NETWORK STRUCTURAL RECONSTRUCTION BASED ON DELAYED TRANSFER ENTROPY AND SYNTHETIC DATA

Daniel Rodrigues de Lima*, Fernando Pasquini Santos*, Carlos Dias Maciel*

*Department of Electrical and Computational Engineering
University of São Paulo
São Carlos, São Paulo, Brazil

Email: {daniel.rodrigues.lima;fernando.pasquini.santos;carlos.maciel}@usp.br

Abstract— The knowledge of how signals are received, processed, and transmitted in neuronal systems is one of the bio-inspired engineering objectives. In this area, not only the physiology of separated neurons is relevant, but also, the connections among them, the neuronal topology. A modeling process of such biological system requires the integration of both, physiological and topological properties. However, there is a limitation in the modeling process due to the impossibility of recording the activity of all the system neurons at the same time. To solve this problem, we propose the usage of simulations and information theoretic measures to infer a network topology. Three test cases were simulated, and the interactions were measured with Transfer Entropy resulting in topology candidates. In two cases we could visually recover the connections from the graphs. In a third case we found a residual connection which allowed us to explore some properties from the network topology.

Keywords— Bio-inspired engineering; neuronal systems; simulation; information theoretic measures; network topology.

Resumo— Compreender o funcionamento dos processos de recepção, processamento e transmissão de informação nos sistemas neuronais é um dos objetivos da engenharia bio-inspirada. Nesse campo de estudo, não somente a fisiologia neuronal é relevante, mas também faz-se necessário o conhecimento sobre a organização da topologia dos neurônios. A modelagem desses sistemas requer a integração de aspectos fisiológicos e topológicos. No entanto, existem limites no processo de modelagem devido à impossibilidade de gravar a atividade de todos os neurônios simultaneamente. Para resolver este problema, propõe-se o uso de simulações e medidas de teoria da informação para inferir uma topologia de redes. Simulou-se três casos de teste e as interações entre variáveis foram medidas com Transferência de Entropia. Em dois casos foi possível recuperar as conexões apenas observando os resultados gráficos. Em um terceiro caso, encontrou-se uma conexão residual, a qual permitiu que algumas propriedades da topologia de rede fossem exploradas.

Palavras-chave— Engenharia bio-inspirada; sistemas neuronais; simulações; medidas de teoria da informação; topologia de rede.

1 Introduction

To propose sound and efficient ways to integrate multimodal sensory inputs in robots, studies are nowadays directing attention to the brain control systems of animals, looking for models and architectural principles that could guide better solutions (Wessnitzer and Webb, 2006). Insects are the most used in these studies due to the way they integrate multiple sensory information with relatively simple and distributed neural systems (when compared to vertebrates) (Wessnitzer and Webb, 2006; Burrows, 1996).

Among the methods used, individual or groups of neurons are recorded and used to train different models whose parameters or structure could then be interpreted and linked to the function of the system represented (Marmarelis, 2004) or some neural coding scheme (Brown et al., 2004). Employed models may range from black-box approaches (Meruelo et al., 2016; Dewhirst et al., 2013), or also directed information theory measures applied to discover the internal structure of neural pathways (Endo et al., 2015; Maciel et al., 2012).

This last approach requires recordings not only from the input and output but also knowledge of the internal states of the system, which will be the nodes of a network model. However, this may become a problem once the number of simultaneous channels available in recording setups is limited (Ritzmann and Büschges, 2007), and also not every neuron or group of neurons can be anatomically identifiable. For example, in the work by Endo et al. (2015), with intracellular recordings, only four channels could be recorded at once.

It becomes difficult to evaluate information transfer between some nodes due to the absence of their simultaneous recordings. Even more because, if the system presents synergistic effects between nodes, higher dimensional information transfer measures, such as conditional transfer entropy (Faes and Porta, 2014; Runge et al., 2012), need to be used, requiring even more and abundant simultaneous data.

Many studies, therefore, argue that simulations are the answer to the complexity and limitation of experimental techniques (Buschmann et al., 2015). The approaches range from simulations of the entire function of the system being studied (Szczecinski et al., 2014) to observations of emergent phenomena from simple and modular neuron networks (Bullmore and Sporns, 2009).

In the present work, we propose a simulation which generates signals in different neural networks, which are used to observe the behavior of information theoretic measures applied between them.

