

Granger causality maps for Langevin systems

Lionel Barnett^{✉,*}, Benjamin Wahl,[†] Nadine Spychala, and Anil K. Seth[✉]*Sussex Centre for Consciousness Science, Department of Informatics, University of Sussex, Falmer, Brighton, United Kingdom*

(Received 7 January 2026; accepted 13 May 2026; published 1 June 2026)

Wahl *et al.* [*Phys. Rev. E* **93**, 022213 (2016)] introduced Granger causality (GC) maps for Langevin systems: Dynamics are localized linearly at each point in phase space as vector Ornstein-Uhlenbeck (VOU) processes, for which GCs may in principle be calculated, thus constructing a GC map on phase space. These maps may, furthermore, be averaged over phase space to yield a systemwide global GC value. The implementation, however, suffered some significant drawbacks: GCs were approximated from models based on discrete-time stroboscopic sampling of local VOU processes, which is not only computationally inefficient but, more seriously, infeasible on regions of phase space where local dynamics are unstable, leaving “holes” in the GC maps. We solve these problems by deriving an analytical expression for GC rates associated with a VOU process which, under quite general conditions, yields a meaningful solution even in the unstable case. Applied to GC maps, this not only “fills in the holes” but also furnishes a computationally efficient method of calculation devolving to solution of algebraic Riccati equations which, in the case of a univariate source, become simple quadratic equations. We show, furthermore, that the GC rate for VOU processes is invariant under rescaling of the overall fluctuations intensity, so that GC maps may effectively be calculated for *deterministic* nonlinear dynamical systems, with a residual “ghost of noise” represented by a variance-covariance map.

DOI: [10.1103/96ch-bxmf](https://doi.org/10.1103/96ch-bxmf)

I. INTRODUCTION

Wiener-Granger causality (henceforth GC), a widely used method for quantifying directed information transfer between stochastic variables, is based on the premise that cause (a) precedes effect and (b) contains unique information about effect [1,2]. While physics, and indeed other branches of science, are traditionally concerned with “mechanism,” in the sense of the structure and parameters of models, the appeal of information theory is that it abstracts away mechanism in favor of causal (in the Wiener-Granger sense) relationships among system variables. As such, it has been applied in fields as diverse as econometrics, the neurosciences, genomics, ecology, and climate science. Most commonly operationalized via linear modeling [3,4], it is widely (if sometimes unfairly [5]) viewed as inappropriate for stochastic systems featuring nonlinear interactions—in contrast to its nonparametric cousin transfer entropy (TE) [6–8]. Transfer entropy is, however, in general analytically intractable for nonlinear systems, especially in continuous time [9], and frequently problematic to estimate empirically [10].

Of especial interest are *diffusion processes* described by Langevin equations—equivalently (in general nonlinear)

stochastic differential equations (SDEs) or Fokker-Planck equations—which are ubiquitous in statistical physics and have further applications in biology, econometrics, machine learning, and beyond. Wahl *et al.* [11,12] present a powerful approach to analysis of information transfer in autonomous Langevin systems. Following standard practice when faced with analytically intractable nonlinearity, they linearize locally in the system phase space. The locally linearized dynamics of a Langevin process may, under mild assumptions, be represented by a vector Ornstein-Uhlenbeck (VOU) process. Noting that stroboscopic (regularly spaced) observation of a VOU process yields a first-order vector autoregressive (VAR) process, now in discrete time, Wahl *et al.* [11] calculate Granger causalities between a given pair of subprocesses from local VAR(1) models obtained from suitably fine-grained subsampling, thus constructing a mapping from phase space to local GC values for those subprocesses. This mapping may, further, be averaged over the stationary distribution of the process in phase space, yielding a global, systemwide GC value for the subprocesses in question.

This procedure, however, has two drawbacks. Firstly, we note that as the sampling interval shrinks, the GC value approaches zero [13], while the GC *rate*—information transfer per unit time—approaches a finite limit [14]. Since the sample time interval in the subsampling procedure is finite, the procedure can thus yield at best approximations to local GCs rates. Secondly, and more problematically, it is commonplace in nonlinear dynamics that, even if globally stable, dynamics may well be locally unstable in large regions of the phase space (indeed, local instability is a fundamental characteristic of chaotic dynamics). Locally linearized dynamics are determined by the Jacobian: The system is locally stable at a point in phase space iff the Jacobian matrix evaluated at that

*Contact author: lionelb@sussex.ac.uk; <https://users.sussex.ac.uk/~lionelb/>

[†]Unaffiliated.

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

point is Hurwitz stable; that is, all its eigenvalues lie strictly in the left half-plane in the complex plane. The problem, then, is that available techniques for deriving GCs from local discrete-time VAR(1) models, e.g., via spectral factorization [15,16], fail in the unstable case, leaving “holes” in phase space where the GC map is undefined. A particular consequence of these holes is that it becomes impossible to calculate meaningful systemwide GC values. Wahl *et al.* [11] propose a scheme whereby local GCs are averaged only over stable regions in phase space; however, the resulting average fails to reflect the contribution to global information transfer of unstable regions.

In this article we resolve both issues, firstly by demonstrating that GC rates for a VOU process [17] may be calculated directly from the VOU model parameters without recourse to subsampling and that, furthermore, under broad conditions results retain an interpretation as local information transfer even when the local VOU model is unstable. Our resolution leans heavily on previous work by the authors: an analytic formulation of GC rates for a class of continuous-time, distributed-lag stochastic processes [14] and calculation of Granger causalities for discrete-time state-space systems [18].

We also demonstrate a seldom-remarked invariance of GC with respect to rescaling of the noise intensity. This has interesting consequences for GC maps, namely that it permits an extension of information transfer analysis to *deterministic* nonlinear dynamical systems described by sets of ordinary differential equations (ODEs), by globally dialling the noise down to zero, while retaining its “ghost” in the form of a noise variance-covariance map on the phase space, which we interpret as specifying notional infinitesimal fluctuations.

In summary, the main contributions of this paper are as follows: (i) calculation of GC rates for VOU processes directly from model parameters, including in the unstable case, allowing (ii) construction of GC maps for Langevin systems over the entire phase space and (iii) the extension of GC maps and global GC rates to classical deterministic dynamics.

II. GRANGER CAUSALITY RATE FOR VECTOR ORNSTEIN-UHLENBECK PROCESSES

Barnett and Seth [14] define the *zero-horizon Granger causality rate* (henceforth GC rate) for a class of distributed-lag vector stochastic processes in continuous time. Although the analysis there only addresses the unconditional case, all results extend straightforwardly to the conditional case. We start with a brief recap of the construction.

We consider the class of stationary, zero-mean, continuous-time vector moving-average (VMA) processes of the form

$$\mathbf{y}(t) = \int_{u=0}^{\infty} B(u) d\mathbf{w}(t-u), \quad -\infty < t < \infty. \quad (1)$$

Here $\mathbf{w}(t)$ is an n -dimensional Wiener process with $d\mathbf{w}(t) \sim \mathcal{N}(0, \Sigma dt)$, where Σ is an $n \times n$ positive-definite covariance matrix and $B(u)$ an $n \times n$ square-integrable moving-average kernel with $B(0) = I$. The integral in (1) is to be interpreted as an Itô integral [19]. It is assumed that the VMA form (1) may be inverted to yield a VAR form as a stochastic

integrodifferential equation,

$$d\mathbf{y}(t) = \left[\int_{u=0}^{\infty} A(u) \mathbf{y}(t-u) du \right] dt + d\mathbf{w}(t), \quad -\infty < t < \infty, \quad (2)$$

with square-integrable autoregressive kernel $A(u)$. Barnett and Seth [14] consider only *stable* and *minimum-phase* models, and it is further assumed that any *subprocess* of $\mathbf{y}(t)$ also has an invertible VMA representation. We refer to such processes as continuous-time vector autoregressive (CTVAR) processes.

