

Article

Residual Predictive Information Flow in the Tight Coupling Limit: Analytic Insights from a Minimalistic Model

Benjamin Wahl ^{1,2,*}, Ulrike Feudel ^{1,3}, Jaroslav Hlinka ^{4,5} , Matthias Wächter ², Joachim Peinke ² and Jan A. Freund ^{1,3}

¹ Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, 26129 Oldenburg, Germany; ulrike.feudel@uni-oldenburg.de (U.F.); jan.freund@uni-oldenburg.de (J.A.F.)

² ForWind-Center for Wind Energy Research, Institute of Physics, University of Oldenburg, 26129 Oldenburg, Germany; matthias.waechter@uni-oldenburg.de (M.W.); joachim.peinke@uni-oldenburg.de (J.P.)

³ Research Center Neurosensory Science, University of Oldenburg, 26129 Oldenburg, Germany

⁴ Institute of Computer Science of the Czech Academy of Sciences, 18207 Prague, Czech Republic; hlinka@cs.cas.cz

⁵ National Institute of Mental Health, 25067 Klecany, Czech Republic

* Correspondence: benjamin.wahl@uni-oldenburg.de

Received: 30 July 2019; Accepted: 11 October 2019; Published: 17 October 2019



Abstract: In a coupled system, predictive information flows from the causing to the caused variable. The amount of transferred predictive information can be quantified through the use of transfer entropy or, for Gaussian variables, equivalently via Granger causality. It is natural to expect and has been repeatedly observed that a tight coupling does not permit to reconstruct a causal connection between causing and caused variables. Here, we show that for a model of interacting social groups, carried from the master equation to the Fokker–Planck level, a residual predictive information flow can remain for a pair of uni-directionally coupled variables even in the limit of infinite coupling strength. We trace this phenomenon back to the question of how the synchronizing force and the noise strength scale with the coupling strength. A simplified model description allows us to derive analytic expressions that fully elucidate the interplay between deterministic and stochastic model parts.

Keywords: time series; information transfer; causality

1. Introduction

In many scientific problems, causal interactions of different processes or process components are clear from a mechanistic understanding of the system dynamics and reflected by model formulation. In other situations, the interaction structure of a large and/or complex system, e.g., the assembly dynamics of a chemical system [1,2], of an ecological [3,4] or social community [5,6] or of neuronal populations [7,8], may not be clear in advance. Then, the empirical reconstruction of causal interactions from multivariate time series can be attempted using a suitable concept of causality. Since the seminal works of Wiener [9] and Granger [10], this reconstruction is based on the rationale that a causal interaction can be inferred from an increased prediction error when excluding the causing process component from the prediction of the caused component. Operationalizing this idea leads to Granger causality (GC), a measure commonly used to compare causal interactions of different process components. The quantification of a predictive information flow is also behind the definition of transfer entropy [11], which quantifies the predictive information transfer from the causing to the caused component. As shown by Barnett et al. [12], Granger causality and transfer entropy are equivalent for Gaussian variables.

From the outlined definition, it is plausible to expect that a stronger coupling will enhance the predictive information flow. On the other hand, it can be expected that an increasing coupling strength will eventually synchronize the caused with the causing variable. The approach to the synchronization manifold can be monitored by the synchronization index (phase locking value) or, for a linear dynamics, by the cross-correlation coefficient. Once the dynamics is tied to the synchronization manifold, a reconstruction is no longer possible because in a synchronous mode of motion leader and follower cannot be told apart—early reports of this insight [13,14] were followed by repeated observations of the phenomenon in the context of various directionality indices and synchronization measures (e.g., [15–19]), which underpins a generic phenomenon. Combining the synchronization triggered decline of directionality measures with the initial increase in the small coupling regime immediately suggests unimodal curve describing the relation between directionality and coupling strength, where any directionality index vanishes for zero (uncoupled) and infinite (tight coupling limit) coupling strength.

We note that, to the best of our knowledge, all reports of this phenomenon are solely based on numerical or empirical data and subsequent estimation of directionality indices. Moreover, directionality indices are mostly constructed as differences (asymmetries) between directed coupling indices, thus quantifying a two-way net directionality. Quite often estimators of the one-way directionality indices are plagued by a considerable bias and, in particular, non-zero values for uncoupled variables. Computing the net directionality index this bias may (or may not) be removed by forming differences. Therefore, an analytic treatment of a one-way directionality index vs. coupling strength relation for a tractable but generic dynamics seems desirable.

