Phase-dynamic causalities within dynamical effects framework

Dmitry A. Smirnov[a] [ID]

AFFILIATIONS

Saratov Branch, Kotelnikov Institute of Radioengineering and Electronics of the Russian Academy of Sciences, 38 Zelyonaya Street, Saratov 410019, Russia

Note: This paper is part of the Focus Issue, In Memory of Vadim S. Anishchenko: Statistical Physics and Nonlinear Dynamics of Complex Systems.
[a]**Author to whom correspondence should be addressed:** smirnovda@yandex.ru

ABSTRACT

This work investigates numerics of several widely known phase-dynamic quantifiers of directional (causal) couplings between oscillatory systems: transfer entropy (TE), differential quantifier, and squared-coefficients quantifier based on an evolution map. The study is performed on the system of two stochastic Kuramoto oscillators within the framework of dynamical causal effects. The quantifiers are related to each other and to an asymptotic effect of the coupling on phase diffusion. Several novel findings are listed as follows: (i) for a non-synchronous regime and high enough noise levels, the TE rate multiplied by a certain characteristic time (called here reduced TE) equals twice an asymptotic effect of a directional coupling on phase diffusion; (ii) "information flow" expressed by the TE rate unboundedly rises with the coupling coefficient even in the domain of effective synchronization; (iii) in any effective synchronization regime, the reduced TE is equal to 1/8 n.u. in each direction for equal coupling coefficients and equal noise intensities, and it is in general a simple function of the ratio of noise intensities and the ratio of coupling coefficients.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0055586

Phase-dynamic model description is often used[1-29] **to detect and quantify directional couplings (also called causalities**[30-38]**) between oscillatory systems. In time series analysis, different estimated values of a quantifier are often interpreted as different "coupling strengths" in that particular sense. However, the sense itself, i.e., meaning of numerical values of a directional coupling quantifier in terms of the coupling impact on dynamics, is not usually studied in more detail and remains largely unknown. In this work, several phase dynamics based quantifiers, or "phase-dynamic causalities," are studied theoretically within the framework of dynamical causal effects (DCEs)**[39-44] **for the system of two stochastic Kuramoto oscillators with constant couplings. The reduced transfer entropy (TE) is related to the coupling effect on the phase diffusion of a coupling recipient. Characteristic values of the reduced transfer entropy and other quantifiers are obtained for typical dynamical regimes to serve as a reference for dynamical interpretations of phase-dynamic causalities estimated from time series.**

I. INTRODUCTION

Phase description of dynamical systems has long been widely and fruitfully used to study spectral properties of self-sustained oscillations,[45,46] chaotic synchronization[47] and other collective phenomena,[48,49] coherent resonance,[50] chronotaxic systems,[51,52] coupling functions,[53] and detection of synchrony from time series,[54-56] including biomedical data.[57-60] Professor Vadim Anishchenko and his colleagues have made a large contribution to this field including all the above-mentioned problems as given in numerous influential publications, e.g., a monograph[61] and papers.[62-74] As his student at the undergraduate level, I learned many of those results from his lectures in Saratov State University.

Among the problems of phase-dynamics analysis, the detection and characterization of directional (also called causal)[30-38] couplings between oscillators from time series are actively studied[1-29] with applications to physics,[5] neuroscience,[6,7,10,12,19,20,23] climate science,[11,22,27] and other fields. To assess how strongly one oscillator influences another oscillator, one uses, e.g., mean-squared derivative

of an evolution map,[1] transfer entropy[3] (TE), and squared-coefficients of an evolution map with different normalizations.[21,28] Having estimated any such quantifier from data, one can check whether its value differs from zero statistically significantly (though this is not trivial, see, e.g., Refs. 4, 12, 14, and 29) and, if yes, detects the respective directional coupling at the corresponding significance level. If a time series is long enough and measurement noise is weak, that numerical value is estimated very accurately. Then, a further question of interest is "What dynamical meaning (effect) is hidden behind an obtained value of a directional coupling quantifier?" Many authors[1–6,10–13,23] use one of the above quantifiers to compute the directionality index in the form $(I_{Y \to X} - I_{X \to Y})/(I_{Y \to X} + I_{X \to Y})$, where $I_{Y \to X}$ is a quantifier of coupling in the direction $Y \to X$. Such an index shows that the coupling in one direction is stronger or weaker than the coupling in the opposite direction in the sense of that particular quantifier. But, again, what is the dynamical meaning of the quantifier $I_{Y \to X}$ itself? For example, having a one-step TE[3] equal to 1 nat or a differential quantifier[1] equal to 0.1 units in the direction $Y \to X$, can one conclude that the presence of the directional coupling $Y \to X$ leads to significant changes in the dynamics of the recipient (sometimes also called target) X of this coupling? In particular, does such coupling strongly change the phase diffusion constant of X as compared to the uncoupled case? Having answers to such questions, one could assess an overall dynamical role of a coupling based on the above causality quantifiers accessible from time series. With respect to TE, the asymptotic effect on variance, and spectral causalities in linear stochastic systems, such questions have been addressed[39–44] within the recently developed framework of dynamical causal effects[39,44] (DCEs). This work presents a similar study of the phase-dynamic causalities in coupled phase oscillators, a contribution in memory of Professor V.S. Anishchenko.

Stochastic phase description of two oscillators and several causality quantifiers are recapitulated in Sec. II. The DCE framework and interpretations of those quantifiers are provided in Sec. III. The numerical and analytic results are presented in Sec. IV and their discussion in Sec. V. Section VI shows the conclusions. Appendixes A–C provide technical details.

II. BASIC SYSTEM AND QUANTIFIERS

Detection and characterization of directional couplings between oscillatory systems is often done[47,75,76] on the basis of a phase-dynamics model in the form of stochastic differential equations,

$$\dot{\phi}_t^X = G_X(\phi_t^X, \phi_t^Y) + \xi_t^X,$$
$$\dot{\phi}_t^Y = G_Y(\phi_t^Y, \phi_t^X) + \xi_t^Y, \qquad (1)$$

where ϕ_t^X and ϕ_t^Y are phases of two oscillatory systems X and Y at a time instant t, ξ_t^X and ξ_t^Y are mutually independent Gaussian white noises with autocovariance functions $\langle \xi_t^X \xi_{t'}^X \rangle = \Gamma_X \delta(t - t')$ and $\langle \xi_t^Y \xi_{t'}^Y \rangle = \Gamma_Y \delta(t - t')$, Γ_X and Γ_Y are noise intensities, and 2π-periodic (in both arguments) smooth functions G_X and G_Y describe both individual dynamics of the systems and their couplings. References 77 and 78 report how to derive a more general phase description for limit cycle oscillators with different amplitude relaxation times and external noise properties. However, to start the

theoretical investigation of directional coupling quantifiers based on phase dynamics, it is necessary first to define and study them for the simplest system (1) taken as an original one. Having such theoretical results, one can later turn to "causality estimation" from time series on a firmer ground and provide richer interpretations of those estimates.

It is appropriate to mention that many researchers in physics and dynamical system communities prefer to avoid the term "causality" since causality in the real world is hard or impossible "to prove." Instead, one can say "directional coupling," "influence," etc. However, in the field of time series analysis originating from engineering, econometrics, and close disciplines, researchers often use this term, see, e.g., Refs. 30–35 and recall "Granger causality" as a famous example. In this work, "directional coupling" and "causality" are used as synonyms with a detailed discussion given in Refs. 39 and 44. In brief, the entire consideration is performed here for the mathematical system (1), which is arbitrarily manipulated and fully observed. That is, one can specify a complete initial state (ϕ_0^X, ϕ_0^Y) and all parameters (if any) and observe future values of both phases. So, it is always possible to show that a certain variation of an initial state or a parameter of the subsystem Y is a cause of the corresponding future response of the subsystem X (an effect) *under other equal conditions*, which are all under control. Such a nonzero response can take place only if $\frac{\partial G_X}{\partial \phi_t^Y} \neq 0$ at least at some (ϕ_t^X, ϕ_t^Y). The latter condition on the mathematical system (1) is equivalent to saying that there is an influence or causal coupling $Y \to X$ in the system. So, there is no ambiguity with the terms "causality" and "directional couplings" in this study. Detailed definitions are given in Sec. III.

To characterize directional couplings in system (1), several quantifiers have been suggested. First, Rosenblum and Pikovsky[1] introduced an evolution map $(F_X^{(t)}, F_Y^{(t)})$ acting over the interval $(0, t)$ relating phase increments $(\Delta\phi_t^X, \Delta\phi_t^Y) = (\phi_t^X - \phi_0^X, \phi_t^Y - \phi_0^Y)$ to an initial state (ϕ_0^X, ϕ_0^Y) as

$$\Delta\phi_t^X = F_X^{(t)}(\phi_0^X, \phi_0^Y) + \varepsilon_t^X,$$
$$\Delta\phi_t^Y = F_Y^{(t)}(\phi_0^Y, \phi_0^X) + \varepsilon_t^Y, \qquad (2)$$

where the functions $F_X^{(t)}$ and $F_Y^{(t)}$ are conditional expectations $F_X^{(t)}(\phi_0^X, \phi_0^Y) = E(\Delta\phi_t^X | \phi_0^X, \phi_0^Y)$ and $F_Y^{(t)}(\phi_0^Y, \phi_0^X) = E(\Delta\phi_t^Y | \phi_0^X, \phi_0^Y)$ over realizations of the noise (ξ_t^X, ξ_t^Y) on the interval $(0, t)$ represented in general as trigonometric polynomials of arbitrary orders, and $(\varepsilon_t^X, \varepsilon_t^Y)$ is a zero-mean random vector representing deviations of $(\Delta\phi_t^X, \Delta\phi_t^Y)$ from the expectation. Since the phases in Eq. (1) are influenced by a stationary noise, the evolution map (2) remains the same if any two time instants $(t', t' + t)$ are used instead of $(0, t)$. A squared partial derivative describing steepness of the dependence of $\Delta\phi_t^X$ on ϕ_0^Y serves[1] as a quantifier of the coupling $Y \to X$,

$$C_{Y \to X}^{(t)} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left(\frac{\partial F_X^{(t)}(\phi_0^X, \phi_0^Y)}{\partial \phi_0^Y} \right)^2 d\phi_0^X d\phi_0^Y. \qquad (3)$$

Later, Paluš and Stefanovska[3] suggested to characterize the dependencies captured by the evolution map (2) via an appropriate conditional mutual information.[79] The latter for the direction $Y \to X$ is defined as a reduction of uncertainty in the value of $\Delta\phi_t^X$ if the value of ϕ_0^Y is taken into account, given ϕ_0^X. So, it is the difference

of the two conditional Shannon entropies,

$$
\begin{aligned}
T_{Y\to X}^{(t)} &= H_{X|X}^{(t)} - H_{X|X,Y}^{(t)} \\
&= \int_{-\pi}^{\pi} H(\Delta\phi_t^X|\phi_0^X)\tilde{p}_X^{st}(\phi_0^X)d\phi_0^X \\
&\quad - \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} H(\Delta\phi_t^X|\phi_0^X,\phi_0^Y)\tilde{p}_{XY}^{st}(\phi_0^X,\phi_0^Y)d\phi_0^X d\phi_0^Y,
\end{aligned}
\tag{4}
$$

where $\tilde{p}_{XY}^{st}$ is the stationary probability distribution function (PDF) of the phases $(\tilde{\phi}_t^X, \tilde{\phi}_t^Y)$ wrapped to an interval $(-\pi, \pi)$, $\tilde{p}_X^{st}$ is the marginal stationary PDF of the wrapped phase $\tilde{\phi}_t^X$,

$$
H(\Delta\phi_t^X|\phi_0^X) = -\int_{-\infty}^{\infty} p(\Delta\phi_t^X|\phi_0^X)\ln p(\Delta\phi_t^X|\phi_0^X)d(\Delta\phi_t^X)
\tag{5}
$$

is the Shannon entropy of the PDF of $\Delta\phi_t^X$ conditioned only on the concrete value ϕ_0^X, and

$$
H(\Delta\phi_t^X|\phi_0^X,\phi_0^Y) = -\int_{-\infty}^{\infty} p(\Delta\phi_t^X|\phi_0^X,\phi_0^Y)\ln p(\Delta\phi_t^X|\phi_0^X,\phi_0^Y)d(\Delta\phi_t^X)
\tag{6}
$$

is the Shannon entropy of the PDF of ϕ_t^X conditioned on both values ϕ_0^X and ϕ_0^Y. The information-theoretic measure (4) is invariant under a nonlinear change of coordinates. Another name of this conditional mutual information is TE,[80] see also the discussion in Refs. 3 and 79.