Given a CTVAR process $\mathbf{y}(t)$ as above, Barnett and Seth [14] consider the optimal least-squares predictor $\hat{\mathbf{y}}(t; h) = \mathbb{E}[\mathbf{y}(t+h) | \mathbf{y}(u) : u \leq t]$ at finite prediction horizon $h > 0$ and show that the covariance matrix of the (stationary) prediction error process $\mathbf{e}(t; h) = \hat{\mathbf{y}}(t; h) - \mathbf{y}(t+h)$ may be expressed as

$$\mathcal{E}(h) = \mathbb{E}[\mathbf{e}(t; h) \mathbf{e}(t; h)^T] = \int_{u=0}^h B(u) \Sigma B(u)^T du, \quad (3)$$

where superscript T denotes matrix transpose. Suppose now that $\mathbf{y}$ is partitioned as [20] $\mathbf{y} = [\mathbf{y}_1^T \mathbf{y}_2^T \mathbf{y}_3^T]^T$. For GC analysis, $\mathbf{y}_1$ will be the *target* variable, $\mathbf{y}_2$ the *conditioning* variable, and $\mathbf{y}_3$ the *source* variable. Throughout, we use subscript $\mathbf{R}$ to denote the multi-index pair (1,2) and superscript $\mathbf{R}$ to denote quantities associated with the “reduced system” $\mathbf{y}_{\mathbf{R}} = [\mathbf{y}_1^T \mathbf{y}_2^T]^T$; i.e., with the source variable $\mathbf{y}_3$ omitted. Analogously to the discrete-time case [3,4], the Granger causality from $\mathbf{y}_3$ to $\mathbf{y}_1$ conditional on $\mathbf{y}_2$ at prediction horizon h is defined to be

$$\mathcal{F}_{\mathbf{y}_3 \rightarrow \mathbf{y}_1 | \mathbf{y}_2}(h) = \log \frac{|\mathcal{E}_{11}^{\mathbf{R}}(h)|}{|\mathcal{E}_{11}(h)|}, \quad (4)$$

where $|\cdots|$ denotes matrix determinant. It is shown that $\mathcal{F}_{\mathbf{y}_3 \rightarrow \mathbf{y}_1 | \mathbf{y}_2}(0) = 0$, and the GC rate is defined as

$$\mathcal{R}_{\mathbf{y}_3 \rightarrow \mathbf{y}_1 | \mathbf{y}_2} = \dot{\mathcal{F}}_{\mathbf{y}_3 \rightarrow \mathbf{y}_1 | \mathbf{y}_2}(0) = \lim_{h \rightarrow 0} \frac{1}{h} \mathcal{F}_{\mathbf{y}_3 \rightarrow \mathbf{y}_1 | \mathbf{y}_2}(h), \quad (5)$$

where the dot denotes time differentiation. Finally, it is shown that

$$\mathcal{R}_{\mathbf{y}_3 \rightarrow \mathbf{y}_1 | \mathbf{y}_2} = \text{trace}[\Sigma_{11}^{-1} (D_{11}^{\mathbf{R}} - D_{11})], \quad (6)$$

where

$$D = \frac{1}{2} \ddot{\mathcal{E}}(0) = \frac{1}{2} [\dot{B}(0) \Sigma + \Sigma \dot{B}(0)^T] \quad (7)$$

and $D^{\mathbf{R}}$ is the corresponding quantity for the reduced process.

Vector Ornstein-Uhlenbeck processes, as defined by a linear SDE of the form:

$$d\mathbf{y}(t) = A\mathbf{y}(t) dt + d\mathbf{w}(t) \quad (8)$$

are a special case of CTVAR processes with autoregressive kernel $A(u) = A\delta(u)$, where A is an $n \times n$ matrix. Stability requires that $|Iz - A| \neq 0$ for z in the right half-plane $\text{Re}(z) \geq 0$ of the complex plane—i.e., A is Hurwitz stable—and the process is always minimum phase. A subprocess of a stable VOU process, while not in general itself a VOU process, will have (stable, minimum-phase) VMA and VAR representations of the form (1) and (2), respectively. Barnett and Seth [14], Appendix F] show that in the general case

$\dot{B}(u) = \int_0^u A(s)B(u-s)ds$, $u > 0$, from which we derive $B(u) = e^{Au}$ for the VOU process (8), leading to $\dot{B}(0) = A$ and

$$D = \frac{1}{2}(A\Sigma + \Sigma A^T). \quad (9)$$

Relaxing the stability requirement (i.e., A may have eigenvalues in the right complex half-plane), the process $y(t)$ may no longer be assumed stationary and consequently may not, as in (1), be taken to extend into the infinite past. Thus, rather than (1), we consider CTVAR processes of the form

$$y(t) = \int_{u=0}^t B(u) d\mathbf{w}(t-u), \quad t \geq 0, \quad (10)$$

initialized at $t = 0$, where $B(u)$ is no longer assumed square integrable. In Appendix A we show that (10) may always be inverted to yield a (not necessarily stable) continuous-time VAR representation, and although the process $y(t)$ itself may not be stationary, the finite-horizon prediction error process $e(t; h)$ is nonetheless (wide-sense) stationary with covariance matrix $\mathcal{E}(h)$ as in (3). The construction of $\mathcal{R}_{y_3 \rightarrow y_1|y_2}$ outlined above thus goes through unchanged, and—noting that $y_R(t)$ will not in general be a VOU process—it remains to calculate $\dot{B}^R(0)$ and thence D^R . Our principal result shows how $\mathcal{R}_{y_3 \rightarrow y_1|y_2}$ may be calculated explicitly from the VOU parameters (A, Σ) under relaxed assumptions on the stability of A (we shall still require Σ to be positive definite).

Theorem 1. For the VOU process (8) with $d\mathbf{w}(t) \sim \mathcal{N}(0, \Sigma dt)$ and a partitioning $\mathbf{y} = [\mathbf{y}_1^T \mathbf{y}_2^T \mathbf{y}_3^T]^T$, if Σ is positive-definite and the matrix pair (A_{33}, A_{R3}) detectable, then the continuous-time algebraic Riccati equation (CARE),

$$A_{33}P_{33} + P_{33}A_{33}^T + \Sigma_{33} = (P_{33}A_{R3}^T + \Sigma_{3R})\Sigma_{RR}^{-1}(P_{33}A_{R3}^T + \Sigma_{3R})^T, \quad (11)$$

has a unique stabilizing solution P_{33} and the Granger causality rate from y_3 to y_1 conditional on y_2 for the VOU (8) is given by

$$\mathcal{R}_{y_3 \rightarrow y_1|y_2} = \text{trace}[\Sigma_{11}^{-1}A_{13}P_{33}A_{13}^T], \quad (12)$$

and is equal to twice the corresponding transfer entropy rate, under an appropriate definition of the latter (cf. Barnett and Bossomaier [21] and Barnett *et al.* [8]).

Proof. See Appendix B. ■

We also state an unconditional counterpart to Theorem 1.:

Corollary 1. Given the partitioning $\mathbf{y} = [\mathbf{y}_1^T \mathbf{y}_2^T \mathbf{y}_3^T]^T$, then assuming the associated CAREs have unique stabilizing solutions, the unconditional Granger causality rate from y_2 to y_1 for the VOU (8) is given by:

$$\mathcal{R}_{y_2 \rightarrow y_1} = \mathcal{R}_{y_{23} \rightarrow y_1} - \mathcal{R}_{y_3 \rightarrow y_1|y_2} \quad (13)$$

where $y_{23} = [\mathbf{y}_2^T \mathbf{y}_3^T]^T$.

Proof. This follows from the corresponding standard result in discrete time [4, Sec. 3], which survives passage to the continuous-time limit [14, Sec. 3.3, Eq. 59]. ■

We note the following:

(1) Under the weaker condition that Σ is positive semidefinite and Σ_{RR} invertible, the CARE (11) has a unique stabilizing solution under the additional condition that $(A_{33} - \Sigma_{3R}\Sigma_{RR}^{-1}A_{R3}, \Sigma_{33} - \Sigma_{3R}\Sigma_{RR}^{-1}\Sigma_{3R})$ is stabilizable; see Appendix B. It is not entirely clear however, how to interpret the

resulting GC rate $\mathcal{R}_{y_3 \rightarrow y_1|y_2}$, so we do not consider this case further.

(2) If either or both of A, A_{33} are Hurwitz stable, then the detectability condition of Theorem 1 holds. The case A_{33} stable is trivial; the case A stable is proved in Appendix C.

(3) Of the VAR coefficients, only A_{13}, A_{23}, A_{33} —those with y_3 as source—appear in the solution (by contrast, all the Σ_{ij} potentially affect the result). In particular, if $A_{13} = 0$, then $\mathcal{R}_{y_3 \rightarrow y_1|y_2}$ vanishes as expected. If A_{R3} (and thus A_{13}) are identically zero and A_{33} is unstable, while the detectability condition is technically violated, then it still makes sense to designate $\mathcal{R}_{y_3 \rightarrow y_1|y_2} = 0$, since the dynamics of y_R are completely unaffected by the source y_3 .

(4) If $\dim(y_3) = 1$, i.e., the source variable is one-dimensional, then P_{33} is scalar and the CARE (11) becomes the quadratic equation in P_{33} :

$$(A_{R3}^T \Sigma_{RR}^{-1} A_{R3})P_{33}^2 - 2(A_{33} - \Sigma_{3R}\Sigma_{RR}^{-1}A_{R3})P_{33} - (\Sigma_{33} - \Sigma_{3R}\Sigma_{RR}^{-1}\Sigma_{3R}) = 0. \quad (14)$$

If $A_{R3} \equiv 0$, then as above we take $\mathcal{R}_{y_3 \rightarrow y_1|y_2} = 0$. Otherwise, the discriminant of the quadratic will be > 0 , and a unique stabilizing solution for P_{33} (i.e., with $P_{33} > 0$) is obtained as the solution of (14) with positive square root.

(5) If we *rescale* the residuals covariance matrix by $\Sigma \rightarrow \nu\Sigma$, then νP_{33} is the solution of the corresponding CARE (11), and from (12) we see that the Granger causality rate $\mathcal{R}_{y_3 \rightarrow y_1|y_2}$ is unchanged. Thus the Granger causality rate is invariant under rescaling of the overall intensity (but not in general under changes to relative variances and covariances) of fluctuations.

