= e^{-bt}, \end{aligned} \tag{50}$$

and $T_{2 \rightarrow 1}(h)$ is obtained from Eq. (4) with $i = 2$ and $j = 1$, where $\sigma_{11}(h)$ and $\sigma'_{11}(h)$ are the variances of $p[X_1(t+h)|\mathbf{X}^-(t)]$ and $p[X_1(t+h)|X_1^-(t)]$, respectively. Likewise, Eq. (34) is replaced by

$$\begin{aligned} \mathcal{T}_{2 \rightarrow 1} &= \lim_{h \rightarrow 0^+} \frac{1}{h} T_{2 \rightarrow 1}(h) \\ &= \frac{1}{2} \left[\dot{H}'_{11}(0^+) - \dot{H}_{11}(0^+) - \frac{D_{12}}{D_{11}} \dot{H}_{12}(0^+) \right] \\ &= \frac{1}{2} [\dot{H}'_{11}(0^+) + a], \end{aligned} \tag{51}$$

as $\dot{H}_{11}(0^+) = -a$ and $\dot{H}_{12}(0^+) = 0$ from Eq. (36).

1. Wiener-Hopf factorization

It should be clear from the preceding section that the main task is to compute the response function $H'_{11}(t)$. This requires the factorization of the PSD $S_{11}(\omega) = \langle X_1(\omega)X_1(-\omega) \rangle$, which reads

$$S_{11}(\omega) = \frac{\omega^2 + b^2 + c^2 + 2\rho cv(\omega, b, \tau)}{(a^2 + \omega^2)(b^2 + \omega^2)}, \tag{52}$$

where $v(\omega, b, \tau) \equiv b \cos \omega\tau - \omega \sin \omega\tau$ (for comparison we use the same notation as in [25]). This turns out to be a nontrivial operation. In [25] it is claimed that the causal factor $H'_{11}(\omega)$ satisfying $S_{11}(\omega) = |H'_{11}(\omega)|^2$ is given by

$$\begin{aligned} H'_{11}(\omega) &= \frac{\sqrt{(1-\rho^2)c^2 + (b + \rho c \cos \omega\tau)^2 - i(\omega - \rho c \sin \omega\tau)}}{(a - i\omega)(b - i\omega)}. \end{aligned} \tag{53}$$

However, this statement is wrong when $\rho \neq 0$ because the inverse Fourier transform of this function is not causal. This can be readily seen by setting $\omega = x + iy$ and considering the limit $y \rightarrow +\infty$, which yields

$$H'_{11}(\omega) \sim \frac{\rho c}{2y^2} (\sqrt{e^{-2i\tau x}} - e^{-i\tau x}) e^{\tau y} + O(1/y). \tag{54}$$

Hence, $H'_{11}(\omega)$ diverges like $-\rho c y^{-2} e^{-i\tau x} e^{\tau y}$ if $\cos \tau x < 0$ and the condition for applying Jordan's lemma is not satisfied. Accordingly, the inverse Fourier transform does not vanish for $t < 0$, as can be checked numerically. Moreover, $H'_1(t = 0^+)$ is not equal to 1, contrary to what it should be (see Fig. 3 below), which implies that $\sigma'_1(h) = 2D_{11} \int_0^h dt H'_1(t)^2 \neq 2D_{11}h[1 + O(h)]$ so that the formula $\mathcal{T}_{2 \rightarrow 1} = \lim_{h \rightarrow 0^+} (2h)^{-1} \ln[\sigma'_{11}(h)/\sigma_{11}(h)]$ gives an infinite result. This is of course a serious shortcoming.

Before presenting our solution to the factorization problem, let us explain why this operation is nontrivial, even from the numerical point of view. First, one could try to apply the standard Wiener-Hopf method [53] and transform the multiplicative factorization problem into an additive one by taking the logarithm of $S_{11}(\omega)$. In order to have a function that goes to 1 as $|\omega| \rightarrow \infty$, one may consider the ratio $K(\omega) = S_{11}(\omega)/S_{11}(\omega, \rho = 0) = [\omega^2 + b^2 + c^2 + 2\rho cv(\omega, b, \tau)]/[\omega^2 + b^2 + c^2]$, and $H'_{11}(\omega)$ is then obtained as

$$H'_{11}(\omega) = \frac{\sqrt{b^2 + c^2} - i\omega}{(a - i\omega)(b - i\omega)} K_+(\omega), \tag{55}$$

where $K_+(\omega)$ is the causal factor of $K(\omega)$ given by

$$K_+(\omega) = \exp \left[\frac{1}{2i\pi} \int_{i\delta - \infty}^{i\delta + \infty} d\zeta \frac{\ln K(\zeta)}{\zeta - \omega} \right]. \tag{56}$$

In this formula, ω must lie above δ and the integration path must belong to a finite-width strip \mathcal{D} around the real axis where $K(\omega)$ is analytic and free of zeros. The problem with this procedure is that the numerator of $K(\omega)$ [i.e., the function $\omega^2 + b^2 + c^2 + 2\rho cv(\omega, b, \tau)$] has infinitely many zeros in the complex ω plane when $\rho \neq 0$. Since there does not seem to be any simple and systematic way of computing these zeros for arbitrary values of the parameters, determining the zero-free strip \mathcal{D} is a daunting task.

Alternatively, one could try to solve the problem directly in the time domain. Recall that in order to compute $T_{2 \rightarrow 1}(h)$, we need to calculate the MMSE estimate $\langle X_1(t+h)|X_1^-(t) \rangle$, which is the orthogonal projection of $X_1(t+h)$ onto the trajectory $X_1^-(t)$. It thus satisfies the equation

$$\langle [X_1(t+h) - \langle X_1(t+h)|X_1^-(t) \rangle] X_1(s) \rangle = 0 \quad \forall s \leq t \tag{57}$$

and is a linear functional of $X_1^-(t)$,

$$\langle X_1(t+h)|X_1^-(t) \rangle = \int_{-\infty}^t ds f_h(t-s) X_1(s), \tag{58}$$

where $f_h(t)$ is an unknown function to be determined from Eq. (57). [To be precise, the kernel $f_h(t)$ must also include a term proportional to the Dirac distribution $\delta(t)$ which singles out the dependence on $X_1(t)$.] Inserting Eq. (58) into Eq. (57) and changing variables yields the Wiener-Hopf integral equation

$$\phi_{11}(t+h) = \int_0^\infty ds \phi_{11}(t-s) f_h(s) \quad \forall t \geq 0, \tag{59}$$

where $\phi_{11}(t) \equiv \langle X_1(0)X_1(t) \rangle$ is the inverse Fourier transform of $S_{11}(\omega)$. Since $\phi_{11}(t)$ is a combination of exponentials [see Eqs. (78) and (79) in [25], where $\phi_{11}(t)$ is denoted by $\Gamma_{xx}(t)$], one could hope to find some systematic procedure to solve Eq. (59) and determine $f_h(t)$, at least numerically. However, this goal cannot be achieved because $\phi_{11}(t)$ has different expressions for $t < \tau$ and $t > \tau$: $f_h(t)$ is then an infinite sum of functions defined in the successive intervals $[0, \tau]$, $[\tau, 2\tau]$, $[2\tau, 3\tau]$, ..., with the function in the interval $n\tau \leq t \leq (n+1)\tau$ depending on the function in the next interval. Therefore, a step-by-step solution of Eq. (59) is impossible.