The third quantifier under study is the sum of squared coupling coefficients[21] in the trigonometric polynomial $F_X^{(t)}$ normalized in different ways. The non-normalized sum is interpreted as prediction improvement (PI) in the mean-square sense for weak enough couplings, see also Appendix A. Let us define that quantifier here in the general form of PI. The non-normalized PI reads

$$
\begin{aligned}
P_{Y\to X}^{(t)} = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} (\mathrm{var}(\Delta\phi_t^X|\phi_0^X) \\
- \mathrm{var}(\Delta\phi_t^X|\phi_0^X,\phi_0^Y))\tilde{p}_{XY}^{st}(\phi_0^X,\phi_0^Y)d\phi_0^X d\phi_0^Y,
\end{aligned}
\tag{7}
$$

where $\mathrm{var}(\phi_t^X|\cdot)$ denotes the conditional variance with $\mathrm{var}(\Delta\phi_t^X|\phi_0^X, \phi_0^Y) = \mathrm{var}(\varepsilon_t^X|\phi_0^X,\phi_0^Y)$. Being divided by the variance of the noise ε_t^X, it becomes the relative PI and reads[21]

$$
G_{Y\to X}^{(t)} = P_{Y\to X}^{(t)} \Big/ \left\langle \mathrm{var}(\Delta\phi_t^X|\phi_0^X,\phi_0^Y) \right\rangle,
\tag{8}
$$

where $\left\langle \mathrm{var}(\Delta\phi_t^X|\phi_0^X,\phi_0^Y) \right\rangle = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \mathrm{var}(\Delta\phi_t^X|\phi_0^X,\phi_0^Y)\tilde{p}_{XY}^{st}(\phi_0^X,\phi_0^Y)d\phi_0^X d\phi_0^Y$. For Gaussian distribution of $\varepsilon_t^X(\phi_0^X,\phi_0^Y)$ with variance independent of (ϕ_0^X,ϕ_0^Y), the relative PI relates to the TE as[43,81]

$$
T_{Y\to X}^{(t)} = \frac{1}{2}\ln(1 + G_{Y\to X}^{(t)}).
\tag{9}
$$

Furthermore, for small (i.e., much less than unity) values of both quantifiers as is the case for nonzero noises and small enough t,

Eq. (9) simplifies to $T_{Y\to X}^{(t)} \approx G_{Y\to X}^{(t)}/2$. So, the right-hand side of Eq. (9) can be considered an approximate value of the TE (4) or as an alternative to the TE similarly to Ref. 39. A nonzero difference between the two sides of Eq. (9) at a given t indicates that the PDF of ε_t^X differs from a constant-variance Gaussian law.

Below, the TE (4) is used as a basic quantifier due to its universality. The relative PI is also studied and turns out to provide often a good approximation of the TE. The differential quantifier (3) is only sometimes similar to the TE and PI as commented in Sec. IV E. Meanings of these quantifiers within the dynamical effects framework[39–44] are described in Sec. III, where another useful quantifier is introduced for inter-comparisons and richer interpretation of the TE numerics.

III. DYNAMICAL CAUSAL EFFECTS

A. Transient dynamical causal effects

Dynamical causal effects are introduced[39] for stochastic dynamical systems (SDSs) where a state vector uniquely determines PDFs of the future states. That is, a time realization of the full state vector of an SDS is a first-order (in general, high-dimensional vector-valued) Markov process. In other words, SDS is understood as a Markovian random dynamical system.[82] System (1) is such an SDS with the two-dimensional state vector (ϕ_t^X, ϕ_t^Y).

Let us specify the values of an initial state (ϕ_0^X, ϕ_0^Y) and a parameter vector a. The latter may include coupling coefficients (a_{xy}, a_{yx}), individual natural frequencies, etc. Then, an ensemble of time realizations is readily obtained via numerical simulations of Eq. (1), e.g., as described in Appendix B. Given the value of a, a dynamical causal effect (DCE) $Y \to X$ is defined[39] according to Pearl's interventional causality viewpoint[30] as a future effect of an initial intervention to the value of ϕ_0^Y. In Ref. 44, the term "initial intervention" was changed to the "initial variation" of ϕ_0^Y.

More generally, an ensemble of realizations may start from an initial distribution of states with some PDF $w(\phi_0^X, \phi_0^Y)$ rather than from a single initial state. Such initial PDF can be called[44] a (functional) initial condition. Let us specify a reference initial condition $w_1(\phi_0^X, \phi_0^Y)$ and an alternative initial condition $w_2(\phi_0^X, \phi_0^Y)$ with the same marginal PDF $w_{1,X}(\phi_0^X) = w_{2,X}(\phi_0^X)$. If there is no coupling $Y \to X$, then the evolutions of X from the two initial conditions are the same. A DCE $Y \to X$ on a finite temporal horizon t is any measure of difference between the PDFs $p_X^{(t)}[\phi_t^X|w_1]$ and $p_X^{(t)}[\phi_t^X|w_2]$, where the square brackets show that the PDF is obtained as a result of the evolution from a functional initial condition and so can be called a "functionally conditional" PDF.[44] It is found via marginalization of the ordinary conditional PDF with the initial condition w as

$$
p_X^{(t)}[\phi_t^X|w] = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} p_X^{(t)}(\phi_t^X|\phi_0^X,\phi_0^Y)w(\phi_0^X,\phi_0^Y)d\phi_0^X d\phi_0^Y.
\tag{10}
$$

Having computed such "elementary" DCEs for many initial variations (i.e., for many pairs of functional initial conditions), one naturally aims at "assembling" them into a single number. For that, one can parameterize initial variations with a vector λ (e.g., locations of Dirac-delta PDFs) and then average over λ with some weight

function. After such assemblage,[44] one finally gets a desired concrete DCE. If the temporal horizon t is not much greater than characteristic times of the coupled system, then the DCE is called short-term or transient.[39–41,44]

B. Transfer entropy and other quantifiers

To see that the TE (4) is a specific DCE,[42] let us take a Dirac-delta reference initial condition $w_1 = \delta(\phi_0^X - \phi_{0,1}^X)\delta(\phi_0^Y - \phi_{0,1}^Y)$, where both initial phases are given, and an alternative initial condition $w_2 = \delta(\phi_0^X - \phi_{0,1}^X)\tilde{p}_{Y|X}^{st}(\phi_0^Y|\phi_{0,1}^X)$, where the phase ϕ_0^Y is "freed" to vary according to the stationary conditional PDF of the wrapped phase $\tilde{\phi}_t^Y$. Define an elementary DCE as the difference of the functionally conditional Shannon entropies $H[\phi_t^X|w_2] - H[\phi_t^X|w_1]$ and average it over the assemblage parameter $\lambda = (\phi_{0,1}^X, \phi_{0,1}^Y)$ with the stationary PDF $\tilde{p}_{XY}^{st}(\phi_{0,1}^X, \phi_{0,1}^Y)$ to get

$$T_{Y \to X}^{(t)} = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} (H[\phi_t^X|w_2] - H[\phi_t^X|w_1])\tilde{p}_{XY}^{st}(\phi_{0,1}^X, \phi_{0,1}^Y)d\phi_{0,1}^X d\phi_{0,1}^Y. \quad (11)$$

Indeed, writing down marginalization over w_1 and w_2 in the right-hand side of Eq. (11) explicitly, one can immediately see that this transient DCE coincides with the TE (4).

The non-normalized PI (8) is formulated as a DCE similarly to the above TE, but with the functionally conditional variances instead of the functionally conditional Shannon entropies

$$P_{Y \to X}^{(t)} = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} (\text{var}[\phi_t^X|w_2] - \text{var}[\phi_t^X|w_1])\tilde{p}_{XY}^{st}(\phi_{0,1}^X, \phi_{0,1}^Y)d\phi_{0,1}^X d\phi_{0,1}^Y. \quad (12)$$

The relative PI $G_{Y \to X}^{(t)}$ is the corresponding relative difference of the same variances.

The differential quantifier (3) is readily seen to compare evolutions from infinitesimally close initial states $w_1 = \delta(\phi_0^X - \phi_{0,1}^X)\delta(\phi_0^Y - \phi_{0,1}^Y)$ and $w_2 = \delta(\phi_0^X - \phi_{0,1}^X)\delta(\phi_0^Y - \phi_{0,1}^Y - \delta\phi)$. One should just take the squared difference of expectations and average it with the uniform PDF to get

$$C_{Y \to X}^{(t)} = \lim_{\delta\phi \to 0} \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \frac{(E[\phi_t^X|w_2] - E[\phi_t^X|w_1])^2}{(\delta\phi)^2}d\phi_{0,1}^X d\phi_{0,1}^Y. \quad (13)$$

So, from the DCE viewpoint, all these quantifiers appear conceptually similar to each other. Appendix A provides further details and other versions of these quantifiers.

For the SDS (1), the finite-time TE $T_{Y \to X}^{(t)}$ is typically small at small response time t, reaches its maximal value at some intermediate t close to some characteristic time of the coupled system, and decreases at greater response times.[39,43] The maximal TE is denoted $T_{Y \to X}^{max} = \sup_{t>0} T_{Y \to X}^{(t)}$ and the respective temporal horizon $t_{Y \to X}^{max} = \arg\sup_{t>0} T_{Y \to X}^{(t)}$. Another useful parameter-free quantifier is the TE rate $\tau_{Y \to X} = \left.\frac{dT_{Y \to X}^{(t)}}{dt}\right|_{t=0}$, which characterizes responses at very

small temporal horizons. The TE rate has the dimension of inverse time. The respective dimensionless quantifier is the TE rate multiplied by some characteristic time t_{char}.[43] The quantity $T_{Y \to X} = \tau_{Y \to X}t_{char}$ is the TE reduced to the time t_{char} called in Ref. 43 just "reduced TE." The same consideration applies to the relative PI and to the differential quantifier. In particular, the relative PI rate $g_{Y \to X} = \left.\frac{dG_{Y \to X}^{(t)}}{dt}\right|_{t=0}$ exactly equals the doubled TE rate $g_{Y \to X} = 2\tau_{Y \to X}$ for SDS (1) with nonzero noises.[37]

C. Asymptotic effect on phase diffusion

The second important type of DCE originates from the initial variations of the parameter a. For example, one can switch the coupling $Y \to X$ "on" or "off" by introducing or zeroing the corresponding coupling coefficient. Then, it is natural to ask about the dynamics as a whole, i.e., about an asymptotic regime or DCEs on the infinite temporal horizon. Such DCEs are called long-term or asymptotic[39,44] (and also long-term, stationary, and equilibrium).[39–41] As an example relevant here, let us define an initial condition $w = \delta(\phi_0^X - \phi_{0,1}^X)\tilde{p}_{Y|X}^{st}(\phi_0^Y|\phi_{0,1}^X)$, set $t \to \infty$, and take the variance divided by t to get the phase diffusion constant. The constant at given coupling coefficients (a_{xy}, a_{yx}) is denoted as $D_X(a_{xy}, a_{yx})$, i.e.,

$$D_X(a_{xy}, a_{yx}) = \lim_{t \to \infty}(\text{var}[\phi_t^X|w, a_{xy}, a_{yx}]/t). \quad (14)$$

Let us denote this constant in the uncoupled regime (i.e., $a_{xy} = 0$) as $D_{X,0} = D_X(0, a_{yx})$. Then, the effect of coupling $Y \to X$ on the phase diffusion constant of X under other equal conditions is

$$D_{Y \to X} = \frac{D_X(a_{xy}, a_{yx}) - D_{X,0}}{D_{X,0}}. \quad (15)$$

This asymptotic DCE exists for system (1), despite a stationary PDF of the non-wrapped phases does not exist: the process of the non-wrapped phase variations is nonstationary with variance linearly growing in time. The quantifier (15) is dynamically meaningful by definition: Its positive value shows how strongly switching the coupling $Y \to X$ on increases the phase diffusion constant of X, i.e., decreases the quality of oscillations[46,47] (recall that the width of the power spectral line of the signal $\cos\phi_t^X$ is often proportional to the phase diffusion constant[45,46]). It is desirable to relate the TE to this asymptotic DCE to have as clear dynamical interpretation of the TE numerics.

