Partial Mutual Information Analysis of Financial Networks

P. Fiedor

Cracow University of Economics, Rakowicka 27, 31-510 Kraków, Poland

The econophysics approach to socio-economic systems is based on the assumption of their complexity. Such assumption inevitably leads to another assumption, namely that underlying interconnections within socio-economic systems, particularly financial markets, are nonlinear, which is shown to be true even in mainstream economic literature. Thus it is surprising to see that network analysis of financial markets is based on linear correlation and its derivatives. An analysis based on partial correlation is of particular interest as it leads to the vicinity of causality detection in time series analysis. In this paper we generalise the planar maximally filtered graphs and partial correlation planar graphs to incorporate nonlinearity using partial mutual information.

1. Introduction

The field of econophysics is perhaps best known for its treatment of financial markets as complex systems. In this treatment network theory plays an important role. Most commonly a correlation-based network structure is created, which quantifies the interrelations between financial instruments on a given market. Such analysis uncovers basic structure of the studied market but can also be useful for practical applications such as portfolio optimisation [1]. There have been numerous studies looking into stock markets in daily [2–7] and intraday [8–10] scales, as well as market indices [11–18] and foreign exchange markets [19]. The well-corroborated results show that markets are structured according to sectors and sub-sectors of economic activities for stock markets and geographical locations for market indices and foreign exchange markets. Thus one can technically predict which sector does a financial instrument belong to, which is practically useless however. We see that there are limitations to this approach. One limitation is related to the above-mentioned limited usefulness. Much more interesting would be an analysis of causal relationships within the markets. This is done recently using either lead-lag effect (asymmetric correlations) or partial correlations. The second limitation is connected with the researchers' insistence on using strictly linear measure of closeness between studied elements of the market, even though it goes against the assumption of complexity of those systems [20, 21] and the solid evidence of nonlinearity on financial markets with regards to stock returns [22–26], market index returns [27–31], and currency exchange rate changes [32–35]. Recently we have been studying the financial networks using information-theoretic approach to account for nonlinearity [36, 37].

In this paper we are trying to address those issues together by using network analysis of financial markets based on partial mutual information. Partial mutual information is a generalisation of partial correlations, which is sensitive to nonlinear dependencies, to which Pearson's correlation and partial correlation are strictly not sensitive. Using partial mutual information allows us to either refine the classical structural analysis of the market by adding nonlinearity and controlling for mediating influence of third instruments. But partial mutual information may also be used to bring the analysis closer to market dynamics and causal relationships, similarly to the analysis performed with partial correlation in [38, 39].

This paper is structured as follows: in Sect. 2 we present the proposed methodology. In Sect. 3 we show the results obtained for NYSE and briefly discuss them. In Sect. 4 we conclude our study and propose further research.

2. Methods

Our analysis is based on time series describing stock log returns, which is the standard way for analysing price movements. Thus data points are the log ratios between consecutive daily closing prices (this can be done at any other scale) [20]:

$$r_t = \ln(p_t/p_{t-1}). \tag{1}$$

For the purpose of estimating mutual information we need to discretize those data points. Thus we transform the data points into 4 distinct states. The states represent 4 quartiles, for the discussion of the significance and robustness of this step see [40, 41].

On this basis we are estimating the partial mutual information (PMI), which is based on Shannon's entropy. For a discrete random variable X with probabilities $p(x)$ of outcomes $\{x\}$, Shannon's entropy is defined as [42]:

$$H(X) = -\sum_x p(x) \ln p(x). \tag{2}$$

Mutual information (MI) of two random variables X and Y is given by

$$I(X,Y) = H(X) + H(Y) - H(X,Y), \tag{3}$$

where $H(X,Y)$ is obtained from the joint distribution of (X,Y). MI is symmetric, and

$$0 \leq I(X,Y) \leq \min\{H(X), H(Y)\}. \tag{4}$$

Mutual information measures information shared between two variables, thus both linear and nonlinear relationships between studied financial instruments are measured this way.