In the following, we address such a system dynamics that follows from a socio-dynamic model originally formulated in the framework of a master equation [20]. In the thermodynamic limit, the description is shifted to the level of a Fokker–Planck equation [21] with related drift and diffusion coefficients. The resulting expressions show that the coupling strength enters the diffusion matrix which means that random fluctuations scale up together with increasing coupling strength. Our analytic derivation leads to the interesting result that even in the tight coupling limit a residual directionality remains. Since the resultant dynamics can effectively be mapped to a vector-autoregressive process (linear stochastic process with Gaussian variables), the residual Granger causality can be interpreted as a residual predictive information flow remaining for tight coupling. Although we use a specific model at the outset of our presentation, we argue that the finding of a residual GC is a generic effect.

2. Granger Causality in a Nutshell

To set the stage, we briefly recall how GC is defined in the framework of a (two-dimensional) vector autoregressive process VAR[∞] (of order infinity). In this case, two time series are fitted by the following full model

$$\begin{aligned} X_{1,t} &= \sum_{j=1}^{\infty} a_j X_{1,t-j} + \sum_{j=1}^{\infty} b_j X_{2,t-j} + \epsilon_t, \\ X_{2,t} &= \sum_{j=1}^{\infty} c_j X_{1,t-j} + \sum_{j=1}^{\infty} d_j X_{2,t-j} + \eta_t \end{aligned} \quad (1)$$

with white noise covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} \text{var}(\epsilon_t) & \text{cov}(\epsilon_t, \eta_t) \\ \text{cov}(\epsilon_t, \eta_t) & \text{var}(\eta_t) \end{pmatrix}. \quad (2)$$

In describing stable (non-divergent) time series by a VAR, the dynamical coefficients a_j, b_j, c_j and d_j will obey a stationarity condition [22].

The prediction error of X_2 , using both its own past *and* the past of X_1 , is quantified by Σ_{22} , the variance of η_t . Predicting X_2 *only* from its own past means to use a reduced univariate autoregressive processes description for the second time series

$$X_{2,t} = \sum_{j=1}^{\infty} \tilde{d}_{rj} X_{2,t-j} + \tilde{\eta}_t, \quad \text{var}(\tilde{\eta}_t) = \Xi_2. \quad (3)$$

In passing, we note that the parameters of this reduced model are not fitted independently but derived from Equation (1) for example through a spectral factorization [23]. Moreover, based on a state-space formulation of GC, a closed form involving partial covariances can be obtained for finite order vector-autoregressive models [24].

Now, the prediction error of X_2 without inclusion of X_1 is quantified by Ξ_2 , the variance of $\tilde{\eta}_t$. Since in this step no information is added (but rather removed), Ξ_2 will never be smaller than Σ_{22} , guaranteeing that Granger causality GC (from $X_1 \rightarrow X_2$) defined as the quantity $\log(\frac{\Xi_2}{\Sigma_{22}})$ is non-negative.

VAR processes belong to the model class of time discrete linear stochastic processes. In a recent work, we extended the concept of GC to diffusion processes which constitute a special class of time continuous stochastic processes. Diffusion processes can be formulated either via a Langevin type stochastic differential equation system or via the related Fokker–Planck equation (FPE) [25]. In the framework of a local linearization of the dynamics and a stroboscopic map (i.e., the temporal integration of the dynamic flow over a time increment Δt), this allows establishing a connection between drift and diffusion coefficients of the FPE and a local GC. Averaging the local GC with respect to the invariant measure yields a global GC for the original diffusion process. For details, we refer the reader to our recent publication [26].

3. Weidlich's Socio-Dynamic Model

Here, we apply this connection to a stochastic socio-dynamics model proposed by Weidlich [20] to describe opinion formation in a system with two groups, *leaders* (Group 1 comprising n_1 members) and *followers* (Group 2 with n_2 members). Individuals of each group $i = 1, 2$ can adopt one of two opinions $j = 1, 2$. Opinion diversity across the community is coded by variables $n_{i,j}$ specifying the number of individuals of group i sharing opinion j . Since it is assumed that everybody has an opinion, $n_{i,1} + n_{i,2} = n_i$. Eventually, the system is considered in the “thermodynamic limit” $n_1, n_2 \rightarrow \infty$ with $n_1/n_2 = \epsilon$, which suggests a change to quasi-continuous variables

$$0 \leq y_{i,j} := \frac{n_{i,j}}{n_i} \leq 1 \quad \text{with} \quad y_{i,1} + y_{i,2} = 1. \quad (4)$$

Members of group i change their opinion from j to j' with a transition rate $p_{j \rightarrow j'}^{(i)}$ that may depend on the present opinion configuration $\{y_{\alpha,\beta}\}$ (with $\alpha, \beta = 1, 2$). In line with Weidlich, we employ the following expressions

$$p_{1 \rightarrow 2}^{(1)} = \frac{3}{4}p, \quad p_{2 \rightarrow 1}^{(1)} = \frac{1}{4}p, \quad p_{1 \rightarrow 2}^{(2)} = \frac{1}{4}p + \mu p \cdot y_{1,2}, \quad p_{2 \rightarrow 1}^{(2)} = \frac{3}{4}p + \mu p \cdot y_{1,1} \quad (5)$$

with a straightforward interpretation: without coupling ($\mu = 0$) leaders and followers have a diametrical preference for one opinion over the alternative one; while leaders flip opinion without reference to the population of followers ($y_{2,j}$), with increasing coupling followers enhance the rate of switching their opinion to the opinion populated by leaders ($y_{1,j}$). This means that followers are uni-directionally forced by leaders with the coupling strength scaled by the parameter μ .

