Exploration of synergistic and redundant information sharing in static and dynamical Gaussian systems

Adam B. Barrett*
Sackler Centre for Consciousness Science, Department of Informatics, University of Sussex, Brighton BN1 9QJ, United Kingdom and Department of Clinical Sciences, University of Milan, Milan 20157, Italy
(Received 13 November 2014; published 8 May 2015)

To fully characterize the information that two source variables carry about a third target variable, one must decompose the total information into redundant, unique, and synergistic components, i.e., obtain a partial information decomposition (PID). However, Shannon's theory of information does not provide formulas to fully determine these quantities. Several recent studies have begun addressing this. Some possible definitions for PID quantities have been proposed and some analyses have been carried out on systems composed of discrete variables. Here we present an in-depth analysis of PIDs on Gaussian systems, both static and dynamical. We show that, for a broad class of Gaussian systems, previously proposed PID formulas imply that (i) redundancy reduces to the minimum information provided by either source variable and hence is independent of correlation between sources, and (ii) synergy is the extra information contributed by the weaker source when the stronger source is known and can either increase or decrease with correlation between sources. We find that Gaussian systems frequently exhibit net synergy, i.e., the information carried jointly by both sources is greater than the sum of information carried by each source individually. Drawing from several explicit examples, we discuss the implications of these findings for measures of information transfer and information-based measures of complexity, both generally and within a neuroscience setting. Importantly, by providing independent formulas for synergy and redundancy applicable to continuous time-series data, we provide an approach to characterizing and quantifying information sharing amongst complex system variables.

DOI: 10.1103/PhysRevE.91.052802 PACS number(s): 89.75.−k, 89.70.−a, 87.19.lo

I. INTRODUCTION

Shannon's information theory [1] has provided extremely successful methodology for understanding and quantifying information transfer in systems conceptualized as receiver and transmitter, or stimulus and response [2,3]. Formulating information as reduction in uncertainty, the theory quantifies the information $I(X;Y)$ that one variable Y holds about another variable X as the average reduction in the surprise of the outcome of X when knowing the outcome of Y compared to when not knowing the outcome of Y. (Surprise is defined by how unlikely an outcome is and is given by the negative of the logarithm of the probability of the outcome. This quantity is usually referred to as the mutual information since it is symmetric in X and Y.) Recently, information theory has become a popular tool for the analysis of so-called complex systems of many variables, for example, for attempting to understand emergence, self-organization, and phase transitions and to measure complexity [4]. Information theory does not, however, in its current form, provide a complete description of the informational relationships between variables in a system composed of three or more variables. The information $I(X;Y,Z)$ that two source variables Y and Z hold about a third target variable X should decompose into four parts:[1] (i) $U(X;Y|Z)$, the unique information that only Y (out of Y and Z) holds about X; (ii) $U(X;Z|Y)$, the unique information that only Z holds about X; (iii) $R(X;Y,Z)$, the redundant information that both Y and Z hold about X; and (iv) $S(X;Y,Z)$, the synergistic information about X that only arises from knowing both Y and Z (see Fig. 1). The set of quantities $\{U(X;Y|Z),U(X;Z|Y),R(X;Y,Z),S(X;Y,Z)\}$ is called a partial information decomposition (PID). Information theory gives us the following set of equations for them:

$$I(X;Y,Z) = U(X;Y|Z) + U(X;Z|Y) + S(X;Y,Z) + R(X;Y,Z), \quad (1)$$

$$I(X;Y) = U(X;Y|Z) + R(X;Y,Z), \quad (2)$$

$$I(X;Z) = U(X;Z|Y) + R(X;Y,Z). \quad (3)$$

However, these equations do not uniquely determine the PID. One cannot obtain synergy or redundancy in isolation, but only the net synergy

$$\Delta I(X;Y,Z) =: I(X;Y,Z) - I(X;Y) - I(X;Z) = S(X;Y,Z) - R(X;Y,Z). \quad (4)$$

An additional ingredient to the theory is required, specifically, a definition that determines one of the four quantities in the PID. A consistent and well-understood approach to PIDs would extend Shannon information theory into a more complete framework for the analysis of information storage and transfer in complex systems.

In addition to the four equations above, the minimal further axioms that a PID of information from two sources should

*Adam.Barrett@sussex.ac.uk

[1]It is our convenient convention of terminology to refer to variables as sources and targets, with Y and Z always being the sources that contribute information about the target variable X. These terms relate to the status of the variables in the informational quantities that we compute and should not be considered as describing the dynamical roles played by the variables.

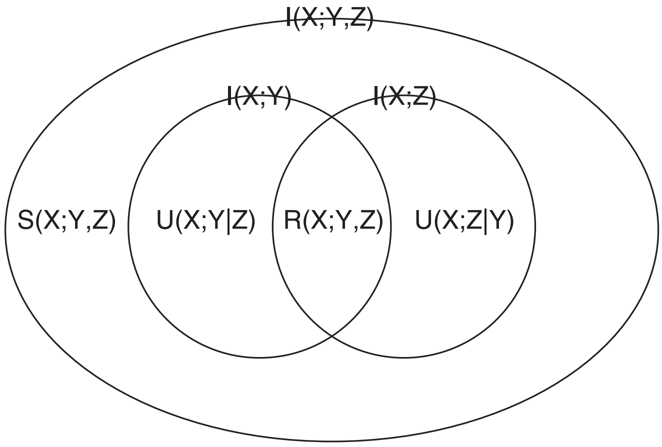

FIG. 1. General structure of the information that two source variables $\boldsymbol{Y}$ and $\boldsymbol{Z}$ hold about a third target variable $\boldsymbol{X}$. The ellipses indicate $I(\boldsymbol{X};\boldsymbol{Y})$, $I(\boldsymbol{X};\boldsymbol{Z})$, and $I(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z})$ as labeled and the four distinct regions enclosed represent the redundancy $R(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z})$, the synergy $S(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z})$, and the unique information $U(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{Z})$ and $U(\boldsymbol{X};\boldsymbol{Z}|\boldsymbol{Y})$ as labeled.

satisfy are (i) that the four quantities $U(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{Z})$, $U(\boldsymbol{X};\boldsymbol{Z}|\boldsymbol{Y})$, $R(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z})$, and $S(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z})$ should always all be greater than or equal to zero and (ii) that redundancy $R(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z})$ and synergy $S(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z})$ are symmetric with respect to $\boldsymbol{Y}$ and $\boldsymbol{Z}$ [5–10]. Interestingly, several distinct PID definitions have been proposed, each arising from a distinct idea about what exactly should constitute redundancy and/or synergy. These previous studies of PIDs have focused on systems composed of discrete variables. Here, by considering PIDs on Gaussian systems, we provide a study of PIDs that focuses on continuous random variables.

One might naively expect that for sources and target being jointly Gaussian, the linear relationship between the variables would imply zero synergy and hence a trivial PID with the standard information theory equations (1)–(3) determining the redundant and unique information. However, this is not the case; net synergy (4), and hence synergy, can be positive [11,12]. We begin this study (Sec. III) by illustrating the prevalence of jointly Gaussian cases for which net synergy (4) is positive. Of particular note is the fact that there can be positive net synergy when sources are uncorrelated. After this motivation for the study, in Sec. IV we introduce three distinct previously proposed PID procedures: (i) that of Williams and Beer [5]; (ii) that of Griffith *et al.* [6,9] and Bertschinger *et al.* [8,10]; and (iii) that of Harder *et al.* [7]. In addition to satisfying the minimal axioms above, these PIDs have the further commonality that redundant and unique information depend only on the pair of marginal distributions of each individual source with the target, i.e., those of $(\boldsymbol{X},\boldsymbol{Y})$ and $(\boldsymbol{X},\boldsymbol{Z})$, while only the synergy depends on the full joint distribution of all three variables $(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z})$. Bertschinger *et al.* [10] have argued for this property by considering unique information from a game-theoretic view point. Our key result, which we then demonstrate, is that for a jointly Gaussian system with a univariate target and sources of arbitrary dimension, any PID with this property reduces to simply taking redundancy as the minimum of the mutual informations $I(\boldsymbol{X};\boldsymbol{Y})$ and $I(\boldsymbol{X};\boldsymbol{Z})$ and letting the other quantities follow from (1)–(3). This common PID, which we call the minimum mutual information (MMI) PID, (i) always assigns the source providing less information about the target as providing zero unique information, (ii) yields redundancy as being independent of the correlation between sources, and (iii) yields synergy as the extra information contributed by the weaker source when the stronger source is known. In Sec. V we proceed to explore partial information in several example dynamical Gaussian systems, examining (i) the behavior of net synergy, which is independent of any assumptions on the particular choice of PID, and (ii) redundancy and synergy according to the MMI PID. We then discuss implications for the transfer entropy measure of information flow (Sec. VI) and measures that quantify the complexity of a system via information flow analysis (Sec. VII). We conclude with a discussion of the shortcomings and possible extensions to existing approaches to PIDs and the measurement of information in complex systems.

This paper provides tools for exploring information sharing in complex systems that go beyond what standard Shannon information theory can provide. By providing a PID for triplets of Gaussian variables, it will enable one to study synergy among continuous time-series variables, independently of redundancy. In Sec. VIII we consider possible application to the study of information sharing among brain variables in neuroscience. More generally, there exists the possibility of application to complex systems in any realm, e.g., climate science, financial systems, and computer networks.

II. NOTATION AND PRELIMINARIES

Let $\boldsymbol{X}$ be a continuous random variable of dimension m. We denote the probability density function by $P_{\boldsymbol{X}}(\boldsymbol{x})$, the mean by $\bar{\boldsymbol{x}}$, and the $m \times m$ matrix of covariances $\mathrm{cov}(X^i,X^j)$ by $\Sigma(\boldsymbol{X})$. Let $\boldsymbol{Y}$ be a second random variable of dimension n. We denote the $m \times n$ matrix of cross covariances $\mathrm{cov}(X^i,Y^j)$ by $\Sigma(\boldsymbol{X},\boldsymbol{Y})$. We define the partial covariance of $\boldsymbol{X}$ with respect to $\boldsymbol{Y}$ as

$$\Sigma(\boldsymbol{X}|\boldsymbol{Y}) =: \Sigma(\boldsymbol{X}) - \Sigma(\boldsymbol{X},\boldsymbol{Y})\Sigma(\boldsymbol{Y})^{-1}\Sigma(\boldsymbol{Y},\boldsymbol{X}). \quad (5)$$

If $\boldsymbol{X} \oplus \boldsymbol{Y}$ is multivariate Gaussian (we use the symbol $\oplus$ to denote vertical concatenation of vectors), then the partial covariance $\Sigma(\boldsymbol{X}|\boldsymbol{Y})$ is precisely the covariance matrix of the conditional variable $\boldsymbol{X}|\boldsymbol{Y}=\boldsymbol{y}$, for any $\boldsymbol{y}$,

$$\boldsymbol{X}|(\boldsymbol{Y}=\boldsymbol{y}) \sim \mathcal{N}[\boldsymbol{\mu}_{\boldsymbol{X}|\boldsymbol{Y}=\boldsymbol{y}}, \Sigma(\boldsymbol{X}|\boldsymbol{Y})], \quad (6)$$

where $\boldsymbol{\mu}_{\boldsymbol{X}|\boldsymbol{Y}=\boldsymbol{y}} = \bar{\boldsymbol{x}} + \Sigma(\boldsymbol{X},\boldsymbol{Y})\Sigma(\boldsymbol{Y})^{-1}(\boldsymbol{y}-\bar{\boldsymbol{y}})$.