D. Design of numerical experiments

The simplest version of system (1) is given by the Kuramoto oscillators[83]

$$\dot{\phi}_t^X = \omega_X + k_{XY}\sin(\phi_t^Y - \phi_t^X) + \xi_t^X,$$
$$\dot{\phi}_t^Y = \omega_Y + k_{YX}\sin(\phi_t^X - \phi_t^Y) + \xi_t^Y, \quad (16)$$

where ω_X and ω_Y are natural frequencies of the oscillators, couplings are given as sine of the phase difference $\Delta\Phi_t = \phi_t^Y - \phi_t^X$, and $k_{XY} \geq 0$ and $k_{YX} \geq 0$ are the coupling coefficients, non-negative in this work. Due to its simplicity, this model is used as a basis for studying synchronization phenomena in ensembles of coupled

oscillators[47,83] and many other problems. Here, it is taken as a basis to study numerics of the phase-dynamic directional coupling quantifiers.

A stationary solution to the FPE for the phase difference $\Delta\Phi$ wrapped to the interval $(-\pi, \pi)$ can be computed in a closed form via taking the following integral:

$$\tilde{p}^{st}(\Delta\Phi) = \frac{1}{C} \int_{\Delta\Phi}^{\Delta\Phi+2\pi} e^{-\tilde{\Delta}(\psi-\Delta\Phi)-\tilde{K}(\cos\psi-\cos\Delta\Phi)} d\psi, \quad (17)$$

where $\tilde{\Delta} = 2\Delta\omega/\Gamma_\Sigma$, $\tilde{K} = 2K_\Sigma/\Gamma_\Sigma$, $\Gamma_\Sigma = \Gamma_X + \Gamma_Y$ is a total noise intensity, $\Delta\omega = \omega_Y - \omega_X$ is a frequency mismatch, $K_\Sigma = k_{XY} + k_{YX}$ is sometimes called a total "coupling strength," and C is a normalizing constant. Equation (17) is derived and discussed in detail in Ref. 45 [see Eq. (18.52)] and briefly in Ref. 47 [see Eq. (9.11)]. For zero noise, transition to synchronization occurs at the threshold $\tilde{K}_c = \tilde{\Delta}$. If $\tilde{\Delta} \gg 1$, the transition is step-like, and otherwise it is smeared (Fig. 6 in Appendix B).

Conditional future PDFs entering the definition of all causality quantifiers under study are computed numerically. To get a usual conditional PDF of the future phase, given an initial state, one can either solve the Fokker–Planck equation (FPE) or simulate an ensemble of stochastic time realizations. The latter method often appears more convenient and fast as shown by Anishchenko et al.,[84–86] and so it is used here as a basic one (see Appendix B for details).

IV. NUMERICAL RESULTS

The TE rate and its relations to the asymptotic effect on phase diffusion are studied for SDS (16) well below the synchronization threshold $\tilde{K}_c = \tilde{\Delta}$, around the threshold, and well above it, i.e., in an effective synchronization regime[47,77,78] (Subsections IV A–IV D). Then, its relations to the finite-time TEs are revealed, and other quantifiers are commented (Subsection IV E).

A. Transfer entropy rates: Strong information flow in effective synchronization regime

Let us first fix $\Gamma_X = \Gamma_Y$ (so $\Gamma_\Sigma = 2\Gamma_X$) and $k_{XY} = k_{YX}$ (so $K_\Sigma = 2k_{XY}$). Then, the TE rates in both directions coincide, $\tau_{Y\to X} = \tau_{X\to Y}$. Just to define time units, let us specify $\Gamma_X = 1$, and so $\Gamma_\Sigma = 2$, $\tilde{\Delta} = \Delta\omega$, and $\tilde{K} = K_\Sigma$. The TE rate for $\Delta\omega = 50$ is presented in Fig. 1(b) with a thick line. It first rises with K_Σ and then quickly decreases near the synchronization threshold due to a more narrow PDF of the phase difference [see Fig. 6(b) in Appendix B]. Such a decrease of any directional coupling quantifier with transition to synchronization is often observed in time series analysis of exemplary systems. A common thought is that such a quantifier must be close to zero in an almost synchronous regime due to weak variations of the simultaneous subsystems' states relative to each other. However, note that the TE rate in Fig. 1(b) rises linearly with a further increase of the coupling coefficients within the domain of effective synchronization, where the mean phase coherence quite slowly changes tending to unity [see Fig. 6(c) in Appendix B].

This behavior of the TE rate is readily explained via linearization of Eq. (16) using the condition of $\left| \phi_t^Y - \phi_t^X - \arcsin\frac{\Delta\omega}{K_\Sigma} \right| \ll 1$, which applies in the case of $K_\Sigma \gg |\Delta\omega|$. Then, one can get

$$\dot{\phi}_t^X = \omega_X + \frac{k_{XY}}{K_\Sigma}\sqrt{K_\Sigma^2 - (\Delta\omega)^2}(\phi_t^Y - \phi_t^X) + \xi_t^X,$$
$$\dot{\phi}_t^Y = \omega_Y + \frac{k_{YX}}{K_\Sigma}\sqrt{K_\Sigma^2 - (\Delta\omega)^2}(\phi_t^X - \phi_t^Y) + \xi_t^Y, \quad (18)$$

and the evolution equation for the phase difference $\Delta\Phi_t = \phi_t^Y - \phi_t^X$ reads

$$\Delta\dot{\Phi}_t = \Delta\omega - \sqrt{K_\Sigma^2 - (\Delta\omega)^2}\Delta\Phi_t + \xi_t, \quad (19)$$

where ξ is the white noise with variance Γ_Σ. The TE rate for the linear system (18) is found as[43]

$$\tau_{Y\to X} = \frac{k_{XY}^2(1 - (\Delta\omega)^2/K_\Sigma^2)\text{var}(\Delta\Phi_t|\phi_t^X)}{2\Gamma_X}, \quad (20)$$
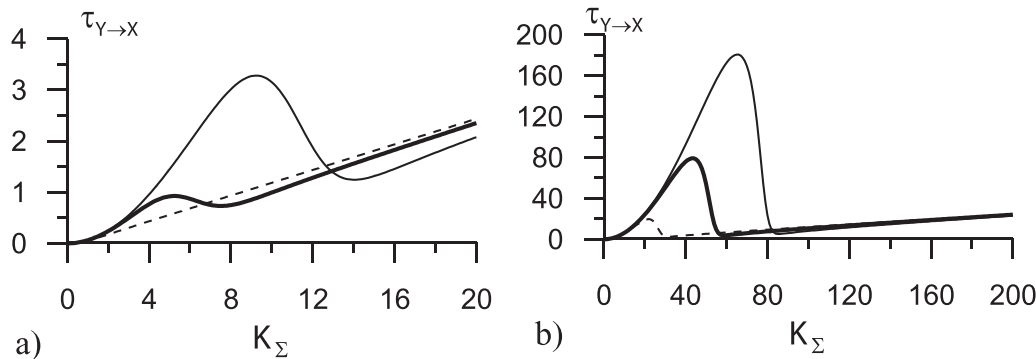

FIG. 1. TE rates for system (16) with $\Gamma_X = \Gamma_Y = 1$ and $k_{xy} = k_{yx} = K_\Sigma/2$: (a) $\Delta\omega = 0$ (dashed line), $\Delta\omega = 5$ (thick solid line), and $\Delta\omega = 10$ (solid line); (b) $\Delta\omega = 25$ (dashed line), $\Delta\omega = 50$ (thick solid line), and $\Delta\omega = 75$ (solid line).

where the numerator is the squared coupling coefficient multiplied by the conditional variance of the phase difference and the denominator is the doubled intensity of the noise in the subsystem X, doubling arises due to the relation $T_{Y \to X}^{(t)} = G_{Y \to X}^{(t)}/2$ for infinitesimally small t. The conditional variance of the phase difference in Eq. (18) equals the stationary variance of the phase difference in Eq. (19) and is much less than unity. So, it holds $\text{var}(\Delta \Phi_t | \phi_t^X) = \text{var}(\Delta \Phi_t) = \frac{\Gamma_\Sigma}{2\sqrt{K_\Sigma^2 - (\Delta\omega)^2}}$. Then,

$$\tau_{Y \to X} = \frac{k_{XY}^2 \Gamma_\Sigma \sqrt{1 - (\Delta\omega)^2/K_\Sigma^2}}{4 \Gamma_X K_\Sigma} \approx \frac{k_{XY}^2 \Gamma_\Sigma}{4 \Gamma_X K_\Sigma}, \qquad (21)$$

and for equal coupling coefficients and noise intensities $\tau_{Y \to X} = \frac{K_\Sigma \sqrt{1 - (\Delta\omega)^2/K_\Sigma^2}}{8} \approx \frac{K_\Sigma}{8}$. So, an interplay of the two factors—increase of the coupling coefficient and decrease of the phase difference variance—provides an unbounded (linear) increase of the TE rate in the effective synchronization domain $\tau_{Y \to X} \underset{K_\Sigma \to \infty}{\to} \frac{K_\Sigma}{8}$, i.e., the TE rate becomes infinitely large when an effective synchrony approaches the strict synchrony.

The phase difference relaxation time is $t_{char} = \frac{1}{\sqrt{K_\Sigma^2 - (\Delta\omega)^2}} = \underset{K \to \infty}{\to} \frac{1}{K_\Sigma}$, which is arbitrarily small for large enough couplings. At large enough times $t \gg t_{char}$, the finite-time TE is arbitrarily small, being inversely proportional to t as illustrated in Sec. IV E [Fig. 3(a)]. So, the finite-time TE estimated in practice is non-small only on small enough temporal horizons. Hence, one needs a very high resolution with respect to phases and time to detect a nonzero TE from a time series. Otherwise, one gets an approximate, coarse-grained[3] TE that effectively compares evolutions from the same initial conditions $w_1 \approx w_2$ and so turns out to be, indeed, zero. The theoretical non-small TE rate is probably not so essential in practical estimation, but it is of a basic interest to realize that the information flow in an effective synchronization regime for the basic Kuramoto model gets infinitely strong with rising coupling coefficients rather than arbitrarily weak.

The reduced TE equals $T_{Y \to X} = \tau_{Y \to X} t_{char} = 1/8$ nats for any sufficiently large $K_\Sigma > |\Delta\omega|$. This is a characteristic value for the basic case of effectively synchronized Kuramoto oscillators with equal coupling strengths and noise intensities. This number can be understood in practically clear terms as follows. The maximal finite-time TE is reasonably close to the reduced TE and is achieved approximately on a temporal horizon $t_{char} = 1/K_\Sigma$ (see Sec. IV E). Note that a finite-time TE is approximately equal to half the relative PI (up to a moderate factor if close to the synchronization transition, Sec. IV E). So, the relative PI (8) of the t_{char}-future phase of X is equal to 25%. If the non-normalized PI is divided by the individual prediction error variance instead of Eq. (8), it gives a normalized PI $P_{Y \to X}/(1 + P_{Y \to X})$ equal here to 20%. The relative PIs on smaller temporal horizons are obtained just as proportionally smaller percentages. It is interesting that in a highly synchronous regime (i.e., a very large K_Σ) leading to arbitrarily small variations of the current phase of Y around the current phase of X, such weakly varying current phase of Y contributes a constant non-negligible relative value of $\sim$20% to the future variance of X. Still, since the temporal horizon of that affected future decreases as $1/K_\Sigma$, the absolute (not relative)

values of the conditional variances and the contribution to variance (i.e., PI) get progressively smaller as well.

