Supplementary Information

for

"The Epidemic Spreading Model and the Direction of Information Flow in Brain Networks"

J. Meier^{*}, X. Zhou^{*}, A. Hillebrand[†], P. Tewarie[‡], C.J. Stam[§] and P. Van Mieghem^{*}

^{*}Email: j.m.meier@tudelft.nl (corresponding author), x.zhou@student.tudelft.nl, p.f.a.vanmieghem@tudelft.nl, Faculty of Electri-

cal Engineering, Mathematics and Computer Science, Delft University of Technology, P.O Box 5031, 2600 GA Delft, The Netherlands. [†]*Email:* a.hillebrand@vumc.nl, Department of Clinical Neurophysiology and Magnetoencephalography Center, VU University Medical Centre, Amsterdam, The Netherlands.

^{*}*Email*: prejaas.tewarie@nottingham.ac.uk, Department of Neurology, VU University Medical Center, Amsterdam, The Netherlands, and Sir Peter Mansfield Imaging Centre, School of Physics and Astronomy, University of Nottingham, University Park, Nottingham, The United Kingdom.

[§]Email: cj.stam@vumc.nl, Department of Clinical Neurophysiology, VU University Medical Center, Amsterdam, The Netherlands.

A Additional Figures



Figure A.1: **Die-out probability of activity spreading.** Approximate formula for the die-out probability $1/x^n$ (Liu and Van Mieghem, 2016a) depending on the number of initially infected nodes *n* and $x = \tau \cdot \lambda_1$ the product between the effective infection rate $\tau = 0.2$ and the largest eigenvalue of the structural adjacency matrix *A*, $\lambda_1 = 10.47$, in our case.



Figure A.2: Fraction of activated nodes over whole simulation time. Mean fraction of activated nodes for every time step averaged over the 100 simulation runs and the whole simulation time for the applied effective infection rate $\tau = 0.2$. The smaller plot shows the first 300 time units. In both figures, the horizontal black line represents the mean fraction of activated nodes in the metastable state.



Figure A.3: **Global pattern in surrogate data.** For different time delays, fraction of surrogate data possessing a smaller value than the observed posterior-anterior (PA) value and the observed correlation between degree and averaged directed transfer entropy (dTE) value, respectively. When the fraction is above 0.95, the observed value is significantly larger based on the 5% significance level. For time delays where the fraction is below the 0.05 line, we observe a significantly smaller value than the surrogate data (P value < 0.05).

B Randomly Reshuffled Matrices



Figure B.1: **Global pattern in randomly reshuffled networks.** Posterior-anterior (PA) value over different time delays. The blue line represents the already known result for the structural adjacency matrix *A* and the other lines indicate the PA values for the randomly reshuffled matrices.



Figure B.2: **Correlation between degree and dTE in randomly reshuffled networks.** Correlation between degree and averaged directed transfer entropy (dTE) over different time delays. The blue line represents the already known result for the structural adjacency matrix *A* and the other lines indicate the correlation values for the randomly reshuffled matrices.



Figure B.3: **Comparison between reshuffled networks and structural network.** For different time delays, fraction of reshuffled networks possessing a smaller value than the observed posterior-anterior (PA) value (in blue) and the observed correlation between degree and averaged directed transfer entropy (dTE) value (in red), respectively. For small time delays, the behavior of the structural adjacency matrix *A* is significantly different than the behavior of the reshuffled versions. For larger time delays, we observe a significantly larger correlation in the structural network (between degree and dTE) and also a quite large PA value with respect to the reshuffled matrices.



Figure B.4: **Significant correlation for reshuffled matrices.** For different time delays, fraction of surrogate data possessing a smaller value than the observed PA value and the observed correlation between degree and averaged directed transfer entropy (dTE) value, respectively. When the fraction is above 0.95, the observed value is significantly larger based on the 5% significance level. Up to a certain time delay, all the randomly reshuffled networks exert a significantly positive correlation between degree and directed transfer entropy towards the rest of the network.