Our solution to the factorization problem consists in replacing the delay term $e^{i\omega\tau}$ in the frequency domain by an

all-pass (i.e., with unit amplitude) rational function of the form $\mathcal{Q}_n(-\omega)/\mathcal{Q}_n(\omega)$, where $\mathcal{Q}_n(\omega)$ is a polynomial with no zeros in the upper half of the complex ω plane. This is a classic procedure in the field of control systems [1], and several choices of \mathcal{Q}_n are possible, in particular Padé approximants (which are also often used to approximate Wiener-Hopf kernels [54]). After various trials, we have found that the simplest and yet effective approximation for the problem at hand is the so-called Laguerre shift formula

$$e^{i\omega\tau} \approx \frac{\left(1 + \frac{i\omega\tau}{2n}\right)^n}{\left(1 - \frac{i\omega\tau}{2n}\right)^n}, \quad (60)$$

which introduces a single pole of multiplicity n at $\omega = -i2n/\tau$. The convergence rate of this approximation for $n \rightarrow \infty$ has been studied in detail in the literature [55,56] and the formula is successfully used in robust control as it is easy to implement with analog filters. Accordingly, the expression (52) of $S_{11}(\omega)$ is now replaced by

$$S_{11,n}(\omega) = \frac{\mathcal{P}_n(\omega)}{(a^2 + \omega^2)(b^2 + \omega^2)\left(1 + \frac{\omega^2\tau^2}{4n^2}\right)^n}, \quad (61)$$

where

$$\begin{aligned} \mathcal{P}_n(\omega) &= (\omega^2 + b^2 + c^2) \left(1 + \frac{\omega^2\tau^2}{4n^2}\right)^n \\ &+ \rho c \left[\left(1 + i\frac{\omega\tau}{2n}\right)^{2n} (b + i\omega) \right. \\ &\left. + \left(1 - i\frac{\omega\tau}{2n}\right)^{2n} (b - i\omega) \right] \end{aligned} \quad (62)$$

is an even polynomial of order $2n + 2$. The factorization problem now boils down to finding all the roots of a polynomial, a standard numerical task. Since $\mathcal{P}_n(\omega)$ has no real roots,⁴ Eq. (61) can be rewritten as

$$S_{11,n}(\omega) = \frac{\prod_{k=1}^{n+1} (\omega - \omega_k)(\omega - \omega_k^*)}{(a^2 + \omega^2)(b^2 + \omega^2)(4n^2/\tau^2 + \omega^2)^n}, \quad (63)$$

where ω_k denotes a root with a negative imaginary part. The causal factor $H'_{11,n}(\omega)$ is then readily obtained as

$$H'_{11,n}(\omega) = \frac{i \prod_{k=1}^{n+1} (\omega - \omega_k)}{(\omega + ia)(\omega + ib)(\omega + 2in/\tau)^n}, \quad (64)$$

where the factor i is included in order that $H'_1(t=0^+) = \lim_{\omega \rightarrow \infty} -i\omega H'_{11}(\omega) = 1$, as it must be. From this we can compute numerically the inverse Fourier transform $H'_{11,n}(t)$ and then

$$T_{2 \rightarrow 1}^{(n)}(h) \equiv \frac{1}{2} \ln \frac{\sigma'_{11,n}(h)}{\sigma_{11,n}(h)}, \quad (65)$$

⁴The polynomial $(1 + \omega^2\tau^2/4n^2)^n \mathcal{P}_n(\omega)$ can be rewritten as $(AB)^n \mathcal{P}_n(\omega) = [\omega(AB)^n + i(\rho c/2)(A^{2n} - B^{2n})]^2 + [b(AB)^n + (\rho c/2)(A^{2n} + B^{2n})]^2 + (1 - \rho^2)c^2(AB)^{2n}$, where $A = 1 + i\omega\tau/2n$ and $B = A^* = 1 - i\omega\tau/2n$. If ω is real, $A^{2n} + B^{2n}$ and $A^2 - B^{2n}$ are real and imaginary quantities, respectively, and $(AA^*)^n \mathcal{P}_n(\omega)$ is then a sum of three positive terms. Therefore, $\mathcal{P}_n(\omega)$ has no real roots.

where

$$\sigma'_{11,n}(h) = \int_0^h dt H_{11,n}^{\prime 2}(t), \quad (66)$$

$$\sigma_{11,n}(h) = \int_0^h [H_{11}^2(t) + H_{12,n}^2(t) + 2\rho H_{11}(t)H_{12,n}(t)], \quad (67)$$

with $H_{12,n}(t)$ the inverse Fourier transform of $H_{12,n}(\omega) = (1 + \frac{i\omega\tau}{2n})^n / [(1 - \frac{i\omega\tau}{2n})^n (a - i\omega)(b - i\omega)]$.

Remarkably, this procedure leads to an explicit and concise expression of the TE rate $\mathcal{T}_{2 \rightarrow 1}^{(n)}$ in terms of the roots ω_k . Similarly to Eq. (51), one has

$$\begin{aligned} \mathcal{T}_{2 \rightarrow 1}^{(n)} &= \lim_{h \rightarrow 0^+} \frac{1}{h} T_{2 \rightarrow 1}^{(n)}(h) \\ &= \frac{1}{2} \left[\dot{H}'_{11,n}(0^+) - \dot{H}_{11}(0^+) - \frac{D_{12}}{D_{11}} \dot{H}_{12,n}(0^+) \right], \end{aligned} \quad (68)$$

where $H'_{11,n}(t)$ for $t \geq 0$ is obtained by using Cauchy's residue theorem as

$$H'_{11,n}(t) = A_n e^{-at} + B_n e^{-bt} + Q_n(t) e^{-2nt/\tau}, \quad (69)$$

with $A_n = \prod_{k=1}^{n+1} (a - i\omega_k)/(a - b)(a - 2n/\tau)^n$, $B_n = \prod_{k=1}^{n+1} (b - i\omega_k)/(b - a)(b - 2n/\tau)^n$, and $Q_n(t) = q_n^{(0)} + q_n^{(1)}t + \dots + q_n^{(n-1)}t^{n-1}$. Moreover, it can be shown that $\dot{H}'_{12,n}(0^+) = (-1)^n c$. As a result,

$$\begin{aligned} \mathcal{T}_{2 \rightarrow 1}^{(n)} &= \frac{1}{2} \left[-aA_n - bB_n + \left(q_n^{(1)} - \frac{2n}{\tau} q_n^{(0)} \right) \right. \\ &\left. + a + (-1)^{n+1} \rho c \right]. \end{aligned} \quad (70)$$

Using $q_n^{(0)} = 1 - A_n - B_n$ [as $H'_{11,n}(0^+) = 1$] and $q_n^{(1)} = -\lim_{\omega \rightarrow \infty} \omega^2 [H'_{11,n}(\omega) - A_n/(a - i\omega) - B_n/(b - i\omega) - q_n^{(0)}/(2n/\tau - i\omega)]$, we find that the coefficients of A_n and B_n cancel, and after some algebra we finally obtain

$$\begin{aligned} \mathcal{T}_{2 \rightarrow 1}^{(n)} &= \frac{1}{2} \left[a + (-1)^{n+1} \rho c - \frac{2n}{\tau} \right. \\ &\left. - \lim_{\omega \rightarrow \infty} \omega^2 \left(H'_{11,n}(\omega) - \frac{1}{2n/\tau - i\omega} \right) \right] \\ &= \frac{1}{2} \left[-b + (-1)^{n+1} \rho c + i \sum_{k=1}^{n+1} \omega_k - \frac{2n^2}{\tau} \right], \end{aligned} \quad (71)$$

which is the explicit expression announced above. Note that $\mathcal{T}_{2 \rightarrow 1}^{(n)}$ and thus $\mathcal{T}_{2 \rightarrow 1} = \lim_{n \rightarrow \infty} \mathcal{T}_{2 \rightarrow 1}^{(n)}$ are invariant in the change $(\rho, c) \rightarrow (-\rho, -c)$ and do not depend on a^{-1} , the intrinsic relaxation time of the process X_1 . This is not the case for $\bar{\mathcal{T}}_{2 \rightarrow 1}$, whose expression is given by Eq. (A16) in Appendix A.

Of course, this is still formal and a problem of practicality remains: How large must n be to provide an accurate estimation of $\mathcal{T}_{2 \rightarrow 1}$ and, more generally, of $T_{2 \rightarrow 1}(h)$? Although we have no rigorous mathematical answer to this

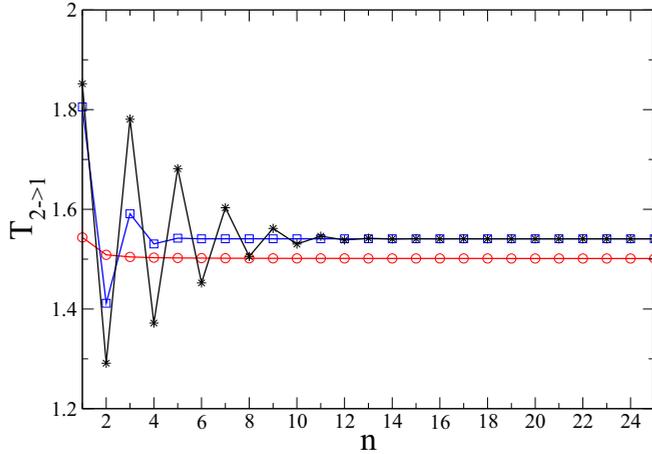


FIG. 1. Evolution of $\mathcal{T}_{2 \rightarrow 1}^{(n)}$ computed from Eq. (71) as a function of n for $\rho = 0.2$, $c = 4$, and three values of the delay: $\tau = 0.5$ (red circles), $\tau = 5$ (blue squares), $\tau = 30$ (black stars) (b^{-1} is taken as the time unit). The exact asymptotic results given by Eq. (73) are 1.501 for $\tau = 0.5$ and 1.541 for $\tau = 5$ and 30.

question,⁵ numerical calculations show that the convergence to the asymptotic limit is quite fast (see Figs. 1 and 3 below). As shown in Appendix B, this is not the case when the delay kernel $\delta(t - \tau)$ is approximated in the time domain by a sequence of gamma distributions [42–44]. In addition, we have another way to assess the accuracy of Eq. (71), which is to compare with the predictions of the spectral formula

⁵The evolution of the roots of the polynomial $\mathcal{P}_n(\omega)$ with n is indeed quite complicated. In particular, some of the roots are spurious in the sense that they are not zeros of the function $\omega^2 + b^2 + c^2 + 2\rho cv(\omega, b, \tau)$ [i.e., the actual zeros of $S_{11}(\omega)$] in the limit $n \rightarrow \infty$. However, they must be included in the calculation to ensure that $i \sum_{k=1}^{n+1} \omega_k - 2n^2/\tau$ goes asymptotically to a finite value.