B. Transfer entropy rates below synchronization threshold

In general, the TE rate for the system (16) reads

$$\tau_{Y \to X} = \frac{k_{XY}^2 \text{var}(\sin \Delta \Phi_t | \phi_t^X)}{2 \Gamma_X}, \qquad (22)$$

extending Eq. (20), which holds only if $\text{var}(\Delta \Phi_t | \phi_t^X) \ll 1$. For the couplings well below the synchronization threshold $\tilde{K}_c = \tilde{\Delta}$, one has $\text{var}(\sin \Delta \Phi_t | \phi_t^X) = \text{var}(\sin \Delta \Phi_t) \approx \frac{1}{2}$ and

$$\tau_{Y \to X} \approx \frac{k_{XY}^2}{4 \Gamma_X}. \qquad (23)$$

For equal coupling coefficients and equal noise intensities, it reads $\tau_{Y \to X} = \frac{K_\Sigma^2}{8 \Gamma_\Sigma}$. For $\tilde{\Delta} = 50$, this relation holds true up to $\tilde{K} \approx 0.8 \tilde{\Delta}$ [Fig. 1(b)] and the upper bound seemingly approaches $\tilde{K} \approx \tilde{\Delta}$ with $\tilde{\Delta} \to \infty$. So, the TE rate reaches its maximum over $\tilde{K}$ roughly right before the threshold $\tilde{K}_c$ with the value roughly assessed as $\tau_{Y \to X} \approx \frac{(\Delta\omega)^2}{8 \Gamma_\Sigma} = \frac{\tilde{\Delta}^2 \Gamma_\Sigma}{32}$. Right after the threshold, it is roughly $\tau_{Y \to X} \approx \frac{\Delta\omega}{8} = \frac{\tilde{\Delta} \Gamma_\Sigma}{16}$. The ratio is $\frac{\tau_{Y \to X, before}}{\tau_{Y \to X, after}} \approx \frac{\tilde{\Delta}}{2}$, i.e., a large jump if $\tilde{\Delta}$ is large. As observed in Fig. 1(b), an actual value of this ratio is about $\frac{0.4 \tilde{\Delta}}{2}$ for $\tilde{\Delta} = 50$ and about $\frac{0.5 \tilde{\Delta}}{2}$ for $\tilde{\Delta} = 75$. So, one must apply an additional factor of about 1/2 to get the ratio at such finitely large $\tilde{\Delta}$.

To get the reduced TEs, the characteristic time can be estimated below $\tilde{K}_c$ as $t_{char} = \frac{1}{\Delta\omega}$, which is the time when the coupling term manifests its nonlinear character. Then, the reduced TE reads $T_{Y \to X} = \frac{K_\Sigma^2}{8 \Gamma_\Sigma \Delta\omega} = \frac{\tilde{K}^2}{16 \tilde{\Delta}}$ nats. So, right before $\tilde{K}_c$, one has roughly $T_{Y \to X} = \frac{\tilde{\Delta}}{16}$ nats. Right after $\tilde{K}_c$, the reduced TE is roughly $T_{Y \to X} = \frac{1}{8}$ nats since the characteristic time is $t_{char} = \frac{1}{K_\Sigma}$.

For a smaller $\tilde{\Delta}$, synchronization transition is smeared stronger, see, e.g., Fig. 1(a) (thick line) for $\tilde{\Delta} = 5$. The above approximations apply only to larger frequency mismatch-to-noise ratios, e.g., $\tilde{\Delta} > 10$. For smaller ones $1 < \tilde{\Delta} < 10$, one should divide the "right before $\tilde{K}_c$" values by larger values of about 4 to 5. For zero $\tilde{\Delta}$, only the effective synchronization domain is present [Fig. 1(a)].

C. Unequal coupling coefficients and noise intensities

If one varies coupling coefficients and noise intensities in such a way that their sums K_Σ and Γ_Σ remain constant, then the stationary PDF of the wrapped phase difference (17) does not change and the TE rates are obtained from the previous expressions via a simple rescaling. Let us first vary only the coupling coefficients and denote their ratio $r_{xy} = \frac{k_{xy}}{k_{yx}}$, so $k_{xy}/(K_\Sigma/2) = \frac{2 r_{xy}}{r_{xy}+1}$ and denote the TE rate for equal couplings found above as $\tau_{Y \to X}^{eq} = \tau_{Y \to X}^{eq}(K_\Sigma, \Gamma_\Sigma)$. Then, the TE rate for unequal couplings reads $\tau_{Y \to X} = \left(\frac{2 r_{xy}}{r_{xy}+1} \right)^2 \tau_{Y \to X}^{eq}$ with the ratio $\frac{\tau_{Y \to X}}{\tau_{X \to Y}} = r_{xy}^2$. For a unidirectional coupling ($r_{xy} = \infty$), the TE rate achieves the maximal value of $\tau_{Y \to X} = 4 \tau_{Y \to X}^{eq}$ with $\tau_{X \to Y} = 0$.

Let us now vary noise levels in such a way that $\Gamma_\Sigma = const$ at equal coupling coefficients and denote $r'_{xy} = \frac{\Gamma_X}{\Gamma_Y}$. Then, $\tau_{Y\to X} = \frac{r'_{xy}+1}{2r'_{xy}}\tau^{eq}_{Y\to X}$, $\frac{\tau_{Y\to X}}{\tau_{X\to Y}} = \frac{1}{r'_{xy}}$, and $\frac{1}{\tau_{Y\to X}} + \frac{1}{\tau_{X\to Y}} = \frac{2}{\tau^{eq}_{Y\to X}}$. So, an increase of the noise level decreases the TE rate twice at maximum, while a decrease of the noise level increases the TE rate arbitrarily strongly because the TE rate is inversely proportional to the noise intensity. If both noise levels and coupling coefficients are varied, then $\frac{\tau_{Y\to X}}{\tau_{X\to Y}} = \frac{r^2_{xy}}{r'_{xy}}$. For finite-time (e.g., maximal) TEs, such scaling with respect to r_{xy} and r'_{xy} applies approximately.

To get a closer "feeling" of the range of numerical values (as in Sec. IV A), consider an effective synchronization regime and the two coupling coefficients differing by the factor of $r_{xy} = a > 1$ for equal noise intensities. Then, the reduced TE is $T_{Y\to X} = \frac{a^2}{2(a+1)^2}$. In the limit of unidirectional coupling $Y \to X$, it reaches $T_{Y\to X} = 1/2$ nats, i.e., the relative PI is $e^{2T_{Y\to X}} - 1 \approx 1.7$ and so the normalized PI or the contribution of the current phase of Y to the future variance of the phase of X is 1.7/2.7 or 63%, while the opposite TE and PIs are zero. If in an effective synchronization regime, the noise intensities differ by the factor of $r'_{xy} = b$ under equal coupling coefficients, then the reduced TE $T_{Y\to X} = \frac{b+1}{16b}$. In the limit of zero noise intensity in the source Y, it reaches $T_{Y\to X} = 1/16$ nats, i.e., the corresponding normalized PI is about 6%, while in the opposite direction, the reduced TE is infinite and the normalized PI is 100% because the joint prediction error on an infinitesimally small temporal horizon is an infinitesimally small quantity of higher order.

D. Transfer entropies vs asymptotic effects on phase diffusion

Figure 2 presents the values of the reduced TE and the asymptotic effect on the phase diffusion (15). The latter is shown with solid lines and obtained as an empirical ensemble variance for the initial condition $w = \delta(\phi^X_0)\tilde{p}^{st}(\phi^Y_0)$ at large time $t_{\mathrm{int}} = 100/\Delta\omega$ divided by t_{int}. Recall that the phase diffusion constants for the uncoupled oscillators are equal precisely to $D_{X,0} = \Gamma_X$ and $D_{Y,0} = \Gamma_Y$.

A simple approximation below the synchronization threshold is that the phase diffusion constant equals roughly the sum $D_X \approx \frac{k^2_{xy}}{2|\Delta\omega|} + \Gamma_X$. The first term is the squared coupling term averaged with the uniform PDF (which gives the factor 1/2) and multiplied by the characteristic time for the non-synchronized region $|\Delta\omega|^{-1}$. So, the asymptotic effect equals just the doubled reduced TE,

$$D_{Y\to X} \approx \frac{k^2_{xy}}{2\Gamma_X|\Delta\omega|} = 2T_{Y\to X}. \qquad (24)$$

This approximation works quite well, e.g., for $\tilde{\Delta} \approx 5$ [Fig. 2(a), dashed line] up to half the threshold value of K_Σ. For weaker noises (namely, for $\tilde{\Delta} > 10$), this formula does not apply since the contribution of the coupling term to the phase diffusion is determined by a more complex interplay of the noise and coupling terms. An additional factor of 1/5 is applied to get an approximation shown with the dashed line in Fig. 2(b). In general, the factor of $10/\tilde{\Delta}$ seems to be applicable for $\tilde{\Delta} > 10$.

For an effective synchronization, the diffusion constants read $D_X = D_Y = \frac{k^2_{xy}\Gamma_{yy}+k^2_{yx}\Gamma_{xx}}{(k_{xy}+k_{yx})^2}$ [see, e.g., Eq. (9.15) in Ref. 47]. Hence, an asymptotic effect on phase diffusion reads

$$D_{Y\to X} = \frac{k^2_{xy}\Gamma_{yy}/\Gamma_{xx} + k^2_{yx}}{(k_{xy}+k_{yx})^2} - 1, \qquad (25)$$

which can be rewritten as

$$D_{Y\to X} = \frac{1 + r^2_{xy}/r'_{xy}}{(1 + r_{xy})^2} - 1. \qquad (26)$$

Recalling the TE rates' ratio $\frac{\tau_{Y\to X}}{\tau_{X\to Y}} = \frac{r^2_{xy}}{r'_{xy}}$, one can see that the asymptotic effect (26) essentially depends on the TE rates. To explore it in more detail, consider for definiteness $r^2_{xy} \gg r'_{xy}$, i.e., the coupling $Y \to X$ is predominant with respect to the TE rate. There are two following cases.
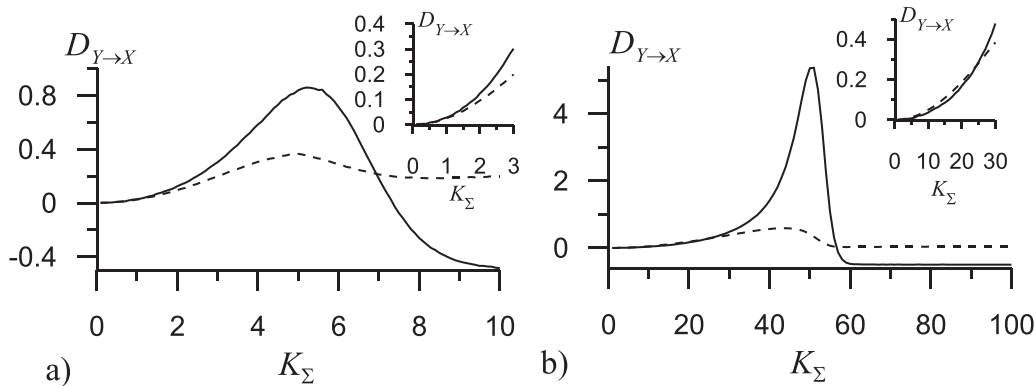

FIG. 2. Asymptotic effect of coupling on the phase diffusion for system (16) with equal noise intensities $\Gamma_X = \Gamma_Y = 1$ and equal coupling coefficients: (a) $\Delta\omega = 5$ and (b) $\Delta\omega = 50$. Each inset shows an amplified fragment of the respective plot. Dashed lines show approximations via the reduced TE (24), with an additional factor of 0.2 in panel (b). The asymptotic DCE equals $-1/2$ units in an effective synchronization regime, i.e., the phase diffusion constant decreases twice due to switching the coupling on.

If $r_{xy} \gg 1$, i.e., the coupling $Y \to X$ is also predominant with respect to the coupling coefficient, one has $D_{Y \to X} \approx -1 + 1/r'_{xy}$, which is very large as $D_{Y \to X} \approx 1/r'_{xy}$ for a relatively weaker noise in X ($r'_{xy} \ll 1$) indicating that the phase diffusion in X strengthens $1/r'_{xy}$ times due to the coupling from the higher diffusion source Y and becomes equal to $D_{Y,0}$. The opposite DCE $D_{X \to Y} \approx -2/r_{xy}$ is much smaller than unity, so the phase diffusion constant of the weakly influenced high-diffusion oscillator Y remains almost the same becoming the constant for both coupled oscillators. For a relatively stronger noise in X ($r'_{xy} \gg 1$), the effect $D_{Y \to X} \approx -1 + 1/r'_{xy}$ is close to -1 indicating that the phase diffusion of X weakens r'_{xy} times due to the coupling from the lower-diffusion source Y and gets equal again to $D_{Y,0}$. The opposite DCE is again $D_{X \to Y} \approx -2/r_{xy}$, so the phase diffusion constant of the weakly influenced low-diffusion oscillator Y remains the same and applies to both synchronized oscillators.

If $r_{xy} \ll 1$, i.e., the coupling $Y \to X$ is deficient with respect to the coupling coefficient, and X is a lower-diffusion oscillator, the asymptotic effect just equals the ratio of TE rates $D_{Y \to X} = \frac{\tau_{Y \to X}}{\tau_{X \to Y}}$, i.e., the phase diffusion in X rises $\frac{r^2_{xy}}{r'_{xy}}$ times due to the coupling from the higher diffusion source Y, which is deficient with respect to r_{xy} but predominant with respect to TE rates. In the opposite direction, one has $D_{X \to Y} \approx -1 + r^2_{xy}$, i.e., the phase diffusion in Y weakens r^2_{xy} times due to the coupling from a lower-diffusion source X. So, the phase diffusion constants of both oscillators in an effective synchronization regime are equal to each other and so to $r^2_{xy} D_{Y,0} \ll D_{Y,0}$ and to $\frac{r^2_{xy} D_{X,0}}{r'_{xy}} \gg D_{X,0}$.