Entropy H characterizes uncertainty and is defined as

$$H(\boldsymbol{X}) =: -\int P_{\boldsymbol{X}}(\boldsymbol{x}) \ln P_{\boldsymbol{X}}(\boldsymbol{x}) d^m\boldsymbol{x}. \quad (7)$$

[Note that, strictly, Eq. (7) is the differential entropy since entropy itself is infinite for continuous variables. However, considering continuous variables as continuous limits of discrete variable approximations, entropy differences and hence information remain well defined in the continuous limit and may be consistently measured using Eq. (7) [2]. Moreover, this equation assumes that $\boldsymbol{X}$ has a density with respect to the Lebesgue measure $d^m\boldsymbol{x}$; this assumption is upheld whenever we discuss continuous random variables.] The conditional entropy $H(\boldsymbol{X}|\boldsymbol{Y})$ is the expected entropy of $\boldsymbol{X}$ given $\boldsymbol{Y}$, i.e.,

$$H(\boldsymbol{X}|\boldsymbol{Y}) =: \int H(\boldsymbol{X}|\boldsymbol{Y}=\boldsymbol{y})P_{\boldsymbol{Y}}(\boldsymbol{y})d^n\boldsymbol{y}. \tag{8}$$

The mutual information $I(\boldsymbol{X};\boldsymbol{Y})$ between $\boldsymbol{X}$ and $\boldsymbol{Y}$ is the average information, or reduction in uncertainty (entropy), about $\boldsymbol{X}$, knowing the outcome of $\boldsymbol{Y}$:

$$I(\boldsymbol{X};\boldsymbol{Y}) = H(\boldsymbol{X}) - H(\boldsymbol{X}|\boldsymbol{Y}). \tag{9}$$

Mutual information can also be written in the useful form

$$I(\boldsymbol{X};\boldsymbol{Y}) = H(\boldsymbol{X}) + H(\boldsymbol{Y}) - H(\boldsymbol{X},\boldsymbol{Y}), \tag{10}$$

from which it follows that mutual information is symmetric in $\boldsymbol{X}$ and $\boldsymbol{Y}$ [2]. The joint mutual information that two sources $\boldsymbol{Y}$ and $\boldsymbol{Z}$ share with a target $\boldsymbol{X}$ satisfies a chain rule

$$I(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z}) = I(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{Z}) + I(\boldsymbol{X};\boldsymbol{Z}), \tag{11}$$

where the conditional mutual information $I(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{Z})$ is the expected mutual information between $\boldsymbol{X}$ and $\boldsymbol{Y}$ given $\boldsymbol{Z}$. For $\boldsymbol{X}$ Gaussian,

$$H(\boldsymbol{X}) = \tfrac{1}{2}\ln[\det\Sigma(\boldsymbol{X})] + \tfrac{1}{2}m\ln(2\pi e), \tag{12}$$

and for $\boldsymbol{X}\oplus\boldsymbol{Y}$ Gaussian,

$$H(\boldsymbol{X}|\boldsymbol{Y}) = \frac{1}{2}\ln[\det\Sigma(\boldsymbol{X}|\boldsymbol{Y})] + \frac{1}{2}m\ln(2\pi e), \tag{13}$$

$$I(\boldsymbol{X};\boldsymbol{Y}) = \frac{1}{2}\ln\left[\frac{\det\Sigma(\boldsymbol{X})}{\det\Sigma(\boldsymbol{X}|\boldsymbol{Y})}\right]. \tag{14}$$

For $\boldsymbol{X}$ a dynamical variable evolving in discrete time, we denote the state at time t by $\boldsymbol{X}_t$ and the infinite past with respect to time t by $\boldsymbol{X}_t^- =: \boldsymbol{X}_{t-1}\oplus\boldsymbol{X}_{t-2},\ldots$. The p past states with respect to time t are denoted by $\boldsymbol{X}_t^{(p)} =: \boldsymbol{X}_{t-1}\oplus\boldsymbol{X}_{t-2}\oplus\ldots\oplus\boldsymbol{X}_{t-p}$.

III. SYNERGY IS PREVALENT IN GAUSSIAN SYSTEMS

In this section we demonstrate the prevalence of synergy in jointly Gaussian systems and hence that the PIDs for such systems are typically nontrivial. We do this by computing the net synergy, i.e., synergy minus redundancy ΔI (4). Since the axioms for a PID impose that S and R are greater than or equal to zero, this quantity provides a lower bound on synergy and in particular a sufficient condition for nonzero synergy is $\Delta I(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z})>0$. Some special cases have previously been considered in [11,12]. Here we consider the most general three-dimensional jointly Gaussian system $(X,Y,Z)^{\mathrm{T}}$ (here we use lightface rather than boldface type for the random variables since they are one dimensional). Setting means and variances of the individual variables to 0 and 1, respectively, preserves all mutual information between the variables and so without loss of generality this system can be specified with a covariance matrix of the form

$$\Sigma = \begin{pmatrix} 1 & a & c \\ a & 1 & b \\ c & b & 1 \end{pmatrix}, \tag{15}$$

where a, b, and c satisfy $|a|,|b|,|c|<1$ and

$$2abc - a^2 - b^2 - c^2 + 1 > 0 \tag{16}$$

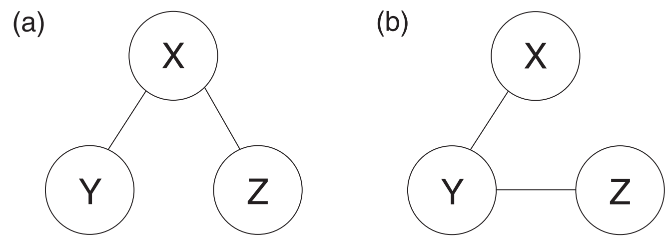

FIG. 2. Correlational structure of two example systems of univariate Gaussian variables for which Y and Z exhibit positive net synergy with respect to information about X. Variables are shown as circles and the variables that are correlated are joined by lines. (a) Y and Z are uncorrelated and yet show synergy. (b) X and Z are uncorrelated and yet Z contributes synergistic information about X in conjunction with Y. See the text for details.

(a covariance matrix must be nonsingular and positive definite).

Using (5) and (14), the mutual information between X and Y and Z is given by

$$I(X;Y) = \frac{1}{2}\ln\left(\frac{1}{1-a^2}\right), \tag{17}$$

$$I(X;Z) = \frac{1}{2}\ln\left(\frac{1}{1-c^2}\right), \tag{18}$$

$$I(X;Y,Z) = \frac{1}{2}\ln\left(\frac{1-b^2}{1-(a^2+b^2+c^2)+2abc}\right) \tag{19}$$

and thus the general formula for the net synergy is

$$\Delta I(X;Y,Z) = \frac{1}{2}\ln\left[\frac{(1-a^2)(1-b^2)(1-c^2)}{1-(a^2+b^2+c^2)+2abc}\right]. \tag{20}$$

This quantity is often greater than zero. Two specific examples illustrate the prevalence of net synergy in an interesting way. Consider first the case $a=c$ and $b=0$, i.e., the sources each have the same correlation with the target, but the two sources are uncorrelated [see Fig. 2(a)]. Then there is net synergy since

$$\Delta I(X;Y,Z) = \frac{1}{2}\ln\left(\frac{1-2a^2+a^4}{1-2a^2}\right) > 0. \tag{21}$$

It is remarkable that there can be net synergy when the two sources are not correlated. However, this can be explained by the concave property of the logarithm function. If one instead quantified information as reduction in covariance, the net synergy would be zero in this case. That is, if we were to define $I_\Sigma(X;Y) =: \Sigma(X) - \Sigma(X|Y)$, etc., and $\Delta I_\Sigma =: I_\Sigma(X;Y,Z) - I_\Sigma(X;Y) - I_\Sigma(X;Z)$, then we would have

$$\Delta I_\Sigma(X;Y,Z) = \frac{(a^2+c^2)b^2-2abc}{1-b^2}, \tag{22}$$

which gives the output of zero whenever the correlation b between sources is zero. This is intuitive: The sum of the reductions in covariance of the target given each source individually equals the reduction in covariance of the target given both sources together, for the case of no correlation between sources. There is net synergy in the Shannon information provided by the sources about the target because this quantity is obtained by combining these reductions in covariance nonlinearly via the concave logarithm function. This suggests that perhaps I_Σ would actually be a better measure of information for Gaussian variables than Shannon information (although unlike standard mutual information I_Σ is not symmetric). Note that Angelini *et al.* [12] proposed a version of Granger causality (which is a measure of information flow for variables that are at least approximately Gaussian [13]) based on straightforward difference of variances without the usual logarithm precisely so that for a linear system the Granger causality from a group of variables equals the sum of Granger causalities from members of the group (see Sec. VI for a recap of the concept of Granger causality).

Second, we consider the case $c = 0$, i.e., in which there is no correlation between the target X and the second source Z [see Fig. 2(b)]. In this case we have

$$\Delta I(X;Y,Z) = \frac{1}{2}\ln\left(\frac{1-a^2-b^2+a^2b^2}{1-a^2-b^2}\right) > 0. \quad (23)$$

Hence, the two sources Y and Z exhibit synergistic information about the target X even though X and Z are uncorrelated and this is modulated by the correlation between the sources Y and Z. Although this is perhaps from a naive point of view counterintuitive, it can be explained by thinking of Z as providing information about why Y has taken the value it has and from this one can narrow down the range of values for X, beyond what was already known about X just from knowing Y. Note that in this case there would be net synergy even if one quantified information as reduction in covariance via $I_\Sigma(X;Y)$ defined above.

Figures 3(a), 3(b), 3(d), and 3(e) show more generally how net synergy depends on the correlation between source variables Y and Z. For correlations a and c between the two sources and the target being equal and positive, net synergy is a decreasing function of the correlation b between the sources, while for correlations a and c being equal but opposite, net synergy is an increasing function of the correlation b between sources [Fig. 3(a)]. Net synergy asymptotically approaches infinity as the correlation values approach limits at which the covariance matrix becomes singular. This makes sense because in those limits X becomes completely determined by Y and Z. More generally, when a and c are unequal, net synergy is a U-shaped function of correlation between sources [Fig. 3(d)]. In Figs. 3(b) and 3(e) the alternative measure ΔI_Σ of net synergy based on information as reduction in variance is plotted. As described above, this measure behaves more elegantly, always taking the value zero when the correlation between sources is zero. Taken together these plots show that net redundancy (negative net synergy) does not necessarily indicate a high degree of correlation between source variables.

This exploration of net synergy demonstrates that it would be useful to obtain explicit measures of synergy and redundancy for Gaussian variables. As mentioned in the Introduction, several measures have been proposed for discrete variables [5–10]. In the next section we will see that, for a broad class of jointly Gaussian systems, these all reduce essentially to redundancy being the minimum of $I(\boldsymbol{X};\boldsymbol{Y})$ and $I(\boldsymbol{X};\boldsymbol{Z})$.