An important special case is the Granger-causal graph [22]—that is, the pairwise-conditional causality rates

$$\mathcal{G}_{ij} = \mathcal{R}_{y_j \rightarrow y_i|y_{[ij]}}, \quad i, j = 1, \dots, n, \quad i \neq j, \quad (15)$$

where subscript $[\dots]$ indicates that the enclosed indices are omitted. Since the source variable is one dimensional, we may apply (14). The subprocess indices 1, 2, 3, and R in the previous analysis then map as $1 \rightarrow i$, $2 \rightarrow [ij]$, $3 \rightarrow j$, and $R \rightarrow [j]$, and we have

$$\mathcal{G}_{ij} = \Sigma_{ii}^{-1}A_{ij}^2P_{jj} \quad (16)$$

with P_{jj} the (positive root) solution of the quadratic equation

$$(A_{[ij]j}^T \Sigma_{[j][j]}^{-1} A_{[ij]j})P_{jj}^2 - 2(A_{jj} - \Sigma_{j[j]} \Sigma_{[j][j]}^{-1} A_{[ij]j})P_{jj} - (\Sigma_{jj} - \Sigma_{j[j]} \Sigma_{[j][j]}^{-1} \Sigma_{[ij]j}) = 0. \quad (17)$$

For the unconditional case, from (13) we have $\mathcal{R}_{y_j \rightarrow y_i} = \mathcal{R}_{y_{[ij]} \rightarrow y_i} - \mathcal{G}_{ij}$. For $\mathcal{R}_{y_{[ij]} \rightarrow y_i}$ the subprocess indices map as $1 \rightarrow i$, $2 \rightarrow \emptyset$, $3 \rightarrow [i]$, and $R \rightarrow i$, and the corresponding $(n-1)$ -dimensional CARE is

$$A_{[i][i]}P_{[i][i]} + P_{[i][i]}A_{[i][i]}^T + \Sigma_{[i][i]} = (P_{[i][i]}A_{[i]i}^T + \Sigma_{[i]i})\Sigma_{ii}^{-1}(P_{[i][i]}A_{[i]i}^T + \Sigma_{[i]i})^T. \quad (18)$$

We then have

$$\mathcal{R}_{y_j \rightarrow y_i} = \Sigma_{ii}^{-1} \text{trace}[A_{[i]i}P_{[i][i]}A_{[i]i}^T] - \mathcal{G}_{ij}. \quad (19)$$

III. GRANGER CAUSALITY MAPS FOR LANGEVIN PROCESSES

Following Wahl *et al.* [11], we consider multivariate Langevin systems specified by autonomous SDEs of the form

$$d\mathbf{y}(t) = f(\mathbf{y}(t))dt + d\mathbf{w}(\mathbf{y}(t), t) \quad (20)$$

on $\mathbb{R}^n$, with *drift function* $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ and Wiener noise $d\mathbf{w}(\mathbf{y}, t) \sim \mathcal{N}(0, \Sigma(\mathbf{y})dt)$, where $\Sigma(\mathbf{y})$ is the *diffusion function*, which maps from $\mathbb{R}^n$ to the manifold of $n \times n$ positive-definite matrices. Again following Wahl *et al.* [11], we linearize (20) around a point $\mathbf{y}_0 \in \mathbb{R}^n$. Setting $\boldsymbol{\eta}(t) = \mathbf{y}(t) - \mathbf{y}_0$ we assume $\|\boldsymbol{\eta}(t)\| < \varepsilon$, and work to $\mathcal{O}(\varepsilon)$. We have

$$d\mathbf{y}(t) = f(\mathbf{y}_0 + \boldsymbol{\eta}(t))dt + d\mathbf{w}(\mathbf{y}_0 + \boldsymbol{\eta}(t), t). \quad (21)$$

Now

$$f(\mathbf{y}_0 + \boldsymbol{\eta}(t)) = f(\mathbf{y}_0) + J(\mathbf{y}_0) \cdot \boldsymbol{\eta}(t) + \mathcal{O}(\varepsilon^2), \quad (22)$$

where $J(\mathbf{y}_0) = \nabla f(\mathbf{y}_0)$ is the *Jacobian matrix* [23] of $f(\cdots)$ evaluated at $\mathbf{y}_0$, and for $\|\boldsymbol{\eta}\| < \varepsilon$

$$\begin{aligned} d\mathbf{w}(\mathbf{y}_0 + \boldsymbol{\eta}, t) \\ \sim \mathcal{N}(0, \{\Sigma(\mathbf{y}_0) + \nabla \Sigma(\mathbf{y}_0) \cdot \boldsymbol{\eta} + \mathcal{O}(\varepsilon^2)\}dt). \end{aligned} \quad (23)$$

We make the further assumption that the fluctuations covariance term $\Sigma(\mathbf{y})$ varies slowly with $\mathbf{y}$; specifically, we assume $\|\nabla \Sigma(\mathbf{y})\| = \mathcal{O}(\varepsilon)$ everywhere (Wahl *et al.* [11] refer to this assumption, common in statistical physics, as “weakly multiplicative” noise) [24]. We then have to $\mathcal{O}(\varepsilon)$

$$d\boldsymbol{\eta}(t) = \{f(\mathbf{y}_0) + J(\mathbf{y}_0) \cdot \boldsymbol{\eta}(t)\}dt + d\mathbf{w}(\mathbf{y}_0, t). \quad (24)$$

Setting

$$\mathbf{z}(t) = J(\mathbf{y}_0)^{-1} \cdot f(\mathbf{y}_0) + \boldsymbol{\eta}(t) \quad (25)$$

(a linear translation in $\mathbb{R}^n$ to the mean reversion level), we find that to first order in ε , $\mathbf{z}(t)$ satisfies the VOU,

$$d\mathbf{z}(t) = J(\mathbf{y}_0) \cdot \mathbf{z}(t)dt + d\mathbf{w}(\mathbf{y}_0, t). \quad (26)$$

The analysis of the previous section may be applied to the locally linearized process $\mathbf{z}(t)$ around $\mathbf{y}_0$, wherever the Jacobian matrix $J(\mathbf{y}_0)$ is nonsingular. Note that local stability requires invertibility of $J(\mathbf{y}_0)$, since its eigenvalues must lie in the complex half-plane $\text{Re}(z) < 0$. Here we do not demand stability of the locally linearized VOUs everywhere but assume that the Jacobian is singular on at most a set of measure zero in $\mathbb{R}^n$. We note that invertibility of $J(\mathbf{y}_0)$ does not imply the detectability condition for Theorem 1; detectability must be satisfied, again on all but a set of measure zero. This is of no consequence for the case of a one-dimensional source, where as we have seen (item 4 following Corollary 1) the existence of a local GC rate is guaranteed. More generally, since the condition is algebraic on the matrix coefficients, the detectability requirement is not likely to be overly restrictive, although some highly specific connectivity structures may conceivably be precluded.

Given a partitioning of the phase space into $\mathbf{y} = [\mathbf{y}_1^T \mathbf{y}_2^T \mathbf{y}_3^T]^T$ as in Sec. II, the mapping $\mathbf{y}_0 \mapsto \mathcal{R}_{\mathbf{z}_3 \rightarrow \mathbf{z}_1 | \mathbf{z}_2}(\mathbf{y}_0)$ as $\mathbf{y}_0$ varies over the system phase space, with $\mathbf{z}$ as in (25), defines the Granger causality map. Note that the map inherits invariance under rescaling of the covariance matrices $\Sigma(\mathbf{y}) \rightarrow \nu(\mathbf{y})\Sigma(\mathbf{y})$; rescaling may be state dependent, as long as the resultant

noise remains weakly multiplicative. Conceptually, the map might thus be regarded as associated with the *deterministic* autonomous ODE $\dot{\mathbf{y}}(t) = f(\mathbf{y}(t))$ along with a given equivalence class of covariance matrix mappings $\mathbf{y} \mapsto \Sigma(\mathbf{y})$ under rescaling.

As suggested by Ref. [11], we may obtain a *global* GC rate, as the expectation of $\mathcal{R}_{\mathbf{z}_3 \rightarrow \mathbf{z}_1 | \mathbf{z}_2}(\mathbf{y}_0)$ over the stationary distribution (assuming it exists) of states defined by the dynamics of (20). Under ergodicity assumptions, in the limit of high noise this tends to the distribution of the Wiener process $\mathbf{w}(\mathbf{y}, t)$ appearing in (20). In the limit of low noise, we might, alternatively, take the expectation over the *attractor(s)* of the autonomous ODE, leaving the covariance function $\Sigma(\mathbf{y})$ as a “ghost” of notional noise fluctuations. Note that by noise scale invariance, the only influence of noise level on the global GC rate is the distribution on phase space with respect to which we calculate the expectation. Under ergodicity assumptions, we may in practice calculate the global GC rate as

$$\mathcal{R}_{\mathbf{y}_3 \rightarrow \mathbf{y}_1 | \mathbf{y}_2} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t=0}^T \mathcal{R}_{\mathbf{z}_3 \rightarrow \mathbf{z}_1 | \mathbf{z}_2}(\mathbf{y}(t)) dt, \quad (27)$$

where $\mathbf{y}(t)$ is a trajectory of (20). Wahl *et al.* [11] show that if the process (20) is actually linear (i.e., it is a VOU), then the global GC rate always corresponds to the VOU Granger causality rate $\mathcal{R}_{\mathbf{y}_3 \rightarrow \mathbf{y}_1 | \mathbf{y}_2}$ of Theorem 1. In the nonlinear case, we stress that, although the (nonparametric) transfer entropy rate from $\mathbf{y}_3$ to $\mathbf{y}_1$ conditioned on $\mathbf{y}_2$ may be well defined [9], we have no reason to expect that the global GC rate (27) will correspond to the TE rate for any covariance mapping $\Sigma(\mathbf{y})$; in other words, while the equivalence of GC with TE applies locally in phase space (as per Theorem 1), it may not be assumed to hold globally for global GC defined, as described above, as the averaged local GC.