C Directed Structural Network of the Macaque Brain

We repeated the analysis on the directed structural networks of the macaque with N = 47 nodes, which has earlier been analyzed by (Honey et al., 2007). The SIS model can easily be extended towards a directed underlying network (Li et al., 2013). Applying the same effective infection rate $\tau = 0.2$ as for the human brain, we ran 10 simulation runs on the macaque neocortex. In a directed network, we can calculate different forms of the nodal degree, the in- and the out-degree but also the sum of them both. We plotted the correlation between the different forms of the degree and the directed transfer entropy for different time delays in Appendix Figure C.1. Overall, we observe a positive correlation for the different degrees with the average directed transfer entropy value. However, regarding different time scales, we can see some fluctuations with regard to in- and out-degree. For small time scales, a high out-degree reaches the maximum positive correlation. With regard to longer time scales, the in- and out-degree differences seem to disappear. This result is in line with intuition since for a higher number of (direct) outgoing links one would expect a more sending property of a node (for short time delays). However, the general notion of hubs cannot be extended in a straightforward manner. For the macaque network, nodes with a high sum of in- and out-degrees can have much more incoming than outgoing links or the other way around, which should result in different sending/receiving properties. The different behavior of those nodes is shown by the correlation between the difference of in- and out-degree and the dTE value (purple dots in Figure C.1). Nodes with a much higher number of incoming than outgoing connections, which can still be hubs related to the overall degree (see Figure C.4), seem to be more receiving with regard to short time delays but more sending when analyzing longer time delays. We also plotted the out-degree and the difference between in- and out-degree against the dTE values for specific time delays (see Figures C.2 and C.3) reaching the maximum and minimum correlation values, respectively.



Figure C.1: **Correlation for macaque brain.** Correlation between different forms of the degree and averaged directed transfer entropy (dTE) over different time delays. For the in-degree, the out-degree and the sum of in- and out-degree the positive correlation values are significant with regard to randomly reshuffled versions of the dTE values. For the difference between in- and out-degree, the three marked points are the only significant correlation values.



Figure C.2: Maximum correlation for macaque brain. The out-degree for each node in the macaque structural brain network against the directed transfer entropy (dTE) for the time delay h = 2 reaching the maximum correlation value.



Figure C.3: Minimum correlation for macaque brain. The difference between in- and out-degree for each node in the macaque structural brain network against the directed transfer entropy (dTE) for the time delay h = 0.5 reaching the minimum correlation value.



Figure C.4: **Degrees in the macaque brain.** Sum of in- and out-degree for each node in the macaque structural brain networks versus the difference between in- and out-degree. Some high-degree nodes have big differences in their number of incoming and outgoing links.

D Correlation versus Transfer Entropy

The transfer entropy is equal to the conditional mutual information (MI)

$$\begin{split} TE_{i \to j}(h) &= MI(X_j(t+h); X_i(t) | X_j(t)) \\ &= \sum_{k,l,m \in \{0,1\}} \Pr[X_j(t+h) = k, X_j(t) = l, X_i(t) = m] \cdot \log\left(\frac{\Pr[X_j(t+h) = k | X_j(t) = l, X_i(t) = m]}{\Pr[X_j(t+h) = k | X_j(t) = l]}\right) \\ &= \sum_{k,l,m \in \{0,1\}} \Pr[X_j(t+h) = k, X_j(t) = l, X_i(t) = m] \cdot \log\left(\frac{\Pr[X_j(t+h) = k, X_j(t) = l, X_i(t) = m]}{\Pr[X_j(t+h) = k, X_j(t) = l]}\right) \\ \end{split}$$

where we applied the law of Bayes for the last equality.

The mutual information and the measure of correlation want to measure the same underlying property of two random variables, their 'distance to independence'. The covariance is defined as

$$cov(i, j, h) = E[X_i(t)X_j(t+h)] - E[X_i(t)]E[X_j(t+h)]$$

and measures the distance in terms of expected values of the random variables itself whereas the mutual information can be written as

$$MI(X_{j}(t+h); X_{i}(t)) = E[\log(Pr[X_{i}(t), X_{j}(t+h)])] - E[\log(Pr[X_{i}(t)]Pr[X_{j}(t+h)])]$$

and measures the distance in terms of the expected value of the logarithm of their probabilities.