Then partial mutual information $I(X, Y|Z)$ denotes the part of mutual information $I(X, Y)$ that is not in Z and is defined as:

$$I(X, Y|Z) = H(X, Z) + H(Y, Z)$$
$$-H(Z) - H(X, Y, Z). \tag{5}$$

PMI is symmetric so that $I(X, Y|Z) = I(Y, X|Z)$ and $0 \leq I(X, Y|Z)$. MI and PMI are only equal to 0 when X and Y are strictly independent.

To estimate PMI we need an estimator of Shannon's entropy. There is an abundance of estimators [43–48], in this study we use the Schurmann–Grassberger estimate of the entropy of a Dirichlet probability distribution, which is thought to be the best choice outside very specific conditions (particularly small samples) [49]. The Schurmann–Grassberger estimator is a Bayesian parametric procedure which assumes samples distributed following a Dirichlet distribution:

$$\hat{H}(X) = \frac{1}{m + |\chi|N} \sum_{x \in \chi} (\#(x) + N)(\psi(m + |\chi|N + 1)$$
$$-\psi(\#(x) + N + 1)), \tag{6}$$

where $\#(x)$ is the number of data points having value x, $|\chi|$ is the number of bins from the discretization step, m is the sample size, and $\psi(z) = \mathrm{d}\ln\Gamma(z)/\mathrm{d}z$ is the digamma function. The Schurmann–Grassberger estimator assumes $N = 1/|\chi|$ as the prior [50].

For comparison we have also used Pearson's correlation and partial correlation, and networks based on it, as defined in [39].

We now turn to the networks we are creating based on partial mutual information. First we refine the structural view of the market offered by the networks based on correlation. It is refined first by swapping correlation with mutual information, which adds nonlinearity. We may further refine this by removing the mediated parts of the interrelations between financial instruments by controlling for a third variable with partial mutual information. We take the minimum of the PMI calculated controlling for all other financial instruments, to show only the part of the mutual information not contained in other studied time series. Thus taking in mind the standard mutual information based metric [36, 37] the distance used for the network topology is defined as:

$$d(X, Y) = H(X, Y) - \min_{Z \neq X, Y} I(X, Y|Z). \tag{7}$$

On this basis we may create a network with topological restraints of our choosing, we calculate minimal spanning trees and planar maximally filtered graphs which we call PMIMST and PMIPMFG. These trees and planar graphs are created based on a list of $d(X, Y)$ sorted in increasing order. By starting from the first entry of the list, we add a corresponding link if and only if the resulting network is still a tree or a forest (PMIMST) or is still planar, i.e. it can be drawn on the surface of a sphere without link crossing (PMIPMFG).

Second we may refine the partial correlation planar graph as defined in [38] by directly swapping partial

correlation with partial mutual information. For this purpose we need a measure of MI influence or influence of an element Z on the pair of elements X and Y. This quantity is large only when a significant fraction of the PM $I(X, Y)$ can be explained in terms of Z. This measure is defined as:

$$d(X, Y|Z) = I(X, Y) - I(X, Y|Z). \tag{8}$$

We define the average influence $d(X|Z)$ of element Z on the MIs between element X and all the other elements in the system as

$$d(X|Z) = \langle d(X, Y|Z) \rangle_{Y \neq X, Z}. \tag{9}$$

In order to construct a planar graph based on PMI we list the $N(N-1)$ values of the average MI influence $d(X|Z)$ in decreasing order. The construction protocol of the network begins by considering an empty network with N vertices. By starting from the first entry of the list we put a link between them if and only if the resulting network is still planar. Similarly we can create a related minimal spanning tree by only adding a link if and only if the resulting network is still a tree or a forest.

Here we note that instead of filtering the information by network topology as we are doing above we may also filter the information by using a threshold to find which links should be entered into the network. Finding an appropriate threshold is not trivial however, but one can test for statistical significance based on the fact that $I(X, Y|Z)$ when X and Y are independent conditioned on Z follows a Gamma distribution with shape parameter $\kappa = |Z|(|X| - 1)(|Y| - 1)/2$ and scale parameter $\Theta = 1/N$ [51].

