Networks in financial markets based on the mutual information rate

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

(Received 12 January 2014; published 1 May 2014)

In the last few years there have been many efforts in econophysics studying how network theory can facilitate understanding of complex financial markets. These efforts consist mainly of the study of correlation-based hierarchical networks. This is somewhat surprising as the underlying assumptions of research looking at financial markets are that they are complex systems and thus behave in a nonlinear manner, which is confirmed by numerous studies, making the use of correlations which are inherently dealing with linear dependencies only baffling. In this paper we introduce a way to incorporate nonlinear dynamics and dependencies into hierarchical networks to study financial markets using mutual information and its dynamical extension: the mutual information rate. We show that this approach leads to different results than the correlation-based approach used in most studies, on the basis of 91 companies listed on the New York Stock Exchange 100 between 2003 and 2013, using minimal spanning trees and planar maximally filtered graphs.

DOI: 10.1103/PhysRevE.89.052801 PACS number(s): 89.65.−s, 05.10.−a, 64.60.aq, 89.70.−a

I. INTRODUCTION

Financial markets are studied in econophysics as complex systems, partly due to the lack of fundamental theory behind their behavior in economics. In particular, network theory is helpful in characterizing the interdependencies of different financial instruments [1–3] or classifying the financial instruments according to their interdependencies. Studies in financial markets use exclusive unsupervised classifications [4–6], which are obtained in clustering. Based on single linkage clustering analysis econophysicists have developed a popular method for creating correlation networks of financial markets.

The crucial problem in any clustering procedure is the choice of the measure of proximity between objects. In analyzing the financial markets researchers are persistently using only Pearson's correlation coefficient and its derivatives. The correlation structure of log returns of financial instruments contains key information for many practical applications such as portfolio optimization, risk management, and option pricing [7,8]. Such correlation structures have been investigated for time series describing stock returns [7–11], market index returns [12–19], and currency exchange rates [20]. The tools for analyzing such correlation structures contain spectral density analysis of the eigenvalues of the correlation matrix, tools of multivariate analysis, and random matrix theory [9–11]. Similarity-based graphs, or in other words networks associated with the similarity matrices [7,8,21–24], are also used.

The insistence of researchers in using Pearson's correlation coefficient is troubling and surprising however. It is well known that financial markets involve terms that are not of the first degree. There is now overwhelming evidence of nonlinear dynamics in stock returns [25–30], market index returns [31–36], and currency exchange rate changes [26,37–40]. Therefore the assumptions that only linear dependencies are relevant in financial markets found in hierarchical clustering methodology used in econophysics is baffling. In this paper we propose to amend the methodology of clustering for financial data so that the measure of similarity takes nonlinear dependencies into account.

Strictly linear correlation analysis can potentially miss important features of any dynamical system, particularly financial systems. Correlation coefficient is then contrasted by the measure of mutual information (MI, I_S), which differs from correlation due to its information theoretic background [41], which makes it a much more general measure. In fact, $I_S = 0$ if and only if the two studied random variables are strictly (statistically) independent. Mutual information is then a natural measure which can be used to extend the similarity measure to make it sensitive to nonlinear dependencies, and has indeed been successfully used in some applications [42–44]. Mutual information can be interpreted as a measure of how much information two studied systems exchange or two studied stochastic processes or data sets share. Mutual information is suitable for many applications [45–54]. The estimation of mutual information in dynamical systems faces some difficulties, however. These need to be understood in practical applications but are not severe [49,55–60].

In this study we propose to use mutual information and the mutual information rate as measures of similarity between financial instruments. MI presents a good extension to the correlation-based studies, while the mutual information rate (MIR, I_R) can be used in a more dynamic analysis as it can be understood as a measurement of all the interdependencies between the spatiotemporal organization of the observed sequences. Mutual information is based on the Shannon's concept of entropy, and consequently the dynamical extension of mutual information, that is the mutual information rate, is in turn based on the dynamical extension of entropy or the entropy rate [61,62]. We will use the concept of Lempel-Ziv complexity [41,63] to obtain accurate estimates of the entropy rate. It has been in many other fields [64–68] and recently also in econophysics [69,70]. Most of these studies use one-dimensional analysis. In this paper we use an extension of Lempel-Ziv complexity to multidimensional signals [71,72] to study the estimate of higher order correlations between pairs of financial instruments. The validity of estimating mutual information using Lempel-Ziv complexity for nonlinear time series has been confirmed in earlier studies [64,73]. The mutual