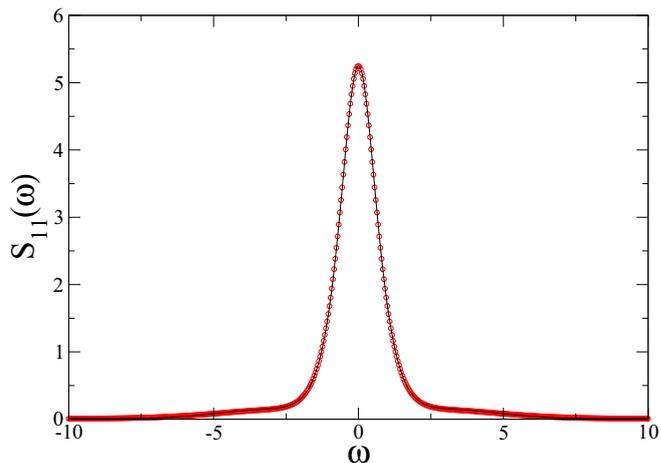


FIG. 2. Comparison between the exact PSD $S_{11}(\omega)$ given by Eq. (52) (solid line) and the approximate PSD $S_{11}^{(n)}(\omega) = |H'_{1,n}|^2$ with $n = 4$ (red circles) for $\rho = 0.5$ and $\tau = 1$. The model parameters are $a = 2$, $b = 1$, and $c = 4$.

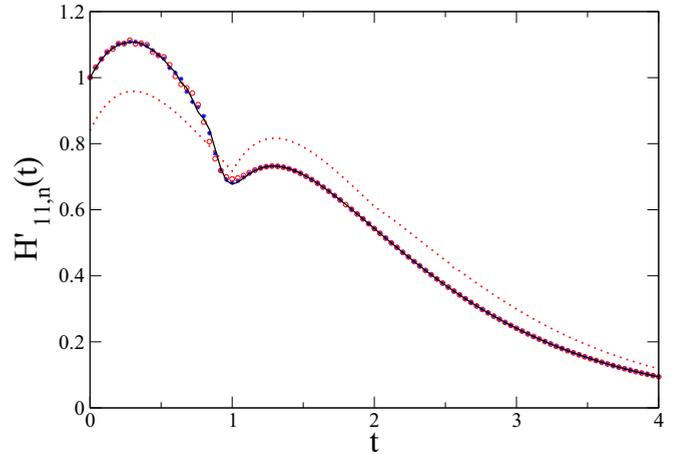


FIG. 3. Function $H'_{11,n}(t)$ for $n = 20$ (red circles), $n = 30$ (blue stars), and $n = 40$ (solid black line), as obtained from the numerical inverse Fourier transform of Eq. (64) for $\rho = 0.5$, $a = 2$, $c = 4$, and $\tau = 1$. The dotted red line represents $H'_1(t)$ computed from the numerical Fourier transform of Eq. (53). Note that this function is not equal to 1 at $t = 0$ (it is also nonzero for $t < 0$).

for $\mathcal{T}_{2 \rightarrow 1}$. The latter is indeed exact in a certain range of the parameters, as we now discuss.

2. Spectral expression of the TE rate

The spectral expression of $\mathcal{T}_{2 \rightarrow 1}$ is obtained by exchanging the labels 1 and 2 in Eq. (49). For the present simplified model, one has

$$\tilde{H}_{11}(\omega) = H_{11}(\omega) + \rho H_{12}(\omega) = \frac{b - i\omega + \rho c e^{i\omega\tau}}{(a - i\omega)(b - i\omega)}, \quad (72)$$

which gives [cf. Eq. (83) in [25]]

$$\mathcal{T}_{2 \rightarrow 1} = \frac{1}{2} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \ln \frac{\omega^2 + b^2 + c^2 + 2\rho cv(\omega, b, \tau)}{\omega^2 + b^2 + \rho^2 c^2 + 2\rho cv(\omega, b, \tau)}. \quad (73)$$

However, as was stressed above in Sec. III A 3, the domain of validity of the spectral expression is limited and must be carefully determined, a task that has been overlooked in [25]. Otherwise, the actual TE rate is underestimated. According to the previous discussion, the correct result is obtained if $\tilde{H}_{11}(\omega)$ has no zeros in the upper half plane. The stationary process $\tilde{X}_1(\omega) = \tilde{H}_{11}(\omega)\xi_1(\omega)$ is then minimum phase, like all stochastic processes considered in this work.

Remarkably, the solution to this problem is already available in the literature. Indeed, it turns out that the equation $f(\omega) \equiv b - i\omega + \rho c e^{i\omega\tau} = 0$ that determines the zeros of $\tilde{H}_{11}(\omega)$ is also the characteristic equation that determines the stability of the linear equation

$$\dot{x}(t) = -bx(t) - \rho cx(t - \tau), \quad (74)$$

which has been widely studied in the literature on delay differential equations. For instance, Ref. [57] (see also [58] and Theorem 8.6 in [43]) tells us that this equation is asymptotically stable when all the roots of $f(\omega)$ have a negative real part, a condition that is always satisfied if $-b < \rho c \leq b$

and always violated if $\rho c < -b$, whatever the value of τ (recall that $-1 < \rho < 1$ and $b > 0$ in the present model). In the case $\rho c > b$, there is a critical value of the delay $\tau^* = \arccos(-b/\rho c)/\sqrt{\rho^2 c^2 - b^2}$ beyond which the condition is violated (when $\tau = \tau^*$, a Hopf bifurcation occurs). Note that this has nothing to do with the stability of the joint process \mathbf{X} itself: As we have already mentioned, the stationary state is stable for all values of τ . Likewise, there is a spectral representation of the n th-order approximant, given by

$$\mathcal{T}_{2 \rightarrow 1}^{(n)} = \frac{1}{2} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \ln \frac{\mathcal{P}_n(\omega)}{\mathcal{P}_n(\omega) - c^2(1 - \rho^2)(1 + \omega^2 \tau^2 / 4n^2)^n}, \quad (75)$$

whose domain of validity depends on n .

IV. NUMERICAL ILLUSTRATION

A. Convergence with n

We first consider the issue of the convergence of $\mathcal{T}_{2 \rightarrow 1}^{(n)}$ with n . A typical example is shown in Fig. 1, where ρ and c are chosen such that the spectral formula (73) gives the exact value of $\mathcal{T}_{2 \rightarrow 1}$ for all values of τ . As indicated, the good news is that $\mathcal{T}_{2 \rightarrow 1}^{(n)}$ converges quite rapidly towards the exact asymptotic value, even when τ is much larger than the relaxation time b^{-1} of the process X_2 . (In this figure and in the following we take b^{-1} as the time unit.) For instance, the relative error between $\mathcal{T}_{2 \rightarrow 1}^{(n)}$ and $\mathcal{T}_{2 \rightarrow 1}$ for $\tau = 30$ is already less than 1% with $n = 10$. Therefore, there is no need to use large values of n , which would have been a practical limitation to our solution of the Wiener-Hopf factorization.⁶ Note that the exact PSD $S_{11}(\omega)$ is very well reproduced by $S_{11,n} = |H'_{11,n}(\omega)|^2$ even for n small, as shown in Fig. 2 where $n = 4$. However, it is well known in the field of Wiener-Hopf factorization (see, e.g., [54]) that by itself this is not a good criterion for assessing the accuracy of the factorization: For instance, the inverse Fourier transform of the function $H'_{11}(\omega)$ given by Eq. (53) [which, by construction, exactly reproduces $S_{11}(\omega)$] is neither causal nor equal to 1 at $t = 0$, which implies that the corresponding TE rate is infinite, as we have already pointed out. This is illustrated in Fig. 3, which also shows the evolution of $H'_{11,n}(t)$ with n . The small wiggles for $t < \tau$ are a consequence of the Laguerre formula (60) and they become negligible for $n \gtrsim 40$. It also seems that a cusp will occur at $t = \tau$ in the limit $n \rightarrow \infty$, as is the case with the function $H'_1(t)$ computed from Eq. (53).