For the transfer entropy, we can apply the chain rule of the mutual information and obtain

$$TE_{i \to j}(h) = MI(X_j(t+h); X_i(t)|X_j(t))$$

$$= MI(X_j(t+h); X_i(t), X_j(t)) - MI(X_j(t+h), X_j(t))$$

$$= \sum_{k,l,m=\{0,1\}} \Pr[X_j(t+h) = k, X_j(t) = l, X_i(t) = m] \cdot \log\left(\frac{\Pr[X_j(t+h) = k, X_j(t) = l, X_i(t) = m]}{\Pr[X_j(t+h)]\Pr[X_i(t) = k, X_j(t) = l]}\right)$$

$$- \sum_{k,l=\{0,1\}} \Pr[X_j(t+h) = k, X_j(t) = l] \cdot \log\left(\frac{\Pr[X_j(t+h) = k, X_j(t) = l]}{\Pr[X_j(t+h)]\Pr[X_j(t) = l]}\right)$$
(D.1)

For the second term we followed the derivations in (Li, 1990) and used that our activation series are binary resulting in an approximative formula

$$MI(X_{j}(t+h), X_{j}(t)) \approx \frac{1}{2} \left(\frac{auto_{j}}{Pr[X_{j}(t) = 1](1 - Pr[X_{j}(t) = 1])} \right)^{2},$$

where *auto_j* denotes the auto-correlation of *j*. To reach this result, we did assume that $Pr[X_j(t) = 1] \approx Pr[X_j(t + h) = 1]$ which could be confirmed by our simulations for small values of the time lag *h*. Thus, the second term of (D.1) can be interpreted as some correction for the auto-correlation that is included in the transfer entropy.

If we apply the Kirkwood superposition approximation to the first term of (D.1) which involves all three entities, we can approximate the joint probability of the three terms

$$MI(X_{j}(t+h); X_{i}(t), X_{j}(t)) = \sum_{k,l,m \in \{0,1\}} \Pr[X_{j}(t+h) = k, X_{j}(t) = l, X_{i}(t) = m] \cdot \log\left(\frac{\Pr[X_{j}(t+h) = k, X_{j}(t) = l]\Pr[X_{j}(t+h) = k, X_{i}(t) = l]}{\Pr[X_{j}(t+h)]^{2}\Pr[X_{i}(t) = k]\Pr[X_{j}(t) = k]}\right)$$

where e.g. in the case of k = l = m we obtain the element of the sum in the logarithm as

$$\log\left(\frac{\left(auto_j + \Pr[X_j(t) = k]^2\right)}{\Pr[X_j(t) = k]^2} \cdot \frac{\left(corr_{del}(i, j, h) + \Pr[X_j(t) = k]\Pr[X_i(t) = k]\right)}{\Pr[X_j(t) = k]\Pr[X_i(t) = k]}\right)$$

where $corr_{del}(i, j, h)$ is the delayed correlation function between node *i* and node *j* which will be further studied in the next section. For the other elements of the sum we can derive similar results in the logarithm reducing the expression to a combination of the auto-correlation of *j* and the delayed correlation between *i* and *j*. For the three-way joint probability in front of the logarithm, we can again use the Kirkwood superposition approximation and obtain e.g. for the element k = l = m

$$\begin{aligned} &\Pr[X_j(t+h) = 1, X_j(t) = 1, X_i(t) = 1] \\ &\approx \frac{\Pr[X_j(t+h) = 1, X_j(t) = 1]\Pr[X_j(t+h) = 1, X_i(t) = 1]\Pr[X_j(t) = 1, X_i(t) = 1]}{\Pr[X_j(t) = 1]\Pr[X_i(t) = 1]\Pr[X_j(t) = 1]} \\ &= \frac{(auto_j + \Pr[X_j(t) = 1]^2)(corr_{del}(i, j, h) + \Pr[X_j(t) = 1]\Pr[X_i(t) = 1])(corr(i, j) + \Pr[X_j(t) = 1]\Pr[X_i(t) = 1])}{\Pr[X_j(t) = 1]\Pr[X_i(t) = 1]\Pr[X_j(t) = 1]} \end{aligned}$$

where corr(i, j) is the correlation (or covariance) of the two nodes' binary time series. For the other elements of the sum, the derivation can be conducted similarly. To sum up, we have demonstrated that the transfer entropy from node *i* to node *j* can be expressed as a combination of the (delayed) correlation between *i* and *j* corrected for the auto-correlation of *j*. In the following section, we will further examine analytically the (delayed) correlation and auto-correlation as 'building blocks' of the transfer entropy.