*s801dok@wizard.uek.krakow.pl

information rate and mutual Lempel-Ziv complexity provide a good account of the spatiotemporal structure of the sequences, as these sequences are then not only joint realizations of two random variables, but instead a joint realization of one random process, rendering the mutual Lempel-Ziv complexity much more meaningful than a collection of quantities computed for singular random variables [72].

The complexity of the financial markets and their behavior in the recent years, together with the very fast dynamics (e.g., the so-called flash crash), means that we no longer can ignore the nonlinearity of financial markets without any loss of important information. Therefore in this paper we extend the known methodology of hierarchical clustering of financial data and creating dependency networks by exchanging the similarity measure from the Pearson's correlation coefficient to the information-theoretic approach using mutual information and mutual information rate. We then apply it to log returns on the New York Stock Exchange in order to show the validity of this approach.

II. SIMILARITY MEASURE

The topological arrangement of the nodes in network-based financial models is most often based on the Pearson's correlation coefficient ρ [74] for a studied period. A metric is necessary to find an approximate distance between the nodes in a network. Usually an Euclidean metric is used [75]:

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

To extend such a measure to include nonlinear dependencies we propose to base the topological arrangement of the nodes in a network on the mutual information and mutual information rate rate between closing prices for two consecutive days for two financial assets. These are built on Shannon's formulation of entropy, entropy rate, and mutual information [61]. The entropy rate is a term derivative to the notion of entropy, which measures the amount of uncertainty in a random variable. We denote Shannon's entropy of a single random variable X as $H(X)$ [61]. Shannon also introduced the entropy rate, which generalizes the notion of entropy for sequences of dependent random variables. We denote it as $h(X)$ and interpret it as a measure of the average uncertainty left in the generation of information in a process at time n having observed the complete history up to that point [41]. For two discrete random variables X and Y we can also find mutual between them, which is defined as

$$I_S(X,Y) = H(X) + H(Y) - H(X,Y), \tag{2}$$

where $H(X)$ and $H(Y)$ are the marginal entropies and $H(X,Y)$ is the joint entropy of X and Y. Mutual information measures the amount of information shared by X and Y, or in other words how much the information about one stochastic process reduces uncertainty about the other. Mutual information is non-negative and $I_S(X,X) = H(X)$.

The mutual information rate (MIR) was also first introduced by Shannon [61] as the rate of actual transmission [76] and was consequently more rigorously defined by other researchers [77,78]. Just as the entropy rate represents entropy per unit of time, the mutual information rate represents the mutual information exchanged between two dynamical variables per unit of time. To simplify the calculation of the MIR, if we have two continuous dynamical variables, we transform them into two discrete symbolic sequences X and Y. For such sequences the mutual information rate is defined by

$$I_R = \lim_{n\to\infty} \frac{I_S(n)}{n}, \tag{3}$$

where $I_S(n)$ represents mutual information between the two sequences X and Y calculated by considering words of length n. It has been showed that it can be reliably estimated with no need for stationarity, statistical stability, or a memoryless source [79].

The above definitions do not actually present an obvious way to calculate the mutual information or mutual information rate in practice. To estimate MI we need an estimator of Shannon's entropy. There is an abundance of estimators [55,80–84], and in this study we use the Schurmann-Grassberger estimate of the entropy of a Dirichlet probability distribution [85], which is thought to be the best choice outside very specific conditions (particularly small samples) [86].

One of the ways to estimate MIR is to use the Lempel-Ziv complexity, which can be used to estimate both entropy rate and mutual information rate. Lempel-Ziv is a data compression algorithm [87], on which basis a number of estimators of entropy rate have been created. In this article we follow Ref. [69] and use the estimator created by Kontoyiannis in 1998 (estimator a) [88]. This estimator is widely used [69,89] and has good statistical properties [88]. For other variants see Ref. [62].