In any case, the TE rates play an important role in determining an asymptotic effect on phase diffusion in a synchronized regime, though the relation is not as simple as Eq. (24).

E. Finite-time transfer entropies

Let us relate the TE rates to finite-time TEs on various temporal horizons as in Refs. 40, 41, and 43. The finite-time TE reaches its maximum near t_{char} both in non-synchronized and effectively synchronized regimes [Figs. 3(a), 3(c), 3(d), and 3(f)]. Most often, the maximum time is twice as large as t_{char}, and it is about five times as large near the synchronization threshold [Figs. 3(b) and 3(e)]. At large $\tilde{\Delta}$ and $\tilde{K}$ near the synchronization threshold, the actual value of the maximal TE is about twice as small as that of the reduced TE [Fig. 3(e)]; otherwise, both values are much closer to each other [Figs. 3(a)–3(d), and 3(f)].

The mean-squared approximation of the TE [right-hand side of Eq. (9)] is very close to the actual value of the TE, apart from the vicinity of the synchronization threshold. The approximate TE reaches its maximum close to t_{char} (Fig. 3, dashed lines). This logarithmic relative PI[21] is an approximation of the TE exactly as the logarithmic Granger causality is an approximation of the infinite-history TE in Refs. 43 and 81. Figure 4 provides the values of the reduced and maximal TEs in a range of coupling coefficient values confirming the above statements. Synchronization threshold and its vicinity differ from other domains by an essentially non-Gaussian character of the conditional PDFs for both phases as illustrated in more detail in Appendix B.
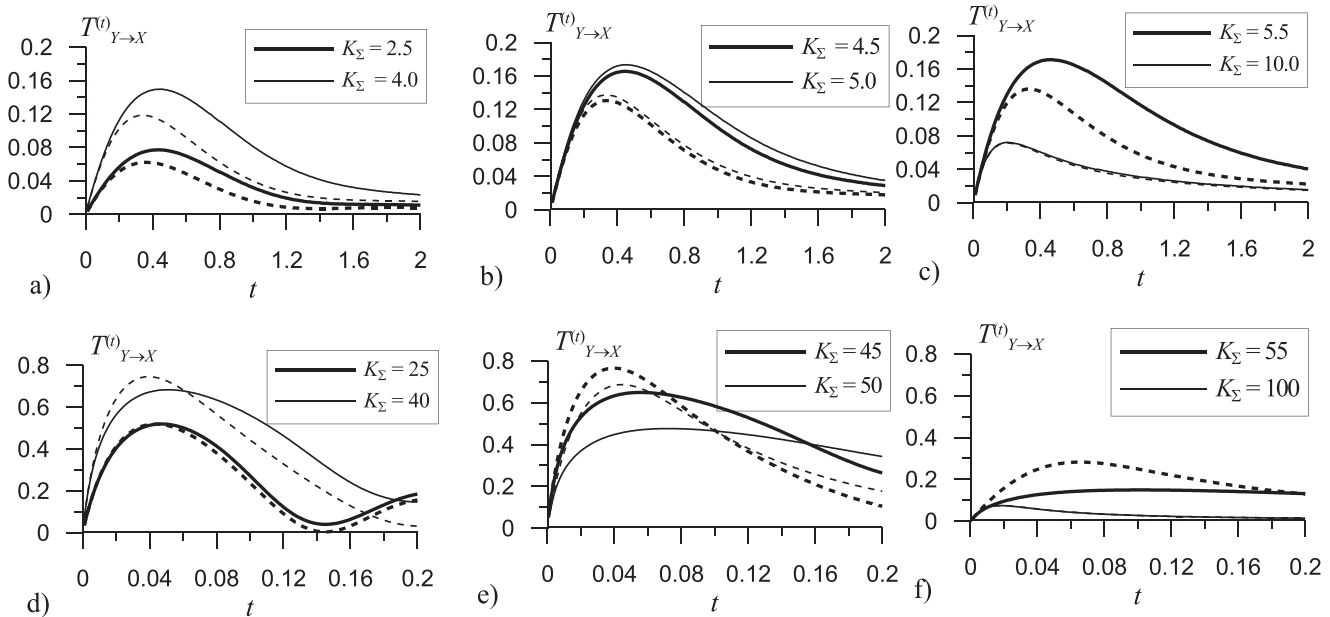

FIG. 3. Finite-time TEs for system (16) with $\Gamma_X = \Gamma_Y = 1$, $\Delta\omega = 5$ (upper row) and $\Delta\omega = 50$ (bottom row): (a) $K_\Sigma = 2.5$ (thick lines) and $K_\Sigma = 4.0$ (thin lines); (b) $K_\Sigma = 4.5$ (thick lines) and $K_\Sigma = 5.0$ (thin lines); (c) $K_\Sigma = 5.5$ (thick lines) and $K_\Sigma = 10.0$ (thin lines); (d) $K_\Sigma = 25$ (thick lines) and $K_\Sigma = 40$ (thin lines); (e) $K_\Sigma = 45$ (thick lines) and $K_\Sigma = 50$ (thin lines); and (f) $K_\Sigma = 55$ (thick lines) and $K_\Sigma = 100$ (thin lines). Logarithmic relative PI approximations given by the rhs of Eq. (9) are shown with dashed lines.
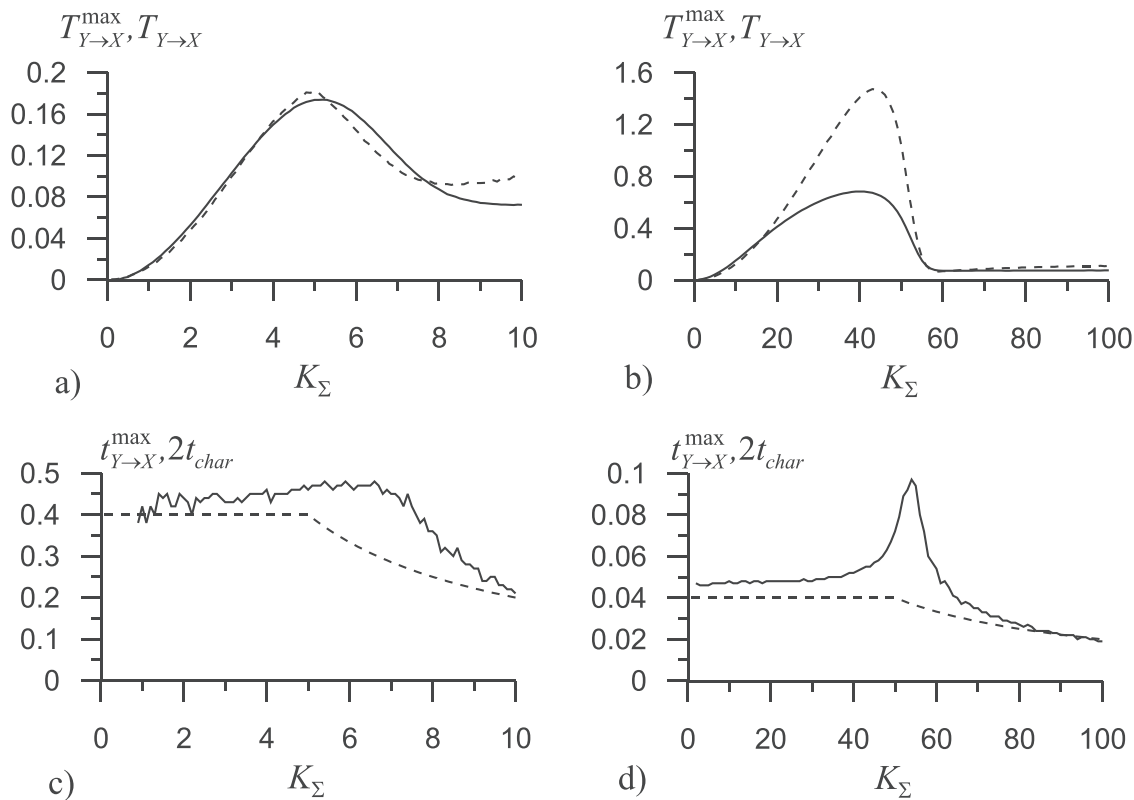
FIG. 4. Maximal TEs [(a) and (b)] and their response times [(c) and (d)] (solid lines) and reduced TEs [(a) and (b)] and estimates of the maximum response times [(c) and (d)] obtained as doubled characteristic times (dashed lines) for system (16) with $\Gamma_X = \Gamma_Y = 1$ and equal coupling coefficients: $\Delta\omega = 5$ [(a) and (c)] and $\Delta\omega = 50$ [(b) and (d)].

For large $\tilde{\Delta}$, the differential quantifier reaches its maximum on a temporal horizon considerably greater than that for the TE [Figs. 5(a) and 5(b)]. The plots for the differential quantifier differ from those for the TE in the domains near and well above the synchronization threshold. In particular, the values of the former differ strongly depending on whether the average in its definition is performed with the uniform [as in the original version, solid lines in Figs. 5(a) and 5(b)] or the stationary [Appendix A, dashed lines in Figs. 5(a) and 5(b)] PDF. This is because for a large temporal horizon t, there is a narrow region of the large derivative of the conditional expectation [Figs. 5(c) and 5(d), solid lines] whose squared value contributes strongly or weakly to the integral (3) depending on its weight [Figs. 5(c) and 5(d), dashed lines]. If this narrow region corresponds to a large enough value of the stationary PDF, the stationary averaged quantifier (A5) is greater [Figs. 5(a) and 5(c)]. Otherwise, the uniformly averaged quantifier (3) is greater [Figs. 5(b) and 5(d)] Nonetheless, the differential quantifier differs from the squared-coefficients quantifier (and, hence, from non-normalized PI) in general. However, in practice, when one uses low-order trigonometric polynomials in an empirical model (2), it means that a "smoothed" approximate PDF is used for a coupling characterization, giving smaller values of the corresponding

approximate differential quantifier and making it closer to the corresponding approximate squared-coefficients quantifier. So, many works based on the differential quantifier[1,2,4–6,9–14,19,20,22,23] have estimated in fact the (approximate) squared-coefficients quantifier or the non-normalized PI (up to some moderate distortion), i.e., again a quantity close to the (approximate) finite-time TE finite-time TE (up to normalization).

V. DISCUSSION

The transfer entropy rate is a dynamically informative characteristic of directional (causal) coupling in the paradigmatic example of Kuramoto oscillators. Within the framework of dynamical causal effects, it is revealed that the reduced TE equals half the value of an asymptotic effect of coupling on the phase diffusion constant of a coupling recipient in a non-synchronized regime for a high enough noise level. Thereby, that asymptotic effect can be often estimated in practice (including meta-analysis of the published works) on the basis of reported values of the finite-time TEs.