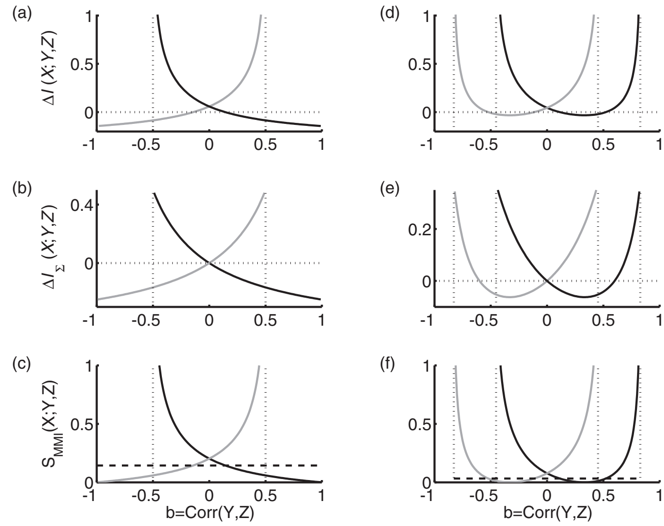

FIG. 3. Illustrative examples of net synergy and synergy S_{MMI} between Gaussian variables. (a) Net synergy in Shannon information that sources Y and Z share about the target X, as a function of the correlation between Y and Z for (black) correlations between X and Y and between X and Z equal and both positive ($a = c = 0.5$) and (gray) correlations between X and Y and between X and Z equal and opposite ($a = -c = 0.5$). (b) Same as (a) but using information defined as reduction in variance instead of reduction in Shannon entropy. (c) Synergy according to the MMI PID for the same parameters as (a). Here the dashed line shows redundancy according to the MMI PID, which does not depend on the correlation between Y and Z. (d) Example of net synergy as a function of the correlation between Y and Z for (black) correlations between X and Y and between X and Z unequal and both positive ($a = 0.25, c = 0.75$) and (gray) correlations between X and Y and between X and Z unequal and of opposite sign ($a = 0.25, c = -0.75$). (e) Same as (d) but using information defined as reduction in variance instead of reduction in Shannon entropy. (f) Synergy according to the MMI PID for the same parameters as (d). Here the dashed line shows redundancy according to the MMI PID, which does not depend on the correlation between Y and Z. See the text for full details of the parameters. In all panels dotted vertical lines indicate boundaries of the allowed parameter space, at which the measures go to infinity, and horizontal dotted lines indicate zero.

IV. PARTIAL INFORMATION DECOMPOSITION ON GAUSSIAN SYSTEMS

In this section we first revise the definitions of three previously proposed PIDs. We note that all of them have the property that redundant and unique information depend only on the pair of marginal distributions of each individual source with the target, i.e., those of $(\boldsymbol{X},\boldsymbol{Y})$ and $(\boldsymbol{X},\boldsymbol{Z})$, while only the synergy depends on the full joint distribution of all three variables $(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z})$. Bertschinger *et al.* [10] have argued for this property by considering unique information from a game-theoretic view point. We then prove our key result, namely, that any PID satisfying this property reduces, for a jointly Gaussian system with a univariate target and sources of arbitrary dimension, to simply taking redundancy as the minimum of $I(\boldsymbol{X};\boldsymbol{Y})$ and $I(\boldsymbol{X};\boldsymbol{Z})$ and letting the other quantities follow from (1)–(3). We term this PID the MMI PID and give full formulas for it for the general fully univariate case considered in Sec. III. In Sec. V we go on to apply the MMI PID to dynamical Gaussian systems.

A. Definitions of previously proposed PIDs

Williams and Beer's proposed PID uses a definition of redundancy as the minimum information that either source provides about each outcome of the target, averaged over all possible outcomes [5]. This is obtained via a quantity called the specific information. The specific information of outcome $\boldsymbol{X} = \boldsymbol{x}$ given the random variable $\boldsymbol{Y}$ is the average reduction in surprise of outcome $\boldsymbol{X} = \boldsymbol{x}$ given $\boldsymbol{Y}$:

$$I(\boldsymbol{X} = \boldsymbol{x}; \boldsymbol{Y}) = \int d\boldsymbol{y}\, p(\boldsymbol{y}|\boldsymbol{x}) \left[\ln \frac{1}{p(\boldsymbol{x})} - \ln \frac{1}{p(\boldsymbol{x}|\boldsymbol{y})} \right]. \tag{24}$$

The mutual information $I(\boldsymbol{X};\boldsymbol{Y})$ is recovered from the specific information by integrating it over all values of $\boldsymbol{x}$. Redundancy is then the expected value over all $\boldsymbol{x}$ of the minimum specific information that $\boldsymbol{Y}$ and $\boldsymbol{Z}$ provide about the outcome $\boldsymbol{X} = \boldsymbol{x}$:

$$R(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z}) = \int d\boldsymbol{x}\, p(\boldsymbol{x}) \min_{\boldsymbol{\Xi} \in \{\boldsymbol{Y},\boldsymbol{Z}\}} I(\boldsymbol{X} = \boldsymbol{x}; \boldsymbol{\Xi}). \tag{25}$$

Griffith *et al.* [6,9] consider synergy to arise from information that is not necessarily present given the marginal distributions of source one and target $(\boldsymbol{X},\boldsymbol{Y})$ and source two and target $(\boldsymbol{X},\boldsymbol{Z})$. Thus

$$S(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z}) =: I(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z}) - \mathcal{U}(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z}) \tag{26}$$

where

$$\mathcal{U}(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z}) =: \min_{(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Z}})} I(\tilde{\boldsymbol{X}};\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Z}}) \tag{27}$$

and $\tilde{\boldsymbol{X}}$, $\tilde{\boldsymbol{Y}}$, and $\tilde{\boldsymbol{Z}}$ are subject to the constraints $P_{\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Y}}} = P_{\boldsymbol{X},\boldsymbol{Y}}$ and $P_{\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}}} = P_{\boldsymbol{X},\boldsymbol{Z}}$. The quantity $\mathcal{U}(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z})$ is referred to as the union information since it constitutes the whole information minus the synergy. Expressed alternatively, $\mathcal{U}(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z})$ is the minimum joint information provided about $\boldsymbol{X}$ by an alternative $\boldsymbol{Y}$ and $\boldsymbol{Z}$ with the same relations with $\boldsymbol{X}$ but different relations to each other. Bertschinger *et al.* [10] independently introduced identically the same PID, but starting from the equation

$$U(\boldsymbol{X};\boldsymbol{Y}|\boldsymbol{Z}) =: \min_{(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Z}})} I(\tilde{\boldsymbol{X}};\tilde{\boldsymbol{Y}}|\tilde{\boldsymbol{Z}}). \tag{28}$$

They then derive (27) via the conditional mutual information chain rule (11) and the basic PID formulas (1) and (3).

Harder, Salge, and Polani's PID [7] defines redundancy via the divergence of the conditional probability distribution $P_{\boldsymbol{X}|\boldsymbol{Z}=\boldsymbol{z}}$ for $\boldsymbol{X}$ given an outcome for $\boldsymbol{Z}$ from linear combinations of conditional probability distributions for $\boldsymbol{X}$ given an outcome for $\boldsymbol{Y}$. Thus, the following quantity is defined:

$$P_{\boldsymbol{X};\boldsymbol{Z}=\boldsymbol{z}\to\boldsymbol{Y}} = \arg\min_{\boldsymbol{y}_1,\boldsymbol{y}_2,\lambda\in[0,1]} D_{\mathrm{KL}}[P_{\boldsymbol{X}|\boldsymbol{Z}=\boldsymbol{z}} || \lambda P_{\boldsymbol{X}|\boldsymbol{Y}=\boldsymbol{y}_1} + (1-\lambda) P_{\boldsymbol{X}|\boldsymbol{Y}=\boldsymbol{y}_2}], \tag{29}$$

where D_{KL} is the Kullback-Leibler divergence, defined for continuous probability density functions P and Q by

$$D_{\mathrm{KL}}(P||Q) =: \int P(\boldsymbol{x}) \ln \left[\frac{P(\boldsymbol{x})}{Q(\boldsymbol{x})}\right] d^m\boldsymbol{x}. \tag{30}$$

Then the projected information $I_{\boldsymbol{X}}^{\pi}(\boldsymbol{Z}\to\boldsymbol{Y})$ is defined as

$$I_{\boldsymbol{X}}^{\pi}(\boldsymbol{Z}\to\boldsymbol{Y}) = I(\boldsymbol{X};\boldsymbol{Z}) - \int d\boldsymbol{z}\, p(\boldsymbol{z}) D_{\mathrm{KL}}[P_{\boldsymbol{X}|\boldsymbol{Z}=\boldsymbol{z}} || P_{\boldsymbol{X};\boldsymbol{Z}=\boldsymbol{z}\to\boldsymbol{Y}}] \tag{31}$$

and the redundancy is given by

$$R(\boldsymbol{X};\boldsymbol{Y},\boldsymbol{Z}) = \min\{I_{\boldsymbol{X}}^{\pi}(\boldsymbol{Z}\to\boldsymbol{Y}), I_{\boldsymbol{X}}^{\pi}(\boldsymbol{Y}\to\boldsymbol{Z})\}. \tag{32}$$

Thus, broadly, the closer the conditional distribution of $\boldsymbol{X}$ given $\boldsymbol{Y}$ is to the conditional distribution of $\boldsymbol{X}$ given $\boldsymbol{Z}$, the greater the redundancy.

B. Common PID for Gaussian systems

While the general definitions of the previously proposed PIDs are quite distinct, one can note that for all of them the redundant and unique information depend only on the pair of marginal distributions of each individual source with the target, i.e., those of $(\boldsymbol{X},\boldsymbol{Y})$ and $(\boldsymbol{X},\boldsymbol{Z})$. Here we derive our key result, namely, the following. Let X, $\boldsymbol{Y}$, and $\boldsymbol{Z}$ be jointly multivariate Gaussian, with X univariate and $\boldsymbol{Y}$ and $\boldsymbol{Z}$ of arbitrary dimensions n and p. Then there is a unique PID of $I(X;\boldsymbol{Y},\boldsymbol{Z})$ such that the redundant and unique information $R(X;\boldsymbol{Y},\boldsymbol{Z})$, $U(X;\boldsymbol{Y}|\boldsymbol{Z})$, and $U(X;\boldsymbol{Z}|\boldsymbol{Y})$ depend only on the marginal distributions of $(X,\boldsymbol{Y})$ and $(X,\boldsymbol{Z})$. The redundancy according to this PID is given by

$$R_{\mathrm{MMI}}(X;\boldsymbol{Y},\boldsymbol{Z}) =: \min\{I(X;\boldsymbol{Y}), I(X;\boldsymbol{Z})\}. \tag{33}$$

The other quantities follow from (1)–(3), assigning zero unique information to the source providing the least information about the target and synergy as the extra information contributed by the weaker source when the stronger source is known. We term this common PID the minimum mutual information PID. It follows that all of the previously proposed PIDs reduce down to the MMI PID for this Gaussian case.