Formally to calculate the entropy rate of a random variable X, the probability of each possible outcome $p(x_i)$ must be known. When these probabilities are not known, entropy can be estimated by replacing the probabilities with relative frequencies from observed data. The mentioned estimator is defined as

$$\hat{h}_{lz} = \frac{n \log_2 n}{\sum_i \Lambda_i}, \tag{4}$$

where n denotes the length of the time series, and Λ_i denotes the length of the shortest substring starting from time i that has not yet been observed prior to time i, i.e., from time 1 to $i-1$. It is known that for stationary ergodic processes, $\hat{h}_{lz}(X)$ converges to the entropy rate $h(X)$ with a probability of 1 as n approaches infinity [88].

It is important that in cases where the original data points are continuous (which is the case for financial markets) we need to discretize the data points for the purpose of the Lempel-Ziv complexity estimator. This procedure can be performed in many ways; the number of bins into which the data is assigned is a matter of convention and researchers' choice, but it is advised that it should not be larger than square root of the sample size, and in fact should presumably be much smaller. In the case of financial markets we propose that the number of bins should be between 4 [70] and 8 [69]; see also comments in the empirical part of this study. It is important, however, that the states represent quartiles or other equal divisions; therefore each state is assigned the same number of data points. This design means that the model has no unnecessary parameters, which could affect the results and conclusions reached while using the data. This experimental setup also proved to be very efficient at revealing the randomness of the original data [90].

Based on this we can also define Lempel-Ziv complexity for multidimensional sequences [91]. Extending the Lempel-Ziv complexity for vectorial data has been proposed in Ref. [71], and here we follow this method. Within this methodology the joint Lempel-Ziv complexity of sequences $X_0,\ldots,X_{k-1}$ is defined as

$$h_{lz}(X_0,\ldots,X_{k-1}) = h_{lz}(Z), \tag{5}$$

where Z is a sequence of n k-uplets $z_j = (x_{0,j},\ldots,x_{k-1,j})$. Then Lempel-Ziv complexity of multidimensional sequences can then be viewed as a joint Lempel-Ziv complexity.

Therefore, analogous with the Shannon information theory [61], mutual Lempel-Ziv complexity can be defined using the joint Lempel-Ziv complexity defined for two sequences X and Y. Then mutual Lempel-Ziv complexity is defined as [72]

$$h_{m,lz}(X,Y) = h_{lz}(X) + h_{lz}(Y) - h_{lz}(X,Y). \tag{6}$$

The mutual Lempel-Ziv complexity (MLZC) can be interpreted as a convergence measure between two sequences. Mutual Lempel-Ziv complexity can be negative transiently for finite N, but for $N \to \infty$ the asymptotic quantity $h_{m,lz}(X,Y)$ is always positive. In fact, the MLZC converges asymptotically to a dynamic extension of the mutual information: the mutual information rate [92,93].

We now know what mutual information rate is and how to estimate it asymptotically using Lempel-Ziv complexity for multidimensional data. But in order to create a topology of the dependence network we would prefer to have an Euclidean metric, which neither the mutual information nor mutual information rate are. Here we will use the mutual information-based metric proposed in Ref. [6]. Since mutual information and the mutual information rate share most of their properties it is therefore possible to use this metric directly exchanging mutual information with mutual information rate. The first such metric is well known [41]. The quantity

$$d(X,Y) = H(X|Y) + H(Y|X) = H(X,Y) - I_S(X,Y), \tag{7}$$

$$d(X,Y) = H(X) + H(Y) - 2I_S(X,Y) \tag{8}$$

satisfies the triangle inequality, is non-negative and symmetric, and satisfies $d(X,X) = 0$. This has been proved in Ref. [6]. Here we note that this metric may be normalized [94] since mutual information depends on the size of the studied sequence, so that the normalized version is defined as

$$D(X,Y) = 1 - \frac{I_S(X,Y)}{H(X,Y)} = \frac{d(X,Y)}{H(X,Y)}, \tag{9}$$

but this is not necessary in our case since we use the same sample size for all studied instruments.