B. Influence of τ and ρ

We are now in position to compare the TE rate estimated from Eq. (71) with the predictions of Eq. (73) and investigate the influence of the delay. This is illustrated in Fig. 4. Since $\rho c > b$ with our choice of the parameters, the discussion in

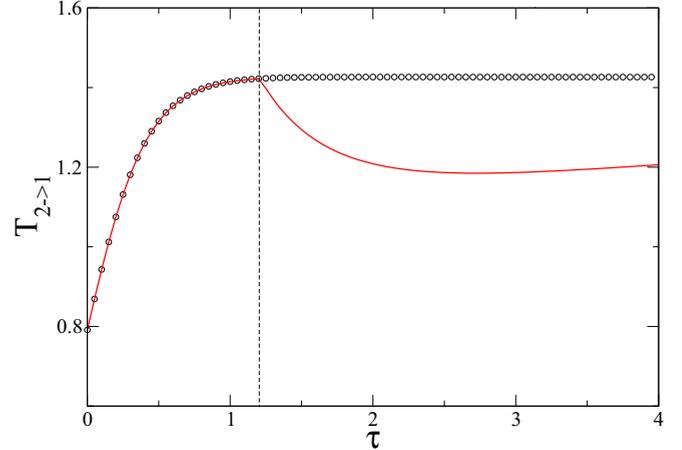


FIG. 4. Comparison between $\mathcal{T}_{2 \rightarrow 1}^{(n)}$ obtained from Eq. (71) for $n = 25$ (black circles) and the spectral formula (73) of $\mathcal{T}_{2 \rightarrow 1}$ (red line) as a function of τ for $\rho = 0.5$ and $c = 4$. Equation (73) is valid for $\tau < 1.209$ only, as indicated by the vertical dashed line.

the preceding section tells us that the spectral formula is valid up to $\tau = \tau^* \approx 1.209$. Indeed, we see in the figure that the agreement is excellent for $\tau \leq \tau^*$ but that the two curves deviate beyond τ^* . The spectral formula then predicts a lower value of the TE rate, in line with the arguments of [47]. More generally, our calculations show that $\mathcal{T}_{2 \rightarrow 1}$ is monotonically increasing with τ for $\rho c > 0$, whereas it first decreases and then increases for $\rho c < 0$. (It can be analytically shown that $d/d\tau \ln \mathcal{T}_{2 \rightarrow 1}|_{\tau=0} = \rho c$.) In both cases, $\mathcal{T}_{2 \rightarrow 1}$ goes to a finite value as $\tau \rightarrow \infty$. Note that the behavior of $\mathcal{T}_{2 \rightarrow 1}$ computed from Eq. (A16) is completely different. Consider, for instance, the simplest case of independent noises ($\rho = 0$). Then $\mathcal{T}_{2 \rightarrow 1} = (\sqrt{b^2 + c^2} - b)/2$ does not depend on τ , as already noticed in [25], whereas $\mathcal{T}_{2 \rightarrow 1}$ decreases with τ .⁷ (We chose not to plot $\mathcal{T}_{2 \rightarrow 1}$ since it depends on the value of a , in contrast with $\mathcal{T}_{2 \rightarrow 1}$.)

To further illustrate the differences between Eqs. (71) and (73), the behavior of $\mathcal{T}_{2 \rightarrow 1}$ as a function of ρ for a fixed value of τ is shown in Fig. 5. The spectral formula is now valid in the interval $-0.25 \leq \rho \leq 0.565$, where the minimal value corresponds to $\rho = -b/c$ and the maximal value is the solution of the equation $\tau \sqrt{\rho^2 c^2 - b^2} - \arccos(-b/\rho c) = 0$. The most striking feature is that $\mathcal{T}_{2 \rightarrow 1}$ goes to a finite value for $\rho = \pm 1$, at variance with the outcome of the spectral formula.⁸ It is a nontrivial fact that the TE rate remains finite even when the noises are strongly correlated or anticorrelated.

⁷This can be easily understood. Since $X_1(t+h)$ for h infinitesimal is determined by $X_1(t)$ and $X_2(t-\tau)$, the additional information provided by the knowledge of $X_2(t)$ becomes less and less relevant as τ increases, which makes $\mathcal{T}_{2 \rightarrow 1}$ decrease. On the other hand, if the whole past of X_1 up to t is already known and one adds information about the whole past of X_2 (that includes its value at $t-\tau$), then τ no longer plays a role and $\mathcal{T}_{2 \rightarrow 1}$ remains constant. For $\rho \neq 0$, however, this reasoning is no longer valid.

⁸The same feature already occurs for $\tau = 0$. Consider, for instance, $\rho = -1$. The exact expression of the TE rate given by Eq. (39) (exchanging the labels 1 and 2) with $r_1 = |b - c|$ yields

⁶In any case, there exist sophisticated polynomial root finders in the literature which allow one to consider very large values of n . Here we simply determine the roots of $\mathcal{P}_n(\omega)$ with the command `fsolve` of MAPLE that works well up to $n \lesssim 100$.

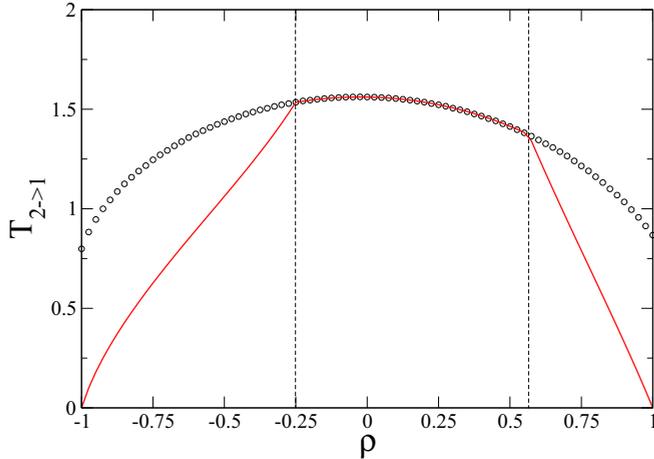


FIG. 5. Same as Fig. 4 as a function of ρ for $\tau = 1$. The spectral formula (73) is only valid in the range of ρ delimited by the vertical dashed lines.

C. Delay detection

Finally, we discuss the issue of delay detection and estimation. As pointed out in the Introduction, this is a potentially important application of transfer entropy, especially in neuroscience [16,25]. Since it is still actively debated whether or not this method is reliable [24], it is sensible to perform numerical tests on well-controlled dynamical systems, even as simple as the present one.

The idea is that the finite-horizon TE, in the present case $T_{2 \rightarrow 1}(h)$, should display a maximum in the vicinity of $h = \tau$. Indeed, as long as $h \lesssim \tau$, the trajectory of X_2 in the time interval $[t - \tau, t + h - \tau]$ provides a useful information about the future of X_1 , which makes $T_{2 \rightarrow 1}(h)$ increase with h . On the other hand, for $h \gtrsim \tau$, the trajectory of X_2 in the time interval $[t, t + h - \tau]$ is no longer taken into account in $T_{2 \rightarrow 1}(h)$ since only the trajectory of X_2 prior to t contributes, by definition. Eventually, as $h \rightarrow \infty$, both $p[X_1(t+h)|\{X(s)\}_{s \leq t}]$ and $p[X_1(t+h)|X_1^-(t)]$ approach the stationary probability distribution function $p(X_1)$ and $T_{2 \rightarrow 1}(h) \rightarrow 0$.

The above argument is only qualitative though, and the accuracy of the estimate of τ must be checked numerically. Typical results are shown in Fig. 6, where the value of ρ is arbitrarily fixed at 0.5. Indeed, whereas the magnitude of $T_{2 \rightarrow 1}(h)$ depends on ρ , the overall behavior remains qualitatively unchanged and in particular the position of maximum varies little.⁹ In line with the qualitative argument above, we observe in the figure that the maximum in $T_{2 \rightarrow 1}(h)$ occurs just

⁹ $T_{2 \rightarrow 1} = (1/2)(|b - c| - b + c) = c - b$ for $c > b$. On the other hand, the spectral formula always gives $T_{2 \rightarrow 1} = (1/2)(|c - b| - |b - c|) = 0$.