These rates enter a master equation describing the evolution of a statistical ensemble of interacting social groups. After another change of variables $x_i = y_{i,2} - y_{i,1}$, reflecting directed imbalance and

opinion diversity within group i , the related Fokker–Planck equation can be written down [20]. It is completely specified by its drift vector

$$D^{(1)} = -p \begin{pmatrix} x_1 - \frac{1}{2} \\ x_2 + \frac{1}{2} + \mu \cdot (x_2 - x_1) \end{pmatrix} \tag{6}$$

and its diffusion matrix

$$D_{11}^{(2)} = \frac{p}{n_1}(2 - x_1), \quad D_{22}^{(2)} = \frac{p}{n_2}(2 + x_2 + 2\mu(1 - x_1x_2)), \quad D_{12}^{(2)} = D_{21}^{(2)} = 0. \tag{7}$$

Notice that the diffusion matrix is positive definite since $-1 \leq x_1, x_2 \leq 1$. The inverse scaling with n_1 and n_2 , respectively, means that diffusion vanishes in the “thermodynamic limit” $n_1, n_2 \rightarrow \infty$ (with $\epsilon = n_1/n_2$). Moreover, since the parameter p (controlling the intrinsic opinion flip rate) scales both drift vector and diffusion matrix in the same way, it may well be absorbed in a reparameterization of time. Finally, we point out that in the step from the master equation to the FPE the coupling parameter μ does not only enter the drift vector but also the diffusion matrix.

4. Simulation Results

Being equipped with drift vector and diffusion matrix, we can compute the GC of the socio-dynamic diffusion process (as outlined in [26]) for various values of the coupling parameter μ . This means we compute GCs locally for subsampled vector-Ornstein–Uhlenbeck processes with local additive noise; global GC is then obtained as the statistical average of local GCs with respect to a stationary distribution estimated from the simulation (1,000,000 data points for each μ value). Additionally, we compute the GC of a global vector-Ornstein–Uhlenbeck process with constant average-noise. The results are plotted with circle and crosses, respectively, in Figure 1.

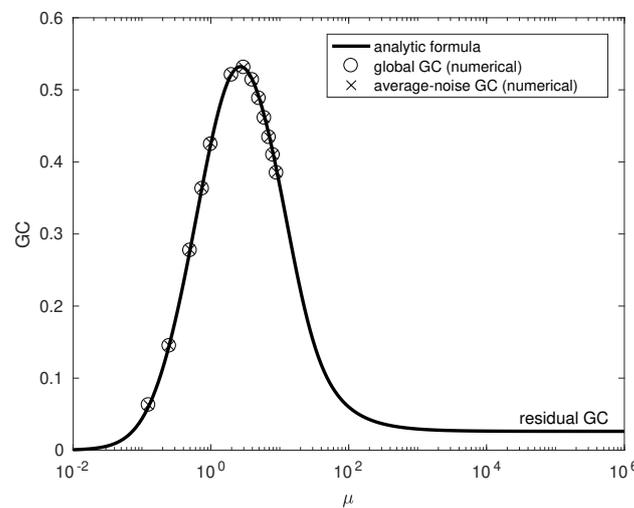


Figure 1. Dependence of Granger causality GC on the coupling parameter μ : circles, numerical simulation results of the Langevin equation based on drift and diffusion coefficients (Equations (6) and (7), respectively); crosses, GC of the global VAR(1) (Equation (8)) with average-noise; solid line, analytic GC (Equation (20)) of the minimalistic model.

As expected, for $\mu = 0$, the GC is zero because in this uncoupled situation inclusion of x_1 cannot reduce the prediction error of x_2 . With increasing coupling, parameter μ GC initially increases up to a maximum around $\mu = 3$ before it decreases again. Surprisingly, we find that a linearized dynamics with average-noise (notice that the only non-linearity $\sim x_1x_2$ resides in $D_{22}^{(2)}$) with state space averaged additive noise gives rise to the same values (plotted in Figure 1 with crosses). Therefore, the non-monotonic characteristic of GC requires neither a non-linear dynamics nor multiplicative noise.