Example: The Lorenz system

The well-known three-variable Lorenz system [25] is defined by the parametrized set of ODEs,

$$\dot{y}_1 = \sigma(y_2 - y_1), \quad (28a)$$

$$\dot{y}_2 = y_1(\rho - y_3) - y_2, \quad (28b)$$

$$\dot{y}_3 = y_1 y_2 - \beta y_3. \quad (28c)$$

In some parameter regimes it exhibits chaotic dynamics, with the iconic “butterfly” strange attractor.

We posit a set of SDEs of the form (20) based on (28), with constant fluctuations covariance matrix $\Sigma(\mathbf{y}) = \nu I$. Fluctuations are then globally scaled down by letting $\nu \rightarrow 0$, so that in the limit the (deterministic) dynamics are those of the ODEs (28). The Jacobian of the system is

$$J(\mathbf{y}) = \begin{bmatrix} -\sigma & \sigma & 0 \\ \rho - y_3 & -1 & -y_1 \\ y_2 & y_1 & -\beta \end{bmatrix}. \quad (29)$$

We have

$$|J(\mathbf{y})| = \sigma[\beta(\rho - 1 - y_3) - y_1(y_1 + y_2)], \quad (30)$$

which vanishes on the quadratic (two-dimensional and hence measure-zero) surface $y_3 = \rho - 1 - \frac{1}{\beta}y_1(y_1 + y_2)$.

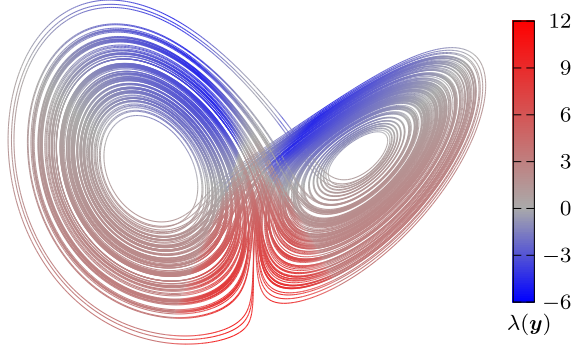


FIG. 1. Local stability of Lorenz dynamics on the attractor. The color scale corresponds to the largest real part $\lambda(\mathbf{y})$ of the eigenvalues of the Jacobian matrix $J(\mathbf{y})$; the system is locally unstable where $\lambda(\mathbf{y}) \geq 0$ (gray $\rightarrow$ red).

We simulated the deterministic equations (28) with canonical parameters $\sigma = 10$, $\rho = 28$, $\beta = 8/3$, using the Runge-Kutta (4,5) method implemented by the MATLAB function `ode45`, for 200 s [26], sampling the trajectory at time increments of 0.01 s [27], with initial value $\mathbf{y}(0) = [1 \ 1 \ 1]^T$, and allowing 100 s for the trajectory to settle into the attractor. Figure 1 displays the largest real part $\lambda(\mathbf{y})$ of the eigenvalues of $J(\mathbf{y})$ plotted on a color scale over the simulated attractor; the system is locally unstable where $\lambda(\mathbf{y}) \geq 0$ (gray $\rightarrow$ red). We

then calculated the Granger causal graph $\mathcal{G}_{ij}(\mathbf{y})$ (15) on the attractor (Fig. 2), estimating global GC rates according to (27) by numerical quadrature. We may confirm from the form of $J(\mathbf{y})$ that, since the $J_{ii}(\mathbf{y})$ are always < 0 , the detectability condition of Theorem 1 is always satisfied. Since $J_{13}(\mathbf{y}) = 0$, $\mathcal{G}_{13}(\mathbf{y}) \equiv 0$ as expected [cf. Eq. (16)], while all other pairwise-conditional GC rates are generally nonzero.

IV. DISCUSSION

Calculating information transfer between components of a nonlinear stochastic system, whether in the parametric sense of Wiener-Granger causality [1,2] or the nonparametric sense of transfer entropy [6,7], has long been perceived as troublesome empirically and challenging, if not intractable, at the analytical level. The notion of Granger causality maps introduced by Wahl *et al.* [11,12] promised an elegant and powerful approach to calculation of information transfer for the important class of nonlinear stochastic dynamics that may be described by Langevin equations. The original formulation of the technique was, however, compromised by computational issues and, more seriously, by the presence of “holes” in the system phase space where the calculation procedure breaks down. In this paper we have presented a comprehensive resolution to both issues, so that Granger causality maps may now be considered a viable and computationally tractable method for calculation of both local and systemwide

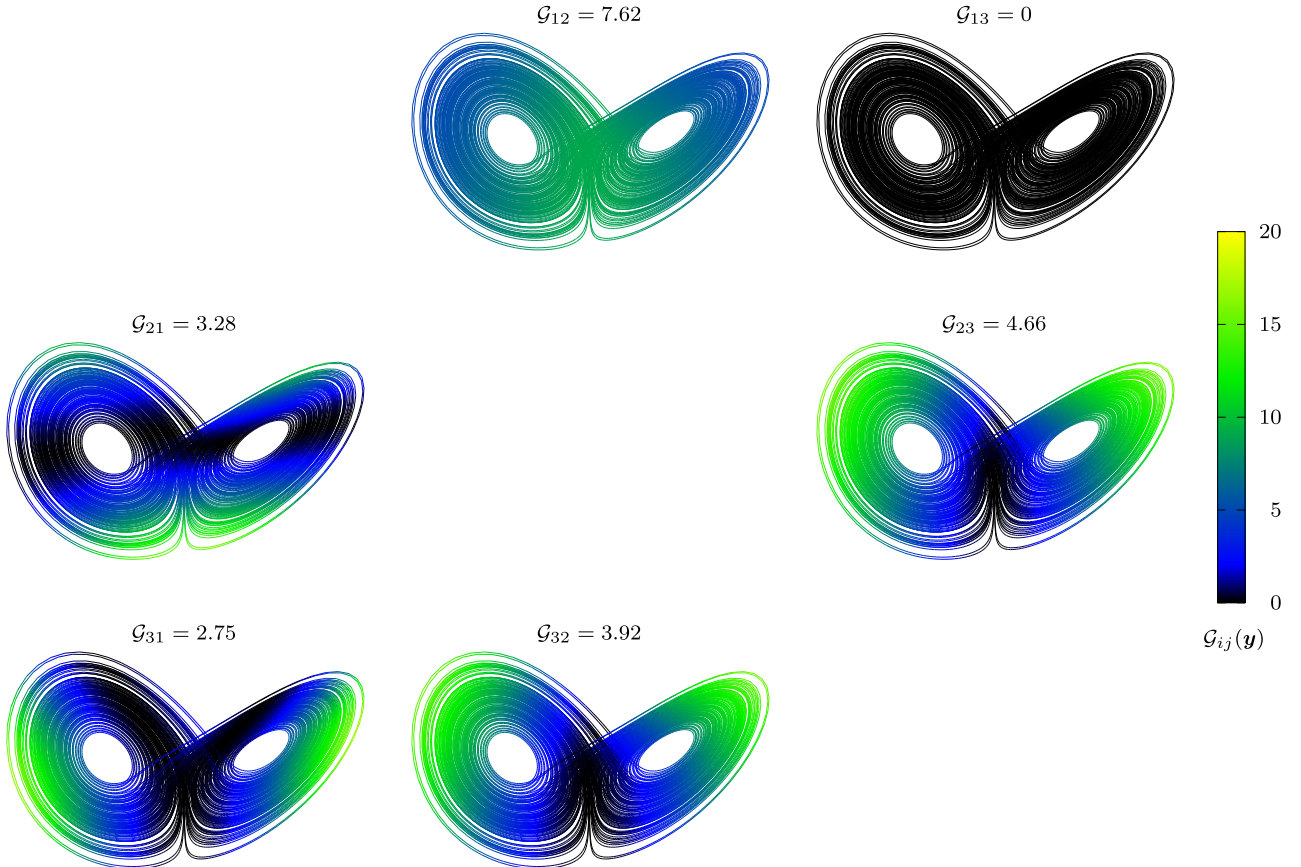


FIG. 2. GC maps of Lorenz dynamics on the attractor. The color scale measures local Granger-causal graph values $\mathcal{G}_{ij}(\mathbf{y})$. The $\mathcal{G}_{ij}$ values above the plots are the global Granger-causal graph values calculated according to (27).

information transfer in Langevin systems. We also describe how the approach extends to calculation of information transfer for *deterministic* dynamical systems described by sets of ordinary differential equations, given a template for infinitesimal fluctuations in the form of a variance-covariance matrix phase-space map.