⁹This was already noticed in [25] (see Fig. 4 therein). Although the results in [25] are obtained by using a flawed Wiener-Hopf factorization, the choice of the model parameters, in particular the very large value of τ , makes the errors rather small. In particular, the fact that the Granger causality $\mathcal{F}_{Y \rightarrow X}(h)$ does not vanish at the origin when $\rho \neq 0$ is invisible on the scale of the figure. This is no longer true for the parameters used in our Fig. 6.

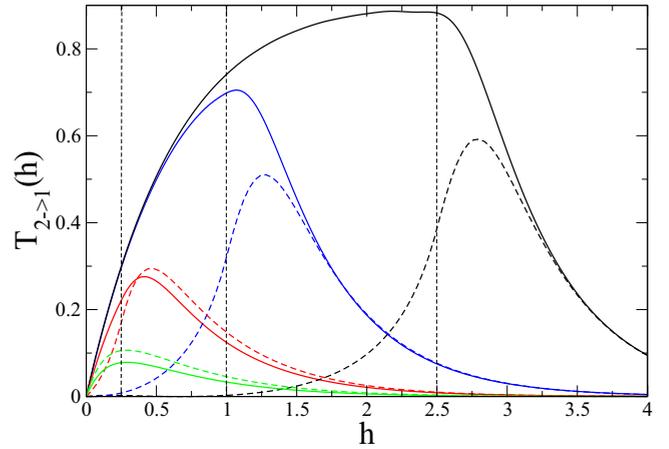


FIG. 6. Finite-horizon TE $T_{2 \rightarrow 1}^{(n)}(h)$ computed from Eq. (65) for $n = 30$, $\rho = 0.5$, and different values of the delay: from top to bottom, $\tau = 2.5$ (black line), $\tau = 1$ (blue line), $\tau = 0.25$ (red line), and $\tau = 0$ (green line). The model parameters are $a = 2$, $b = 1$, and $c = 4$. The corresponding dashed lines represent $\bar{T}_{2 \rightarrow 1}(h)$.

beyond τ for the two largest values of the delay (we recall that b^{-1} , the relaxation time of the process X_2 , is here taken as the natural time scale in the problem). On the other hand, the agreement is not so good for the smallest values of τ . The obvious problem is that $T_{2 \rightarrow 1}(h)$ exhibits a maximum at a certain time $h = h_0$ even when $\tau = 0$. This time depends in a complicated way on the two relaxation times a^{-1} and b^{-1} and on the coupling strength c . It also differs from the time h_{\max} associated with the maximum of the cross-correlation function $\phi_{21}(h)$ (for the case considered in Fig. 6, $h_0 \approx 0.29$ whereas $h_{\max} \approx 0.15$). We may tentatively regard h_0 as the time it would take for X_2 to effectively influence X_1 if there were no inherent delay in the coupling between the two processes. Therefore, the lesson to be drawn from the example in Fig. 6 is that τ must be significantly larger than h_0 to be properly estimated by scanning the horizon h in $T_{2 \rightarrow 1}(h)$. This is certainly a limitation because the value of h_0 is unknown in practice (e.g., in biochemical processes), although its order of magnitude may possibly be guessed.

On the positive side, we wish to stress that τ can be detected via $T_{2 \rightarrow 1}(h)$ even when there is no clear signature of a delayed interaction in the cross-correlation function $\phi_{21}(h)$. [Otherwise, there would indeed be no profit in using $T_{2 \rightarrow 1}(h)$, which is much less easily extracted from time-series data than $\phi_{21}(h)$.] This occurs, for instance, when the coupling parameter c is small, as illustrated in Fig. 7. In this case, there is no maximum in $\phi_{12}(h)$, except in $h = 0$, whereas the maximum in $T_{2 \rightarrow 1}(h)$ still takes place in the vicinity of $h = \tau$.¹⁰

We end this section with a short comment about $\bar{T}_{2 \rightarrow 1}(h)$, which is also represented in Figs. 6 and 7 [this function is computed from Eq. (A7) with the labels 1 and 2 exchanged].

¹⁰One could object that there is a kink at $h = \tau$ in both $\phi_{11}(h)$ and $\phi_{21}(h)$ which signals the presence of a delay in the interaction. However, it is very difficult in general to recognize such a feature in experimental data.

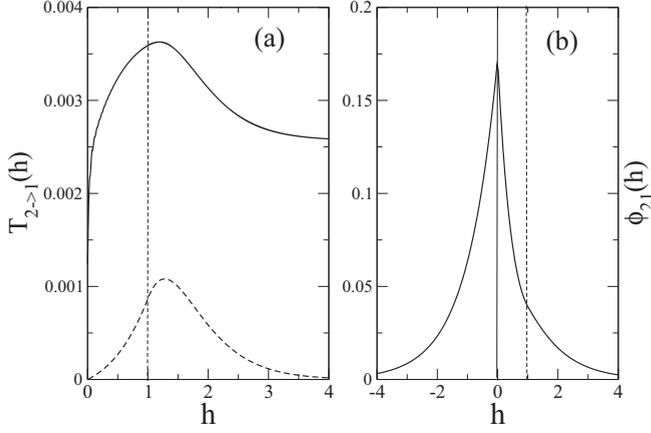


FIG. 7. (a) Finite-horizon TE $T_{2 \rightarrow 1}^{(n)}(h)$ for $n = 30$, $\rho = 0.5$, and $\tau = 1$. The model parameters are $a = 2$, $b = 1$, and $c = 0.1$. The dashed line represents $\bar{T}_{2 \rightarrow 1}(h)$. (b) Corresponding cross-correlation function $\phi_{21}(h)$.

Although the initial behavior of $\bar{T}_{2 \rightarrow 1}(h)$ with h differs from that of $T_{2 \rightarrow 1}(h)$, in relation to the fact that the rates $\bar{T}_{2 \rightarrow 1}$ and $T_{2 \rightarrow 1}$ are quite different, $\bar{T}_{2 \rightarrow 1}(h)$ also displays a maximum in the vicinity of $h = \tau$ when τ is sufficiently larger than h_0 . This is interesting because this function can be more easily estimated from time-series data than $T_{2 \rightarrow 1}(h)$ as it only requires the knowledge of the stationary correlation functions or the corresponding power spectral densities. One then circumvents the numerically challenging problem of estimating high-dimensional probability distributions (the so-called curse of dimensionality [59]). Note, however, that the delay is better estimated with $T_{2 \rightarrow 1}(h)$: For instance, in Fig. 6 and $\tau = 2.5$, the maxima of $T_{2 \rightarrow 1}(h)$ and $\bar{T}_{2 \rightarrow 1}(h)$ are located at $h \approx 2.52$ and $h \approx 2.79$, respectively. In Fig. 7, where $\tau = 1$, the maxima are located at 1.19 and 1.27 respectively.¹¹

V. SUMMARY

Is an information-theoretic measure such as the transfer entropy able to detect interaction delays in coupled systems? This question, still debated, has prompted us to revisit the recent calculation performed in [25] for a linear stochastic process with a delayed coupling. By focusing on a simple model that can be solved analytically in continuous time, thus

¹¹We have not considered in this work the time-delayed variant of the TE introduced in [16] (see also [15,60]) which is claimed to fulfill the so-called self-prediction optimality condition required by Wiener's causality principle. The continuous-time version of this quantity would be $\bar{T}_{i \rightarrow j}(u) = \lim_{h \rightarrow 0^+} (h^{-1}) I[X_j(t+h) : X_i(t-u) | X_j(t)]$, which is parametrized by the delay variable u . For Gaussian stationary processes, it is rather easy to guess, by extrapolating the calculation of $\bar{T}_{1 \rightarrow 2} \equiv \bar{T}_{1 \rightarrow 2}(u=0)$ [see Eq. (8) in [60]], that the expression of $\bar{T}_{i \rightarrow j}(u)$ involves the derivative of the correlation function $\phi_{ij}(u)$ at time u [cf. Eq. (A8)]. In the case of a discrete delay τ , this implies that $\bar{T}_{i \rightarrow j}(u)$ has a jump discontinuity at $u = \tau$. In a sense, one might say that this piece of information is already

provided by the kink in the correlation functions, as noticed above (preceding footnote).

avoiding the difficulties arising from time discretization, one may hope to get a better understanding of the issue. However, even in the simple case of stationary Gaussian processes, the calculation of the finite-horizon TE (or equivalently Granger causality) in the presence of delay requires the solution of a nontrivial Wiener-Hopf factorization problem that was not properly treated in [25]. The main contribution of the present work is to provide an efficient solution to this problem in the case where the stochastic noises are correlated, as is often required in the modeling of real networks. As a by-product, we have derived a compact expression of the zero-horizon TE rate. We have also clarified the conditions under which the spectral representation the TE rate is valid, an issue that seems to be overlooked in the literature. Our numerical results for a bivariate model with unidirectional delayed coupling show that the finite-horizon TE is indeed able to detect and estimate the delay (under some conditions though), even when there is no clear signature in the cross-correlation function. Interestingly, this is also true for the much simpler version of TE that only takes into account the immediate past of the source and the target.