It is also found that the information flow captured by the TE rate can be arbitrarily large in an effective synchronization regime and rises with coupling coefficients, rather than being zero or very
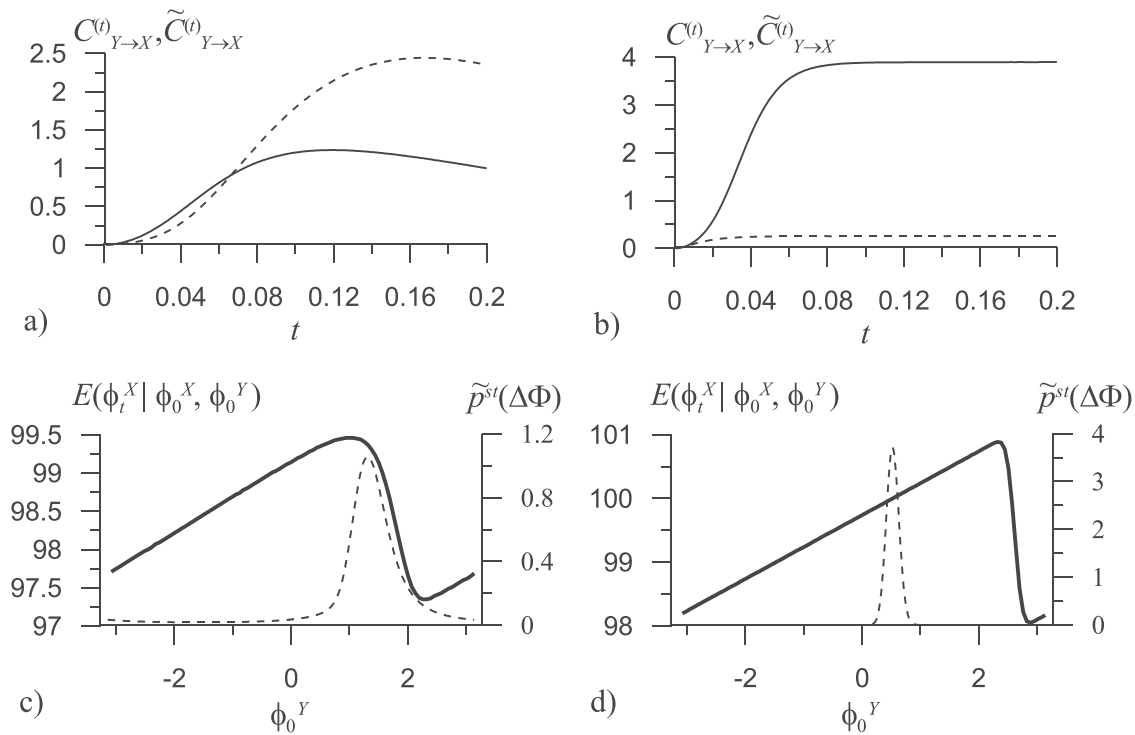
FIG. 5. Differential quantifiers with uniform [solid lines in (a) and (b)] and stationary [dashed lines in (a) and (b)] weighting PDF for system (16) with $\Gamma_X = \Gamma_Y = 1$, equal coupling coefficients, and $\Delta\omega = 50$: (a) and (c) $K_\Sigma = 50$; (b) and (d) $K_\Sigma = 100$. In panels (c) and (d), conditional PDFs at $t = 0.2$ are shown with solid lines and stationary PDFs (17) with dashed lines.

small as it is often thought. The reason is that the coarse-grained TE is indeed small in such regime, while a sufficiently high resolution with respect to phases and time reveals a strong information flow occurring at small scales. The latter is necessary to maintain a high degree of an effective synchronization. A characteristic value of the reduced TE in an effective synchronization regime is found to be 1/8 nats for equal noise intensities and coupling strengths (Sec. IV A). The range of the reduced TE values for unequal coupling coefficients and noise intensities are given in Sec. IV C. Such values for synchronization regimes are of theoretical interest, not of great practical importance. Still, the entire picture of these characteristic values and the TE plots for non-synchronous regimes (Figs. 3 and 4) can serve a reference for dynamically meaningful interpretations of the TE numerical values encountered in practice.

Note also that a relatively large value of the reduced (or finite-time) TE is not compulsorily a sign of so large role of the coupling in dynamics. In particular, consider an effective synchronization regime where the noise level in a coupling recipient X is close to zero ($r'_{xy} \ll 1$), the respective coupling coefficient is also very small ($r_{xy} \ll 1$), but the ratio $\frac{r^2_{xy}}{r'_{xy}}$ is very large. The latter is just equal to the ratio of the reduced TEs (Sec. IV C) and so $T_{Y \to X}$ is much greater than $T_{X \to Y}$. However, synchronization is established entirely due to the coupling $X \to Y$ whose coefficient is much

greater while the opposite coupling coefficient is negligibly small. If one switches the coupling $X \to Y$ (with arbitrarily small $T_{X \to Y}$) off, the dynamical regime gets non-synchronous, while switching the coupling $Y \to X$ (with arbitrarily large $T_{Y \to X}$) off changes nothing. In general, a large TE value implies a large *short-term* effect of the coupling, and only under certain conditions (studied in Sec. IV D), it evidences an important overall dynamical role of the coupling.

Other coupling quantifiers including the normalized squared-coefficients quantifier (relative PI) and the differential quantifier have also a clear interpretation as transient DCEs. These quantifiers are shown to be close to the TE well below the synchronization threshold for the low-order nonlinearity in the system under study. The quantifiers may differ from each other stronger (especially the differential quantifier) for higher-order nonlinearities. Further studies of these quantifiers for the general phase oscillators (1) seem to be of interest. Moreover, it would be relevant to study these quantifiers for some physical model oscillators (as, e.g., in Ref. 46) whose dynamics is only approximately described with phase equations with corresponding constraints on parameter values, noise properties, etc., Thereby, one could also get physical interpretations of the phase-dynamic causalities. From a dynamical systems side, this research can be continued for more general limit cycle oscillators[77,78] and for more than two coupled oscillators in the same manner

as done here promising practically useful results generalizing the results presented in this study.

Note that the parameter dependencies of the coupling quantifiers differing between the non-synchronous and effectively synchronous regimes (e.g., Fig. 1, below and above the synchronization threshold) are not observed for linear systems. Thus, Appendix C shows that in a two-dimensional linear SDS formally quite similar to Eq. (16), such dependence may look either like that in a non-synchronous regime [Fig. 8(a), proportional coupling] or like that in a synchronous regime [Fig. 8(b), difference coupling]. In the linear system, this character depends on coupling parameterization and the two types of the dependence are not observed in the same plot, contrary to the phase oscillators (16). So, the nonlinear system (16) exhibits a richer set of relations between coupling quantifiers and coupling coefficients.

Finally, it is possible to assess a time series length necessary to obtain sufficiently accurate estimates of the coupling quantifiers. It definitely depends on how densely all bins of the histogram of the phase difference $\Delta\Phi$ constructed from a time series are populated. As a rough estimate, one can take that $\Delta\Phi_t$ must wrap around the 2π interval at least ten times, i.e., the phase difference must change by 20π within a time series. Its expected increment is the sum of Γt and $|\Delta\omega|t$. If the frequency mismatch term is greater, the minimal time series length is assessed as $\approx 20\pi/|\Delta\omega|$ or $\approx 10 T_{beat}$, where $T_{beat} = 2\pi/|\Delta\omega|$ is a "beating" period. Note that such time series length does not depend on the values of both periods themselves contrary to multiple previous attempts to relate it to the number of basic periods. Still, if the frequency mismatch is about 0.2 of the mean frequency (as that in the basic example of Kuramoto oscillators in Refs. 1 and 4), then the beating period is about five times as great as the mean period T, so one can assess the necessary time series length as $\approx 50T$, exactly as obtained from numerical simulations and statistical estimation experiments in Refs. 4, 10, and 14. For a too small (e.g., zero) frequency mismatch, the time series length required is determined by the noise level rather than by $\Delta\omega$. Going a step further, if a coupling function in the direction $Y \to X$ includes strong "modulating" terms like $\cos(\phi_t^Y + \delta)$, the necessary time series length is determined by the frequency of the coupling source Y itself.

VI. CONCLUSIONS

Various quantifiers of directional couplings based on phase-dynamics description are fruitfully used in time series analysis to detect couplings and estimate temporal variations of their "strength" or "importance" in terms of these quantifiers. However, it is quite desirable to have meaningful dynamical interpretations of their numerical values. Such an opportunity is provided within the framework of dynamical causal effects.[39–44] A particularly vivid result of such consideration is a quantitative relation between certain quantifier of interest and some asymptotic effect of the directional coupling on the entire dynamics of the coupling recipient.

Here, several phase-dynamic coupling quantifiers (mainly the TE rate and the TE reduced to a characteristic time) are related to an asymptotic effect of a coupling on phase diffusion of the coupling recipient for a paradigmatic system of two stochastic Kuramoto oscillators with constant couplings. In particular, in a

non-synchronized regime, the reduced TE equals half the asymptotic effect on phase diffusion for a high enough ratio of the total noise intensity to the frequency mismatch. In an effectively synchronized regime, the dependency is more complicated but still strongly determined by the ratio of the two reduced TEs. It is also revealed that the information flow expressed by the TE rate (for high enough resolution with respect to phases and time) unboundedly rises with the coupling coefficients in the effective synchronization domain contrary to a widespread thought that the information flow decreases to zero above the synchronization threshold. A set of characteristic values of the reduced TEs is also reported to serve as a reference for practical interpretations.

Similar questions concerning oscillators with higher-order nonlinearities and amplitude dynamics including, more than two interacting oscillators, various noise properties deserve further studies that should make a practical analysis of directional couplings from time series more informative.

ACKNOWLEDGMENTS

APPENDIX A: FORMAL DETAILS

The function $F_X^{(t)}$ in Eq. (2) can be generally written as

$$F_X^{(t)}(\phi_0^X, \phi_0^Y) = a_0 + \sum_{m,n}(a_{m,n}\cos(m\phi_0^X + n\phi_0^Y) + b_{m,n}\sin(m\phi_0^X + n\phi_0^Y)), \quad (A1)$$

where m and n are arbitrary integers and the polynomial coefficients depend on t. For system (16) and t small enough, one can get $a_0 = \omega_X t$, $b_{-1,1} = k_{xy}t$, and all other coefficients are zero. However, the polynomial (A1) at large t even for system (16) may include many considerably nonzero coefficients exhibiting a saw-like plot [Figs. 5(c) and 5(d)]. The squared-coefficients quantifier[21] reads

$$c_{Y \to X}^{(t)} = \frac{1}{2} \sum_{(m,n), n \neq 0} (a_{m,n}^2 + b_{m,n}^2), \quad (A2)$$

i.e., it sums over all coupling terms. For weak enough couplings, all terms on the right-hand side of Eq. (A1) are mutually orthogonal, so Eq. (A2) represents the total contribution of all coupling terms to the variance of the phase increment,[21]

$$\text{var}(\Delta\phi_t^X|\phi_0^X) = c_{Y \to X}^{(t)} + \left\langle \text{var}(\varepsilon_t^X|\phi_0^X, \phi_0^Y) \right\rangle_{\phi_0^Y}. \quad (A3)$$

As suggested in Ref. 21 it can be divided by the noise variance to give the relative PI $G_{Y \to X}^{(t)}$ (8). For stronger couplings, the sum of squared-coefficients (A1) may differ from the mean-squared PI. So, the latter quantifier is more widely applicable to characterize an effect of coupling on the dynamics. As another possibility, it was suggested[21] (similarly to Ref. 18) to divide $c_{Y \to X}^{(t)}$ by the mean phase increment $\left\langle F_X^{(t)}(\phi_0^X, \phi_0^Y) \right\rangle$, which is equal to $\omega_X t$ for weak enough couplings. However, contrary to TE and PI, such a quantifier does not simply relate to any dynamical characteristics of system (16) and is not further discussed here.

Concerning the differential quantifier (3), it reads

$$C_{Y \to X}^{(t)} = \frac{1}{2} \sum_{(m,n), n \neq 0} n^2 (a_{m,n}^2 + b_{m,n}^2). \qquad (A4)$$

It selects the same squared-coefficients of coupling, but higher-order coefficients enter with greater weights. Therefore, if a low-order approximating polynomial is used in practice,[1,2,4–6,9–13,19,20,23] then the differential quantifier gets very close to the squared-coefficients' one and, hence, to the non-normalized PI. For system (16) and small t, both quantifiers coincide. In general, the differential quantifier may strongly differ from $c_{Y \to X}^{(t)}$, depending on whether $F_X^{(t)}$ has steep local slopes.

As for the weighting functions in the TE (11) and the differential quantifier (13), the former is averaged with the stationary PDF of the phase difference and the latter with the uniform PDF. One could perform either choice in each case. Thus, another version of the differential quantifier reads

$$\tilde{C}_{Y \to X}^{(t)} = \lim_{\delta \phi \to 0} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{(E[\phi_t^X | w_2] - E[\phi_t^X | w_1])^2}{(\delta \phi)^2} \tilde{p}_{XY}^{st}(\phi_{0,1}^X, \phi_{0,1}^Y) d\phi_{0,1}^X d\phi_{0,1}^Y.$$

$$(A5)$$

It is briefly compared to the original version (3) in Fig. 5 to demonstrate that the differential quantifier is sensitive to the location of a steep slope, and it is quite difficult to relate it to any long-term dynamical characteristics. Similarly, uniformly averaged versions of TE and PI are possible. They are not discussed here in more detail.