We also note a few areas in which this methodology can be extended. First we note that one could add market (index) as a second variable which is being controlled in PMI, thus eliminating the effects of general market trends on the relationships between financial instruments. Further we note that another approach to extending the analysis may be taken using transfer entropy, which is a measure closely related to partial mutual information [52]. Transfer entropy is a measure quantifying causal information transfer between systems evolving in time, based on appropriately conditioned transition probabilities, thus it uses time lags. Time lagged causal analysis can also be performed using correlation [53] or mutual information [37] based method for filtering similarity measures into a network of statistically-validated directed links.

3. Results

To find the networks presented above we have taken log returns for 91 securities out of 100 which constitute the NYSE 100, excluding those with incomplete data. These log returns are based on daily closing prices. The data has been downloaded from Google Finance database available at http://www.google.com/finance/ and was up to date as of the 11th of November 2013, going 10 years back. The data is transformed in the standard way for analysing price movements, that is so

that the data points are the log ratios between consecutive daily closing prices, as defined above, and those data points are, for the purpose of estimating mutual information, discretized into 4 distinct states.

Since there is no theory of financial markets we do not have any frame of reference to definitely state whether moving from correlation to mutual information and partial mutual information is bringing the analysis forward. We may only compare these two approaches and note how different the results are, and based on this we can risk an educated guess on the quality and usefulness of this method based on the fact that Shannon's mutual information is a more general measure than Pearson's correlation. We need to remember that while changing the particular methodology for creating networks of financial markets is changing the results, these changes will not be dramatic if the method is well-defined, thus graphically we will not see a completely different network. In other words the stocks should be clustered by economic sectors in all cases. But even small changes in these measurements may be of importance, particularly if this method is to be used in practice, especially automated analysis or trading. Then there is no reason to use correlation instead of mutual information, according to this study. This is particularly important for intraday price changes, where nonlinearity is significantly more prevalent [41], which is important with the rising popularity of intraday and automated algorithmic trading.

Thus to analyse the differences we have created 6 similarity measures for the networks, and for each of them created a tree and a planar graph, creating 12 networks together. The distances used are as follows:

1. $d(X,Y) = \sqrt{2(1 - \rho_{X,Y})}$;

2. $d(X,Y) = H(X,Y) - I(X,Y)$;

3. $d(X,Y) = \sqrt{2(1 - \min_{Z \neq X,Y} \rho_{X,Y})}$;

4. $d(X,Y) = H(X,Y) - \min_{Z \neq X,Y} I(X,Y|Z)$;

5. $d(X|Z) = \langle d(X,Y|Z) \rangle_{Y \neq X,Z}$ based on ρ;

6. $d(X|Z) = \langle d(X,Y|Z) \rangle_{Y \neq X,Z}$ based on I.

We can compare the networks on a node level, cluster level, and network level. First at node level we calculate Markov centrality (for detailed description see [54–56]) for each node in all those networks and compare them between networks obtaining correlations presented in Figs. 1 (for trees) and 2 (for planar graphs). We note that Markov centrality is basically a function of node degree, thus an analysis of node degree would be virtually identical. Similar analysis may be based also on node betweenness and other measures. Analyses for trees and graphs present similar data so we will comment collectively. We can see that the refinement of the standard methods by using minimal partial correlation/mutual information controlling for third variables is only changing