Pairwise information theoretic measures, such as mutual information (McDonnell et al., 2011) or transfer entropy (Schreiber, 2000), present different peaks of maximum information transfer according to differences in direct or indirect connections. It is expected that, the more indirect the connection, the lower the information transfer, due to the data processing inequality (Cover and Thomas, 2006). Thus, by different values of information theoretic measures, direct and indirect connections can be inferred and a network structure can be found even when lacking simultaneous recordings of some nodes. The behavior of measures applied to the simulations is then compared with the real cases measures, and the corresponding network structure is then found.

This work is then organized as follows. The simulation framework, Section 2, describing the steps and presenting the background to our method of simulation. Section 3 presents some aspects related to Bayesian Networks (BN) and Delayed Transfer Entropy (DTE), which is used to reconstruct the network structure. In Section 4, we present our results with the simulated experiment to generate data and analyze it. At last, we conclude this work in Section 5 discussing the results, its applications, and our next steps.

2 Simulation Framework

In this work, we use a process of simulation based on a combination of Mutual Information (MI) and Surrogate Time Series (STS). This simulation approach performs calculations over joint probability tables (JPDs) and generates the correspondent synthetic data.

2.1 Mutual Information

McDonnell et al. (2011) said that MI indicates the shared information between two random variables. Cover and Thomas (2006) presented it as in the following equation:

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}. \quad (1)$$

Where X and Y are two random variables; $p(x)$ and $p(y)$ are their probability distributions; and $p(x,y)$ is the joint probability distribution of those random variables.

If two random variables are independent, MI assumes its minimum, which is zero, because

$$p(x,y) = p(x)p(y). \quad (2)$$

In the opposite, if two variables are equal, MI assumes its maximum, which is the Shannon's Entropy

$$H(X) = -\sum_{x \in X} p(x) log_2 p(x). \quad (3)$$

2.2 Surrogate Time Series

The word "surrogate" stands for something that is used instead of something else. In the case of surrogate signals, the synthetic data used is randomly generated, but it also presents some characteristics of the signal that it is taking place (Dolan and Spano, 2001). A surrogate time series presents the same power spectrum as the original signal, but these two signals are uncorrelated.

One technique of generation for surrogate time series consists of making changes in the signal's phase while keeping its power spectrum and probability distribution. Such technique is the "Iterative Amplitude Adjusted Fourier Transform" (Schreiber and Schmitz, 2000; Venema et al., 2006). It consists on a phase shifting of the original signal, where the phase is substituted by a Gaussian white noise in an iterative process.

Suppose a signal x that is represented in both domains, time and frequency, by $x(t)$ and $X(j\omega)$. Also, suppose a surrogate signal x_S, that is represented in time and frequency by $x_S(t)$ and $X_S(j\omega)$. These signals present the following relations

$$\begin{cases} x(t) \leftrightarrow X(j\omega) = A(\omega)e^{j\phi(\omega)}, \\ x_S(t) \leftrightarrow X_S(j\omega) = A(\omega)e^{j\phi(GWN)}, \\ p(x_S) = p(x) \end{cases} \quad (4)$$

where $p(x)$, and $p(x_S)$ are the probability mass functions of the variables x, and x_S; the signals' power spectrum is given by $A(j\omega)$; and $e^{j\phi(\omega)}$ and $e^{j\phi(GWN)}$ stand for the signals' phase.

2.3 Simulation

Considering two random variables, X and Y, the simulation takes their probability distributions, $p(x)$ and $p(y)$, and builds a JPD, $p(x,y) = p(x)p(y)$. It proceeds, then, by generating the signals $x(t)$, $y(t)$, $x_S(t)$, and $y_S(t)$ to calculate the MI of $[x_S(t), y_S(t)]$. This will determine a threshold (MI_T) for noisy interaction, i.e., above MI_T we consider that X and Y are not independent. These proceedings are the state A of the diagram shown in Figure 1.

The second state, B, is a process that changes values inside the JPD and is repeated while the MI calculated over the JPD is smaller than MI_T. The algorithm subtracts some mass γ (a percentage of the chosen cell) from a random cell $JPD(x_{i1}, y_{j1})$ and adds it to another cell $JPD(x_{i1}, y_{j2})$. This process keeps the integral over the JPD equal to one, but in order to also keep the integral of the marginals distribution equal to one, the algorithm subtracts γ from a cell $JPD(x_{i2}, y_{j2})$ and adds it to another cell $JPD(x_{i2}, y_{j1})$.

After state B, the next step, state C, is to generate data following the JPD from state B. This is done by a random sampling algorithm based on a method described by Peebles (1993) to generate synthetic data for a specific probability distribution.