It is clear, however, that more analytical and numerical work remains to be done before reaching a comprehensive picture. A natural extension of the present work would be to consider multiple delays occurring in both directions (not to mention the case of multivariate systems). It would also be interesting to investigate the behavior of the TE in an oscillatory regime and in the vicinity of a Hopf bifurcation. We leave this to future investigations.

APPENDIX A: EXPRESSIONS OF $\bar{T}_{i \rightarrow j}(h)$ AND $\bar{T}_{j \rightarrow i}$ IN THE PRESENCE OF TIME DELAY

In this appendix we derive the expressions of the simplified TE $\bar{T}_{i \rightarrow j}(h)$ ($i, j = 1, 2$) defined by Eq. (7) and of the corresponding rate $\bar{T}_{i \rightarrow j}$ for the bivariate process governed by Eq. (10). Our starting point is the expression of the conditional probability distribution function $p(\mathbf{x}, t+h | \mathbf{x}', t)$ of a Gaussian stationary process in terms of the matrix $\Phi(t)$ of the correlation functions $\phi_{ij}(h) \equiv \langle X_i(t) X_j(t+h) \rangle$,

$$p(\mathbf{x}, t+h | \mathbf{x}', t) = \frac{1}{2\pi \sqrt{\text{Det} \Sigma(h)}} e^{-\{[\mathbf{x} - \mathbf{G}(h)\mathbf{x}']^T \cdot \Sigma(h)^{-1} \cdot [\mathbf{x} - \mathbf{G}(h)\mathbf{x}']\}/2}, \quad (\text{A1})$$

where

$$\mathbf{G}(h) = \Phi(h)^T \cdot \Sigma^{-1} \quad (\text{A2})$$

and

$$\Sigma(h) = \Sigma - \mathbf{G}(h) \cdot \Phi(h). \quad (\text{A3})$$

We recall that $\Sigma = \Sigma(\infty)$ is the stationary covariance matrix with elements $\sigma_{ij} = \phi_{ij}(0)$ and that

$$p(\mathbf{x}) = \frac{1}{2\pi \sqrt{\text{Det} \Sigma}} e^{-\mathbf{x}^T \cdot \Sigma^{-1} \cdot \mathbf{x}/2}. \quad (\text{A4})$$

One can check that the correlation functions are indeed the second moments of $p(\mathbf{x}, t+h; \mathbf{x}', t) = p(\mathbf{x}, t+h | \mathbf{x}', t) p(\mathbf{x}')$, i.e., $\phi_{ij}(h) = \int d\mathbf{x} d\mathbf{x}' x'_i x_j p(\mathbf{x}, t+h; \mathbf{x}', t)$.

Consider, for instance, $\bar{T}_{1 \rightarrow 2}(h)$. By integrating Eq. (A1) over x_1 and then $p(x_2, t + h; \mathbf{x}', t)$ over x'_1 , we successively obtain

$$p(x_2, t + h | \mathbf{x}', t) = \frac{1}{\sqrt{2\pi\sigma_{22}(h)}} \exp\left[-\frac{(x_2 - G_{21}(h)x'_1 - G_{22}(h)x'_2)^2}{2\sigma_{22}(h)}\right] \quad (\text{A5})$$

and

$$p(x_2, t + h | x_2, t) = \sqrt{\frac{\sigma_{22}}{2\pi[G_{21}^2(h)\text{Det}\Sigma + \sigma_{22}(h)\sigma_{22}]}} \exp\left[-\frac{\sigma_{22}(x_2 - \frac{G_{21}(h)\sigma_{12} + G_{22}(h)}{\sigma_{22}}x'_2)^2}{2[G_{21}^2(h)\text{Det}\Sigma + \sigma_{22}(h)\sigma_{22}]}\right]. \quad (\text{A6})$$

This readily yields

$$\begin{aligned} \bar{T}_{1 \rightarrow 2}(h) &\equiv \frac{1}{2} \int d\mathbf{x} d\mathbf{x}' \ln \frac{p(x_2, t + h | \mathbf{x}', t)}{p(x_2, t + h | x_2, t)} = \frac{1}{2} \ln \left(1 + \frac{G_{21}^2(h)\text{Det}\Sigma}{\sigma_{22}(h)\sigma_{22}}\right) \\ &= -\frac{1}{2} \ln \frac{\sigma_{22}}{\text{Det}\Sigma} + \frac{1}{2} \ln \frac{\sigma_{22}^2 - \phi_{22}^2(h)}{\sigma_{22}\text{Det}\Sigma - \phi_{22}^2(h)\sigma_{11} + 2\phi_{22}(h)\phi_{12}(h)\sigma_{12} - \phi_{12}^2(h)\sigma_{22}}. \end{aligned} \quad (\text{A7})$$

Expanding $\phi_{12}(h)$ and $\phi_{22}(h)$ in powers of h , we then obtain the expression of the rate $\bar{T}_{1 \rightarrow 2} = \lim_{h \rightarrow 0^+} \bar{T}_{1 \rightarrow 2}(h)/h$,

$$\bar{T}_{1 \rightarrow 2} = -\frac{1}{4\sigma_{22}\text{Det}\Sigma} \frac{[\dot{\phi}_{22}(0^+)\sigma_{12} - \dot{\phi}_{12}(0^+)\sigma_{22}]^2}{\phi_{22}(0^+)}. \quad (\text{A8})$$

The corresponding expressions of $\bar{T}_{2 \rightarrow 1}(h)$ and $\bar{T}_{2 \rightarrow 1}$ are obtained by exchanging the labels 1 and 2.

To proceed further and express $\bar{T}_{1 \rightarrow 2}$ and $\bar{T}_{2 \rightarrow 1}$ in terms of the σ_{ij} 's only, we need to compute the derivatives of the correlation functions at $t = 0^+$. This can be done without fully solving the dynamics by using together the Fokker-Planck equation for the time-dependent probability distribution $p(\mathbf{x}, t)$ and the differential equations satisfied by the ϕ_{ij} 's. The Fokker-Planck equation is obtained as usual by starting from the definition $p(\mathbf{x}, t) = \langle \delta(X_1(t) - x_1)\delta(X_2(t) - x_2) \rangle$, inserting the Langevin equations, and using Novikov's theorem [61]. This yields

$$\begin{aligned} \frac{\partial p(\mathbf{x}, t)}{\partial t} &= \frac{\partial}{\partial x_1} \left[a_{11}x_1 p(\mathbf{x}, t) + a_{12} \int dy y p(\mathbf{x}, t; y, t - \tau) \right] \\ &+ \frac{\partial}{\partial x_2} [(a_{21}x_1 + a_{22}x_2)p(\mathbf{x}, t)] + D_{11} \frac{\partial^2}{\partial x_1^2} p(\mathbf{x}, t) \\ &+ D_{22} \frac{\partial^2}{\partial x_2^2} p(\mathbf{x}, t) + 2D_{12} \frac{\partial^2}{\partial x_1 \partial x_2} p(\mathbf{x}, t), \end{aligned} \quad (\text{A9})$$

where $p(\mathbf{x}, t; y, t - \tau) = \langle \delta(X_1(t) - x_1)\delta(X_2(t) - x_2)\delta(X_2(t - \tau) - y) \rangle$ is a two-time probability density. [In passing, note that Eq. (A9) is not a closed equation, which is a characteristic feature of time-delayed stochastic systems [62,63].] Multiplying this equation by x_1^2, x_2^2 , and x_1x_2 , respectively, and integrating over \mathbf{x} , we obtain the relations

$$\begin{aligned} a_{11}\sigma_{11} + a_{12}\phi_{21}(\tau) &= D_{11}, \\ a_{21}\sigma_{12} + a_{22}\sigma_{22} &= D_{22}, \\ (a_{11} + a_{22})\sigma_{12} + a_{21}\sigma_{11} + a_{12}\phi_{22}(\tau) &= 2D_{12}. \end{aligned} \quad (\text{A10})$$