5. Analytic Description

To explain the non-monotonic behavior of the GC, we connect the diffusion process specified by Equations (6) and (7) with a VAR(1), i.e., vector-autoregressive process of first order. The general connection is elaborated in [27] and proceeds via the VAR(1) written as

$$X_t = AX_{t-1} + \theta_t \tag{8}$$

where

$$A = \exp \left[\begin{pmatrix} -1 & 0 \\ \mu & -(1 + \mu) \end{pmatrix} p\Delta t \right] = e^{-p\Delta t} \begin{pmatrix} 1 & 0 \\ 1 - e^{-\mu p\Delta t} & e^{-\mu p\Delta t} \end{pmatrix} \tag{9}$$

and θ_t is a two-component vector of zero-mean Gaussian white noise residuals with a covariance matrix Σ given by the following matrix integral

$$\Sigma = \int_0^{\Delta t} e^{\Gamma t} \begin{pmatrix} \hat{D}_{11}^{(2)} & 0 \\ 0 & \hat{D}_{22}^{(2)} \end{pmatrix} e^{\Gamma^+ t} dt, \tag{10}$$

where Γ is the Jacobian matrix of the diffusion process, i.e., of its drift part.

The above connection is derived for the standard vector-Ornstein-Uhlenbeck process with additive noise, which corresponds to a diffusion matrix $\hat{D}_{ij}^{(2)}$ that is constant in state space. By contrast, in Weidlich’s model we see that the diffusion coefficients $D_{11}^{(2)}$ and $D_{22}^{(2)}$ (cf. Equation(7)) vary with x_1 and x_2 . Since the gradients scale with the inverse of n_1 and n_2 , respectively, for rather large n_1, n_2 , the noise is weakly multiplicative [26]. Therefore, we may consider $\Sigma(x_1, x_2)$ as a local covariance matrix. A rigorous derivation would now strive to compute local GC maps and perform a subsequent average over state space (with respect to the stationary probability distribution, essentially a bivariate Gaussian). This route seems feasible but resultant expressions appear far from being handy. Therefore, we approximate state dependent diffusion coefficients by substituting mean values for state variables, i.e.,

$$D_{ij}^{(2)}(x_1, x_2) \rightarrow D_{ij}^{(2)}(\langle x_1 \rangle, \langle x_2 \rangle). \tag{11}$$

The mean values can be read from Equation (6)

$$\langle x_1 \rangle = \frac{1}{2} \quad \text{and} \quad \langle x_2 \rangle = \frac{1}{2} \frac{\mu - 1}{\mu + 1} \tag{12}$$

with the consequence that we approximate

$$\sigma_1^2 := \langle D_{11}^{(2)}(x_1, x_2) \rangle \approx \frac{p}{n_1} \frac{3}{2}, \quad \sigma_2^2(\mu) := \langle D_{22}^{(2)}(x_1, x_2) \rangle \approx \frac{p}{n_2} \frac{1}{2} \frac{3\mu^2 + 10\mu + 3}{\mu + 1}. \tag{13}$$

In line with these simplifications, we now consider the case of a VAR(1) with the following parameterization

$$X_t = \begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix} = \begin{pmatrix} b & 0 \\ c & a \end{pmatrix} \begin{pmatrix} X_{1,t-1} \\ X_{2,t-1} \end{pmatrix} + \begin{pmatrix} \theta_{1,t} \\ \theta_{2,t} \end{pmatrix}, \tag{14}$$

where the parameters a, b, c are to be identified with the entries in Equation (9), i.e.,

$$a = e^{-(1+\mu)h}, \quad b = e^{-h}, \quad c = e^{-h} - e^{-(1+\mu)h} \tag{15}$$

where we have used the abbreviation $h = p\Delta t$.

The covariance matrix of residuals is obtained by inserting σ_1^2 and $\sigma_2^2(\mu)$ of Equation (13) into Equation (10). The integration can be performed analytically and leads to the following expressions

$$\begin{aligned} \Sigma_{11} &= \frac{\sigma_1^2}{2p}(1 - e^{-2h}), & \Sigma_{12} = \Sigma_{21} &= \frac{\sigma_1^2}{2p} \frac{\mu(1 - e^{-2h}) - 2(e^{-2h} - e^{-(2+\mu)h})}{\mu + 2} \\ \Sigma_{22} &= \frac{\sigma_2^2(\mu)}{2p} \frac{1 - e^{-2(1+\mu)h}}{\mu + 1} + \frac{\sigma_1^2}{2p} \times \\ &\times \frac{\mu^2(1 - e^{-2h}) - \mu(3e^{-2h} - 4e^{-(2+\mu)h} + e^{-2(1+\mu)h}) - 2(e^{-2h} - 2e^{-(2+\mu)h} + e^{-2(1+\mu)h})}{\mu^2 + 3\mu + 2}. \end{aligned} \tag{16}$$

Notice that

$$\lim_{\mu \rightarrow \infty} \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{11} \\ \Sigma_{11} & \left[1 + \frac{n_1/n_2}{1 - e^{-2h}}\right] \Sigma_{11} \end{pmatrix}. \tag{17}$$