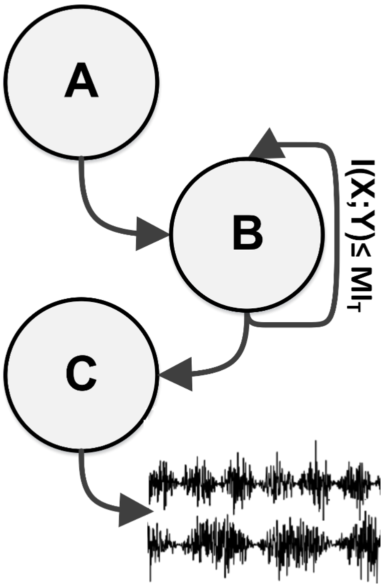

Figure 1: Schematic representation of the simulation process with three states. The first state, A, is the search for MI_T, which is given as a parameter for the JPD calculation, state B, which gives a JPD with specific mutual information to the data sampling state, C.

3 Reconstruction of the Network Structure

The structure of a network model represents the interaction of its random variables. To determine a network topology, we have to test if two nodes are dependents and some times if it is due to a conditional dependency. Here we used a theoretic information measure, the Delayed Transfer Entropy and its properties to characterize connections, find independences, causal relations, and time delays.

DTE has been widely applied to neuroscience data to determine network structures describing dependencies and independencies between regions of the brain or neurons (Wibral et al., 2014). It is based on the transfer entropy, which, given two processes X and Y, is defined as the "amount of information that a past state of X contains about a future observation of Y, given the past state of Y" (Wibral et al., 2014), and is proposed by Schreiber (2000) as:

$$TE = \sum_{\substack{\forall x \in \mathcal{X} \\ \forall y \in \mathcal{Y}}} p(y, y^k, x^l) \log \frac{p(y|y^k, x^l)}{p(y|y^k)} \qquad (5)$$

where p are probability functions for processes X and Y with alphabets $\mathcal{X}$ and $\mathcal{Y}$; $\mathbf{y}^k$ stands for k past states of Y; and $\mathbf{x}^l$ for l past states of X.

By assuming that the underlying process is Markovian (i.e. present states depend only on the immediate past states), we can fix $k = 1$ and $l = 1$. Then, we define the delay of immediate past of Y as α, which, as indicated by Kantz and Schreiber (2004), can be found in the first local minimum of the mutual information of the signal with itself. The delay of the past of X is defined as β. Thus, we can define *delayed transfer entropy* as the transfer entropy measure considering a delay β in the source process X,

$$DTE = \sum_{\substack{\forall x \in \mathcal{X} \\ \forall y \in \mathcal{Y}}} p(y, y_\alpha, x_\beta) \log \frac{p(y|y_\alpha, x_\beta)}{p(y|y_\alpha)}. \qquad (6)$$

We then obtain the time-delay between the variables in a complex system by considering the maximum peak of information transfer in a sweeping of different values of β, i.e., $\delta = \max(DTE(X \to Y))$ (Pampu et al., 2013).

4 Simulation, Measures, and Discussion

Intending to test the limits of our methods, we selected three random variables, two representing a spiking signal ($Spk1$ and $Spk2$) and one a Gaussian White Noise (GWN). We also selected the three configurations (A, B, and C) shown in Figure 2. Our intention is to generate simulated data for each case (A, B, and C), calculate the Delayed Transfer Entropy for all combinations and then reconstruct the network.

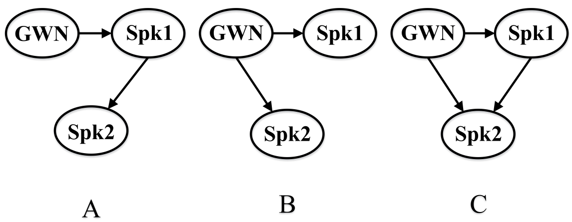

Figure 2: Setup with three tests cases (A, B, and C) for three random variables: one Gaussian White Noise (GWN), and two spiking signals ($Spk1$ and $Spk2$).

To ensure that the reconstruction process would not be influenced by the analysts, we divided the two main tasks, the data generation and its analysis, between two analysts in a double blind process. The person that generated the data followed all the proceedings from Section 2 providing the data set to the person that analyzed the simulated data. This one did not know at the time the structures from Figure 2 and inferred the topologies from DTE measures, following Section 3 proceedings.