APPENDIX B: NUMERICAL TECHNIQUES

Formula (17) is used with a normalizing constant C, which provides $\int_{-\pi}^{\pi} \tilde{p}^{st}(\Delta\Phi) d(\Delta\Phi) = 1$. The integral on the right-hand side of Eq. (17) is numerically computed here and its value is just the actual value of C. The stationary PDF obtained in this way coincides with

several exemplary PDFs reported in Ref. 45. Figures 6(a) and 6(b) present the plots of the stationary PDF of the wrapped phase difference for different frequency mismatches and coupling coefficients. Figure 6(c) shows the mean phase coherence $\rho = |\langle e^{i\Delta\Phi} \rangle|$, i.e., the amplitude of the first Fourier mode of $\tilde{p}^{st}(\Delta\Phi)$, obtained directly via integration of $\tilde{p}^{st}(\Delta\Phi) \cos \Delta\Phi$ and $\tilde{p}^{st}(\Delta\Phi) \sin \Delta\Phi$. The synchronization transition at $K_\Sigma = \Delta\omega$ is smeared due to the noise. A full stationary PDF of the two wrapped phases can be found from the condition that both marginal PDFs are uniform $\tilde{p}_X^{st}(\tilde{\phi}_0^X) = \tilde{p}_Y^{st}(\tilde{\phi}_0^Y) = 1/(2\pi)$ as follows from the invariance of Eq. (16) under any translation of the full initial state $(\tilde{\phi}_0^X, \tilde{\phi}_0^Y)$.

Conditional PDFs of the future states are obtained via numerical integration of stochastic differential equations with the Euler–Maruyama technique with an appropriately small step size of $0.01/\Delta\omega$. For each initial state $(\phi_{0,1}^X, \phi_{0,1}^Y)$, 10^5 time realizations are generated to obtain a histogram with 100 equidistant bins. It is set to $\phi_{0,1}^X = 0$, while $\phi_{0,1}^Y$ is changed from $-\pi$ to π in steps of 0.02π. To obtain functionally conditional quantities, the usual conditional quantities are averaged over $\phi_{0,1}^Y$ using either the stationary PDF of the phase difference $\Delta\Phi$ (for the TE and PI) or uniform PDF (for the differential quantifier). Averaging over different $\phi_{0,1}^X$ is not necessary since the results are independent of this value. As an alternative approach, a solution to the Fokker–Planck equation (FPE) gives a PDF $p_{XY}^{(t)}(\phi_t^X, \phi_t^Y | \phi_0^X, \phi_0^Y)$ at any time $t > 0$. Two approximate methods were used here to solve the FPE for a wider comparison: Gaussian approximation and second-order moment approximation.[87] It allows us to see the difference of a conditional PDF from Gaussian.

Figure 7 presents conditional PDFs on different temporal horizons. Even for the simple system (16) with lowest-order nonlinearity, those PDFs appear to be essentially non-Gaussian on non-small temporal horizons near the synchronization threshold and often exhibit steep slopes [Fig. 5(c)]. Then, the two approximate FPE solutions diverge. In other cases, those PDFs look either slowly varying like a wider Gaussian or strongly localized like a narrower Gaussian. Then, the TE is close to its mean-squared approximation (9) and both approximate FPE solutions are quite accurate.
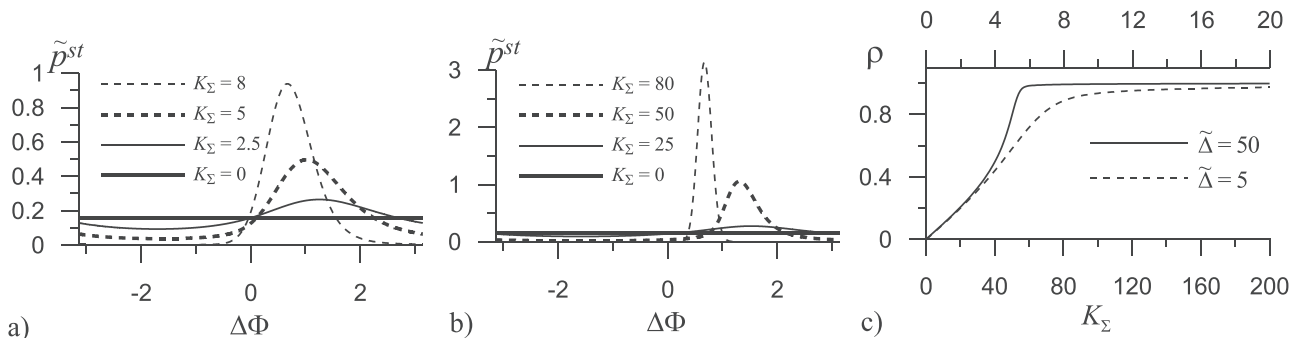

FIG. 6. Stationary PDFs of the wrapped phase difference for system (16) with $\Gamma_X = \Gamma_Y = 1$ and equal coupling coefficients: (a) $\Delta\omega = 5$ and (b) $\Delta\omega = 50$. Mean phase coherence (c) is shown for $\Delta\omega = 5$ (dashed line, upper abscissa axis) and $\Delta\omega = 50$ (solid line, lower abscissa axis).
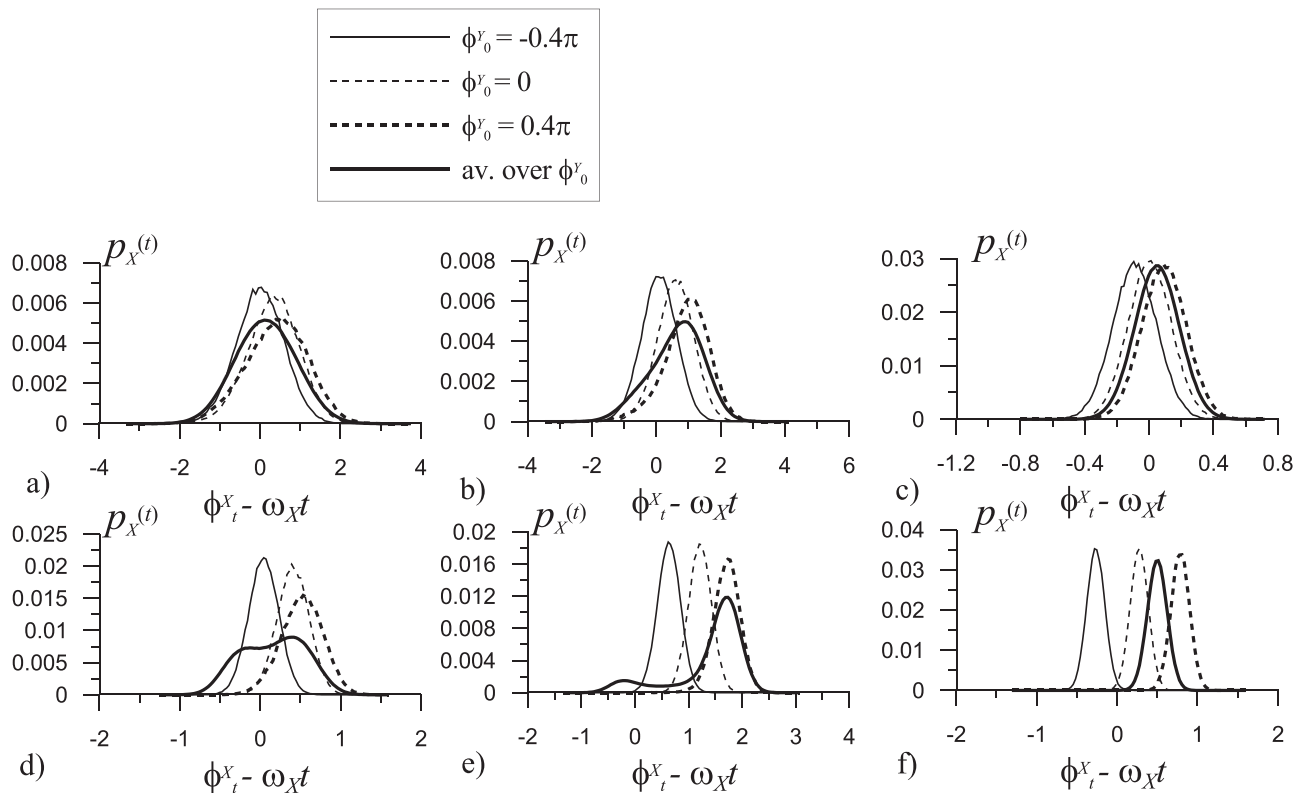
FIG. 7. Conditional PDFs on temporal horizons corresponding to the maximum of the finite-time TE (Fig. 3) for system (16) with $\Gamma_X = \Gamma_Y = 1$ and equal coupling coefficients. Upper row is shown for $\Delta\omega = 5$: (a) $K_\Sigma = 2.5$, (b) $K_\Sigma = 5.0$, and (c) $K_\Sigma = 10.0$; bottom row is shown for $\Delta\omega = 50$: (d) $K_\Sigma = 25$, (e) $K_\Sigma = 50$, and (f) $K_\Sigma = 100$.

APPENDIX C: COMPARISON TO LINEAR EXAMPLE

Figure 8 presents dependencies of the TE rates on the coupling coefficient for linear relaxation systems. The latter are taken with two parameterizations of couplings: a proportional coupling,

$$\dot{x} = -\alpha_x x + k_{xy} y + \xi_x(t),$$
$$\dot{y} = -\alpha_y y + k_{yx} x + \xi_y(t), \tag{C1}$$

and a difference coupling,

$$\dot{x} = -\alpha_x x + k_{xy}(y - x) + \xi_x(t),$$
$$\dot{y} = -\alpha_y y + k_{yx}(x - y) + \xi_y(t), \tag{C2}$$

with $\alpha_x = 1.1, \alpha_y = 0.9$, $\Gamma_{xx} = \Gamma_{yy} = 1$, and $k_{xy} = k_{yx} = k$. Concrete values of the relaxation rates do not influence the conclusions.
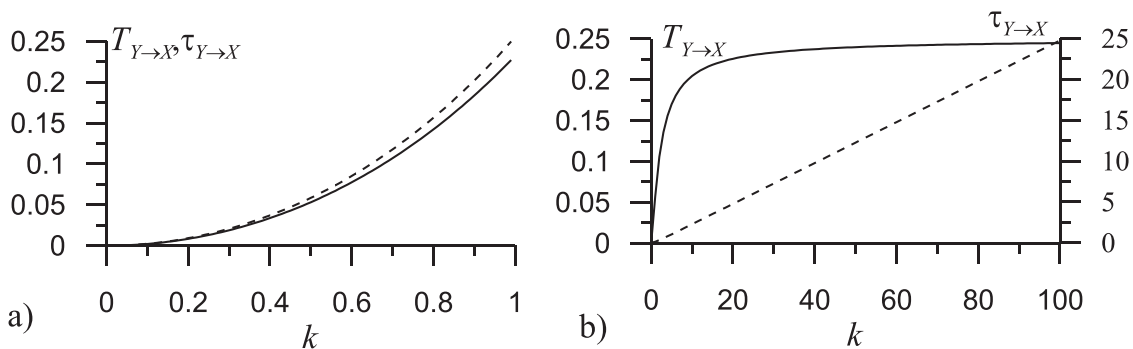

FIG. 8. TE rates (dashes) and reduced TEs (solid lines) for the linear systems (C1) and (C2): (a) proportional coupling (C1) and (b) difference coupling (C2).

One can see that the phase oscillators' curve [e.g., Fig. 1(b)] combines the linear systems' curves of the two types. The proportional coupling curve [Fig. 8(a)] is similar to the non-synchronized phase oscillators' curve [left part of Fig. 1(b)], and the difference coupling curve [Fig. 8(b)] is analogous to the synchronized oscillators' curve [right part of Fig. 1(b)]. Both analogies are explained based on the asymptotic conditions of either almost independent or very close phases of the two phase oscillators as discussed in Secs. IV A and IV B. So, nonlinearity of system (16) makes these parameter dependencies qualitatively different from the plots for any separate linear system (C1) or (C2) despite apparently quite similar forms of the evolution equations.