Let us once again define D, this time in terms of the mutual information rate:

$$D(X,Y) = \frac{d(X,Y)}{h(X,Y)}, \tag{10}$$

where

$$d(X,Y) = h(X,Y) - I_R(X,Y), \tag{11}$$

$$d(X,Y) = h(X) + h(Y) - 2I_R(X,Y). \tag{12}$$

We now have a metric allowing us to quantify distance between nodes in hierarchical networks describing interdependencies on financial markets. Therefore we can turn briefly to summarizing the procedures used for creating such networks.

III. HIERARCHICAL NETWORKS

Having defined the distance measure we now briefly turn to the construction methods for filtered graphs. Networks can be constructed by either topological restraints or setting a threshold on the similarity measure. We are using the first approach to create minimal spanning trees (MSTs) and planar maximally filtered graphs (PMFGs). These methods are well known in literature; hence we will only briefly define them. The distance matrix $\mathcal{D}$ containing $d(X,Y)$ for all studied pairs is used to determine the minimal spanning tree and planar maximally filtered graph [95] connecting n financial instruments in the studied set. On the basis of the distance matrix $\mathcal{D}$ we create an ordered list $\mathcal{S}$, in which the distances are listed in decreasing order. Then, to create a minimal spanning tree, starting from the first element of the list the corresponding link is added to the network if and only if the resulting graph is still a forest or a tree [23]. Similarly a planar maximally filtered graph can be constructed in the same way by adding the corresponding link if and only if the resulting graph is still a planar graph (with genus equal 0). For a detailed description see Refs. [23,96].

IV. EMPIRICAL APPLICATION

To apply mutual information and mutual information rate-based networks in practice and find out their properties we have taken log returns for 91 stocks belonging to the New York Stock Exchange 100 index which were traded continuously between 11 November 2003 and 7 November 2013. The data are transformed in the standard way for analyzing price movements, that is, so that the data points are the log ratios between consecutive daily closing prices, $r_t = \ln(p_t/p_{t-1})$, and those data points are, for the purpose of the mutual information and Lempel-Ziv complexity estimators, discretized into four distinct states. Here we need to briefly discuss the choice of four bins for the discretization stage. In general the number of bins should be chosen according to how granular patterns the researcher want to study; we believe that since financial markets on a daily scale are almost random [97] a large number of bins would not be a good choice as even in granular price changes there are not many patterns to be found. In other studies four or eight have been chosen though in those applications this is irrelevant [69,97], but networks are sensitive to even small changes. So to our first mentioned reason we also add a second, pragmatic one, as we have observed that a larger number of bins provides us with networks of worse characteristics; thus we have settled with quartiles. A further study should nonetheless be performed to scrutinize this choice and compare it with an approach based on differential entropy.

In Figs. 1 and 2 we have presented PMFGs for mutual information and the mutual information rate, respectively. These are illustrative, and we will now turn to the analysis of these networks and corresponding MSTs.

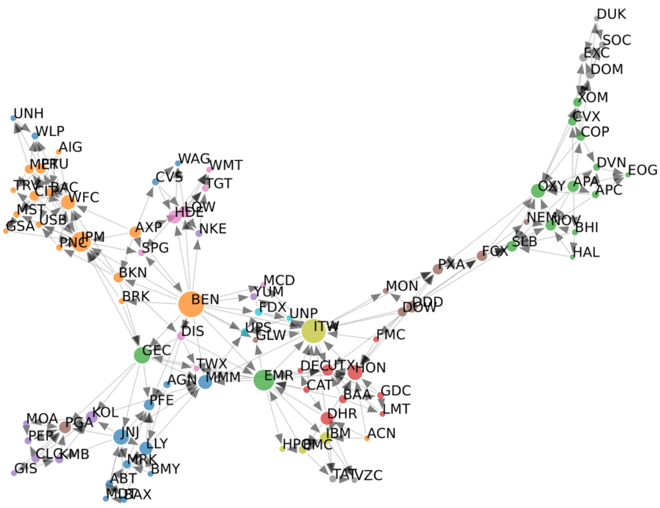

FIG. 1. (Color online) PMFG based on I_S.