With $\eta := \frac{n_1/n_2}{2(1 - e^{-2h})} \sim \epsilon$ being small (since $n_1 \ll n_2$ and $h \ll 1$), this matrix has the eigenvalues

$$\lambda_{1,2} = \left(1 + \eta \pm \sqrt{1 + \eta^2}\right) \Sigma_{11} = \begin{cases} [2 + \eta + \mathcal{O}(\eta^2)] \Sigma_{11} \\ [\eta + \mathcal{O}(\eta^2)] \Sigma_{11} \end{cases}, \tag{18}$$

which shows that even in the tight coupling limit the covariance ellipsoid of residuals possesses two finite semi-axes.

Equation (14) defines a VAR(1) that was recently considered by Barnett and Seth as a minimal model that allows deriving an analytic expression for GC that can be found in [28] and was restricted to the special case that Σ was the identity matrix, i.e., $\theta_{1,t}$ and $\theta_{2,t}$ were uncorrelated ($\Sigma_{12} = \Sigma_{21} = 0$) and both with unit variance ($\Sigma_{11} = \Sigma_{22} = 1$). Relaxing these restrictions and following the same steps outlined by Barnett and Seth we can derive an explicit formula for GC that is structurally identical to Barnett’s formula even for the most general, bivariate VAR(1) process in Equation (8) with bidirectional coupling ($A_{12} = d \neq 0$)

$$GC = \frac{1}{\pi} \int_0^\pi \ln \left(1 + \frac{c'^2}{1 - 2b' \cos \lambda + b'^2}\right) d\lambda \tag{19}$$

$$= \ln \left(\frac{1 + b'^2 + c'^2 + \sqrt{(1 + b'^2 + c'^2)^2 - 4b'^2}}{2}\right) \tag{20}$$

but with the replacement

$$b' = b - \frac{\Sigma_{12}}{\Sigma_{22}}c, \quad c' = \frac{\sqrt{\det \Sigma}}{\Sigma_{22}}c. \tag{21}$$

Through Equations (20) and (21) in conjunction with Equation (15), we have everything at hand to evaluate the dependence of GC on the coupling parameter μ . Note that the GC does not depend on the parameter a , and for a fixed covariance matrix is a monotonically increasing function of c growing to infinity; indeed, in a simple autoregressive model, the parameter c would be naturally considered (a monotonous function of) the coupling strength μ . However, in our example, the dependence of the c' on coupling parameter μ is more complicated, as it also affects the covariance matrix. Moreover, the dependence is such that there is a delicate balance for infinite coupling strength μ between the synchronizing influence of the driver and the diffusion term leading to the finite residual GC. The result is plotted in Figure 1 (with solid line) and shows that the analytic formula perfectly matches the numerical values (plotted with circles and crosses) extracted from the simulation. The

matching of our analytical results, based on Equation (11), and global GC is surprising but fortunate. The coincidence of global GC and average-noise GC, on the other hand, is not surprising since, for weakly multiplicative noise, the latter is in line with our analytic approach.

The analytic formula allows considering the asymptotic limit $\mu \rightarrow \infty$, which is given by inserting

$$\lim_{\mu \rightarrow \infty} b' = \frac{\epsilon e^{-h}}{1 + \epsilon - e^{-2h}}, \quad \lim_{\mu \rightarrow \infty} c' = \frac{\sqrt{\epsilon e^{-2h}(1 - e^{-2h})}}{1 + \epsilon - e^{-2h}} \tag{22}$$

into Barnett’s formula (Equation (20)). Interestingly, n_1 and n_2 again enter this expression only through the ratio $\epsilon = n_1/n_2$. For the chosen parameters, the resulting asymptotic value $\lim_{\mu \rightarrow \infty} GC$ is approximately 0.027, which is larger than zero. As the scaling with $\epsilon \sim \eta$ indicates, the asymptotic residual GC is a consequence of the residual variance in the direction of the second semi-axes of the covariance ellipsoid in Equation (18). Residual fluctuations remaining in the tight coupling limit enable an improvement of prediction.

These residual fluctuations reflect the fact that even for infinite coupling strength synchronization is not complete. To show this, we compute the cross-correlation function by solving the Yule–Walker equations [22]. As can be seen in Figure 2, the maxima of the curves rise with increasing coupling strength while the related lag approaches zero (in discrete steps) reflecting instantaneous response of the caused variable. However, our analytic treatment reveals that even in the tight coupling limit $\mu \rightarrow \infty$ the maximal cross-correlation is 0.995 and not one; the reason is, again, the balance between the synchronizing force (exponential relaxation) and the increasing noise strength $\sigma_{22}(\mu)$. We note that this value is not plagued by estimation errors.