the networks by about 10% (noise from mediating relationships), while moving from correlation to mutual information changes the networks by about 30%, hinting that around a third of relationships between financial instruments is nonlinear. Further we note that the dynamic approach is about 20–30% different from the previous approaches, in this case the correlation and mutual information are giving much more varying results, thus indicating nonlinearity is more important in this dynamic assessment. In general we can conclude that the networks are significantly different (formal tests omitted due to space limitations) and thus the presented approach is valid, if the cluster level confirms this.
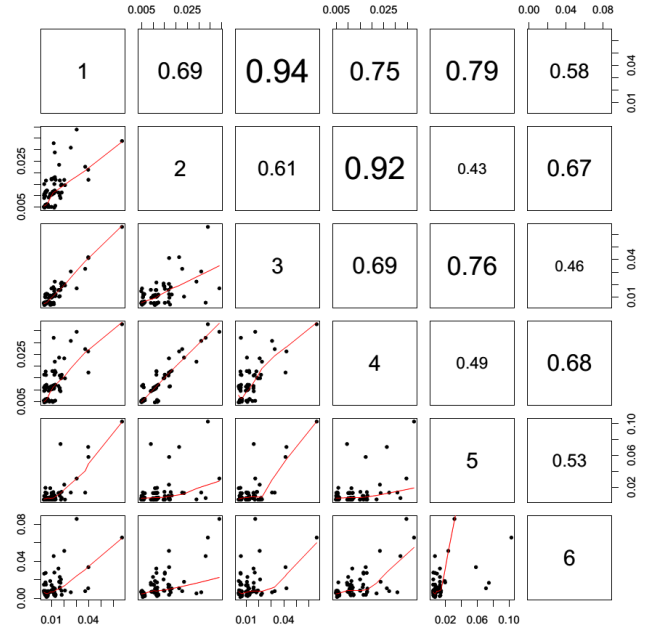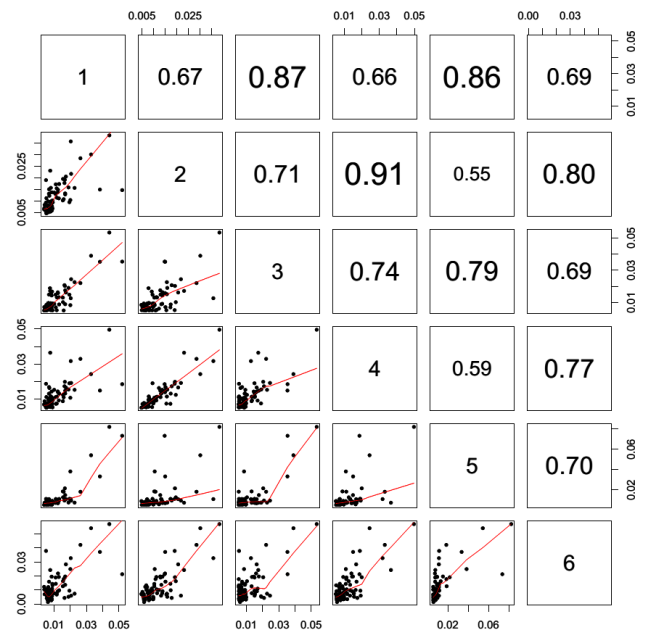

Fig. 1.   MC correlations for trees.