E The Covariance $\widetilde{\rho}(X_i(t), X_j(t+h))$ for a Small Time Lag h

The functional connectivity between two nodes *i* and *j* is defined as the correlation of their activation series

$$\rho(X_i(t), X_j(t)) = \frac{E[X_i(t)X_j(t)] - E[X_i(t)]E[X_j(t)]}{\sqrt{Var[X_i(t)]}\sqrt{Var[X_i(t)]}}$$
(E.1)

over the whole simulation time (Stam et al., 2016).

The numerator is also referred to as the covariance between node *i* and *j*. We compute the expectation $E[X_i(t)X_j(t+h)]$ for a very small time h > 0. Although the derivative of a Bernoulli random variable does not exist, we follow the framework in (Van Mieghem, 2014) and we agree to *formally* define the derivative by the random variable equation

$$\frac{dX_{j}(t)}{dt} = -\delta X_{j}(t) + (1 - X_{j}(t))\beta \sum_{k=1}^{N} a_{kj}X_{k}(t)$$
(E.2)

For small h, the first order expansion of the Taylor series yields

$$X_{j}(t+h) = X_{j}(t) + h \frac{dX_{j}(t)}{dt} + o(h)$$
(E.3)

E.1 Deductions

With the definition (E.1) of the covariance, we find for $i \neq j$ that the *j*-delayed covariance satisfies

$$\widetilde{\rho}\left(X_{i}\left(t\right), X_{j}\left(t+h\right)\right) = (1-\delta h)\widetilde{\rho}\left(X_{i}\left(t\right), X_{j}\left(t\right)\right) + \beta h \sum_{k=1}^{N} a_{kj}\widetilde{\rho}\left(X_{i}\left(t\right), X_{k}\left(t\right)\right) -\beta h \sum_{k=1}^{N} a_{kj}\left\{E\left[X_{i}\left(t\right)X_{j}\left(t\right)X_{k}\left(t\right)\right] - E\left[X_{i}\left(t\right)\right]E\left[X_{j}\left(t\right)X_{k}\left(t\right)\right]\right\} + o\left(h\right)$$
(E.4)

In general, the *j*-delayed covariance $\tilde{\rho}(X_i(t), X_j(t+h))$ is different from the *i*-delayed covariance $\tilde{\rho}(X_j(t), X_i(t+h))$. Indeed, (E.4) demonstrates that

$$\widetilde{\rho}\left(X_{i}\left(t\right), X_{j}\left(t+h\right)\right) - \widetilde{\rho}\left(X_{j}\left(t\right), X_{i}\left(t+h\right)\right) = \beta h \sum_{k=1}^{N} \left(a_{kj}\widetilde{\rho}(X_{i}(t), X_{k}(t)) - a_{ki}\widetilde{\rho}(X_{j}(t), X_{k}(t))\right)$$
$$-\beta h \sum_{k=1}^{N} \left(a_{kj}E[X_{i}(t)X_{j}(t)X_{k}(t)] - \beta h \sum_{k=1}^{N} \left(a_{kj}E[X_{i}(t)]E[X_{j}(t)X_{k}(t)] - a_{ki}E[X_{j}(t)]E[X_{i}(t)X_{k}(t)]\right)$$