Proof. We first show that the PID of Griffith *et al.* [6,9] (equivalent to that of Bertschinger *et al.* [10]) reduces to the MMI PID. Without loss of generality, we can rotate and normalize components of X, $\boldsymbol{Y}$, and $\boldsymbol{Z}$ such that the general case is specified by the block covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \boldsymbol{a}^{\mathrm{T}} & \boldsymbol{c}^{\mathrm{T}} \\ \boldsymbol{a} & I_n & B^{\mathrm{T}} \\ \boldsymbol{c} & B & I_p \end{pmatrix}, \tag{34}$$

where I_n and I_p are, respectively, the n- and p-dimensional identity matrices. We can also without loss of generality just consider the case $|\boldsymbol{a}| \leqslant |\boldsymbol{c}|$. From (5) we have

$$\Sigma(X|\boldsymbol{Y}) = 1 - \boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}, \tag{35}$$

$$\Sigma(X|\boldsymbol{Z}) = 1 - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{c}, \tag{36}$$

and hence $I(X;\boldsymbol{Y}) \leqslant I(X;\boldsymbol{Z})$. Note then that for a $(\tilde{X},\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Z}})$ subject to $P_{\tilde{X},\tilde{\boldsymbol{Y}}} = P_{X,\boldsymbol{Y}}$ and $P_{\tilde{X},\tilde{\boldsymbol{Z}}} = P_{X,\boldsymbol{Z}}$,

$$\begin{aligned} I(\tilde{X};\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Z}}) &\geqslant \max\{I(\tilde{X};\tilde{\boldsymbol{Y}}), I(\tilde{X};\tilde{\boldsymbol{Z}})\} \\ &= \max\{I(X;\boldsymbol{Y}), I(X;\boldsymbol{Z})\} = I(X;\boldsymbol{Z}). \end{aligned} \tag{37}$$

Now the covariance matrix of a $(\tilde{X},\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Z}})$ is given by

$$\tilde{\Sigma} = \begin{pmatrix} 1 & \boldsymbol{a}^{\mathrm{T}} & \boldsymbol{c}^{\mathrm{T}} \\ \boldsymbol{a} & I_n & \tilde{B}^{\mathrm{T}} \\ \boldsymbol{c} & \tilde{B} & I_p \end{pmatrix}, \tag{38}$$

where $\tilde{B}$ is a $p \times n$ matrix. The residual (partial) covariance of $\tilde{X}$ given $\tilde{\boldsymbol{Y}}$ and $\tilde{\boldsymbol{Z}}$ can thus be calculated using (5) as

$$\Sigma(\tilde{X}|\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Z}}) = 1 - (\boldsymbol{a}^{\mathrm{T}}\boldsymbol{c}^{\mathrm{T}}) \begin{pmatrix} I_n & \tilde{B}^{\mathrm{T}} \\ \tilde{B} & I_p \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{c} \end{pmatrix} \tag{39}$$

$$= 1 - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{c} + (\boldsymbol{c}^{\mathrm{T}}\tilde{B} - \boldsymbol{a}^{\mathrm{T}})(I_n - \tilde{B}^{\mathrm{T}}\tilde{B})^{-1}(\boldsymbol{a} - \tilde{B}^{\mathrm{T}}\boldsymbol{c}). \tag{40}$$

It follows from (40) and (36) that if we could find a $\tilde{B}$ that satisfied $\tilde{B}^{\mathrm{T}}\boldsymbol{c} = \boldsymbol{a}$ and for which the corresponding $\tilde{\Sigma}$ were a valid covariance matrix, then $\Sigma(\tilde{X}|\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Z}})$ would reduce to $\Sigma(X|\boldsymbol{Z})$ and hence we would have $I(\tilde{X};\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Z}}) = I(X;\boldsymbol{Z})$ and thus we would have

$$\mathcal{U}(X;\boldsymbol{Y},\boldsymbol{Z}) = \max\{I(X;\boldsymbol{Y}),I(X;\boldsymbol{Z})\} \tag{41}$$

by (37) and the definition (27) of $\mathcal{U}$.

We now demonstrate that there does indeed exist a $\tilde{B}$ satisfying $\tilde{B}^{\mathrm{T}}\boldsymbol{c} = \boldsymbol{a}$ and for which the corresponding $\tilde{\Sigma}$ is positive definite and hence a valid covariance matrix. First note that since $|\boldsymbol{a}| \leqslant |\boldsymbol{c}|$ there exists a $\tilde{B}$ satisfying $\tilde{B}^{\mathrm{T}}\boldsymbol{c} = \boldsymbol{a}$ for which $|\tilde{B}^{\mathrm{T}}\boldsymbol{v}| \leqslant |\boldsymbol{v}|$ for all $\boldsymbol{v} \in \mathbb{R}^p$. Suppose we have such a $\tilde{B}$. Then the matrix

$$\begin{pmatrix} I_p & \tilde{B} \\ \tilde{B}^{\mathrm{T}} & I_n \end{pmatrix} \tag{42}$$

is positive definite: For any $\boldsymbol{v} \in \mathbb{R}^p$, $\boldsymbol{w} \in \mathbb{R}^n$,

$$(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w}^{\mathrm{T}}) \begin{pmatrix} I_p & \tilde{B} \\ \tilde{B}^{\mathrm{T}} & I_n \end{pmatrix} \begin{pmatrix} \boldsymbol{v} \\ \boldsymbol{w} \end{pmatrix} = \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} + 2\boldsymbol{v}^{\mathrm{T}}\tilde{B}\boldsymbol{w} + \boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} \tag{43}$$

$$\geqslant \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} - 2\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w} + \boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} = (\boldsymbol{v} - \boldsymbol{w})^2 \geqslant 0. \tag{44}$$

Since it is also symmetric, it therefore has a Cholesky decomposition

$$\begin{pmatrix} I_p & \tilde{B} \\ \tilde{B}^{\mathrm{T}} & I_n \end{pmatrix} = \begin{pmatrix} I_p & 0 \\ \tilde{B}^{\mathrm{T}} & P \end{pmatrix} \begin{pmatrix} I_p & \tilde{B} \\ 0 & P^{\mathrm{T}} \end{pmatrix}, \tag{45}$$

where P is lower triangular. Hence, from equating blocks (2,2) on each side of this equation, we deduce that there exists a lower triangular matrix P satisfying

$$\tilde{B}^{\mathrm{T}}\tilde{B} + PP^{\mathrm{T}} = I_n. \tag{46}$$

We use this to demonstrate that the corresponding $\tilde{\Sigma}$ is positive definite by constructing the Cholesky decomposition for a rotated version of it. Rotating $(X,\boldsymbol{Y},\boldsymbol{Z}) \to (\boldsymbol{Z},X,\boldsymbol{Y})$ leads to the candidate covariance matrix $\tilde{\Sigma}$ becoming

$$\tilde{\Sigma}_{\mathrm{rot}} = \begin{pmatrix} I_p & \boldsymbol{c} & \tilde{B} \\ \boldsymbol{c}^{\mathrm{T}} & 1 & \boldsymbol{a}^{\mathrm{T}} \\ \tilde{B}^{\mathrm{T}} & \boldsymbol{a} & I_n \end{pmatrix}. \tag{47}$$

The Cholesky decomposition would then take the form

$$\tilde{\Sigma}_{\mathrm{rot}} = \begin{pmatrix} I_p & \boldsymbol{0} & 0 \\ \boldsymbol{c}^{\mathrm{T}} & q & \boldsymbol{0}^{\mathrm{T}} \\ \tilde{B}^{\mathrm{T}} & \boldsymbol{r}^{\mathrm{T}} & S \end{pmatrix} \begin{pmatrix} I_p & \boldsymbol{c} & \tilde{B} \\ \boldsymbol{0}^{\mathrm{T}} & q & \boldsymbol{r} \\ 0 & \boldsymbol{0} & S^{\mathrm{T}} \end{pmatrix}, \tag{48}$$

where S is a lower triangular matrix, q is a scalar, and $\boldsymbol{r}$ is a vector satisfying

$$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{c} + q^2 = 1, \tag{49}$$

$$\boldsymbol{c}^{\mathrm{T}}\tilde{B} + q\boldsymbol{r} = \boldsymbol{a}^{\mathrm{T}}, \tag{50}$$

$$\tilde{B}^{\mathrm{T}}\tilde{B} + \boldsymbol{r}^{\mathrm{T}}\boldsymbol{r} + SS^{\mathrm{T}} = I_n, \tag{51}$$

these equations coming, respectively, from equating blocks (2,2), (2,3), and (3,3) in (47) and (48) (the other block equations are satisfied trivially and do not constrain S, q, and $\boldsymbol{r}$). There exists a q to satisfy Eq. (49) since $1 - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{c} \geqslant 0$ by virtue of it being $\Sigma(X|\boldsymbol{Z})$ (36) and the original Σ being a valid covariance matrix. Equation (50) is satisfied by $\boldsymbol{r} = \boldsymbol{0}$ since $\tilde{B}^{\mathrm{T}}\boldsymbol{c} = \boldsymbol{a}$. Finally, Eq. (51) is then satisfied by $S = P$, where P is that of (46). It follows that the Cholesky decomposition exists, and hence $\tilde{\Sigma}$ is a valid covariance matrix, and thus (41) holds.

Now, given the definition (26) for the union information and our expression (41) for it we have

$$I(X;\boldsymbol{Y},\boldsymbol{Z}) - S(X;\boldsymbol{Y},\boldsymbol{Z}) = \max\{I(X;\boldsymbol{Y}),I(X;\boldsymbol{Z})\}. \tag{52}$$

Thus, by the expression (4) for synergy minus redundancy in terms of mutual information we have

$$R(X;\boldsymbol{Y},\boldsymbol{Z}) = S(X;\boldsymbol{Y},\boldsymbol{Z}) - I(X;\boldsymbol{Y},\boldsymbol{Z}) + I(X;\boldsymbol{Y}) + I(X;\boldsymbol{Z}) \tag{53}$$

$$= -\max\{I(X;\boldsymbol{Y}),I(X;\boldsymbol{Z})\} + I(X;\boldsymbol{Y}) + I(X;\boldsymbol{Z}) \tag{54}$$

$$= \min\{I(X;\boldsymbol{Y}),I(X;\boldsymbol{Z})\} \tag{55}$$

and hence we have reduced this PID to the MMI PID.

Now to show that this is the only PID for this Gaussian case satisfying the given conditions on the marginals of $(X,\boldsymbol{Y})$ and $(X,\boldsymbol{Z})$ we invoke Lemma 3 in Ref. [10]. In the notation of Bertschinger *et al.* [10], the specific PID that we have been considering is denoted by tildes, while possible alternatives are written without tildes. It follows from (55) that the source that shares the smaller amount of mutual information with the target has zero unique information. However, according to Lemma 3 this provides an upper bound on the unique information provided by that source on alternative PIDs. Thus alternative PIDs give the same zero unique information between this source and the target. However, according to Lemma 3, if the unique information is the same, then the whole PID is the same. Hence, there is no alternative PID. Q.E.D.

Note that this common PID does not extend to the case of a multivariate target. For a target with dimension greater than 1, the vectors $\boldsymbol{a}$ and $\boldsymbol{c}$ above are replaced with matrices A and C with more than one column [these being, respectively, $\Sigma(\boldsymbol{Y},\boldsymbol{X})$ and $\Sigma(\boldsymbol{Z},\boldsymbol{X})$]. Then, to satisfy (41) one would need to find a $\tilde{B}$ satisfying $\tilde{B}^{\mathrm{T}}C = A$, which does not in general exist. We leave consideration of this more general case to future work.