This work opens up information flow analysis not only for standard applications of Langevin equations in statistical physics, quantum physics, noisy electrical circuits, soft-matter physics, chemical kinetics, biophysics, neural systems, and econometrics but also extends the ambit of the analysis to deterministic dynamics in Hamiltonian systems, Newtonian and relativistic dynamics, classical dynamical systems and chaos theory (cf. Sec. III A), and also to artificial neural networks in brain modeling (e.g., neural mass models), and as increasingly deployed in machine-learning and AI applications. Taking the latter case as an illustration of potential for application, we propose that analysis of information transfer within and between layers in deep-learning, predictive coding, and transformer architectures may assist in understanding the mechanisms underlying the functionality achieved by such systems—to illuminate the “black box,” as such systems are frequently regarded—and to inform network design and AI safety issues [28].

A. Limitations

A question we have not addressed in this study is the relationship between the global GC rates obtained by averaging local GC rates (see remarks at the end of Sec. III and the global transfer entropy). While GC maps may be viewed as furnishing a more fine-grained picture of information transfer between system components through localization in phase space, we do not know whether, for a given system, the averaged local information transfer agrees with the corresponding transfer entropy. As it stands, we are only assured of a positive answer to this question in the case of linear dynamics.

A limitation for the case of neural systems, both natural and technological, is that such systems in general (certainly in the biological case) feature finite signal propagation delays in information transmission between network nodes. This may well apply too for some financial systems [29]. Neither standard Langevin equations, nor indeed Ornstein-Uhlenbeck equations, accommodate finite lags [14]. Incorporating distributed lags could in principle be achieved through generalized Langevin equations [30]. While this significantly complicates the analysis, it may well be tractable to some differential-delay methods introduced in Barnett and Seth [14] in combination with state-space methods [18,31].

B. Computational considerations

There are two distinct issues pertaining to the computational demands of calculating GC maps and global GC rates: (i) first, the scaling of computational complexity of solution of the individual CAREs (11) with problem size (the linear dimension of P_{33} is the dimension of the source variable y_{33}) and, second, (ii) the number of samples of the equilibrium distribution in phase space (or attractor in the deterministic limit) required to build GC maps at a satisfactory resolution

and in particular to derive good estimates for the global GC rate (27).

Regarding (i), in the case of one-dimensional source variables (and in particular the Granger-causal graph) computation of GC maps and global GC rates is undemanding, involving solution of simple quadratic equations. For higher-dimensional source variables, the industry standard for solution of general (dense) CAREs uses QZ (generalized Schur) decomposition of the associated Hamiltonian pencil [32]. These algorithms have computational complexity $O(d^3)$, where $d = n_3$ is the dimension of the source variable and are thus potentially computationally demanding if high-dimensional source variables are required (see, for example, Barnett *et al.* [33], where a high-dimensional source information flow is calculated, albeit in a discrete-time or state scenario). In mitigation, for higher-dimensional sources computation is “embarrassingly parallel,” since CARE solutions at different states are independent. At the time of writing, there is limited support for GPU-enabled CARE solution [34], and this situation is likely to improve. There are also potentially computationally cheaper alternative algorithms—especially when the Jacobian matrices are sparse (as will be the case, e.g., for sparsely connected network dynamics)—but this takes us beyond the scope of this article [35].

As for point (ii), it is less clear what will constitute sufficient coverage and resolution of the equilibrium distribution or attractor in a given scenario (again, this is of little concern for the quadratic one-dimensional source case). For higher-dimensional sources, we recommend an *ad hoc* approach, for example, to estimate the global GC rate one might evaluate (27) at successively higher sample sizes or time resolutions until the result stabilizes and the standard error over multiple realizations is acceptable.

C. Future directions

We have not addressed deployment of GC maps for inference of information transfer from empirical time-series data. While perhaps of reduced importance in physics, which tends to proceed from a theoretical standpoint (i.e., the dynamical equations are motivated on theoretical grounds), this is of greater importance for analysis of biophysical (especially neural) and econometric systems, where there are frequently weaker theoretical motivations for detailed specification of dynamics, thus entailing the identification of an appropriate model from the data. Here we merely remark that there is a well-developed literature on methods for estimating Langevin models from discretely sampled time-series data [36–38]; GC maps may then be calculated by the method described here.

A significant feature of (parametric) Granger causality is that the time-domain GC may be decomposed in the spectral domain [3]. This decomposition is especially pertinent to functional analysis of biological neural systems, where functional (statistical) relations are frequently associated with specific frequency bands. This spectral decomposition extends, in principle, to the (zero-horizon) continuous-time case and to VOU processes in particular. While tractable in the unconditional case [14], Sec. 3.3, Eq. 62], the conditional case is more mathematically intricate [4] and requires further work to extend to continuous time.

Finally, information transfer also forms the basis of some current theories in the recently reinvigorated field of emergence theory in complex systems [39]. Here typical systems featuring emergence, e.g., stochastic and deterministic models for flocking and swarming, are potential targets for analysis of emergence using Granger causality maps.

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant No. 101019254, project CONSCIOUS).

DATA AVAILABILITY

MATLAB® code implementing calculation of Granger causality rates for VOU processes may be downloaded [40].

APPENDIX A: STATIONARITY OF THE FINITE-HORIZON PREDICTION ERROR

In discrete time, consider a VMA process

$$\mathbf{y}_k = \sum_{\ell=0}^k B_{\ell} \boldsymbol{\varepsilon}_{k-\ell}, \quad k = 0, 1, 2, \dots, \\ B_0 = I, \quad \boldsymbol{\varepsilon}_k \sim \mathcal{N}(0, \Sigma), \quad (\text{A1})$$

with white-noise innovations process $\boldsymbol{\varepsilon}_k$. We do not assume stability, so the B_{ℓ} are not assumed square summable, and the process $\mathbf{y}_k$ may not be assumed stationary; thus, in contrast to Barnett and Seth [14], rather than an infinite past we take $\mathbf{y}_k$ as initialized at time step $k = 0$, with $\mathbf{y}_0 = \boldsymbol{\varepsilon}_0$.

We show first that without restrictions on the B_{ℓ} , (A1) may always be inverted to yield a (possibly unstable) VAR representation for the process $\mathbf{y}_k$ initialized at $k = 0$. To see this, let us set $\mathbf{Y}_k = [\mathbf{y}_0^T, \dots, \mathbf{y}_k^T]^T$ and $\mathbf{E}_k = [\boldsymbol{\varepsilon}_0^T, \dots, \boldsymbol{\varepsilon}_k^T]^T$. From (A1) we then have

$$\mathbf{Y}_k = \mathbf{B}_k \cdot \mathbf{E}_k \quad k = 0, 1, 2, \dots, \quad (\text{A2})$$

where

$$\mathbf{B}_k = \begin{bmatrix} I & 0 & \cdots & 0 & 0 \\ B_1 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ B_{k-1} & B_{k-2} & \cdots & I & 0 \\ B_k & B_{k-1} & \cdots & B_1 & I \end{bmatrix}, \quad k = 0, 1, 2, \dots \quad (\text{A3})$$

Since $\mathbf{B}_k$ is a lower-triangular block-Toeplitz matrix with identity matrices on the diagonal, $|\mathbf{B}_k| = 1$, so the matrix $\mathbf{B}_k$ is invertible, and we have

$$\mathbf{E}_k = \mathbf{B}_k^{-1} \cdot \mathbf{Y}_k, \quad k = 0, 1, 2, \dots \quad (\text{A4})$$

Setting $\mathbf{A}_k = I - \mathbf{B}_k^{-1}$, we have

$$\mathbf{Y}_k = \mathbf{A}_k \cdot \mathbf{Y}_k + \mathbf{E}_k, \quad k = 0, 1, 2, \dots \quad (\text{A5})$$

$\mathbf{A}_k$ is also lower-triangular block-Toeplitz, now with zero matrices on the diagonal; it is easy to see that (A5) thus specifies a (not necessarily stable) VAR representation for the process $\mathbf{y}_k$, initialized at $k = 0$.