DATA AVAILABILITY

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

REFERENCES

[1] M. G. Rosenblum and A. S. Pikovsky, Phys. Rev. E **64**, 045202R (2001).
[2] M. G. Rosenblum, L. Cimponeriu, A. Bezerianos, A. Patzak, and R. Mrowka, Phys. Rev. E **65**, 041909 (2002).
[3] M. Paluš and A. Stefanovska, Phys. Rev. E **67**, 055201(R) (2003).
[4] D. A. Smirnov and B. P. Bezruchko, Phys. Rev. E **68**, 046209 (2003).
[5] B. P. Bezruchko, V. I. Ponomarenko, A. S. Pikovsky, and M. G. Rosenblum, Chaos **13**, 179 (2003).
[6] L. Cimponeriu, M. Rosenblum, T. Fieseler, J. Dammers, M. Schiek, M. Majtanik, P. Morosan, A. Bezerianos, and P. Tass, Prog. Theor. Phys. **150**, 22 (2003).
[7] T. Kiemel, K. M. Gormley, L. Guan, T. L. Williams, and A. H. Cohen, J. Comput. Neurosci. **15**, 233 (2003).
[8] L. Cimponeriu, M. Rosenblum, and A. Pikovsky, Phys. Rev. E **70**, 046213 (2004).
[9] D. A. Smirnov and R. G. Andrzejak, Phys. Rev. E **71**, 036207 (2005).
[10] D. A. Smirnov, M. B. Bodrow, J. L. P. Velazquez, R. A. Wennberg, and B. P. Bezruchko, Chaos **15**, 024102 (2005).
[11] I. I. Mokhov and D. A. Smirnov, Geophys. Res. Lett. **33**, L03708, https://doi.org/10.1029/2005GL024557 (2006).
[12] J. Brea, D. F. Russell, and A. B. Neiman, Chaos **16**, 026111 (2006).
[13] D. Smirnov, B. Schelter, M. Winterhalder, and J. Timmer, Chaos **17**, 013111 (2007).
[14] D. A. Smirnov, I. A. Karpeev, and B. P. Bezruchko, Tech. Phys. Lett. **33**(2), 147 (2007).
[15] M. Timme, Phys. Rev. Lett. **98**, 224101 (2007).
[16] I. T. Tokuda, S. Jain, I. Z. Kiss, and J. L. Hudson, Phys. Rev. Lett. **99**, 064101 (2007).
[17] B. Kralemann, L. Cimponeriu, M. Rosenblum, A. Pikovsky, and R. Mrowka, Phys. Rev. E **76**, 055201 (2007).
[18] B. Kralemann, L. Cimponeriu, M. Rosenblum, A. Pikovsky, and R. Mrowka, Phys. Rev. E **77**, 066205 (2008).
[19] D. A. Smirnov, U. B. Barnikol, T. T. Barnikol et al., Europhys. Lett. **83**, 20003 (2008).
[20] B. P. Bezruchko, V. I. Ponomarenko, M. D. Prokhorov, D. A. Smirnov, and P. A. Tass, Phys. Usp. **51**, 304 (2008).
[21] D. A. Smirnov and B. P. Bezruchko, Phys. Rev. E **79**, 046204 (2009).
[22] S. S. Kozlenko, I. I. Mokhov, and D. A. Smirnov, Izv. Atmos. Ocean. Phys. **45**, 704 (2009).
[23] P. Tass, D. Smirnov, A. Karavaev et al., J. Neural Eng. **7**, 016009 (2010).
[24] B. Kralemann, M. Rosenblum, and A. Pikovsky, Chaos **21**, 025104 (2011).
[25] Z. Levnajić and A. Pikovsky, Phys. Rev. Lett. **107**, 034101 (2011).
[26] B. Kralemann, M. Fruehwirth, A. Pikovsky et al., Nat. Commun. **4**, 2418 (2013).
[27] M. Palus, Phys. Rev. Lett. **112**, 078702 (2014).
[28] B. Kralemann, A. Pikovsky, and M. Rosenblum, New J. Phys. **16**, 085013 (2014).
[29] E. V. Sidak, D. A. Smirnov, and B. P. Bezruchko, J. Commun. Technol. Electron. **62**, 241 (2017).
[30] J. Pearl, Causality: Models, Reasoning, and Inference (Cambridge University Press, Cambridge, 2000).
[31] P. Spirtes, C. Glymour, and R. Scheines, Causation, Prediction, and Search (The MIT Press, Cambridge, 2000).
[32] N. Ay and D. Polani, Adv. Complex Syst. **11**, 17 (2008).
[33] J. T. Lizier and M. Prokopenko, Eur. Phys. J. B **73**, 605 (2010).
[34] M. Eichler, in Causality Statistical Perspectives and Applications, edited by C. Berzuini, P. Dawid, and L. Bernardinelli (John Wiley & Sons Ltd, Chichester, 2012), Chap. 22, p. 327.
[35] D. Janzing, D. Balduzzi, M. Grosse-Wentrup, and B. Schoelkopf, Ann. Stat. **41**, 2324 (2013).
[36] J. Runge, Phys. Rev. E **92**, 062829 (2015).
[37] A. Hannart, J. Pearl, F. Otto, P. Naveau, and M. Ghil, Bull. Am. Meteorol. Soc. **97**, 99 (2016).
[38] J. Runge, S. Bathiany, E. Bollt, G. Camps-Valls, D. Coumou, E. Deyle, C. Glymour, M. Kretschmer, M. D. Mahecha, J. Munoz-Mari, E. H. van Nes, J. Peters, R. Quax, M. Reichstein, M. Scheffer, B. Scholkopf, P. Spirtes, G. Sugihara, J. Sun, K. Zhang, and J. Zscheischler, Nat. Commun. **10**, 2553 (2019).
[39] D. A. Smirnov, Phys. Rev. E **90**, 062921 (2014).
[40] D. A. Smirnov and I. I. Mokhov, Phys. Rev. E **92**, 042138 (2015).
[41] D. A. Smirnov, Chaos **28**, 075303 (2018).
[42] D. A. Smirnov, Europhys. Lett. **128**, 20006 (2019).
[43] D. A. Smirnov, Phys. Rev. E **102**, 062139 (2020).
[44] D. A. Smirnov, "Generative formalism of causality quantifiers for processes," Phys. Rev. E (submitted).
[45] R. L. Stratonovich, Topics in Theory of Random Noise (Gordon and Breach, New York, 1963).
[46] A. N. Malakhov, Fluctuation in Self-Oscillatory Systems (Nauka, Moscow, 1968) (in Russian).
[47] A. S. Pikovsky, M. G. Rosenblum, and J. Kurths, Synchronization: A Universal Concept in Nonlinear Sciences (Cambridge University Press, Cambridge, 2001).
[48] E. Ott and T. M. Antonsen, Chaos **18**, 037113 (2008).
[49] I. V. Tyulkina, D. S. Goldobin, L. S. Klimenko, and A. Pikovsky, Phys. Rev. Lett. **120**, 264101 (2018).
[50] A. S. Pikovsky and J. Kurths, Phys. Rev. Lett. **78**, 775 (1997).
[51] Y. F. Suprunenko, P. T. Clemson, and A. Stefanovska, Phys. Rev. Lett. **111**, 024101 (2013).
[52] P. T. Clemson, Y. F. Suprunenko, T. Stankovski, and A. Stefanovska, Phys. Rev. E **89**, 032904 (2014).
[53] T. Stankovski, T. Pereira, P. V. E. McClintock, and A. Stefanovska, Rev. Mod. Phys. **89**, 045001 (2017).
[54] C. Allefeld and J. Kurths, Int. J. Bifurcat. Chaos **14**, 405 (2004).
[55] B. Schelter, M. Winterhalder, J. Timmer, and M. Peifer, Phys. Lett. A **366**, 382 (2007).
[56] A. E. Hramov and A. A. Koronovskii, Chaos **14**, 603 (2004).
[57] M. D. Prokhorov, V. I. Ponomarenko, V. I. Gridnev, M. B. Bodrov, and A. B. Bespyatov, Phys. Rev. E **68**, 041913 (2003).
[58] A. E. Hramov, A. A. Koronovskii, V. I. Ponomarenko, and M. D. Prokhorov, Phys. Rev. E **75**, 056207 (2007).
[59] A. S. Karavaev, M. D. Prokhorov, V. I. Ponomarenko, A. R. Kiselev, V. I. Gridnev, E. I. Ruban, and B. P. Bezruchko, Chaos **19**, 033112 (2009).
[60] A. S. Karavaev, A. R. Kiselev, A. E. Runnova, M. O. Zhuravlev, E. I. Borovkova, M. D. Prokhorov, V. I. Ponomarenko, S. V. Pchelintseva, T. Y. Efremova, A. A. Koronovskii, and A. E. Hramov, Chaos **28**, 081102 (2018).
[61] V. S. Anishchenko, V. Astakhov, A. Neiman, T. Vadivasova, and L. Schimansky-Geier, Nonlinear Dynamics of Chaotic and Stochastic Systems: Tutorial and Modern Developments (Springer, Berlin, 2007).
[62] A. Neiman, A. Silchenko, V. Anishchenko, and L. Schimansky-Geier, Phys. Rev. E **58**, 7118 (1998).
[63] D. E. Postnov, T. E. Vadivasova, O. V. Sosnovtseva, A. G. Balanov, V. S. Anishchenko, and E. Mosekilde, Chaos **9**, 227 (1999).
[64] V. Astakhov, A. Shabunin, and V. Anishchenko, Int. J. Bifurcat. Chaos **10**, 849 (2000).
[65] T. E. Vadivasova, G. I. Strelkova, and V. S. Anishchenko, Phys. Rev. E **63**, 036225 (2001).

[66] N. B. Janson, A. G. Balanov, V. S. Anishchenko, and P. V. E. McClintock, Phys. Rev. Lett. **86**, 1749 (2001).

[67] N. B. Janson, A. G. Balanov, V. S. Anishchenko, and P. V. E. McClintock, Phys. Rev. E **65**, 036211 (2002).

[68] N. B. Janson, A. G. Balanov, V. S. Anishchenko, and P. V. E. McClintock, Phys. Rev. E **65**, 036212 (2002).

[69] V. S. Anishchenko, T. E. Vadivasova, and G. I. Strelkova, Fluct. Noise Lett. **4**, L219 (2004).

[70] V. S. Anishchenko, T. E. Vadivasova, G. A. Okrokvertskhov, and G. I. Strelkova, Phys. Usp. **48**, 151 (2005).

[71] V. Anishchenko, S. Astakhov, and T. Vadivasova, Europhys. Lett. **86**, 30003 (2009).

[72] A. Zakharova, T. Vadivasova, V. Anishchenko, A. Koseska, and J. Kurths, Phys. Rev. E **81**, 011106 (2010).

[73] N. I. Semenova, G. I. Strelkova, V. S. Anishchenko, and A. Zakharova, Chaos **27**, 061102 (2017).

[74] S. A. Bogomolov, A. V. Slepnev, G. I. Strelkova, E. Schöll, and V. S. Anishchenko, Commun. Nonlinear Sci. Numer. Simul. **43**, 25 (2017).

[75] J.-N. Teramae, H. Nakao, and G. B. Ermentrout, Phys. Rev. Lett. **102**, 194102 (2009).

[76] D. S. Goldobin, J.-N. Teramae, H. Nakao, and G. B. Ermentrout, Phys. Rev. Lett. **105**, 154101 (2010).

[77] M. G. Rosenblum, A. S. Pikovsky, J. Kurths, C. Schafer, and P. A. Tass, in *Neuro-Informatics*, edited by F. Moss and S. Gielen, *Handbook of Biological Physics* 4, 279 (Elsevier Science, New York, 2000).

[78] A. S. Pikovsky, M. G. Rosenblum, and J. Kurths, Int. J. Bifurcat. Chaos **10**, 2291 (2000).

[79] M. Palus, V. Komarek, Z. Hrncir, and K. Sterbova, Phys. Rev. E **63**, 046211 (2001).

[80] T. Schreiber, Phys. Rev. Lett. **85**, 461 (2000).

[81] L. Barnett, A. B. Barrett, and A. K. Seth, Phys. Rev. Lett. **103**, 238701 (2009).

[82] L. Arnold, *Random Dynamical Systems* (Springer-Verlag, Berlin, 1998).

[83] Y. Kuramoto, *Chemical Oscillations, Waves and Turbulence* (Springer-Verlag, Berlin, 1984).

[84] V. S. Anishchenko, A. S. Kopeikin, T. E. Vadivasova, G. I. Strelkova, and J. Kurths, Phys. Rev. E **62**, 7886 (2000).

[85] V. S. Anishchenko, T. E. Vadivasova, A. S. Kopeikin, J. Kurths, and G. I. Strelkova, Phys. Rev. Lett. **87**, 054101 (2001).

[86] V. S. Anishchenko, T. E. Vadivasova, A. S. Kopeikin, J. Kurths, and G. I. Strelkova, Phys. Rev. E **65**, 036206 (2002).

[87] V. I. Tikhonov and M. A. Mironov, *Markov Processes* (Sov. Radio, Moscow, 1977) (in Russian).