C. The MMI PID for the univariate jointly Gaussian case

It is straightforward to write down the MMI PID for the univariate jointly Gaussian case with covariance matrix given by (15). Taking without loss of generality $|a| \leqslant |c|$, we have from (17)–(19) and (33)

$$R_{\mathrm{MMI}}(X;Y,Z) = I(X;Y) = \frac{1}{2}\ln\left(\frac{1}{1-a^2}\right), \quad (56)$$

$$U_{\mathrm{MMI}}(X;Y) = 0, \quad (57)$$

$$U_{\mathrm{MMI}}(X;Z) = I(X;Z) - I(X;Y) = \frac{1}{2}\ln\left(\frac{1-a^2}{1-c^2}\right), \quad (58)$$

$$S_{\mathrm{MMI}}(X;Y,Z) = \frac{1}{2}\ln\left(\frac{(1-b^2)(1-c^2)}{1-(a^2+b^2+c^2)+2abc}\right). \quad (59)$$

It can then be shown that $S_{\mathrm{MMI}} \to \infty$ (and also $\Delta I \to \infty$) at the singular limits $b \to ac \pm \sqrt{(1-a^2)(1-c^2)}$ and also that, at $b = a/c$, S_{MMI} reaches the minimum value of 0. For all in between values there is positive synergy. It is intuitive that synergy should grow largest as one approaches the singular limit, because in that limit X is completely determined by Y and Z. On this PID, plots of synergy against correlation between sources take the same shape as plots of net synergy against correlation between sources, because of the independence of redundancy from correlation between sources [Figs. 3(c) and 3(f)]. Thus, for equal (same sign) a and c, S_{MMI} decreases with correlation between sources, for equal magnitude but opposite sign a and c, S_{MMI} increases with correlation between sources and for unequal magnitude a and c, S_{MMI} has a U-shaped dependence on correlation between sources.

V. DYNAMICAL SYSTEMS

In this section we explore synergy and redundancy in some example dynamical Gaussian systems, specifically multivariate autoregressive (MVAR) processes, i.e., discrete-time systems in which the present state is given by a linear combination of past states plus noise.[2] Having demonstrated (Sec. IV) that the MMI PID is valid for multivariate sources, we are able to derive valid expressions for redundancy and synergy in the information that arbitrary length histories of sources contain about the present state of a target. We also compute the more straightforward net synergy.

A. Example 1: Synergistic two-variable system

The first example we consider is a two-variable MVAR process consisting of two variables X and Y, with X receiving equal inputs from its own past and from the past of Y [see Fig. 4(a)]. The dynamics are given by the following

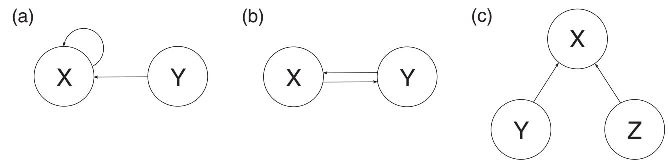

FIG. 4. Connectivity diagrams for example dynamical systems. Variables are shown as circles and directed interactions as arrows. The systems are animated as Gaussian MVAR processes of order 1. (a) Example 1. In this system X receives input from its own past and from the past of Y. There is positive net synergy between the information that the immediate pasts of X and Y provide about the future of X, but zero net synergy between the information provided by the infinite pasts of X and Y about the future of X. (b) Example 2. In this system there is bidirectional connectivity between X and Y. There is zero net synergy between the information provided by the immediate pasts of X and Y about the future of X and negative net synergy (i.e., positive net redundancy) between the information provided by the infinite pasts of X and Y about the future of X. (c) Example 3. Here Y and Z are sources that influence the future of X. Depending on the correlation between Y and Z, there can be synergy between the information provided by the pasts of Y and Z about the future of X (independent of the length of history considered).

equations:

$$X_t = \alpha X_{t-1} + \alpha Y_{t-1} + \epsilon_t^X, \quad (60)$$

$$Y_t = \epsilon_t^Y, \quad (61)$$

where the ϵ are all independent and identically distributed Gaussian variables of mean 0 and variance 1. The variables X and Y have a stationary probability distribution as long as $|\alpha| < 1$. The information between the immediate pasts of X and Y and the present of X can be computed analytically as follows. First, the stationary covariance matrix $\Sigma(X_t \oplus Y_t)$ satisfies

$$\Sigma(X_t \oplus Y_t) = A\Sigma(X_t \oplus Y_t)A^{\mathrm{T}} + I_2, \quad (62)$$

where I_2 is the two-dimensional identity matrix and A is the connectivity matrix

$$A = \begin{pmatrix} \alpha & \alpha \\ 0 & 0 \end{pmatrix}. \quad (63)$$

This is obtained by taking the covariance matrix of both sides of (60) and (61). Hence

$$\Sigma(X_t \oplus Y_t) = \frac{1}{1-\alpha^2}\begin{pmatrix} 1+\alpha^2 & 0 \\ 0 & 1-\alpha^2 \end{pmatrix}. \quad (64)$$

The one-lag covariance matrix $\Gamma_1(X_t \oplus Y_t) =: \Sigma(X_t \oplus Y_t, X_{t-1} \oplus Y_{t-1})$ is given by

$$\Gamma_1(X_t \oplus Y_t) = A\Sigma(X_t \oplus Y_t) = \frac{\alpha}{1-\alpha^2}\begin{pmatrix} 1+\alpha^2 & 1-\alpha^2 \\ 0 & 0 \end{pmatrix}. \quad (65)$$

From these quantities we can obtain the following variances:

$$\Sigma(X_t) = \frac{1+\alpha^2}{1-\alpha^2}, \quad (66)$$

$$\Sigma(X_t|X_{t-1}) = 1+\alpha^2, \quad (67)$$

[2] These are the standard stationary dynamical Gaussian systems. In fact, they are the only stationary dynamical Gaussian systems if one assumes that the present state is a continuous function of the past state [14].

$$\Sigma(X_t|Y_{t-1}) = \frac{1+\alpha^4}{1-\alpha^2}, \tag{68}$$

$$\Sigma(X_t|X_{t-1},Y_{t-1}) = 1. \tag{69}$$

Then from these we can compute the mutual information between the present of X and the immediate pasts of X and Y:

$$I(X_t;X_{t-1}) = \frac{1}{2}\ln\left(\frac{1}{1-\alpha^2}\right), \tag{70}$$

$$I(X_t;Y_{t-1}) = \frac{1}{2}\ln\left(\frac{1+\alpha^2}{1+\alpha^4}\right), \tag{71}$$

$$I(X_t;X_{t-1},Y_{t-1}) = \frac{1}{2}\ln\left(\frac{1+\alpha^2}{1-\alpha^2}\right). \tag{72}$$

Thus, from these we see that there is net synergy between the immediate pasts of X and Y in information about the present of X:

$$\Delta I(X_t;X_{t-1},Y_{t-1}) = \tfrac{1}{2}\ln(1+\alpha^4) > 0. \tag{73}$$

The infinite pasts of X and Y do not, however, exhibit net synergistic information about the present of X. While $\Sigma(X_t|\boldsymbol{X}_t^-) = \Sigma(X_t|X_{t-1})$ and $\Sigma(X_t|\boldsymbol{X}_t^-,\boldsymbol{Y}_t^-) = \Sigma(X_t|X_{t-1},Y_{t-1})$, we have $\Sigma(X_t|\boldsymbol{Y}_t^-) \neq \Sigma(X_t|Y_{t-1})$. This is because the restricted regression of X on the past of Y is of infinite order:

$$X_t = \sum_{n=1}^{\infty}\alpha^n Y_{t-n} + \sum_{n=0}^{\infty}\alpha^n \epsilon_{t-n}^X. \tag{74}$$

Hence,

$$\Sigma(X_t|\boldsymbol{Y}_t^-) = \mathrm{Var}\big(\alpha^n\epsilon_{t-n}^X\big) = \sum_{n=0}^{\infty}\alpha^{2n} = \frac{1}{1-\alpha^2}. \tag{75}$$

Therefore,

$$I(X_t;\boldsymbol{X}_t^-) = \frac{1}{2}\ln\left(\frac{1}{1-\alpha^2}\right), \tag{76}$$

$$I(X_t;\boldsymbol{Y}_t^-) = \frac{1}{2}\ln(1+\alpha^2), \tag{77}$$

$$I(X_t;\boldsymbol{X}_t^-,\boldsymbol{Y}_t^-) = \frac{1}{2}\ln\left(\frac{1+\alpha^2}{1-\alpha^2}\right), \tag{78}$$

and

$$\Delta I(X_t;\boldsymbol{X}_t^-,\boldsymbol{Y}_t^-) = 0. \tag{79}$$

Thus the synergy equals the redundancy between the infinite pasts of X and Y in providing information about the present state of X.

According to the MMI PID, at infinite lags synergy is the same compared to that for one lag, but redundancy is less. We have the following expressions for redundancy and synergy:

$$R_{\mathrm{MMI}}(X_t;X_{t-1},Y_{t-1}) = \frac{1}{2}\ln\left(\frac{1+\alpha^2}{1+\alpha^4}\right), \tag{80}$$

$$S_{\mathrm{MMI}}(X_t;X_{t-1},Y_{t-1}) = \frac{1}{2}\ln(1+\alpha^2), \tag{81}$$

$$R_{\mathrm{MMI}}(X_t;\boldsymbol{X}_t^-,\boldsymbol{Y}_t^-) = S_{\mathrm{MMI}}(X_t;\boldsymbol{X}_t^-,\boldsymbol{Y}_t^-) = \frac{1}{2}\ln(1+\alpha^2). \tag{82}$$

B. Example 2: An MVAR model with no net synergy

Not all MVAR models exhibit positive net synergy, for example [see Fig. 4(b)],

$$X_t = \alpha Y_{t-1} + \epsilon_t^X, \tag{83}$$

$$Y_t = \beta X_{t-1} + \epsilon_t^Y, \tag{84}$$

where again the ϵ are all independent identically distributed random variables of mean 0 and variance 1 and $|\alpha|,|\beta| < 1$ for stationarity. A calculation similar to that for example 1 shows that the one-lag mutual information satisfies

$$I(X_t;X_{t-1}) = 0, \tag{85}$$

$$I(X_t;Y_{t-1}) = \frac{1}{2}\ln\left(\frac{1+\alpha^2}{1-\alpha^2\beta^2}\right), \tag{86}$$

$$I(X_t;X_{t-1},Y_{t-1}) = \frac{1}{2}\ln\left(\frac{1+\alpha^2}{1-\alpha^2\beta^2}\right) \tag{87}$$

and thus synergy and redundancy are the same for one-lag mutual information

$$\Delta I(X_t;X_{t-1},Y_{t-1}) = 0. \tag{88}$$

For infinite lags one has

$$I(X_t;\boldsymbol{X}_t^-) = \frac{1}{2}\ln\left(\frac{1}{1-\alpha^2\beta^2}\right), \tag{89}$$

$$I(X_t;\boldsymbol{Y}_t^-) = \frac{1}{2}\ln\left(\frac{1+\alpha^2}{1-\alpha^2\beta^2}\right), \tag{90}$$

$$I(X_t;\boldsymbol{X}_t^-,\boldsymbol{Y}_t^-) = \frac{1}{2}\ln\left(\frac{1+\alpha^2}{1-\alpha^2\beta^2}\right), \tag{91}$$

and thus

$$\Delta I(X_t;\boldsymbol{X}_t^-,\boldsymbol{Y}_t^-) = -\frac{1}{2}\ln\left(\frac{1}{1-\alpha^2\beta^2}\right) < 0, \tag{92}$$

so there is greater redundancy than synergy.