4.1 Simulation

We defined each node with a discrete probability distribution; the GWN parameters are zero mean and unitary variance. There is also a Shannon's Entropy ($H(X)$) calculated over each variable probability distribution: $H(GWN) = 4,39$, and $H(Spk1) = H(Spk2) = 3,93$. Based on thirty MI measures of STS pairs $\{x_S; y_S\}$, the simulation calculated $MI_T = 1.0$, and we proceeded to the calculation of two JPDs: $p(GWN, Spk)$, for connections $GWN \leftrightarrow Spk1$ and $GWN \leftrightarrow Spk2$; and $p(Spk1, Spk2)$ for connection $Spk1 \leftrightarrow Spk2$.

For each structure in Figure 2 we generated a different data set with $900k$ samples per node, and also inputted artificial time delays between each pair of variables as shown in Table 1.

Table 1: Artificial Time Delays (in ms) between nodes of the structures from Figure 2.

Connection	BN A	BN B	BN C
GWN to Spk1	50	—	20
GWN to Spk2	50	70	—
Spk1 to Spk2	50	70	20

4.2 Measures

As shown by Pampu et al. (2013), calculating the delayed transfer entropy between two processes, X and Y, we will find the interaction time-delay between them in the instant that DTE assumes its maximum value. Since we assume that we are working with causal systems, i.e., the future does not affects the past, only positive delays are considered. Thence, we used DTE to test all the combinations of the possible connections, which are: $GWN \rightarrow Spk1$, $GWN \leftarrow Spk1$, $GWN \rightarrow Spk2$, $GWN \leftarrow Spk1$, $Spk1 \rightarrow Spk2$, and $Spk1 \leftarrow Spk2$.

The DTE measures calculated to test the connections $GWN \rightarrow Spk1$ and $GWN \leftarrow Spk1$ are shown in Figure 3. All configurations showed a connection between GWN and $Spk1$ with time delays equal to $20ms$. Since the tests only showed DTE peaks for connections GWN to $Spk1$, we assume the existence of a connection $GWN \rightarrow Spk1$ with a time delay of $20ms$ in all cases.

To test the connections $GWN \rightarrow Spk2$ and $GWN \leftarrow Spk2$ we calculated their DTE, which is shown in Figure 4. Cases B and C showed a connection between GWN and $Spk2$ with time delays equal to $70ms$. There is also a peak in the time $70ms$ of case A, but this specific peak is small if compared to the rest of the DTE and requires further analysis.

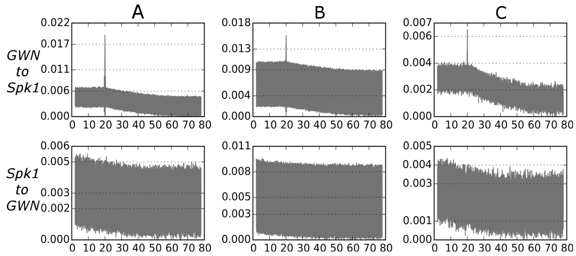

Figure 3: DTE measures testing connections, find the interaction $GWN \rightarrow Spk1$, and showing time delays of $20ms$ in all cases.

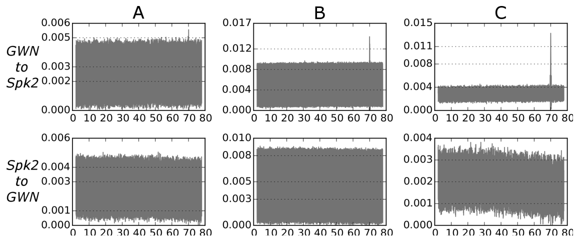

Figure 4: DTE measures testing connections, find the interaction $GWN \rightarrow Spk2$ for cases B and C with a time delay of $70ms$, and also an indirect connection for case A (small peak).

Calculating DTE between $Spk1$ and $Spk2$ in Figure 5, we tested the connections $Spk1 \rightarrow Spk2$ and $Spk1 \leftarrow Spk2$. We identified connections between $Spk1$ and $Spk2$ in cases A and C with time delays equal to $50ms$, but nothing was found in case B.

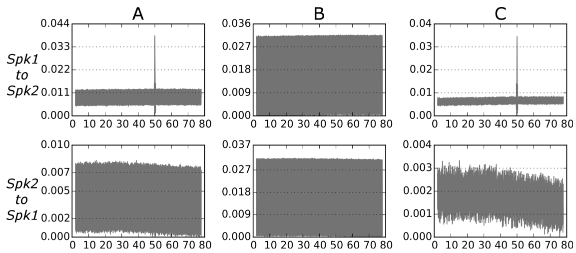

Figure 5: DTE measures testing connections, find the interaction $Spk1 \rightarrow Spk2$ in the cases A and C with a time delay of $50ms$.