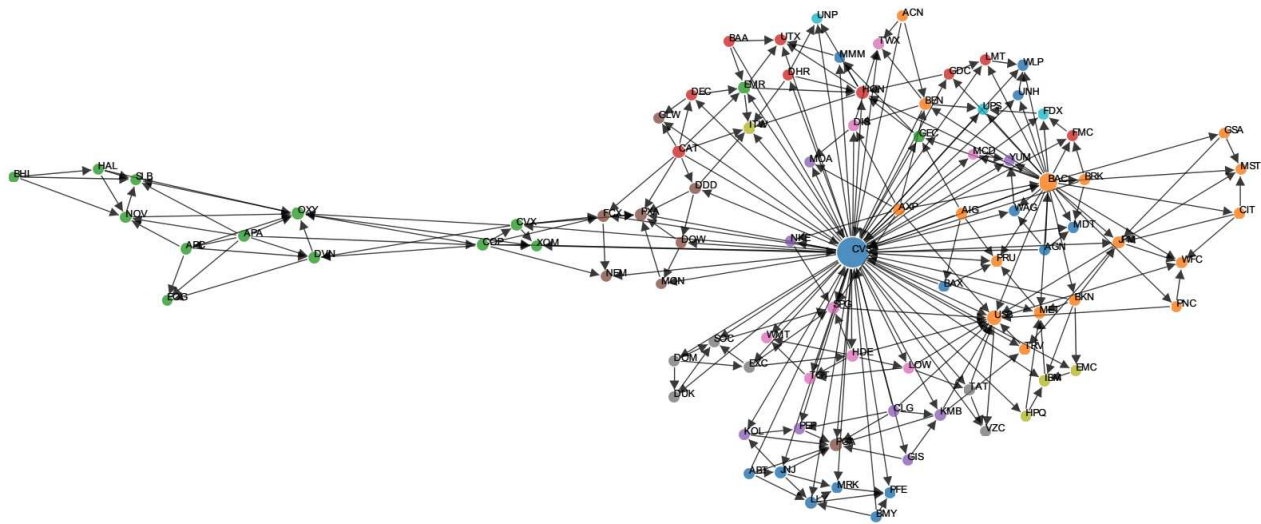Fig. 2.   MC correlations for planar graphs.

Fig. 3.   Partial mutual information PMFG.

On cluster level we confirm that in all created networks the clusters have been aligned largely according to economic sectors, which means that this important characteristic of the network approach to financial markets has been preserved. Preserving this information is important because it cannot be reproduced by simulating a virtual market [57]. This can be shown numerically as the ratio of arcs between stocks of the same sector to all nodes as presented in Table. As can be seen without any topological restraints the full market has 11.58% links within sectors, but for our networks it is between 42% and 67%, thus we see that this important feature is preserved. In general we can also see that mutual information suits this task better, which further corroborates the usefulness of our approach.

TABLE

Network comparision.

Network	Tree ratio	Graph ratio	Clustering
1	62.22%	49.06%	17.60%
2	66.67%	55.81%	20.80%
3	64.44%	51.69%	14.60%
4	65.56%	54.68%	17.70%
5	48.89%	42.70%	5.30%
6	57.78%	47.19%	14.70%
Ref.	11.28%	11.28%	50.00%

On network level there is a limited possibility of investigation as most network-wide measures would be constrained by the common topological restraints we are using. Nonetheless we have calculated clustering coefficients (ratio of the number of triangles observed to the number of possible triangles in the network) for the planar graphs (there can not be any triangles in trees), which are presented in Table below. As can be seen, mutual information produces significantly more clustering than correlation. Nonetheless, partial correlation/mutual information analysis creates slightly less clustered networks, hinting that some clustering happens due to mediating noise, but also due to nonlinearity.
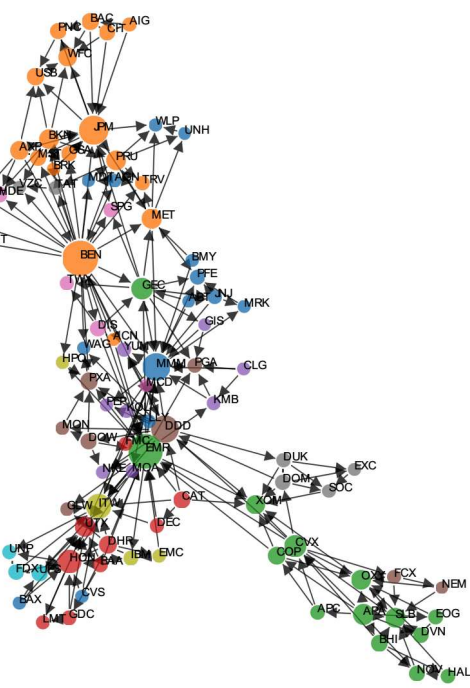

Fig. 4.   Partial mutual information planar graph.

The reference values in Table are calculated for a network with no topological restraints.

Strictly as an example we have provided network number 4 in Fig. 3 and 6 in Fig. 4. The node size is based on its centrality in the network.