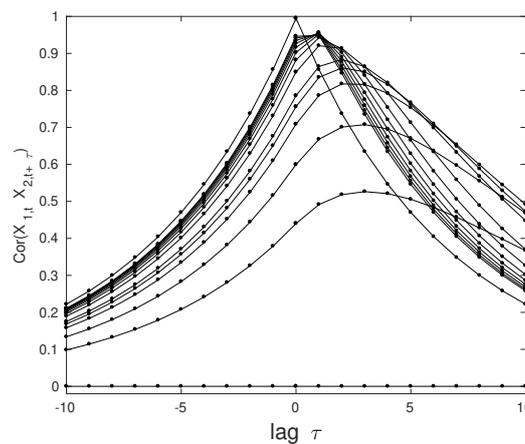


Figure 2. Cross-correlations for couplings $\mu = 0, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1, \dots, 9, \infty$ (bottom to top for all non-positive lags) and various lags τ .

The analytic approach also puts us in the position to sort out how the different system properties contribute to the formation of a maximum in the GC. In particular, we can investigate the effect of coupling dependent diffusivity through replacing $\sigma_2^2(\mu)$ (cf. Equation (13)) by the constant value $\sigma_2^2(\mu = 0) = \epsilon\sigma_1^2$. Through these investigations, we arrive at the following conclusions:

- For uncoupled systems ($\mu = 0$), GC is zero because $c' = 0$.
- The initial increase of GC is caused by the exponential increase of c towards b with increasing coupling c (cf. Equation (15)). Residual variance of the process component X_2 is reduced via information transferred from the driver X_1 .
- Simultaneously with the information transferred from the driver intrinsic fluctuations, modeled by increasing diffusivity $\sigma_2^2(\mu)$, scale up (asymptotically $\sim \mu$). Nevertheless, the residual variance Σ_{22} approaches a constant because tighter coupling to X_1 (as expressed by the denominator $\mu + 1$ in Equation (16)) compensates increasing intrinsic noise. In the limit $\mu \rightarrow \infty$, the residual variance

Σ_{22} is the residual variance Σ_{11} of the driver plus an extra term resulting from the balance between intrinsic noise and tight coupling, i.e., $\lim_{\mu \rightarrow \infty} \Sigma_{22} = \left[1 + \frac{n_1/n_2}{1-e^{-2h}}\right] \Sigma_{11}$

- From the imbalance of diffusivities follows the ratio

$$\frac{\Sigma_{22}(\mu \rightarrow \infty)}{\Sigma_{22}(\mu = 0)} = \frac{1}{\epsilon} \left(1 + \frac{\epsilon}{1 - e^{-2h}}\right) \sim \frac{1}{\epsilon} \quad (23)$$

which means that for $\epsilon = 0.01$ the residual variance Σ_{22} is increasing roughly by the factor 100. As can be read from Equation (21), this tends to reduce c^2 with increasing μ . From the opposing trends - c^2 increases with μ while $1/\Sigma_{22}$ decreases with increasing μ - the peak results as the point where the dominance of both trends changes. We note that analytic expressions can be derived since we have all ingredients at hand, however, explicit formulas are quite lengthy and do not provide deeper insights. In all cases considered, we found that $\sqrt{\det \Sigma}$ varied only mildly with μ and the variation of b^2 is of minor importance for Barnett's formula (Equation (20)).

6. Discussion

The increase of GC with μ is a natural consequence of enhanced information transfer with tighter coupling to the driver. The asymptotic decrease and the formation of a maximum, on the other hand, follow from the fact that the caused variable X_2 asymptotically adopts the residual variance of the driver which is two orders larger because $n_2/n_1 = \epsilon^{-1} = 100$, meaning that "every leader has one hundred followers". This allows predicting that no such maximum can be found if $n_1 \approx n_2$ or for $n_1 > n_2$, i.e., if the leader group is of comparable or even larger size than the follower group—a statement that we can confirm through our analytic analysis. It is interesting to note that the asymptotic residual GC, i.e., an improvement of prediction in the tight-coupling limit, remains even in the thermodynamic limit $n_1, n_2 \rightarrow \infty$, i.e., when the diffusion constants approach zero, provided $\epsilon = n_1/n_2 > 0$. This results from the scaling behavior of the variance Σ_{22} .

The question whether the found residual GC is a property of the specific socio-dynamic model or a generic property of a larger class of systems can be answered in the following way: our results follow from a minimalistic linear system (vector-Ornstein-Uhlenbeck process or VAR(1)), which means it is not a fancy effect of non-linear dynamics. In this setup, the rGC results from the linear asymptotic scaling behavior of $\sigma_2^2(\mu)$ (Equation (13)), which balances the μ in the denominator of the first term in Equation (16) resulting from integrating the exponentials (in Equation (10)) and reflecting the faster relaxation with increasing coupling strength. The question as to which other systems might give rise to a residual GC thus leads back to the linear scaling of the diffusion coefficient $D_{22}^{(2)}(\mu)$. As the derivation in Weidlich's original paper [20] shows, this is a consequence of the ansatz of coupling groups via a coupling parameter that enters the flipping rates (Equation (5)) additively and linearly. Therefore, we conclude that a residual GC will not be restricted to the specific details of Weidlich's model but be a quite generic aspect of a stochastic dynamics described by a similar master equation.