By a standard result, the optimal least-squares predictor $\hat{\mathbf{y}}_k(m) = \mathbb{E}[\mathbf{y}_{k+m} | \mathbf{y}_0, \dots, \mathbf{y}_k]$ for $\mathbf{y}_{k+m}$ given $\mathbf{y}_0, \dots, \mathbf{y}_k$ is given by [41]

$$\hat{\mathbf{y}}_k(m) = \sum_{\ell=m}^{k+m} B_{\ell} \boldsymbol{\varepsilon}_{k+m-\ell}, \quad k = 0, 1, 2, \dots, \\ m = 1, 2, \dots \quad (\text{A6})$$

The prediction error $\mathbf{e}_k(m) = \hat{\mathbf{y}}_k(m) - \mathbf{y}_{k+m}$ is then

$$\hat{\mathbf{e}}_k(m) = - \sum_{\ell=0}^{m-1} B_{\ell} \boldsymbol{\varepsilon}_{k+m-\ell}, \quad (\text{A7})$$

which, we may check, has covariance matrix

$$\mathcal{E}_m = \sum_{\ell=0}^{m-1} B_{\ell} \Sigma B_{\ell}^T. \quad (\text{A8})$$

The crucial point to note is that this expression does not depend on the process time step k , so that the prediction error process is wide-sense stationary.

The above goes through in the continuous-time limit. Suppose given a continuous-time VMA process

$$\mathbf{y}(t) = \int_{u=0}^t B(u) d\mathbf{w}(t-u), \quad t \geq 0, \quad B(0) = I, \\ d\mathbf{w}(t) \sim \mathcal{N}(0, \Sigma dt), \quad (\text{A9})$$

where the VMA kernel $B(u)$ is no longer assumed square integrable, initialized at $t = 0$. Suppose now given a stroboscopic sampling $\bar{\mathbf{y}}_k = \mathbf{y}(k\Delta)$ with period Δ . Working to lowest order in Δ , we have

$$\bar{\mathbf{y}}_k = \int_{u=0}^{(k+1)\Delta} B(u) d\mathbf{w}(k\Delta - u) \\ = \sum_{\ell=0}^k \int_{u=\ell\Delta}^{(\ell+1)\Delta} B(u) d\mathbf{w}(k\Delta - u) \\ \approx \sum_{\ell=0}^k B(\ell\Delta) \int_{u=\ell\Delta}^{(\ell+1)\Delta} d\mathbf{w}(k\Delta - u) \\ \approx \sum_{\ell=0}^k B(\ell\Delta) \bar{\boldsymbol{\varepsilon}}_{k-\ell},$$

where, by the independent increments property of the Wiener process, the $\bar{\boldsymbol{\varepsilon}}_k$ are independent and identically distributed $\sim \mathcal{N}(0, \Sigma \Delta)$. Thus $\bar{\mathbf{y}}_k$ has a VMA representation $\bar{\mathbf{y}}_k = \sum_{\ell=0}^k \bar{B}_{\ell} \bar{\boldsymbol{\varepsilon}}_{k-\ell}$ with $\bar{B}_k = B(k\Delta)$ to lowest order in Δ . Taking the limit $\Delta \rightarrow 0$ (cf. Barnett and Seth [14], Sec. 3.2)], we see that, as in the discrete-time case, (A9) may always be inverted to yield a (possibly unstable) continuous-time VAR representation for the process $\mathbf{y}(t)$ initialized at $t = 0$, and for $h > 0$ the optimal finite-horizon least-squares prediction $\hat{\mathbf{y}}(t; h)$ of $\mathbf{y}(t+h)$ given $\{\mathbf{y}(u) : 0 \leq u \leq t\}$ is

$$\hat{\mathbf{y}}(t; h) = \int_{u=h}^{t+h} B(u) d\mathbf{w}(t+h-u). \quad (\text{A10})$$

The prediction error $\mathbf{e}(t; h) = \hat{\mathbf{y}}(t; h) - \mathbf{y}(t + h)$ is given by

$$\mathbf{e}(t; h) = - \int_{u=0}^h B(u) d\mathbf{w}(t + h - u). \quad (\text{A11})$$

By an Itô isometry [19], we find that (3) still obtains. Thus, as for the discrete-time case, $\mathcal{E}(h) = \mathbb{E}[\mathbf{e}(t; h)\mathbf{e}(t; h)^\top]$ does not depend on the time stamp t , so that the prediction error process $\mathbf{e}(t; h)$ is again wide-sense stationary.

APPENDIX B: PROOF OF THEOREM 1

As in Appendix A, we sample the VOU process (8) stroboscopically with period Δ . The process is not assumed stable, so that A may have eigenvalues in the right half-plane of the complex plane. Using an overbar to denote quantities associated with the subsampled process (with an implied dependency on Δ), we find [14, Section 3.2] that the process $\bar{\mathbf{y}}_k = \mathbf{y}(k\Delta)$ is VAR(1) [42],

$$\bar{\mathbf{y}}_k = \bar{A}\bar{\mathbf{y}}_{k-1} + \bar{\mathbf{e}}_k, \quad (\text{B1})$$

where $\bar{\mathbf{e}}_k$ is a white noise process with covariance matrix $\bar{\Sigma}$, and

$$\bar{A} = e^{A\Delta}, \quad (\text{B2a})$$

$$\bar{\Sigma} = \Omega - e^{A\Delta}\Omega e^{A^\top\Delta}, \quad (\text{B2b})$$

with Ω the solution of the continuous-time Lyapunov equation,

$$A\Omega + \Omega A^\top = -\Sigma. \quad (\text{B3})$$

From (B2) and (B3) we may calculate that to $\mathcal{O}(\Delta)$

$$\bar{A} = I + A\Delta, \quad (\text{B4a})$$

$$\bar{\Sigma} = \Sigma\Delta. \quad (\text{B4b})$$

To calculate the reduced quantities $\Sigma^R, B^R(u)$ and thence D^R , we shall work through the state-space method detailed in Barnett and Seth [18] for the VAR(1) process with parameters (B4) and then take the limit $\Delta \rightarrow 0$.

The VAR(1) (B1) is equivalent to the discrete-time innovations-form state-space (ISS) model,

$$\bar{\mathbf{z}}_{k+1} = \bar{A}\bar{\mathbf{z}}_k + \bar{\mathbf{e}}_k, \quad (\text{B5a})$$

$$\bar{\mathbf{y}}_k = \bar{A}\bar{\mathbf{z}}_k + \bar{\mathbf{e}}_k, \quad (\text{B5b})$$

with state variable $\bar{\mathbf{z}}_k$. The reduced state-space system—now no longer in innovations form—is then given by

$$\bar{\mathbf{z}}_{k+1} = \bar{A}\bar{\mathbf{z}}_k + \bar{\mathbf{e}}_k, \quad (\text{B6a})$$

$$\bar{\mathbf{y}}_{R,k} = \bar{A}_{R*}\bar{\mathbf{z}}_k + \bar{\mathbf{e}}_{R,k}, \quad (\text{B6b})$$

where subscript $*$ denotes the full indices (1,2,3), and R the reduced indices (1,2). Following Barnett and Seth [18], an ISS model for $\bar{\mathbf{y}}_{R,k}$ may be derived by solution of the discrete-time

algebraic Riccati equation (DARE),

$$\begin{aligned} \bar{P} = & \bar{A}\bar{P}\bar{A}^\top + \bar{\Sigma} - (\bar{A}\bar{P}\bar{A}_{R*}^\top + \bar{\Sigma}_{*R})(\bar{A}_{R*}\bar{P}\bar{A}_{R*}^\top + \bar{\Sigma}_{RR})^{-1} \\ & \times (\bar{A}\bar{P}\bar{A}_{R*}^\top + \bar{\Sigma}_{*R})^\top. \end{aligned} \quad (\text{B7})$$

Specifically, if a unique stabilizing solution $\bar{P}$ of (B7) exists, then $\bar{\mathbf{y}}_{R,k}$ satisfies the innovations-form SS model,

$$\bar{\mathbf{z}}_{k+1} = \bar{A}\bar{\mathbf{z}}_k + \bar{K}^R\bar{\mathbf{e}}_k^R, \quad (\text{B8a})$$

$$\bar{\mathbf{y}}_{R,k} = \bar{A}_{R*}\bar{\mathbf{z}}_k + \bar{\mathbf{e}}_k^R, \quad (\text{B8b})$$

where $\bar{\mathbf{e}}_k^R$ is a Gaussian white noise innovations process (note that this will in general *not* be the same process as $\bar{\mathbf{e}}_{R,k}$). The covariance matrix $\bar{\Sigma}^R$ of $\bar{\mathbf{e}}_k^R$ and the Kalman gain matrix $\bar{K}^R$ are given respectively by

$$\bar{\Sigma}^R = \bar{A}_{R*}\bar{P}\bar{A}_{R*}^\top + \bar{\Sigma}_{RR}, \quad (\text{B9a})$$

$$\bar{K}^R = (\bar{A}\bar{P}\bar{A}_{R*}^\top + \bar{\Sigma}_{*R})(\bar{\Sigma}^R)^{-1}. \quad (\text{B9b})$$

Gutknecht and Barnett [43, Section B] show, furthermore, that $\bar{P}_{33}$ is the only nonzero block in $\bar{P}$ and is the unique stabilizing solution (if it exists) of the lower-dimensional DARE,

$$\begin{aligned} \bar{P}_{33} = & \bar{A}_{33}\bar{P}_{33}\bar{A}_{33}^\top + \bar{\Sigma}_{33} - (\bar{A}_{33}\bar{P}_{33}\bar{A}_{R3}^\top + \bar{\Sigma}_{3R}) \\ & \times (\bar{A}_{R3}\bar{P}_{33}\bar{A}_{R3}^\top + \bar{\Sigma}_{RR})^{-1} (\bar{A}_{33}\bar{P}_{33}\bar{A}_{R3}^\top + \bar{\Sigma}_{3R})^\top. \end{aligned} \quad (\text{B10})$$