4. Conclusions

We have analysed a method for producing networks of financial markets based on partial mutual information and how are those different from networks based on correlation, partial correlation and mutual information. The analysis leads us to believe that mutual information should be used in network analysis of financial markets as it provides different and likely more accurate networks, and that partial mutual information may be

used to refine this analysis even further, particularly in controlling for mediating inference of third parties, thus making the network closer to real structure of the market. Further studies should be used based on other information theoretic measures such as transfer entropy, and also with controlling for more than one variable (particularly controlling for the whole market), as well as with other markets, and intraday price data.

References

[1] H. Markowitz, *J. Financ.* **7**, 77 (1952).

[2] R. Mantegna, *Eur. Phys. J. B* **11**, 193 (1999).

[3] P. Cizeau, M. Potters, J. Bouchaud, *Quant. Financ.* **1**, 217 (2001).

[4] K. Forbes, R. Rigobon, *J. Financ.* **57**, 2223 (2002).

[5] B. Podobnik, H. Stanley, *Phys. Rev. Lett.* **100**, 084102 (2008).

[6] T. Aste, W. Shaw, T.D. Matteo, *New J. Phys.* **12**, 085009 (2010).

[7] D. Kenett, T. Preis, G. Gur-Gershgoren, E. Ben-Jacob, *Europhys. Lett.* **99**, 38001 (2012).

[8] G. Bonanno, F. Lillo, R. Mantegna, *Quant. Financ.* **1**, 96 (2001).

[9] M. Tumminello, T.D. Matteo, T. Aste, R. Mantegna, *Eur. Phys. J. B* **55**, 209 (2007).

[10] M. Munnix, R. Schafer, T. Guhr, *Physica A* **389**, 4828 (2010).

[11] G. Bonanno, N. Vandewalle, R.N. Mantegna, *Phys. Rev. E* **62**, 7615 (2000).

[12] S. Maslov, *Physica A* **301**, 397 (2001).

[13] S. Drożdż, F. Grummer, F. Ruf, J. Speth, *Physica A* **294**, 226 (2001).

[14] R. Coelho, C. Gilmore, B. Lucey, P. Richmond, S. Hutzler, *Physica A* **376**, 455 (2007).

[15] C.G. Gilmore, B.M. Lucey, M. Boscia, *Physica A* **387**, 6319 (2008).

[16] M. Eryigit, R. Eryigit, *Physica A* **388**, 3551 (2009).

[17] D.M. Song, M. Tumminello, W.X. Zhou, R.N. Mantegna, *Phys. Rev. E* **84**, 026108 (2011).

[18] L. Sandoval, I. Franca, *Physica A* **391**, 187 (2012).

[19] M. McDonald, O. Suleman, S. Williams, S. Howison, N.F. Johnson, *Phys. Rev. E* **72**, 046110 (2005).

[20] R.N. Mantegna, H.E. Stanley, *Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge University Press, Cambridge 1999.

[21] B. Rosser, *Adv. Complex Syst.* **11**, 745 (2008).

[22] W.A. Brock, D.A. Hsieh, B. LeBaron, *Nonlinear Dynamics, Chaos, and Instability. Statistical Theory and Economic Evidence*, MIT Press, Cambridge 1991.

[23] M. Qi, *J. Bus. Econ. Stat.* **17**, 419 (1999).

[24] D. McMillan, *Int. Rev. Econ. Financ.* **10**, 353 (2001).

[25] D. Sornette, J. Andersen, *Int. J. Mod. Phys. C* **13**, 171 (2002).

[26] K. Oh, K. Kim, *Expert Syst. Appl.* **22**, 249 (2002).

[27] P.H. Franses, D.V. Dijk, *J. Forecasting* **15**, 229 (1996).

[28] A. Abhyankar, L. Copeland, W. Wong, *Econ. J.* **105**, 864 (1995).

[29] P. Chen, *Stud. Nonlinear Dyn. E.* **1**, 87 (1996).