The observed non-monotonic behavior of GC with coupling strength does not rely on non-linear dynamics or multiplicative noise. A peak of the information flow was pointed out and a finite value of the information flow, at large couplings, was also reported in [29]. However, it is interesting to consider how our insights can be transferred to a general diffusion process. Our conclusions drawn from the minimalistic system (see Equation (14) with Equation (15)), i.e., a linear dynamics with additive noise, can be carried over to the class of nonlinear diffusion processes with weakly multiplicative noise in the framework of a local linearization generating a GC map in state space. Subsequent averaging with respect to the invariant measure on the attractor leads to the global GC of the system, a concept that we put forward in a recent publication [26]. In this way, the global GC vs. coupling relation will be computable and the resultant curve will depend on the details of the considered system. The question whether a non-monotonic GC vs. coupling relation can result for a system with diffusion constants that do *not* vary with coupling strength remains open and will be dealt with in a forthcoming publication.

Finally, one may ask whether the non-monotonic GC curve plays any role in natural systems. We speculate that coupling enhanced predictability might be useful for the performance of cognitive functions and that certain brain regions are tuned to optimal synaptic coupling. In this context, a scale up of effects through coupling large neuronal ensembles can also be imagined.

7. Summary

We analytically investigated the non-monotonic behavior of Granger causality, defined on a socio-dynamic model of leaders and followers, with respect to coupling strength. Our derivations finally substantiate frequent observations of this phenomenon for various directionality indices and agree with the common sense that synchronization impedes a reconstruction of causal relationships. To make this point, it was necessary to extend Granger causality of Barnett's minimal model to a general vector-autoregressive process of order one. Fortunately, we were able to demonstrate that a tight coupling in the master equation can map not only to the dynamics but also to the noise and therefore does not automatically imply synchronization and a vanishing Granger causality.

Author Contributions: All authors have contributed substantially to this work: Conceptualization, B.W.; Methodology, B.W. and J.A.F.; Software, B.W.; Formal analysis, B.W. and J.A.F.; Investigation, B.W. and J.A.F.; Resources, U.F.; Data curation, B.W.; Writing—original draft preparation, B.W.; Writing—review and editing, U.F., J.H., M.W., J.P. and J.A.F.; Visualization, B.W. and J.A.F.; Supervision, U.F. and J.P.; Project administration, J.P. and M.W.; and Funding acquisition, M.W., U.F. and J.P.

Acknowledgments: Benjamin Wahl thanks the presidial chair of the University Oldenburg and the EWE research institute NextEnergy for partly supporting this work. Jaroslav Hlinka was supported by the Czech Science Foundation projects GA19-11753S and the Czech Health Research Council project NV17-28427A and project Nr. LO1611 with a financial support from the MEYS under the NPU I program.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Srividhya, J.; Mourão, M.A.; Crampin, E.J.; Schnell, S. Enzyme catalyzed reactions: From experiment to computational mechanism reconstruction. *Comput. Biol. Chem.* **2010**, *34*, 11–18. [[CrossRef](#)] [[PubMed](#)]
2. Srividhya, J.; Crampin, E.J.; McSharry, P.E.; Schnell, S. Reconstructing biochemical pathways from time course data. *Proteomics* **2007**, *7*, 828–838. [[CrossRef](#)] [[PubMed](#)]
3. Fuhrman, J.A.; Cram, J.A.; Needham, D.M. Marine microbial community dynamics and their ecological interpretation. *Nat. Rev. Microbiol.* **2015**, *13*, 133–146. [[CrossRef](#)] [[PubMed](#)]
4. Faisal, A.; Dondelinger, F.; Husmeier, D.; Beale, C.M. Inferring species interaction networks from species abundance data: A comparative evaluation of various statistical and machine learning methods. *Ecol. Inf.* **2010**, *5*, 451–464. [[CrossRef](#)]
5. Wasserman, S.; Faust, K. Structural analysis in the social sciences. In *Social Network Analysis: Methods and Applications*; Cambridge University Press: Cambridge, UK, 1994.
6. Scott, J. *Social Network Analysis: A Handbook*, 2nd ed.; SAGE: New York, NY, USA, 2000.
7. Quinn, C.J.; Coleman, T.P.; Kiyavash, N.; Hatsopoulos, N.G. Estimating the directed information to infer causal relationships in ensemble neural spike train recordings. *J. Comput. Neurosci.* **2011**, *30*, 17–44. [[CrossRef](#)]
8. Gerhard, F.; Pipa, G.; Lima, B.; Neuenschwander, S.; Gerstner, W. Extraction of Network Topology From Multi-Electrode Recordings: Is there a Small-World Effect? *Front. Comput. Neurosci.* **2011**, *5*, 4. [[CrossRef](#)]
9. Wiener, N. The theory of prediction. In *Modern Mathematics for Engineers, Series I*; Beckenham, E.F., Ed.; McGraw-Hill: New York, NY, USA, 1956.
10. Granger, C.W. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* **1969**, *37*, 424–438. [[CrossRef](#)]
11. Schreiber, T. Measuring Information Transfer. *Phys. Rev. Lett.* **2000**, *85*, 461–464. [[CrossRef](#)]
12. Barnett, L.; Barrett, A.B.; Seth, A.K. Granger Causality and Transfer Entropy Are Equivalent for Gaussian Variables. *Phys. Rev. Lett.* **2009**, *103*. [[CrossRef](#)]
13. Paluš, M.; Komárek, V.; Hrnčíř, Z.; Štěrbová, K. Synchronization as adjustment of information rates: Detection from bivariate time series. *Phys. Rev. E* **2001**, *63*, 046211. [[CrossRef](#)]