For the MMI decomposition we have for one lag

$$R_{\mathrm{MMI}}(X_t;X_{t-1},Y_{t-1}) = S_{\mathrm{MMI}}(X_t;X_{t-1},Y_{t-1}) = 0, \tag{93}$$

while for infinite lags

$$R_{\mathrm{MMI}}(X_t;X_t^-,Y_t^-) = \frac{1}{2}\ln\left(\frac{1}{1-\alpha^2\beta^2}\right), \tag{94}$$

$$S_{\mathrm{MMI}}(X_t;X_t^-,Y_t^-) = 0. \tag{95}$$

It is intuitive that for this example there should be zero synergy. All the information contributed by the past of X to the present of X is mediated via the interaction with Y, so no extra information about the present of X is gained from knowing the past of X given knowledge of the past of Y.

It is interesting to note that for both this example and example 1 above,

$$\Delta I(X_t;X_{t-1},Y_{t-1}) > \Delta I(X_t;\boldsymbol{X}_t^-,\boldsymbol{Y}_t^-). \tag{96}$$

That is, there is less synergy relative to redundancy when one considers information from the infinite past compared with information from the immediate past of the system. This can be understood as follows. The complete MVAR model is order 1 in each example (that is, the current state of the system depends only on the immediate past), so $I(X_t;\boldsymbol{X}_t^-,\boldsymbol{Y}_t^-) = I(X_t;X_{t-1},Y_{t-1})$, but restricted effective regressive models of X on just the past of X or just the past of Y are generally of infinite order (that is, one can often obtain lower residual noise in X when regressing on the entire infinite past of just X or just Y compared to when regressing on just the immediate past of just X or just Y). Hence $I(X_t;\boldsymbol{X}_t^-) \geqslant I(X_t;X_{t-1})$ and $I(X_t;\boldsymbol{Y}_t^-) \geqslant I(X_t;Y_{t-1})$ for such two-variable order-1 MVAR systems. For the two examples, both of these inequalities are strict and hence the relation (96) follows.

An interesting question is whether there exists an MVAR model for two variables X_t and Y_t for which the infinite-lag net synergy is greater than zero. It is straightforward to demonstrate that no such system can be found by simple perturbations of the systems considered here. However, a full consideration of the most general MVAR model of order greater than 1 is beyond the scope of the present paper. In any case, in the next example, we see that for an MVAR system with three variables, the infinite past of two variables can provide net synergistic information about the future of the third variable.

C. Example 3: Synergy between two variables influencing a third variable

The third example we consider is an MVAR process with Y and Z being (possibly) correlated sources that are each influencing X [see Fig. 4(c)]:

$$X_t = \frac{1}{\Delta}\left(\alpha Y_{t-1} + \gamma Z_{t-1} + \epsilon_t^X\right), \tag{97}$$

$$Y_t = \epsilon_t^Y, \tag{98}$$

$$Z_t = \epsilon_t^Z, \tag{99}$$

where $\Delta = \sqrt{1+\alpha^2+2\alpha\gamma\rho+\gamma^2}$ and the ϵ are Gaussian noise sources all of zero mean, with zero correlation in time, but with instantaneous correlation matrix

$$\Sigma(\boldsymbol{\epsilon}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \rho \\ 0 & \rho & 1 \end{pmatrix}. \tag{100}$$

Here there is no restriction on connection strengths α or γ; stationarity is satisfied for all values. Following the same method as in examples 1 and 2, we have

$$\Sigma(X_t \oplus Y_t \oplus Z_t) = A\Sigma(X_t \oplus Y_t \oplus Z_t)A^{\mathrm{T}} + \Sigma(\boldsymbol{\epsilon}) \tag{101}$$

and

$$A = \frac{1}{\Delta}\begin{pmatrix} 0 & \alpha & \gamma \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{102}$$

$$\Sigma(X_t \oplus Y_t \oplus Z_t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \rho \\ 0 & \rho & 1 \end{pmatrix}, \tag{103}$$

$$\Gamma_1(X_t \oplus Y_t \oplus Z_t) = \frac{1}{\Delta}\begin{pmatrix} 0 & \alpha+\rho\gamma & \gamma+\rho\alpha \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{104}$$

From these quantities we can compute the mutual information

$$I(X_t;Y_{t-1}) = \frac{1}{2}\ln\left(\frac{1+\alpha^2+2\alpha\gamma\rho+\gamma^2}{1+\gamma^2(1-\rho^2)}\right), \tag{105}$$

$$I(X_t;Z_{t-1}) = \frac{1}{2}\ln\left(\frac{1+\alpha^2+2\alpha\gamma\rho+\gamma^2}{1+\alpha^2(1-\rho^2)}\right), \tag{106}$$

$$I(X_t;Y_{t-1},Z_{t-1}) = \frac{1}{2}\ln(1+\alpha^2+2\alpha\gamma\rho+\gamma^2). \tag{107}$$

Hence, assuming without loss of generality that $|\alpha| \leqslant |\gamma|$,

$$\Delta I(X_t;Y_{t-1},Z_{t-1}) = \frac{1}{2}\ln\left(\frac{[1+\alpha^2(1-\rho^2)][1+\gamma^2(1-\rho^2)]}{1+\alpha^2+2\alpha\gamma\rho+\gamma^2}\right), \tag{108}$$

$$R_{\mathrm{MMI}}(X_t;Y_{t-1},Z_{t-1}) = \frac{1}{2}\ln\left(\frac{1+\alpha^2+2\alpha\gamma\rho+\gamma^2}{1+\gamma^2(1-\rho^2)}\right), \tag{109}$$

$$S_{\mathrm{MMI}}(X_t;Y_{t-1},Z_{t-1}) = \frac{1}{2}\ln(1+\alpha^2[1-\rho^2]). \tag{110}$$

Note that we do not consider the PID for the information provided by the infinite pasts of Y and Z because it is the same as that provided by the immediate pasts for this example.

For the case of no correlation between Y and Z, i.e., $\rho = 0$, we have

$$\Delta I(X_t;Y_{t-1},Z_{t-1}) = \frac{1}{2}\ln\left(\frac{[1+\alpha^2][1+\gamma^2]}{1+\alpha^2+\gamma^2}\right) > 0, \tag{111}$$

i.e., there is net synergy. For the case $\rho = 1$ of Y and Z being perfectly correlated, there is, however, net redundancy, since

$$\Delta I(X_t;Y_{t-1},Z_{t-1}) = \frac{1}{2}\ln\left(\frac{1}{1+(\alpha+\gamma)^2}\right) < 0. \tag{112}$$

This is a dynamical example in which two uncorrelated sources can contribute net synergistic information to a target. The MMI PID synergy S_{MMI} behaves in an intuitive way here, increasing with the square of the weaker connection α, decreasing as the correlation ρ between the sources Y and Z increases, and going to zero when $\alpha = 0$ or $\rho = 1$, reflecting the strength and independence of the weaker link.

Considering this system further for the case $\rho = 0$ and $\alpha = \gamma$, for small α the net synergy is approximately $\alpha^4/2$ and for large α the net synergy is approximately $\ln(\alpha/\sqrt{2})$ (as stated above, α can be arbitrarily large in this model, since the spectral radius, i.e., largest absolute value of the eigenvalues of the connectivity matrix, is zero independent of α). Hence net synergy can be arbitrarily large. The proportion $\Delta I(X_t;Y_{t-1},Z_{t-1})/I(X_t;Y_{t-1},Z_{t-1})$ also grows with connection strength α, reaching, for example, approximately 0.1 for $\alpha = 0.5$.

VI. TRANSFER ENTROPY

The net synergy in the example systems of Sec. V affect transfer entropy and its interpretation. Pairwise (one-lag) transfer entropy is defined as

$$\mathcal{T}^{(1)}_{Y\to X} =: H(X_t|X_{t-1}) - H(X_t|X_{t-1},Y_{t-1}) \equiv I(X_t;X_{t-1},Y_{t-1}) - I(X_t;X_{t-1}). \quad (113)$$

Typically transfer entropy is interpreted straightforwardly as the information that the past of Y contributes to the present of X over and above that already provided by the past of X [15]. It has sometimes been implicitly assumed to be less than the lagged mutual information $I(X_t;Y_{t-1})$ for simple linear systems, for example, in constructing measures of the overall causal interactivity of a system [14]. However, this is not the case when there is net synergy, since transfer entropy measures the unique information provided by the past of Y plus the synergistic information between the pasts of X and Y,

$$\mathcal{T}^{(1)}_{Y\to X} = U(X_t;Y_{t-1}|X_{t-1}) + S(X_t;X_{t-1},Y_{t-1}), \quad (114)$$

whereas the lagged mutual information $I(X_t;Y_{t-1})$ measures the unique information provided by the past of Y plus the redundant information provided by the pasts of X and Y:

$$I(X_t;Y_{t-1}) = U(X_t;Y_{t-1}|X_{t-1}) + R(X_t;X_{t-1},Y_{t-1}). \quad (115)$$

Specifically, for example 1,

$$\mathcal{T}^{(1)}_{Y\to X} = \frac{1}{2}\ln(1+\alpha^2), \quad (116)$$

$$\mathcal{T}^{(1)}_{Y\to X} - I(X_t;Y_{t-1}) = \frac{1}{2}\ln(1+\alpha^4) > 0. \quad (117)$$

The situation can be different when infinite lags are considered:

$$\mathcal{T}^{(\infty)}_{Y\to X} =: H(X_t|\boldsymbol{X}_t^-) - H(X_t|\boldsymbol{X}_t^-,\boldsymbol{Y}_t^-). \quad (118)$$

For example 1, considering infinite lags, the transfer entropy $\mathcal{T}^{(\infty)}_{Y\to X}$ and the lagged mutual information $I(X_t;\boldsymbol{Y}_t^-)$ are equal because the net synergy between complete past histories of X and Y is zero. From (70)–(72) and (76)–(78) we have

$$\mathcal{T}^{(\infty)}_{Y\to X} = \mathcal{T}^{(1)}_{Y\to X}, \quad (119)$$

$$\mathcal{T}^{(\infty)}_{Y\to X} - I(X_t;\boldsymbol{Y}_t^-) = 0. \quad (120)$$

Conditional transfer entropy $\mathcal{T}^{(\infty)}_{Y\to X|Z}$ (infinite lags) is defined as

$$\mathcal{T}^{(\infty)}_{Y\to X|Z} =: H(X_t|\boldsymbol{X}_t^-,\boldsymbol{Z}_t^-) - H(X_t|\boldsymbol{X}_t^-,\boldsymbol{Y}_t^-,\boldsymbol{Z}_t^-). \quad (121)$$

It has sometimes been assumed that the conditional transfer entropy is less than nonconditional transfer entropy, i.e., $\mathcal{T}^{(\infty)}_{Y\to X|Z}$ is less than $\mathcal{T}^{(\infty)}_{Y\to X}$ [14,16]. This is because the pasts of Y and Z might contribute redundant information to the future of X, but as for pairwise nonconditional transfer entropy, synergy is usually not considered important for continuous, linear unimodal systems such as those considered in this paper. However, for example 3 this is not always true. Considering the net synergistic case of $\rho = 0, \alpha = \gamma$,

$$\mathcal{T}_{Y\to X} = \frac{1}{2}\ln\left(\frac{1+2\alpha^2}{1+\alpha^2}\right), \quad (122)$$

$$\mathcal{T}_{Y\to X|Z} = \frac{1}{2}\ln(1+\alpha^2), \quad (123)$$

$$\mathcal{T}_{Y\to X|Z} - \mathcal{T}_{Y\to X} = \frac{1}{2}\ln\left(1+\frac{\alpha^4}{1+2\alpha^2}\right) > 0. \quad (124)$$

Here the number of lags is left unspecified because these quantities are the same for any number of lags. Thus conditional transfer entropy can be affected by synergy even when infinite lags are considered. In this example, because X has no self-connection and thus the past of X contributes no information to the future of X, $\mathcal{T}_{Y\to X}$ reduces to $U(X_t;Y_{t-1}|Z_{t-1}) + R(X_t;Y_{t-1},Z_{t-1})$ and $\mathcal{T}_{Y\to X|Z}$ to $U(X_t;Y_{t-1}|Z_{t-1}) + S(X_t;Y_{t-1},Z_{t-1})$. Nonconditional minus conditional transfer entropy has been applied to assess the balance between synergy and redundancy (i.e., net synergy) among neuroelectrophysiological variables in [17].