On the other hand, from the differential equations for the correlation functions for $t \in [0^+, \tau]$,

$$\begin{aligned} \dot{\phi}_{11}(t) &= -a_{11}\phi_{11}(t) - a_{12}\phi_{21}(\tau - t), \\ \dot{\phi}_{21}(t) &= -a_{11}\phi_{21}(t) - a_{12}\phi_{22}(\tau - t), \end{aligned}$$

$$\begin{aligned} \dot{\phi}_{12}(t) &= -a_{21}\phi_{11}(t) - a_{22}\phi_{12}(t), \\ \dot{\phi}_{22}(t) &= -a_{21}\phi_{21}(t) - a_{22}\phi_{22}(t), \end{aligned} \quad (\text{A11})$$

we obtain

$$\begin{aligned} \dot{\phi}_{11}(0^+) &= -a_{11}\sigma_{11} - a_{12}\phi_{21}(\tau), \\ \dot{\phi}_{21}(0^+) &= -a_{11}\sigma_{12} - a_{12}\phi_{22}(\tau), \\ \dot{\phi}_{12}(0^+) &= -a_{21}\sigma_{11} - a_{22}\sigma_{12}, \\ \dot{\phi}_{22}(0^+) &= -a_{21}\sigma_{12} - a_{22}\sigma_{22}. \end{aligned} \quad (\text{A12})$$

Combining Eqs. (A10) and (A12) then gives

$$\begin{aligned} \dot{\phi}_{11}(0^+) &= -D_{11}, \\ \dot{\phi}_{21}(0^+) &= -2D_{12} + a_{21}\sigma_{11} + a_{22}\sigma_{12}, \\ \dot{\phi}_{12}(0^+) &= -a_{21}\sigma_{11} - a_{22}\sigma_{12}, \\ \dot{\phi}_{22}(0^+) &= -D_{22}. \end{aligned} \quad (\text{A13})$$

Inserting these expressions into Eq. (A8) and into the corresponding equation for $\bar{T}_{2 \rightarrow 1}$, we finally obtain

$$\bar{T}_{1 \rightarrow 2} = \frac{a_{21}^2 \text{Det}\Sigma}{4D_{22}\sigma_{22}} \quad (\text{A14})$$

and

$$\bar{T}_{2 \rightarrow 1} = \frac{[D_{11}\sigma_{12} + (a_{21}\sigma_{11} + a_{22}\sigma_{12} - 2D_{12})\sigma_{11}]^2}{4D_{11}\sigma_{11}\text{Det}\Sigma}. \quad (\text{A15})$$

Note that $\bar{T}_{1 \rightarrow 2}$ depends on τ and ρ only through the covariances σ_{ij} , so Eq. (A14) is formally the same equation as the one derived in [64] for a simple bipartite Ornstein-Uhlenbeck process. Finally, for the model studied in Sec. III B, Eq. (A15) reduces to

$$\bar{T}_{2 \rightarrow 1} = \frac{[\frac{1}{2}\sigma_{12} + (b\sigma_{12} - \rho)\sigma_{11}]^2}{2\sigma_{11}\text{Det}\Sigma}, \quad (\text{A16})$$

with

$$\begin{aligned} \sigma_{11} &= \frac{2bc\rho e^{-a\tau} + ab + b^2 + c^2}{2ab(a + b)}, \\ \sigma_{12} &= \frac{ce^{-b\tau} + 2b\rho}{2b(a + b)}, \\ \sigma_{22} &= \frac{1}{2b}. \end{aligned} \quad (\text{A17})$$

APPENDIX B: GAMMA-DISTRIBUTED DELAY

An approximation often used in the context of biological modeling [42–44] consists in replacing the discrete delay kernel $\delta(t - \tau)$ in the time domain by a sequence of Γ distributions $\delta(t - \tau) \approx g_n(t, \tau/n)$, where $g_n(t, T) = [(n-1)!T^n]^{-1}t^{n-1}e^{-t/T}$. For instance, at the lowest order $n = 1$, the memory kernel reduces to a low-pass filter with bandwidth τ^{-1} and $X_2(t - \tau)$ in Eq. (10) is replaced by $(1/\tau) \int_{-\infty}^t ds e^{-(t-s)/\tau} X_2(s)$. In the frequency domain, this approximation amounts to replacing $e^{i\omega\tau}$ by $(1 - i\omega\tau/n)^{-n}$. Equation (61) is then replaced by

$$S_{11,n}(\omega) = \frac{\mathcal{P}_n(\omega)}{(a^2 + \omega^2)(b^2 + \omega^2)(1 + \frac{\omega^2\tau^2}{n^2})^n}, \quad (\text{B1})$$

where

$$\begin{aligned} \mathcal{P}_n(\omega) &= (\omega^2 + b^2 + c^2) \left(1 + \frac{\omega^2\tau^2}{n^2}\right)^n + \rho c \left[\left(1 + i\frac{\omega\tau}{n}\right)^n \right. \\ &\quad \left. \times (b + i\omega) + \left(1 - i\frac{\omega\tau}{n}\right)^n (b - i\omega) \right], \end{aligned} \quad (\text{B2})$$

and the Wiener-Hopf causal factor is

$$H'_{11,n}(\omega) = \frac{i \prod_{k=1}^{n+1} (\omega - \omega_k)}{(\omega + ia)(\omega + ib)(\omega + in/\tau)^n}. \quad (\text{B3})$$

Noting that $\dot{H}'_{12,n}(0^+) = 0$ with this approximation of $e^{i\omega\tau}$, we finally arrive at

$$\begin{aligned} \mathcal{T}_{2 \rightarrow 1}^{(n)} &= \frac{1}{2} \left[a - \frac{n}{\tau} - \lim_{\omega \rightarrow \infty} \omega^2 \left(H'_{11,n}(\omega) - \frac{1}{n/\tau - i\omega} \right) \right] \\ &= \frac{1}{2} \left[-b + i \sum_{k=1}^{n+1} \omega_k - \frac{n^2}{\tau} \right], \end{aligned} \quad (\text{B4})$$

which replaces Eq. (71).

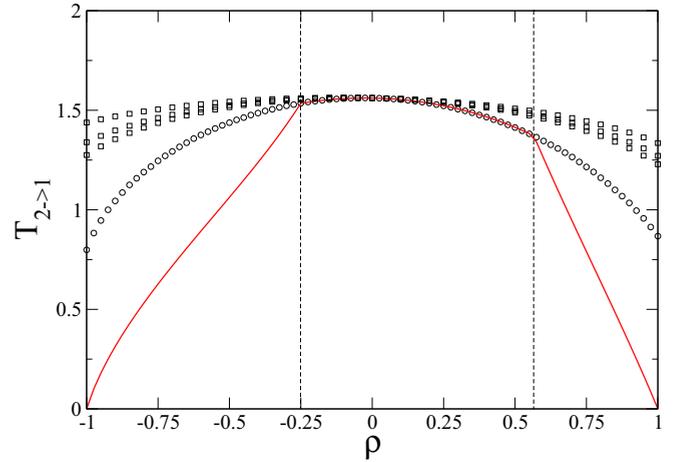


FIG. 8. Same as Fig. 5 with, in addition, the predictions of Eq. (B4) for $n = 30, 60, 90$ (black squares, from top to bottom).

The advantage of this representation of $\delta(t - \tau)$ is that a finite value of n may provide a good description of a given physical or biological process [whereas taking n finite in the Laguerre shift formula (60) does not correspond to a bona fide Langevin process with distributed delay]. Equations (B1)–(B3) then give the exact solution of the corresponding Wiener-Hopf factorization and in turn the exact expression of the TE. On the other hand, as illustrated in Fig. 8, the convergence with n is very slow. Therefore, this approximation is not appropriate for dealing with a true discrete delay.