14. Rosenblum, M.; Cimponeriu, L.; Bezerianos, A.; Patzak, A.; Mrowka, R. Identification of coupling direction: Application to cardiorespiratory interaction. *Phys. Rev. E* **2002**, *65*. [[CrossRef](#)] [[PubMed](#)]
15. Ancona, N.; Marinazzo, D.; Stramaglia, S. Radial basis function approach to nonlinear Granger causality of time series. *Phys. Rev. E* **2004**, *70*, 056221. [[CrossRef](#)] [[PubMed](#)]
16. Krug, D.; Osterhage, H.; Elger, C.E.; Lehnertz, K. Estimating nonlinear interdependences in dynamical systems using cellular nonlinear networks. *Phys. Rev. E* **2007**, *76*, 041916. [[CrossRef](#)] [[PubMed](#)]
17. Prusseit, J.; Lehnertz, K. Measuring interdependences in dissipative dynamical systems with estimated Fokker-Planck coefficients. *Phys. Rev. E* **2008**, *77*, 041914. [[CrossRef](#)]
18. Staniek, M.; Lehnertz, K. Symbolic transfer entropy. *Phys. Rev. Lett.* **2008**, *100*, 158101. [[CrossRef](#)]
19. Rings, T.; Lehnertz, K. Distinguishing between direct and indirect directional couplings in large oscillator networks: Partial or non-partial phase analyses? *Chaos* **2016**, *26*, 093106. [[CrossRef](#)]
20. Weidlich, W. Fokker-Planck Equation Treatment of Interacting Social Groups. In *Synergetics: Cooperative Phenomena in Multi-Component Systems*; Haken, H., Ed.; Vieweg+Teubner Verlag: Wiesbaden, Germany, 1973; pp. 269–279.
21. Van Kampen, N.G. *Stochastic Processes in Physics and Chemistry*, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 1992.
22. Lütkepohl, H. *New Introduction to Multiple Time Series Analysis*; Springer Science & Business Media: Berlin, Germany, 2005.
23. Barnett, L.; Seth, A.K. The MVGC Multivariate Granger Causality Toolbox: A New Approach to Granger-causal Inference. *J. Neurosci. Methods* **2014**, *223*, 50–68. [[CrossRef](#)]
24. Barnett, L.; Seth, A.K. Granger causality for state-space models. *Phys. Rev. E* **2015**, *91*, 040101. [[CrossRef](#)]
25. Gardiner, C.W. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*; Springer: Berlin, Germany, 1985.
26. Wahl, B.; Feudel, U.; Hlinka, J.; Wächter, M.; Peinke, J.; Freund, J.A. Granger-causality maps of diffusion processes. *Phys. Rev. E* **2016**, *93*, 022213. [[CrossRef](#)]
27. Neumaier, A.; Schneider, T. Multivariate Autoregressive and Ornstein-Uhlenbeck Processes: Estimates for Order, Parameters, Spectral Information, and Confidence Regions. Available online: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.26.9407> (accessed 30 July 2019).
28. Barnett, L.; Seth, A.K. Behaviour of Granger causality under filtering: Theoretical invariance and practical application. *J. Neurosci. Methods* **2011**, *201*, 404–419. [[CrossRef](#)]
29. Marinazzo, D.; Pellicoro, M.; Stramaglia, S. Kernel Method for Nonlinear Granger Causality. *Phys. Rev. Lett.* **2008**, *100*, 144103. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).