Since transfer entropy is equivalent to the linear formulation of Granger causality for jointly Gaussian variables [13], the above conclusions pertain also to interpretations of Granger causality. Granger causality quantifies the extent to which the past of one variable Y predicts the future of another variable X over and above the extent to which the past of X (and the past of any conditional variables) predicts the future of X [18,19]. In the usual linear formulation, the prediction is implemented using the framework of linear autoregression. Thus, to measure the Granger causality from predictor Y to predictee X given conditional variables $\boldsymbol{Z}$, one compares the following MVAR models:

$$X_t = A\cdot\left[\boldsymbol{X}_t^{(p)} \oplus \boldsymbol{Z}_t^{(r)}\right] + \boldsymbol{\epsilon}_t, \quad (125)$$

$$X_t = A'\cdot\left[\boldsymbol{X}_t^{(p)} \oplus \boldsymbol{Y}_t^{(q)} \oplus \boldsymbol{Z}_t^{(r)}\right] + \boldsymbol{\epsilon}'_t. \quad (126)$$

Thus the predictee variable X is regressed first on the previous p lags of itself plus r lags of the conditioning variables $\boldsymbol{Z}$ and second, in addition, on q lags of the predictor variable Y (p, q, and r can be selected according to the Akaike or Bayesian information criterion [20]). The magnitude of the Granger causality interaction is then given by the logarithm of the ratio of the residual variances

$$\mathcal{F}_{Y\to X|\boldsymbol{Z}} =: \ln\left(\frac{\Sigma(\boldsymbol{\epsilon}_t)}{\Sigma(\boldsymbol{\epsilon}'_t)}\right) = \ln\left(\frac{\Sigma(X_t|\boldsymbol{X}_t^-,\boldsymbol{Z}_t^-)}{\Sigma(X_t|\boldsymbol{X}_t^-,\boldsymbol{Y}_t^-,\boldsymbol{Z}_t^-)}\right), \quad (127)$$

where the final term expresses Granger causality in terms of partial covariances and hence illustrates the equivalence with transfer entropy for Gaussian variables (up to a factor of 2) [13]. It follows that pairwise Granger causality $\mathcal{F}_{Y\to X}$ (no conditional variables) should be considered as a measure of the unique (with respect to the past of X) predictive power that the past of Y has for the future of X plus the synergistic predictive power that the pasts of X and Y have in tandem for the future of X. Meanwhile, conditional Granger causality $\mathcal{F}_{Y\to X|\boldsymbol{Z}}$ should be considered as a measure of the unique (with respect to the pasts of X and $\boldsymbol{Z}$) predictive power that the past of Y has for the future of X plus the synergistic predictive power that the pasts of X and $Y \oplus \boldsymbol{Z}$ have in tandem for the future of X.

VII. IMPLICATIONS FOR MEASURES OF OVERALL INTERACTIVITY AND COMPLEXITY

The prevalence of synergistic contributions to information sharing has implications for how to sensibly construct measures of overall information transfer sustained in a complex system, or the overall *complexity* of the information transfer. One such measure is causal density [14,16,21]. Given a set of Granger causality values among elements of a system $\boldsymbol{M}$, a simple version of causal density can be defined as the average of all pairwise Granger causalities between elements (conditioning on all remaining elements)

$$\mathrm{cd}(\boldsymbol{M}) =: \frac{1}{n(n-1)} \sum_{i \neq j} \mathcal{F}_{M_j \to M_i | \boldsymbol{M}_{[ij]}}, \tag{128}$$

where $\boldsymbol{M}_{[ij]}$ denotes the subsystem of $\boldsymbol{M}$ with variables M_i and M_j omitted and n is the total number of variables. Causal density provides a principled measure of dynamical complexity inasmuch as elements that are completely independent will score zero, as will elements that are completely integrated in their dynamics. High values will only be achieved when elements behave somewhat differently from each other, in order to contribute novel potential predictive information, and at the same time are globally integrated, so that the potential predictive information is in fact useful [16,22]. In the context of the current discussion, however, causal density counts synergistic information multiple times, while neglecting redundant information. For instance, in example 3 above, the nonzero contributions to causal density are

$$\mathrm{cd} = \tfrac{1}{6}[\mathcal{F}_{Z \to X|Y} + \mathcal{F}_{Y \to X|Z}] \tag{129}$$

$$= \tfrac{1}{3}[U(X_t; Y_{t-1}|Z_{t-1}) + U(X_t; Z_{t-1}|Y_{t-1}) + 2S(X_t; Y_{t-1}, Z_{t-1})]. \tag{130}$$

In spite of this apparent overcounting of synergistic information, the resultant formula is

$$\mathrm{cd} = \tfrac{1}{6}\{\ln[1 + \alpha^2(1-\rho^2)] + \ln[1 + \gamma^2(1-\rho^2)]\}, \tag{131}$$

which is after all a sensible formula for the overall level of transfer of novel predictive information, increasing with connection strengths α and γ, decreasing with the correlation ρ between the source variables, and going to zero if either both α and γ are zero or if $\rho \to 1$.

An alternative to causal density is the global transfer entropy [23,24] $\mathcal{T}_{\mathrm{gl}}$, defined as

$$\mathcal{T}_{\mathrm{gl}}(\boldsymbol{M}) =: \frac{1}{n} \sum_i \mathcal{T}_{\boldsymbol{M} \to M_i}, \tag{132}$$

i.e., the average information flow from the entire system to individual elements. This may be considered a measure of gross past-conditional statistical dependence of the elements of the system, insofar as it vanishes if and only if each system element, conditional on its own past, does not depend on the past of other system elements. Unlike causal density, this measure assigns equal weight to contributions from unique, redundant, and synergistic information flow. However, it is not sensitive to whether the information flow occurs homogeneously or inhomogeneously; it is not concerned with the distribution among sources of the information that flows into the targets. It should thus be interpreted as operationalizing a different conceptualization of complexity to causal density. For example 3 above, the only nonzero contribution to this global transfer entropy arises from $I(X_t; Y_{t-1}, Z_{t-1})$. Thus, from Eq. (107), it is given by

$$\mathcal{T}_{\mathrm{gl}} = \tfrac{1}{6} \ln(1 + \alpha^2 + 2\alpha\gamma\rho + \gamma^2). \tag{133}$$

This quantity is actually increasing with correlation ρ between sources, reflecting explicitly here that this is not a measure of complexity that operationalizes inhomogeneity of information sources. That the information flow into the target is greatest when sources are strongly positively correlated is explained as follows: Fluctuations of the sources cause fluctuations of the target and fluctuations coming from positively correlated sources will more often combine to cause greater fluctuations of the target than of sources, whereas fluctuations coming from uncorrelated sources will more often cancel out at the target. Thus the relative variance of the target before compared with after knowing the pasts of the sources is greatest when sources are strongly positively correlated.

Conceptualizing complexity as having to do with a whole system being greater than the sum of its parts, average synergistic information contributed by the past of a pair of variables to the present of a third variable could form a measure of complexity, by measuring the extent to which joint information contributed by two sources exceeds the sum of information contributed by individual sources. Thus we could define the synergistic complexity $\mathcal{C}_S$ as

$$\mathcal{C}_S(\boldsymbol{M}) =: \frac{2}{n(n-1)(n-2)} \sum_{i,j,k} S(M_{i,t}; \boldsymbol{M}^-_{j,t}, \boldsymbol{M}^-_{k,t}). \tag{134}$$

For example 3, this leads via Eq. (110) to

$$\mathcal{C}_S(\boldsymbol{M}) =: \tfrac{1}{6} \ln(1 + \alpha^2[1-\rho^2]) \tag{135}$$

for the case $|\alpha| \leqslant |\gamma|$, reflecting the strength and level of independence of the weakest connection. This is in the spirit of what the Φ measures of integrated information [21,25,26] are supposed to capture (in some cases of high synergy Φ measures are unsuccessful at doing this [27]). One could also conceive an analogous measure based on net synergy, but this does not lead to a formula that summarizes the complexity of example 3 in any straightforward conceptualization [see Eq. (108) for the nonzero term].

To fully understand the pros and cons of these various measures of complexity, they should be considered on systems composed of many (i.e., much greater than three) elements. While there have been studies of causal density [21] and global transfer entropy [24], the synergistic complexity is a different measure, which should be explored in controlled comparison with the other measures. One could further imagine, for general systems of n variables, a complexity measure based on the synergistic information contributed to one variable from the pasts of all $n-1$ other variables. We do not attempt to consider such a measure here, since consideration of PIDs for more than two source variables is beyond the scope of this paper. This will also be an avenue for future research.

VIII. DISCUSSION

A. Summary

In this paper we have carried out analyses of partial information decompositions for Gaussian variables. That is, we have explored how the information that two source variables carry about a target variable decomposes into unique, redundant, and synergistic information. Previous studies of PIDs have focused on systems of discrete variables and this study focuses on continuous random variables. We have demonstrated that net synergy (i.e., the combined information being greater than the sum of the individual informations) is prevalent in systems of Gaussian variables with linear interactions and hence that PIDs are nontrivial for these systems. We illustrated two interesting examples of a jointly Gaussian system exhibiting net synergy: (i) a case in which the target is correlated with both sources, but the two sources are uncorrelated [Fig. 2(a)], and (ii) a case in which the target is only correlated with one of two sources, but the two sources are correlated [Fig. 2(b)]. Further we have shown that, depending on the signs of the correlations between sources and target, net synergy can either increase or decrease with (absolute) correlation strength between sources (Fig. 3). Thus, redundancy should not be considered a reflection of correlation between sources.