Starting from the original definition of $\widetilde{\rho}(X_i(t), X_j(t+h))$, we can deduce the parts as

$$E\left[X_{i}(t)X_{j}(t+h)\right] = E\left[X_{i}(t)X_{j}(t)\right] + hE\left[-\delta X_{j}(t)X_{i}(t) + (1-X_{j}(t))X_{i}(t)\beta\sum_{k=1}^{N}a_{kj}X_{k}(t)\right] + o(h)$$
$$E\left[X_{j}(t)X_{i}(t+h)\right] = E\left[X_{j}(t)X_{i}(t)\right] + hE\left[-\delta X_{i}(t)X_{j}(t) + (1-X_{i}(t))X_{j}(t)\beta\sum_{k=1}^{N}a_{ki}X_{k}(t)\right] + o(h)$$

Subtraction yields

$$T = E \left[X_{i}(t) X_{j}(t+h) \right] - E \left[X_{j}(t) X_{i}(t+h) \right]$$

$$= \beta h E \left[X_{i}(t) \left(1 - X_{j}(t) \right) \sum_{k=1}^{N} a_{kj} X_{k}(t) - X_{j}(t) \left(1 - X_{i}(t) \right) \sum_{k=1}^{N} a_{ki} X_{k}(t) \right]$$

$$= \beta h E \left[X_{i}(t) \sum_{k=1}^{N} a_{kj} X_{k}(t) - X_{j}(t) \sum_{k=1}^{N} a_{ki} X_{k}(t) + X_{i}(t) X_{j}(t) \sum_{k=1}^{N} \left(a_{ki} - a_{kj} \right) X_{k}(t) \right]$$

where $\sum_{k=1}^{N} a_{kj} X_k(t)$ are the infected neighbors of node *j*. The first equation tells that *T* is the balance of two cases: either an infection at node *i* and all infected neighbors of node *j* try to infect node *j* or an infection at node *j* and all infected neighbors of node *i*.

The other parts of the difference of covariances can be written as

$$\begin{split} E[X_{j}(t)]E[X_{i}(t+h)] - E[X_{i}(t)]E[X_{j}(t+h)] &= \beta h E[X_{j}(t)] \sum_{k=1}^{N} a_{ki} E[X_{k}(t)] - \beta h E[X_{i}(t)] \sum_{k=1}^{N} a_{kj} E[X_{k}(t)] \\ &-\beta h \sum_{k=1}^{N} a_{ki} E[X_{i}(t)X_{k}(t)]E[X_{j}(t)] + \beta h \sum_{k=1}^{N} a_{kj} E[X_{j}(t)X_{k}(t)]E[X_{i}(t)] \end{split}$$

Together we get for the probability flux

$$\widetilde{\rho}\left(X_{i}\left(t\right), X_{j}\left(t+h\right)\right) - \widetilde{\rho}\left(X_{j}\left(t\right), X_{i}\left(t+h\right)\right) = \beta h \sum_{k=1}^{N} a_{kj}\left(E[X_{i}(t)]E[X_{j}(t)X_{k}(t)] - E[X_{i}(t)X_{j}(t)X_{k}(t)]\right)$$
$$-\beta h \sum_{k=1}^{N} a_{ki}\left(E[X_{j}(t)]E[X_{i}(t)X_{k}(t)] - E[X_{i}(t)X_{j}(t)X_{k}(t)]\right)$$

This information flow has also been researched by (Hillebrand et al., 2016), where instead of the delayed correlation, (Hillebrand et al., 2016) used the *Phase Transfer Entropy* as a measure of causality. This difference in coavariance is positive if and only if the fraction from (Hillebrand et al., 2016) is larger than 0.5.

E.2 Just above the Epidemic Threshold

Just above the epidemic threshold (Van Mieghem et al., 2009), the probability of infection is $E[X_i(t)] = \epsilon(x_1)_i$, where $\epsilon > 0$ is small and where x_1 is the principal eigenvector of the adjacency matrix A belonging to the largest eigenvalue λ_1 .