-
- [1] S.-I. Niculescu, *Delay Effects on Stability*, Lecture Notes in Control and Information Sciences Vol. 269 (Springer, Berlin, 2001).
- [2] W. Just, A. Pelster, M. Schanz, and E. Schöll, Delayed complex systems: An overview, *Philos. Trans. R. Soc. A* **368**, 303 (2010).
- [3] *Complex Time-Delay Systems*, edited by F. M. Atay (Springer, Berlin, 2010).
- [4] K. J. Åström and R. M. Murray, *Feedback Systems: An Introduction for Scientists and Engineers* (Princeton University Press, Princeton, 2008).
- [5] J. Bechhoefer, *Rev. Mod. Phys.* **77**, 783 (2005).
- [6] For a recent review, see L. S. Tsimring, *Rep. Prog. Phys.* **77**, 026601 (2014), and references therein.
- [7] M. J. Dunlop, R. S. Cox, J. H. Levine, R. M. Murray, and M. B. Elowitz, *Nat. Genet.* **40**, 1493 (2008).
- [8] D. J. Kiviet, P. Nghe, N. Walker, S. Boulineau, V. Sunderlikova, and S. J. Tans, *Nature (London)* **514**, 376 (2014).
- [9] S. Klein, B. Soden, and N. Lau, *J. Climate* **12**, 917 (1999).
- [10] G. Gu and R. F. Adler, *J. Climate* **24**, 2258 (2011).
- [11] J. Runge, V. Petoukhov, and J. Kurths, *J. Climate* **27**, 720 (2014); J. Runge, Ph.D. thesis, Humboldt University, Berlin, 2014.
- [12] R. Vicente, M. Wibrál, M. Lindner, and G. Pipa, *J. Comput. Neurosci.* **30**, 45 (2011).
- [13] V. A. Vakorin, B. Mišić, O. Krakovska, and A. R. McIntosh, *Front. Syst. Neurosci.* **5**, 96 (2011).
- [14] B. Pompe and J. Runge, *Phys. Rev. E* **83**, 051122 (2011).
- [15] S. Ito, M. E. Hansen, R. Heiland, A. Lumsdaine, A. M. Litke, and J. M. Beggs, *PLoS One* **6**, e27431 (2011).
- [16] M. Wibrál, N. Pampu, V. Priesemann, F. Siebenhühner, H. Seiwert, M. Lindner, and J. T. Lizier, *PLoS One* **8**, e55809 (2013); P. Wollstadt, M. Martínez-Zaruela, R. Vicente, F. J. Díaz-Pernas, and M. Wibrál, *ibid.* **9**, e102833 (2014).
- [17] T. Schreiber, *Phys. Rev. Lett.* **85**, 461 (2000).
- [18] M. Paluš, V. Komárek, Z. Hrnčíř, K. Štěrbová, *Phys. Rev. E* **63**, 046211 (2001).
- [19] N. Wiener, in *Modern Mathematics for the Engineer*, edited by E. F. Beckenbach (McGraw-Hill, New York, 1956).
- [20] C. W. J. Granger, *Inf. Control* **6**, 28 (1963); *Econometrica* **37**, 424 (1969).

- [21] M. Wibral, R. Vicente, and M. Lindner, *Directed Information Measures in Neuroscience* (Springer, Berlin, 2014), pp. 3–32.
- [22] T. Bossomaier, L. Barnett, M. Harr, and J. T. Lizier, *An Introduction to Transfer Entropy: Information Flow in Complex Systems* (Springer, Berlin, 2016).
- [23] S. Cekik, D. Grandjean, and O. Renaud, *Stat. Med.* **37**, 1910 (2018).
- [24] D. Coufal, J. Jakubik, N. Jajcay, J. Hlinka, A. Krakovska, and M. Paluš, *Chaos* **27**, 083109 (2017). Note that this paper only focuses on the case of dynamical systems.
- [25] L. Barnett and A. K. Seth, *J. Neurosci. Methods* **275**, 93 (2017).
- [26] D. Zhou, Y. Zhang, Y. Xiao, and D. Cai, *Front. Comput. Neurosci.* **8**, 75 (2014).
- [27] L. Barnett, A. B. Barrett, and A. K. Seth, *Phys. Rev. Lett.* **103**, 238701 (2009).
- [28] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, *Science* **297**, 1183 (2002).
- [29] M. Hinczewski and D. Thirumalai, *J. Phys. Chem. B* **120**, 6166 (2016).
- [30] S. Lahiri, P. Nghe, S. J. Tans, and M. L. Rosinberg, and D. Lacoste, *PLoS One* **12**, e0187431 (2017).
- [31] T. Sagawa and M. Ueda, *Phys. Rev. E* **85**, 021104 (2012).
- [32] S. Ito and T. Sagawa, *Phys. Rev. Lett.* **111**, 180603 (2013).
- [33] D. Hartich, A. C. Barato, and U. Seifert, *J. Stat. Mech.* (2014) P02016.
- [34] J. M. Horowitz and H. Sandberg, *New J. Phys.* **16**, 125007 (2014).
- [35] D. Hartich, A. C. Barato, and U. Seifert, *Phys. Rev. E* **93**, 022116 (2016).
- [36] F. Comte and E. Renault, *Economet. Theor.* **12**, 215 (1996).
- [37] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley, New York, 2006).
- [38] T. Matsumoto and T. Sagawa, *Phys. Rev. E* **97**, 042103 (2018).
- [39] R. E. Spinney, J. T. Lizier, and M. Prokopenko, *Phys. Rev. E* **94**, 022135 (2016); R. E. Spinney, M. Prokopenko, and J. T. Lizier, *ibid.* **95**, 032319 (2017).
- [40] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis* (Springer, Berlin, 2005).
- [41] J.-M. Dufour and E. Renault, *Econometrica* **66**, 1099 (1998).
- [42] N. MacDonald, *Biological Delay Systems: Linear Stability Theory*, Cambridge Studies in Mathematical Biology Vol. 8 (Cambridge University Press, Cambridge, 1989).
- [43] F. Crauste, *Complex Time-Delay Systems* (Springer, Berlin, 2010), pp. 263–296.
- [44] H. Smith, *An Introduction to Delay Differential Equations with Applications to the Life Sciences*, Texts in Applied Mathematics Vol. 57 (Springer, Berlin, 2011).
- [45] T. J. McKetterick and L. Giuggioli, *Phys. Rev. E* **90**, 042135 (2014); L. Giuggioli, T. J. McKetterick, V. M. Kenkre, and M. Chase, *J. Phys. A: Math. Theor.* **49**, 384002 (2016).
- [46] D. Hunt, B. K. Szymanski, and G. Korniss, *Phys. Rev. E* **86**, 056114 (2012).
- [47] J. Geweke, *J. Am. Stat. Assoc.* **77**, 304 (1982); **79**, 907 (1984).
- [48] See, e.g., A. Fasoula, Y. Attala, and D. Schwartz, *J. Neurosci. Methods* **215**, 170 (2013), and references therein.
- [49] See, however, D. Chicharro, *Biol. Cybern.* **105**, 331 (2011) for a critical view on the spectral representation of the transfer entropy.
- [50] M. Ding, Y. Chen, and S. L. Bressler, in *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*, edited by B. Schelter, M. Winterhalder, and J. Timmer (Wiley-VCH, Weinheim, 2006).
- [51] A. B. Barrett, L. Barnett, and A. K. Seth, *Phys. Rev. E* **81**, 041907 (2010); L. Barnett and A. K. Seth, *J. Neurosci. Methods* **201**, 404 (2011); **223**, 50 (2014).
- [52] D. Hathcock, J. Sheehy, C. Weisenberger, E. Ilker, and M. Hinczewski, *IEEE Trans. Mol. Biol. Multi-Scale Commun.* **2**, 16 (2016).
- [53] B. Noble, *Methods Based on the Wiener-Hopf Technique*, 2nd ed. (Chelsea, New York, 1988).
- [54] I. D. Abrahams, *IMA J. Appl. Math.* **65**, 257 (2000).
- [55] J. Lam, *IEEE Trans. Automat. Contr.* **39**, 1517 (1994).
- [56] P. M. Mäkilä and J. R. Partington, *SIAM J. Control Optim.* **37**, 1897 (1999); *Int. J. Control* **72**, 932 (1999).
- [57] K. L. Cooke and Z. Grossman, *J. Math. Anal. Appl.* **86**, 592 (1982).
- [58] U. Küchler and B. Mensch, *Stoch. Stoch. Rep.* **40**, 23 (1992).
- [59] J. Runge, J. Heitzig, V. Petoukhov, and J. Kurths, *Phys. Rev. Lett.* **108**, 258701 (2012).
- [60] J. M. Nichols, F. Bucholtz, and J. V. Michalowicz, *Entropy* **15**, 3186 (2013).
- [61] E. A. Novikov, *Zh. Eksp. Theor. Fiz.* **47**, 1919 (1964) [*Sov. Phys. JETP* **20**, 1290 (1965)].
- [62] S. Guillouzic, I. L’Heureux, and A. Longtin, *Phys. Rev. E* **59**, 3970 (1999).
- [63] T. D. Frank, *Phys. Rev. E* **72**, 011112 (2005).
- [64] S. Ito and T. Sagawa, *Nat. Commun.* **6**, 7498 (2015).