Our key result is that for a broad class of Gaussian systems, a broad class of PIDs leads to (i) a definition of redundancy as the minimum of the mutual informations between the target and each individual source, and hence they take redundancy as totally independent of the correlation between sources, and (ii) synergy being the extra information contributed by the weaker source when the stronger source is known. Specifically, this holds for a jointly Gaussian system with a univariate target and sources of arbitrary dimension, and any PID for which the redundant and unique information depend only on the pair of marginal distributions of target and source 1 and target and source 2. This property has been argued for in [10] and covers three previously proposed PIDs [5–7,10], which all operationalize distinct conceptualizations of redundancy (see Sec. IV A). Thus it would be reasonable to apply this formula for redundancy to any data that are approximately Gaussian. Note however, there is still debate about the list of axioms a PID should satisfy beyond the minimal ones described in the Introduction [9], so it is still possible that an alternative PID is constructed for which the formula does not hold. We have termed the obtained decomposition the MMI PID. Most usefully, it is applicable in a multivariate-time-series analysis to the computation of synergistic and redundant information arising in an arbitrary length past history of two variables about the present state of a third variable, i.e., to analyses of information transfer.

That there can be net synergy when sources are uncorrelated implies that simple dynamical Gaussian systems can exhibit net synergy when considering the past of two variables as the sources and the present of one variable as the target. Indeed, we have demonstrated this explicitly via some simple examples. We analyzed an MVAR model on which the pasts of two sources influence the present of a target [Fig. 4(c)] and showed that the synergistic information of the past of the sources about the target, as obtained via the MMI PID, increases monotonically with the weaker connection strength and decreases monotonically with correlation between sources (110). Thus, while redundancy does not provide us with distinct knowledge of the system, above and beyond mutual information between individual sources and target, synergy provides an intuitive formula for the extent of simultaneous differentiation (between sources) and integration (of information from both sources).

B. Application to neuroscience

Information-theoretic analyses are increasingly popular in neuroscience, notably for analyzing the neural encoding of stimuli or for analyzing brain connectivity via quantification of information transfer between pairs of brain variables (especially if one considers Granger causality [18,19] as a measure of information transfer based on its correspondence with transfer entropy [13,28,29]); see [30] for a recent review. There have been several studies in which net synergy or redundancy has been computed empirically on neurophysiological data sets, e.g., [17,31–36]. In neural coding, net synergy ($\Delta I > 0$) has been observed in the information successive action potentials carry about a stimulus [31]. In most studies, information transfer between electroencephalogram (EEG) variables has tended to exhibit net redundancy (i.e., $\Delta I < 0$), although recently net synergy ($\Delta I > 0$) has been observed in information transfer among some intracranial EEG variables in an epileptic patient [37]. A pair of recent studies has associated certain pathological brain states with increased net redundancy in information transfer: among electrocorticographic time series (contacts placed intracranially on the surface of the cortex) during seizure onset in an epileptic patient [17] and among scalp EEG time series from traumatic brain injury patients in the vegetative state, compared to analogous recordings from healthy controls [36].

Usually net redundancy has been assumed to arise due to common sources and hence correlation between variables. However, as mentioned above, the results here suggest that this is not always the case. For the Gaussian case we have considered, this holds for positive correlation between sources and an equal correlation between the target and each of the sources, but not more generally (see Fig. 3).

The canonical example scenario for net synergy takes one of the sources to be a suppressor variable, entering a regression via a multiplicative term with the other source [37]. Such nonlinear systems are non-Gaussian, so PID on systems with suppressor variables is beyond the scope of this paper. However, our demonstration of cases of net synergy for linear Gaussian systems suggests that observing net synergy does not necessarily imply the presence of a suppressor variable. Further, in concordance with the nonstraightforward relationship found here between net synergy and correlation between sources, it has been shown in [38,39] that net synergy is not a useful measure for assessing the importance of correlations between neurons (or neural populations) for successful stimulus decoding. Using the MMI PID, redundancy and synergy can now be computed separately on neurophysiological data sets on which a Gaussian approximation is valid to bring more detailed insight into information-theoretic analyses.

C. Final remarks

We found that if one were to quantify information as reduction in variance rather than reduction in entropy for jointly univariate Gaussian variables, then the net synergy would be precisely zero for uncorrelated sources (see Sec. III). Since it is counterintuitive that synergy should arise in the absence of interactions between sources, this suggests that perhaps reduction in variance is a better measure of information for Gaussian variables than mutual information based on Shannon entropy, which results in information being based on the concave logarithmic function and leads to a distorting effect when comparing combined information from two sources with the sum of information from each source on its own in the formula for net synergy. Since Shannon information between continuous random variables is more precisely based on differential, as opposed to absolute, entropy (see Sec. II), its interpretation in terms of reduction of uncertainty is in any case somewhat ambiguous, in spite of being widely used. One would, however, lose the symmetry of information if redefining it as reduction in variance. Angelini *et al.* [12] made a similar observation for Granger causality: A formula based solely on variances, without taking logarithms, results in the Granger causality from a group of independent variables being equal to the sum of Granger causalities from the individual variables (assuming linearity). Future studies of synergy might benefit from further consideration of alternative measures of basic mutual information for continuous random variables.

The MMI PID constitutes a viable candidate PID for information sharing and transfer among a group of three jointly Gaussian variables. This will be useful given that the Gaussian approximation is so widely used when analyzing continuous time-series. There is, therefore, the possibility of application of the MMI PID to a broad range of complex systems, opening up the opportunity to explore relations between any macroscopic phenomenon and the distinct categories of information sharing (redundant, unique, and synergistic) among triplets of continuous time-series variables. The isolation of synergistic information from the other categories could be useful for measuring complexity, by quantifying more correctly than difference in mutual information alone the extent to which information from multiple sources taken together is greater than that from individual sources taken separately (see Sec. VII). A challenge for future work is to obtain a more general framework for PIDs on continuous random variables: for variables following other distributions and for the scenario of more than two source variables.

ACKNOWLEDGMENTS

I thank Lionel Barnett, Joseph Lizier, and Michael Wibral for invaluable discussions during the writing of this paper, Anil Seth for very useful comments on drafts of the manuscript, Michael Schartner for a read-through of the final draft before submission, and Johannes Rauh for comments on the first arXiv version. I thank Marcello Massimini and his laboratory for hosting me as a visiting fellow at the University of Milan. A.B.B. was funded by Engineering and Physical Sciences Research Council Grant No. EP/L005131/1.

[1] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).
[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley-Interscience, New York, 2006).
[3] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms* (Cambridge University Press, Cambridge, 2003).
[4] M. Prokopenko, F. Boschietti, and A. J. Ryan, Complexity **15**, 11 (2009).
[5] P. L. Williams and R. D. Beer, arXiv:1004.2515.
[6] V. Griffith and C. Koch, in *Emergence, Complexity and Computation Volume 9: Guided Self-Organization: Inception*, edited by M. Prokopenko (Springer, Berlin, 2014), pp. 159–190.
[7] M. Harder, C. Salge, and D. Polani, Phys. Rev. E **87**, 012130 (2013).
[8] N. Bertschinger, J. Rauh, E. Olbrich, and J. Jost, in *Proceedings of the European Conference on Complex Systems 2012*, edited by T. Gilbert, M. Kirkilionis, and G. Nicolis (Springer, Berlin, 2013), pp. 251–269.
[9] V. Griffith, E. K. P. Chong, R. G. James, C. J. Ellison, and J. P. Crutchfield, Entropy **16**, 1985 (2014).
[10] N. Bertschinger, J. Rauh, E. Olbrich, J. Jost, and N. Ay, Entropy **16**, 2161 (2014).
[11] I. Kontoyiannis and B. Lucena, *IEEE Proceedings of the International Symposium on Information Theory* (IEEE, Piscataway, 2005), p. 1651.
[12] L. Angelini, M. de Tommaso, D. Marinazzo, L. Nitti, M. Pellicoro, and S. Stramaglia, Phys. Rev. E **81**, 037201 (2010).
[13] L. Barnett, A. B. Barrett, and A. K. Seth, Phys. Rev. Lett. **103**, 238701 (2009).
[14] A. B. Barrett, L. Barnett, and A. K. Seth, Phys. Rev. E **81**, 041907 (2010).
[15] J. T. Lizier and M. Prokopenko, Eur. Phys. J. B **73**, 605 (2010).
[16] A. K. Seth, E. Izhikevich, G. N. Reeke, and G. M. Edelman, Proc. Natl. Acad. Sci. U.S.A. **103**, 10799 (2006).
[17] S. Stramaglia, G. R. Wu, M. Pellicoro, and D. Marinazzo, Phys. Rev. E **86**, 066211 (2012).
[18] N. Wiener, in *Modern Mathematics for Engineers*, edited by E. F. Beckenbach (McGraw-Hill, New York, 1956).
[19] C. W. J. Granger, Econometrica **37**, 424 (1969).
[20] M. Ding, M. Y. Chen, and S. Bressler, in *Handbook of Time Series Analysis*, edited by S. Schelter, M. Winterhalder, and J. Timmer (Wiley, Weinheim, 2006), pp. 438–460.
[21] A. K. Seth, A. B. Barrett, and L. Barnett, Philos. Trans. R. Soc. London Ser. A **369**, 3748 (2011).
[22] M. Shanahan, Phys. Rev. E **78**, 041924 (2008).
[23] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, Chaos **20**, 037109 (2010).
[24] L. Barnett, J. T. Lizier, M. Harré, A. K. Seth, and T. Bossomaier, Phys. Rev. Lett. **111**, 177203 (2013).
[25] D. Balduzzi and G. Tononi, PLoS Comput. Biol. **4**, e1000091 (2008).
[26] A. B. Barrett and A. K. Seth, PLoS Comput. Biol. **7**, e1001052 (2011).

[27] V. Griffith, arXiv:1401.0978.

[28] S. Bressler and A. K. Seth, Neuroimage **58**, 323 (2011).

[29] K. Friston, R. Moran, and A. K. Seth, Curr. Opin. Neurobiol. **23**, 172 (2013).

[30] M. Wibral, J. T. Lizier, and V. Priesemann, Front. Robot. AI **2**, 5 (2015).

[31] N. Brenner, S. P. Strong, R. Koberle, and W. Bialek, Neural Comput. **12**, 1531 (2000).

[32] C. K. Machen, M. B. Stemmler, P. Prinz, R. Krahe, B. Ronacher, and A. V. M. Herz, J. Neurosci. **21**, 3215 (2001).

[33] L. M. A. Bettencourt, G. J. Stephens, M. I. Ham, and G. W. Gross, Phys. Rev. E **75**, 021915 (2007).

[34] L. M. A. Bettencourt, V. Gintautas, and M. I. Ham, Phys. Rev. Lett. **100**, 238701 (2008).

[35] Q. Gaucher, C. Huetz, B. Gourévitch, and J. M. Edeline, J. Neurosci. **33**, 10713 (2013).

[36] D. Marinazzo, O. Gosseries, M. Boly, D. Ledoux, M. Rosanova, M. Massimini, Q. Noirhomme, and S. Laureys, Clin. EEG Neurosci. **45**, 33 (2014).

[37] S. Stramaglia, J. M. Cortes, and D. Marinazzo, New J. Phys. **16**, 105003 (2014).

[38] S. Nirenberg and P. E. Latham, Proc. Natl. Acad. Sci. U.S.A. **100**, 7348 (2003).

[39] P. E. Latham and S. Nirenberg, J. Neurosci. **25**, 5195 (2005).