[30] A. Abhyankar, L. Copeland, W. Wong, *J. Bus. Econ. Stat.* **15**, 1 (1997).

[31] P.A. Ammermann, D.M. Patterson, *Pac. Bas. Financ. J.* **11**, 175 (2003).

[32] D. Hsieh, *J. Bus.* **62**, 339 (1989).

[33] R. Meese, A. Rose, *Rev. Econ. Stud.* **58**, 603 (1991).

[34] C. Brooks, *Appl. Financ. Econ.* **6**, 307 (1996).

[35] M. Qi, Y. Wu, *J. Empir. Financ.* **10**, 623 (2003).

[36] P. Fiedor, *Phys. Rev. E* **89**, 052801 (2014).

[37] P. Fiedor, *Eur. Phys. J. B* **87**, 168 (2014).

[38] D. Kenett, M. Tumminello, A. Madi, G. Gur-Gershgoren, R. Mantegna, E. Ben-Jacob, *PloS one* **5**, e15032 (2010).

[39] D. Kenett, X. Huang, I. Vodenska, S. Havlin, H. Stanley, arXiv:1402.1405 (2014) *Quant. Financ.* **15**, 569 (2015).

[40] N. Navet, S.H. Chen, in: *Natural Computing in Computational Finance*, Eds. T. Brabazon, M. O'Neill, Vol. 100 of *Studies in Computational Intelligence*, Springer, Berlin-Heidelberg 2008, p. 197.

[41] P. Fiedor, in: *Proc. IEEE Computational Intelligence for Financial Engineering and Economics 2014*, Eds.: A. Serguieva, D. Maringer, V. Palade, R.J. Almeida, IEEE, London 2014, p. 247.

[42] C.E. Shannon, *Bell Syst. Tech. J.* **27**, 379 (1948).

[43] L. Paninski, *Neural Comput.* **15**, 1191 (2003).

[44] C. Stevens, A. Zador, *Advances in Neural Information Processing Systems 8*, Eds.: D.S. Touretzky, M.C. Mozer, M.E. Hasselmo, MIT Press, Boston 1995, p. 75.

[45] S. Strong, R. Koberle, R. de Ruyter van Steveninck, W. Bialek, *Phys. Rev. Lett.* **80**, 197 (1998).

[46] P. Reinagel, *Curr. Biol.* **10**, 542 (2000).

[47] W. Nemenman, W. Bialek, R. de Ruyter van Steveninck, *Phys. Rev. E* **69**, 056111 (2004).

[48] D. Warland, P. Reinagel, M. Meister, *J. Neurophysiol.* **78**, 2336 (1997).

[49] J. Bonachela, H. Hinrichsen, M. Munoz, *J. Phys. A Math. Theor.* **41**, 202001 (2008).

[50] T. Schurmann, P. Grassberger, *Chaos* **6**, 414 (1996).

[51] D. Kugiumtzis, *Eur. Phys. J. Spec. Top.* **222**, 401 (2013).

[52] K. Hlavackova-Schindler, M. Palus, M. Vejmelka, J. Bhattacharya, *Phys. Rep.* **441**, 1 (2007).

[53] C. Curme, M. Tumminello, R. Mantegna, H. Stanley, D. Kenett, arXiv:1401.0462, (2014).

[54] H. Jeong, S.P. Mason, A.L. Barabasi, Z.N. Oltvai, *Nature* **411**, 41 (2001).

[55] S. White, P. Smyth, in: *Proc. Ninth ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, Eds.: L. Getoor, T. Senator, P. Domingos, Ch. Faloutsos, ACM, New York 2003, p. 266.

[56] U. Brandes, T. Erlebach, *Network Analysis: Methodological Foundations, Lecture Notes in Computer Science*, Springer, Berlin-Heidelberg 2005.

[57] G. Bonanno, G. Caldarelli, F. Lillo, R. Mantegna, *Phys. Rev. E* **68**, 046130 (2003).