4.3 Discussion

Comparing the connections found in DTE measures with the structures from Figure 2, it is possible to reconstruct the original topologies. Since the DTE peaks found for test cases B and C matches with the connections presented in the original structure, there was no doubt in the reconstruction them. However, in the test case A, DTE measures presented in Figure 4 suggested the existence of a link between GWN and $Spk2$ with a small peak.

We can compare the peak that suggests a connection $GWN \rightarrow Spk2$ with the maximum value of the noise floor (the rest of the DTE). This proceeding shows that there is not so much difference between them. Table 2 summarizes the results from Figures 3, 4, and 5. It shows the percentage difference between peaks and the maximum ground noise for each particular case.

Table 2: Percentage difference between the DTE peak and the maximum of its ground noise.

Connection	BN A	BN B	BN C
GWN to $Spk1$	17,30%	5,51%	3,75%
$Spk1$ to GWN	-	-	-
GWN to $Spk2$	0,84%	5,48%	10,00%
$Spk2$ to GWN	-	-	-
$Spk1$ to $Spk2$	14,62%	-	9,79%
$Spk2$ to $Spk1$	-	-	-

Although the peak of case A connection $GWN \rightarrow Spk2$ is close in value to the ground noise (small value in Table 2), we consider it as an indirect influence of the GWN to $Spk2$. This consideration is made because, in the case A from Figure 2, the node $Spk2$ is directly linked to the node $Spk1$, which by its turn is directly linked to the node GWN. Here we have a case of D-Separation when two variables are independent if the state of an intermediate variable is known (see Hayduk et al. (2003) for further references).

5 Conclusion

The purpose of this paper is to develop, test, and establish a method to reconstruct structural interactions in non-simultaneous recordings. We used simulations and information theoretic measures, such as Delayed Transfer Entropy and Mutual Information.

In a double-blind process of simulation and data analysis, we generated data for three test cases, each one with a different topology. Further, we estimated the structure based on the connections found using DTE measures. In two of those cases (Figure 2, cases B and C), we reconstructed the network structure only by visually identifying a peak in the DTE plot (Figures 3, 4, and 5). In a third case, we could observe an indirect influence, which was already expected due to the data processing inequality (Cover and Thomas, 2006).

It was already expected that DTE measures would point to a connection between the signals GWN and $Spk2$, because Figure's 2 test case A presents an structure with two variables d-separated (Hayduk et al., 2003) by a third. The percentage differences between peaks and maximum ground noises presented in Table 2 reinforce these expectations.

Table 2 shows that the peak found for $GWN \rightarrow Spk2$ is less relevant when compared to the others. As a result, we interpreted the low peak for connection GWN to $Spk2$ case A of Figure 4 as a reflex of an indirect dependency. This means that we recovered the same topology presented in Figure 2 case A.

In future works, we want to extend these methods using them to analyze non-simultaneous neuronal recordings in order to investigate interactions. To represent the neurons dynamics, we intend to train a Dynamic Bayesian Network (Meyer-Baese and Schmid, 2014). The discussion about DTE peaks and if they represent a directed or indirected connection is considered as a challenge that we are expecting to find in the neuronal analysis. However, we could see in this work that indirect connections present a low difference when their DTE peaks are compared with their ground.

Acknowledgment

The authors would like to thank CAPES - Brazilian Federal Agency for Support and Evaluation of Graduate Education within the Ministry of Education of Brazil, São Paulo Research Foundation FAPESP (grant 2012/24272-7), and also thank CNPq - National Counsel of Technological and Scientific Development - Project 475064/2013-5 - for supporting this work.

Moreover, we would like to thank CBA, specially the reviewer whom presented high quality comments and suggestions.

References

Brown, E. N., Kass, R. E. and Mitra, P. P. (2004). Multiple neural spike train data analysis: state-of-the-art and future challenges, *Nature neuroscience* **7**(5): 456–461.

Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems, *Nature Reviews Neuroscience* **10**(3): 186–198.

Burrows, M. (1996). *The neurobiology of an insect brain*, Oxford University Press.

Buschmann, T., Ewald, A., von Twickel, A. and Bueschges, A. (2015). Controlling legs for locomotion-insights from robotics and neurobiology, *Bioinspiration & Biomimetics* **10**(4).

Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*, Wiley-Interscience.

Dewhirst, O. P., Angarita-Jaimes, N., Simpson, D. M., Allen, R. and Newland, P. L. (2013). A system identification analysis of neural adaptation dynamics and nonlinear responses in the local reflex control of locust hind limbs, *Journal of computational neuroscience* **34**(1): 39–58.

Dolan, K. T. and Spano, M. L. (2001). Surrogate for nonlinear time series analysis, *Phys. Rev. E* **64**: 046128.

Endo, W., Santos, F. P., Simpson, D., Maciel, C. D. and Newland, P. L. (2015). Delayed mutual information infers patterns of synaptic connectivity in a proprioceptive neural network, *Journal of Computational Neuroscience* **38**(2): 427–438.

Faes, L. and Porta, A. (2014). Conditional entropy-based evaluation of information dynamics in physiological systems, *in* M. Wibral, R. Vicente and J. T. Lizier (eds), *Directed Information Measures in Neuroscience*, Understanding Complex Systems, Springer Berlin Heidelberg, pp. 61–86.

Hayduk, L., Cummings, G., Stratkotter, R., Nimmo, M., Grygoryev, K., Dosman, D., Gillespie, M., Pazderka-Robinson, H. and Boadu, K. (2003). Pearl's d-separation: One more step into causal thinking, *Structural Equation Modeling: A Multidisciplinary Journal* **10**(2): 289–311.

Kantz, H. and Schreiber, T. (2004). *Nonlinear Time Series Analysis*, Cambridge nonlinear science series, Cambridge University Press.

Maciel, C. D., Simpson, D. M. and Newland, P. L. (2012). Inference about multiple pathways in motor control limb in locust., *in* S. V. Huffel, C. M. B. A. Correia, A. L. N. Fred and H. Gamboa (eds), *Biosignals and Biorobotics Conference (BRC)*, SciTePress, pp. 69–75.

Marmarelis, V. Z. (2004). *Nonlinear dynamic modeling of physiological systems*, Vol. 10, John Wiley & Sons.

McDonnell, M., Ikeda, S. and Manton, J. (2011). An introductory review of information theory in the context of computational neuroscience, *Biological Cybernetics* **105**(1): 55–70.

Meruelo, A. C., Simpson, D. M., Veres, S. M. and Newland, P. L. (2016). Improved system identification using artificial neural networks and analysis of individual differences in responses of an identified neuron, *Neural Networks* **75**: 56 – 65.

Meyer-Baese, A. and Schmid, V. J. (2014). *Pattern Recognition and Signal Analysis in Medical Imaging*, Elsevier.

Pampu, N., Vicente, R., Muresan, R., Priesemann, V., Siebenhuhner, F. and Wibral, M. (2013). Transfer entropy as a tool for reconstructing interaction delays in neural signals, *Signals, Circuits and Systems (ISSCS), 2013 International Symposium on*, pp. 1–4.

Peebles, P. Z. (1993). *Probability, Random Variables, and Rabdom Signal Principles*, 3rd edition edn, McGraw-Hill Inc.

Ritzmann, R. E. and Büschges, A. (2007). 10 insect walking: From reduced preparations to natural terrain, *Cold Spring Harbor Monograph Archive* **49**: 229–250.

Runge, J., Heitzig, J., Petoukhov, V. and Kurths, J. (2012). Escaping the curse of dimensionality in estimating multivariate transfer entropy, *Phys. Rev. Lett.* **108**: 258701.

Schreiber, T. (2000). Measuring information transfer, *Physical review letters* **85**(2): 461.

Schreiber, T. and Schmitz, A. (2000). Surrogate time series, *Physica D: Nonlinear Phenomena* **142**(3): 346–382.

Szczecinski, N. S., Brown, A. E., Bender, J. A., Quinn, R. D. and Ritzmann, R. E. (2014). A neuromechanical simulation of insect walking and transition to turning of the cockroach blaberus discoidalis, *Biological Cybernetics* **108**(1): 1–21.

Venema, V., Ament, F. and Simmer, C. (2006). A stochastic iterative amplitude adjusted fourier transform algorithm with improved accuracy, *Nonlinear Processes in Geophysics* **13**(3): 321–328.

Wessnitzer, J. and Webb, B. (2006). Multimodal sensory integration in insects–towards insect brain control architectures, *Bioinspiration & biomimetics* **1**(3): 63.

Wibral, M., Vicente, R. and Lizier, J. T. (2014). *Directed information measures in neuroscience*, Springer.