The discrete-time reduced innovations covariance matrix and Kalman gain matrix then become, respectively,

$$\bar{\Sigma}^R = \bar{A}_{R3}\bar{P}_{33}\bar{A}_{R3}^\top + \bar{\Sigma}_{RR}, \quad (\text{B11a})$$

$$\bar{K}^R = \begin{bmatrix} I \\ \bar{K}^R \end{bmatrix}, \quad (\text{B11b})$$

where

$$\bar{K}^R = (\bar{A}_{33}\bar{P}_{33}\bar{A}_{R3}^\top + \bar{\Sigma}_{3R})(\bar{\Sigma}^R)^{-1}. \quad (\text{B12})$$

We now consider the limit $\Delta \rightarrow 0$. Letting $P_{33} = \lim_{\Delta \rightarrow 0} \bar{P}_{33}$, we find that P_{33} satisfies the CARE,

$$\begin{aligned} & A_{33}P_{33} + P_{33}A_{33}^\top + \Sigma_{33} \\ & = (P_{33}A_{R3}^\top + \Sigma_{3R})\Sigma_{RR}^{-1}(P_{33}A_{R3}^\top + \Sigma_{3R})^\top \end{aligned} \quad (\text{B13})$$

(we assume from here on that Σ_{RR} is positive definite and hence invertible). Given positive-semidefinite Σ , sufficient conditions for existence of a unique stabilizing solution of (B13) are [32,44]

- (1): (A_{33}, A_{R3}) detectable and
- (2): $(A_{33} - \Sigma_{3R}\Sigma_{RR}^{-1}A_{R3}, \Sigma_{33} - \Sigma_{3R}\Sigma_{RR}^{-1}\Sigma_{R3})$ stabilizable.

If Σ is positive definite, then S2 is trivially satisfied. If A_{33} is stable, then S1 is trivially satisfied, and in Appendix C we show that S1 is satisfied if A is stable. Assuming S1 and S2, we have

$$\Sigma^R = \lim_{\Delta \rightarrow 0} \bar{\Sigma}_{RR} = \Sigma_{RR}, \quad (\text{B14a})$$

$$K^R = \lim_{\Delta \rightarrow 0} \bar{K}^R = \begin{bmatrix} I \\ K^R \end{bmatrix}, \quad (\text{B14b})$$

with

$$\kappa^R = (P_{33}A_{R3}^T + \Sigma_{RR})\Sigma_{RR}^{-1}. \quad (\text{B15})$$

Now let $\bar{\mathbf{y}}_{R,k} = \mathbf{y}_R(k\Delta)$ be the discretized reduced process (B8b). As demonstrated in Appendix A, to lowest order in Δ the VMA coefficient matrices of $\bar{\mathbf{y}}_{R,k}$ are $\bar{B}_k^R = B^R(k\Delta)$, where $B^R(u)$ is the VMA kernel of the continuous-time reduced process $\mathbf{y}_R(t)$. From (B6), it is easily calculated that $\bar{B}_0^R = I$ and

$$\bar{B}_k^R = \bar{A}_{R*}\bar{A}^{k-1}\bar{K}^R, \quad k > 0. \quad (\text{B16})$$

Let us fix $u > 0$. Making explicit the dependency of $\bar{B}_k^R, \bar{A}^R, \bar{K}^R$, etc., on $\Delta = u/k$, we have

$$\begin{aligned} B^R(u) &= \lim_{k \rightarrow \infty} \bar{B}_k^R(u/k) \\ &= \lim_{k \rightarrow \infty} \bar{A}_{R*}(u/k) \cdot \bar{A}(u/k)^{k-1} \cdot \bar{K}^R(u/k) \quad \text{from (B16)} \\ &= \lim_{k \rightarrow \infty} [I + A_{RR}(u/k) A_{R3}(u/k)] \cdot [I + A(u/k)] \\ &\quad \cdot [I + A(u/k)]^k \cdot \begin{bmatrix} I \\ \kappa^R \end{bmatrix} \quad \text{from (B4a) and (B14b)} \\ &= [I \ 0] \cdot e^{Au} \cdot \begin{bmatrix} I \\ \kappa^R \end{bmatrix}, \end{aligned}$$

where in the last step we use $\lim_{k \rightarrow \infty} [I + A(u/k)]^k = e^{Au}$. In particular, we have

$$\bar{B}^R(0) = [I \ 0] \begin{bmatrix} A_{RR} & A_{R3} \\ A_{3R} & A_{33} \end{bmatrix} \begin{bmatrix} I \\ \kappa^R \end{bmatrix} = A_{RR} + A_{R3}\kappa^R. \quad (\text{B17})$$

Putting all this together, from (7), (9), (B15), and (B17), some straightforward algebra yields

$$D^R - D_{RR} = A_{R3}P_{33}A_{R3}^T, \quad (\text{B18})$$

so that in particular

$$D_{11}^R - D_{11} = A_{13}P_{33}A_{13}^T \quad (\text{B19})$$

and from (6) we have

$$\mathcal{R}_{y_3 \rightarrow y_1|y_2} = \text{trace}[\Sigma_{11}^{-1}A_{13}P_{33}A_{13}^T] \quad (\text{B20})$$

as stated in Theorem 1.

Finally, for a (not necessarily stable) CTVAR process $\mathbf{y}(t)$ initialized at $t = 0$, we define the transfer entropy rate analogously to the GC rate as

$$\mathcal{T}_{y_3 \rightarrow y_1|y_2} = \lim_{h \rightarrow 0} \frac{1}{h} \mathcal{T}_{y_3 \rightarrow y_1|y_2}(h), \quad (\text{B21})$$

with

$$\begin{aligned} \mathcal{T}_{y_3 \rightarrow y_1|y_2}(h) &= \mathbf{H}[\mathbf{y}_1(t+h) | \mathbf{y}_R(u) : 0 \leq u \leq t] \\ &\quad - \mathbf{H}[\mathbf{y}_1(t+h) | \mathbf{y}(u) : 0 \leq u \leq t], \quad (\text{B22}) \end{aligned}$$

the TE at finite prediction horizon $h > 0$, where $\mathbf{H}[\cdots | \cdots]$ denotes conditional differential entropy. This is equivalent to the continuous-time TE rate as set out in Cooper and Edgar [45]. Note now, that for multivariate-normal variables $\mathbf{y}, \mathbf{x}$, the conditional differential entropy $\mathbf{H}[\mathbf{y} | \mathbf{x}]$ is given by $\mathbf{H}[\boldsymbol{\varepsilon}]$, where $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \Sigma)$ is the residual error of the projection $\mathbb{E}[\mathbf{y} | \mathbf{x}]$ of $\mathbf{y}$ on $\mathbf{x}$, i.e., $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbb{E}[\mathbf{y} | \mathbf{x}]$, and we have [46] $\mathbf{H}[\boldsymbol{\varepsilon}] = \frac{1}{2} \log |\Sigma|$. By a similar stroboscopic discretization argument to that in Appendix A, and since increments of the Wiener noise $\mathbf{w}(t)$ are multivariate normal, the conditional entropies on the right-hand side of (B22) are seen to be $\frac{1}{2} \log |\mathcal{E}_{11}^R(h)|$ and $\frac{1}{2} \log |\mathcal{E}_{11}(h)|$, respectively [47], with $\mathcal{E}(h)$ as in (3) and $\mathcal{E}^R(h)$ its reduced-process counterpart. Thus from (4) we have $\mathcal{F}_{y_3 \rightarrow y_1|y_2}(h) = 2\mathcal{T}_{y_3 \rightarrow y_1|y_2}(h)$, and from (5) $\mathcal{F}_{y_3 \rightarrow y_1|y_2} = 2\mathcal{T}_{y_3 \rightarrow y_1|y_2}$ as stated in Theorem 1.

APPENDIX C: PROOF OF DETECTABILITY FOR STABLE FULL MATRIX

Given an $n \times n$ matrix,

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad (\text{C1})$$

we show that if A is Hurwitz stable, then (A_{11}, A_{21}) is detectable.