Assuming that the effective infection rate $\tau = \frac{\beta}{\delta} = \tau_c + \epsilon$, then we may discard the last sum with triple expectations (of order $O(\epsilon^2)$) in (E.4) so that

$$\widetilde{\rho}\left(X_{i}\left(t\right),X_{j}\left(t+h\right)\right)\approx\left(1-\delta h\right)\widetilde{\rho}\left(X_{i}\left(t\right),X_{j}\left(t\right)\right)+\beta h\sum_{k=1}^{N}a_{kj}\widetilde{\rho}\left(X_{i}\left(t\right),X_{k}\left(t\right)\right)+o\left(h\right)$$

Since $\tilde{\rho}(X_i(t), X_j(t)) \ge 0$ and assuming that $\tilde{\rho}(X_i(t), X_k(t))$ is of about the same magnitude as $\tilde{\rho}(X_i(t), X_j(t))$ for any node *k* that is a neighbor of *j* (this is possible because node *k* as a neighbor of *j* can only be one hop further away from or nearer to *i* than *j*. If *i* and *j* are directly connected and *k* is a common neighbor, then this approximation is even more accurate), then

$$\widetilde{\rho}\left(X_{i}\left(t\right),X_{j}\left(t+h\right)\right)\approx\left(1+\left(\tau d_{j}-1\right)\delta h\right)\widetilde{\rho}\left(X_{i}\left(t\right),X_{j}\left(t\right)\right)+o\left(h\right)$$

and

$$\frac{\rho\left(X_{i}\left(t\right),X_{j}\left(t+h\right)\right)}{\rho\left(X_{i}\left(t\right),X_{j}\left(t\right)\right)}\approx\left(1+\left(\tau d_{j}-1\right)\delta h\right)\frac{\sqrt{\operatorname{Var}\left[X_{j}\left(t\right)\right]}}{\sqrt{\operatorname{Var}\left[X_{j}\left(t+h\right)\right]}}+o\left(h\right)$$

Assuming $\operatorname{Var} \left[X_j(t+h) \right] \approx \operatorname{Var} \left[X_j(t) \right]$ (which is reasonable for small *h*) the fraction on the right hand side can be approximated by 1 and we obtain

$$\frac{\rho\left(X_{i}\left(t\right), X_{j}\left(t+h\right)\right)}{\rho\left(X_{i}\left(t\right), X_{j}\left(t\right)\right)} \approx 1 + \delta h\left(\tau d_{j}-1\right) + o\left(h\right)$$

Finally, since $\tau > \tau_c \ge \frac{1}{\lambda_1}$, we have that $\tau d_j - 1 > \frac{d_j}{\lambda_1} - 1$. Since the spectral radius is bounded (Van Mieghem, 2011) by max $(d_{av}, \sqrt{d_{max}}) \le \lambda_1 \le d_{max}$, where the average degree $d_{av} = \frac{2L}{N}$, the factor $\tau^* d_j - 1$ is positive for a node *j* with more than average degree, but possibly negative for a node *j* with low degree.

References

preprint arXiv:1402.1731.

- Hillebrand, A., Tewarie, P., van Dellen, E., Yu, M., Carbo, E. W., Douw, L., Gouw, A. A., van Straaten, E. C., and Stam, C. J. (2016). Direction of information flow in large-scale resting-state networks is frequency-dependent. *Proceedings of the National Academy* of Sciences, 113(14):3867–3872.
- Honey, C. J., Kötter, R., Breakspear, M., and Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences*, 104(24):10240–10245.
- Li, C., Wang, H., and Van Mieghem, P. (2013). Epidemic threshold in directed networks. *Physical Review E*, 88(6):062802.
- Li, W. (1990). Mutual information functions versus correlation functions. Journal of Statistical Physics, 60(5-6):823-837.
- Stam, C. J., Hillebrand, A., van Dellen, E., Meier, J., Tewarie, P., van Straaten, E., and Van Mieghem, P. (2016). The relation between structural and functional connectivity patterns in complex brain networks. *International Journal of Psychophysiology*, 103:149–60.
 Van Mieghem, P. (2011). *Graph Spectra for Complex Networks*. Cambridge University Press.
- Van Mieghem, P. (2014). Exact markovian SIR and SIS epidemics on networks and an upper bound for the epidemic threshold. arXiv
- Van Mieghem, P., Omic, J., and Kooij, R. (2009). Virus spread in networks. IEEE/ACM Transactions on Networking, 17(1):1-14.