We use a Popov-Belevitch-Hautus test [48]. Suppose that (A_{11}, A_{21}) is *not* detectable. This implies that A_{11} has an eigenvalue λ with $\text{Re}(\lambda) \geq 0$ such that

$$\text{rank} \begin{bmatrix} \lambda I - A_{11} \\ A_{21} \end{bmatrix} < n_1, \quad (\text{C2})$$

where n_1 is the size of A_{11} , so that there exists a nonzero vector $\mathbf{v}$ such that

$$\begin{bmatrix} \lambda I - A_{11} \\ A_{21} \end{bmatrix} \mathbf{v} = 0, \quad (\text{C3})$$

or

$$A_{11}\mathbf{v} = \lambda\mathbf{v}, \quad (\text{C4a})$$

$$A_{21}\mathbf{v} = 0. \quad (\text{C4b})$$

Now

$$\begin{aligned} A \begin{bmatrix} \mathbf{v} \\ 0 \end{bmatrix} &= \begin{bmatrix} A_{11}\mathbf{v} \\ A_{21}\mathbf{v} \end{bmatrix} \\ &= \begin{bmatrix} \lambda\mathbf{v} \\ 0 \end{bmatrix} \quad \text{by (C4)} \\ &= \lambda \begin{bmatrix} \mathbf{v} \\ 0 \end{bmatrix}, \end{aligned}$$

so that λ is an eigenvalue of A . But by assumption λ is unstable, which contradicts Hurwitz stability of A . Thus (A_{11}, A_{21}) must be detectable.

- [1] N. Wiener, The theory of prediction, in *Modern Mathematics for Engineers*, edited by E. F. Beckenbach (McGraw-Hill, New York, 1956), pp. 165–190.
- [2] C. W. J. Granger, Economic processes involving feedback, *Inform. Control* **6**, 28 (1963).
- [3] J. Geweke, Measurement of linear dependence and feedback between multiple time series, *J. Am. Stat. Assoc.* **77**, 304 (1982).
- [4] J. Geweke, Measures of conditional linear dependence and feedback between time series, *J. Am. Stat. Assoc.* **79**, 907 (1984).
- [5] It is underappreciated that even if a (stationary) stochastic process has a nonlinear generative mechanism, it may nonetheless be amenable to linear modelling. Specifically, under mild conditions it will have a Wold representation as a linear moving-average which may, if minimum-phase conditions hold, be suitable for Granger-causal analysis.
- [6] T. Schreiber, Measuring information transfer, *Phys. Rev. Lett.* **85**, 461 (2000).
- [7] M. Paluš, V. Komárek, Z. Hrnčíř, and K. Štěrbová, Synchronization as adjustment of information rates: Detection from bivariate time series, *Phys. Rev. E* **63**, 046211 (2001).
- [8] L. Barnett, A. B. Barrett, and A. K. Seth, Granger causality and transfer entropy are equivalent for Gaussian variables, *Phys. Rev. Lett.* **103**, 238701 (2009).
- [9] R. E. Spinney, M. Prokopenko, and J. T. Lizier, Transfer entropy in continuous time, with applications to jump and neural spiking processes, *Phys. Rev. E* **95**, 032319 (2017).
- [10] P. Shahsavari Baboukani, C. Graversen, E. Alickovic, and J. Østergaard, Estimating conditional transfer entropy in time series using mutual information and nonlinear prediction, *Entropy* **22**, 1124 (2020).
- [11] B. Wahl, U. Feudel, J. Hlinka, M. Wächter, J. Peinke, and J. A. Freund, Granger-causality maps of diffusion processes, *Phys. Rev. E* **93**, 022213 (2016).
- [12] B. Wahl, U. Feudel, J. Hlinka, M. Wächter, J. Peinke, and J. A. Freund, Conditional Granger causality of diffusion processes, *Eur. Phys. J. B* **90**, 197 (2017).
- [13] J.-P. Florens and D. Fougère, Noncausality in continuous time, *Econometrica* **64**, 1195 (1996).
- [14] L. Barnett and A. K. Seth, Detectability of Granger causality for subsampled continuous-time neurophysiological processes, *J. Neurosci. Methods* **275**, 93 (2017).
- [15] G. T. Wilson, The factorization of matricial spectral densities, *SIAM J. Appl. Math.* **23**, 420 (1972).
- [16] M. Dhamala, G. Rangarajan, and M. Ding, Estimating Granger causality from Fourier and wavelet transforms of time series data, *Phys. Rev. Lett.* **100**, 018701 (2008).
- [17] That is, global GC rates; as VOU processes are linear, there is nothing to gain from local linearization – see Section III.
- [18] L. Barnett and A. K. Seth, Granger causality for state-space models, *Phys. Rev. E* **91**, 040101(R) (2015).
- [19] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications* (Springer-Verlag, Berlin, 2003).
- [20] Throughout, subscripts 1,2,3 on vectors and matrices are multi-indices corresponding to the given partitioning.
- [21] L. Barnett and T. Bossomaier, Transfer entropy as a log-likelihood ratio, *Phys. Rev. Lett.* **109**, 138105 (2012).
- [22] A. K. Seth, Causal networks in simulated neural systems, *Cogn. Neurodyn.* **2**, 49 (2008).
- [23] We assume all requisite derivatives of $f(y)$ and $\Sigma(y)$ exist.
- [24] For Langevin systems, “multiplicative” refers to noise with state-dependent magnitude [49]). It is unclear to us how restrictive the weakly multiplicative assumption might be in the context of the various physical scenarios typically modelled by Langevin equations; while it would be of interest to consider whether the theory might be extended to relax this assumption, we leave that for future work.
- [25] E. N. Lorenz, Deterministic nonperiodic flow, *J. Atmos. Sci.* **20**, 130 (1963).
- [26] The time unit is arbitrary labelled as “seconds.”
- [27] The ode45 solver uses an adaptive step size which may not correspond to the sampling increment h .
- [28] L. Bereska and S. Gavves, Mechanistic interpretability for AI safety—A review, [arXiv:2404.14082](https://arxiv.org/abs/2404.14082).
- [29] F. Comte and E. Renault, Noncausality in continuous time models, *Econ. Theory* **12**, 215 (1996).
- [30] K. Kawasaki, Simple derivations of generalized linear and nonlinear Langevin equations, *J. Phys. A* **6**, 1289 (1973).
- [31] V. Solo, State-space analysis of Granger-Geweke causality measures with application to fMRI, *Neural Comput.* **28**, 914 (2016).
- [32] W. Arnold and A. J. Laub, Generalized eigenproblem algorithms and software for algebraic Riccati equations, *Proc. IEEE* **72**, 1746 (1984).
- [33] L. Barnett, J. T. Lizier, M. Harré, A. K. Seth, and T. Bossomaier, Information flow in a kinetic Ising model peaks in the disordered phase, *Phys. Rev. Lett.* **111**, 177203 (2013).
- [34] See, for example, NVIDIA’s cuSOLVER, which provides GPU platform support for various matrix decompositions (QR, SVD, LU, eigensolvers, etc.) from which CARE solvers may be constructed.
- [35] We have encountered one curious edge case for which we do not as yet have an adequate explanation. Here, though the requisite Jacobians are nonsingular and the appropriate detectability criteria are indeed satisfied, QZ-based CARE solvers fail; solution matrices blow up in magnitude, yielding inaccurate results. This may possibly be a frailty of the Schur decomposition algorithm which, while apparently rare, merits future investigation.
- [36] R. Hindriks, R. Jansen, F. Bijma, H. D. Mansvelder, M. C. M. de Gunst, and A. W. van der Vaart, Unbiased estimation of Langevin dynamics from time series with application to hippocampal field potentials in vitro, *Phys. Rev. E* **84**, 021133 (2011).
- [37] M. Tabar and R. Rahimi, Kramers-Moyal expansion and Fokker-Planck equation, in *Analysis and Data-Based Reconstruction of Complex Nonlinear Dynamical Systems: Using the Methods of Stochastic Processes* (Springer International, Cham, 2019), pp. 19–29.
- [38] P. P. Lin, M. Wächter, J. Peinke, M. Tabar, and R. Rahimi, Assessing the suitability of the Langevin equation for analyzing measured data through downsampling, *J. Phys.: Complexity* **6**, 015016 (2025).

- [39] L. Barnett and A. K. Seth, Dynamical independence: Discovering emergent macroscopic processes in complex dynamical systems, *Phys. Rev. E* **108**, 014304 (2023).
- [40] <https://github.com/lcbarnett/VOUGC>.
- [41] J. D. Hamilton, *Time Series Analysis* (Princeton University Press, Princeton, NJ, 1994).
- [42] Although Barnett and Seth [14] consider only the stable case, the result may be shown to hold in the unstable case.
- [43] A. J. Gutknecht and L. Barnett, Sampling distribution for single-regression Granger causality estimators, *Biometrika* **110**, 933 (2023).
- [44] M.-L. Ni, Existence condition on solutions to the algebraic Riccati equation, *Acta Automat. Sinica* **34**, 85 (2008).
- [45] J. N. Cooper and C. D. Edgar, A development of continuous-time transfer entropy, [arXiv:1905.06406](https://arxiv.org/abs/1905.06406).
- [46] Up to an additive constant that depends only on the dimension of the variable $\mathbf{y}$.
- [47] This confirms that, as our notation suggests, the expression for $\mathcal{T}_{y_3 \rightarrow y_1 | y_2}(h)$ in (B22) does not depend on the time stamp t .
- [48] T. Kailath, *Linear Systems* (Prentice-Hall, Hoboken, NJ, 1980).
- [49] P. Arnold, Symmetric path integrals for stochastic equations with multiplicative noise, *Phys. Rev. E* **61**, 6099 (2000).