As economics has no theory of financial markets behavior there is no frame of reference to definitely state whether moving from correlation to mutual information and mutual information rate is rendering the networks closer to reality. In this sense we can only analyze the resulting networks and find out if they are significantly different. If they are indeed, then we can find out some of their characteristics and add an educated guess based on the properties of the used metrics (MI being a more general measure). In particular we will look at whether the stocks are clustered by economic sectors in all cases. We also note that we produce networks based on daily prices, while this may be even more important for intraday price changes, where nonlinearity is significantly more prevalent [97].

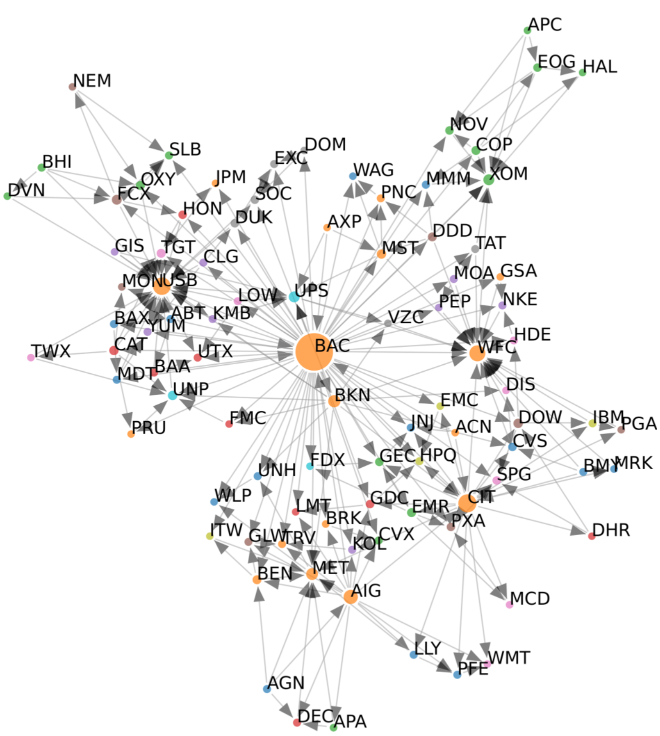

FIG. 2. (Color online) PMFG based on MIR.

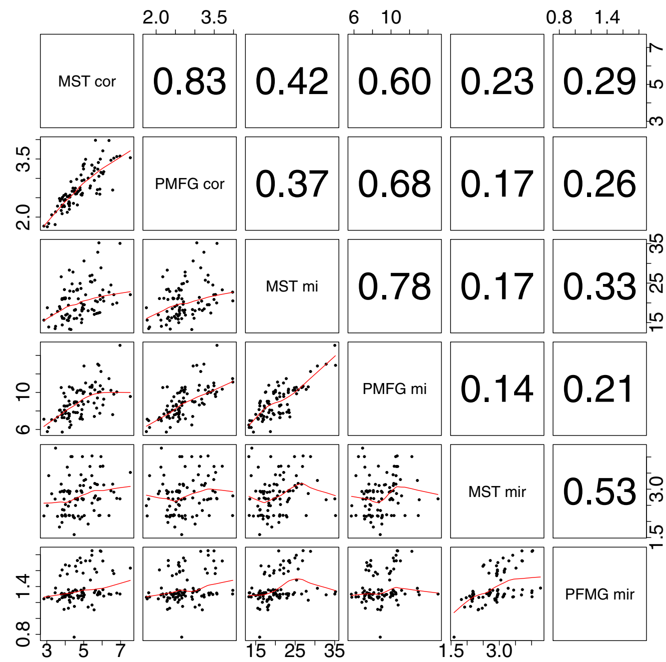

FIG. 3. (Color online) Correlations for average shortest path.

We will compare the networks on the node level, cluster level, and network level. First, on the node level we calculate average shortest path (ASP), betweenness centrality, node degree, and Markov centrality (MC) for each node in all those networks and compare them between networks, obtaining correlations presented in Figs. 3 (for average shortest path) and 4 (for Markov centrality). We note that the results for Markov centrality, node degree, and betweenness centrality are virtually identical; thus the redundant ones have been ignored. In Fig. 5 we present node degree distributions (with fitted power laws) for the created networks.

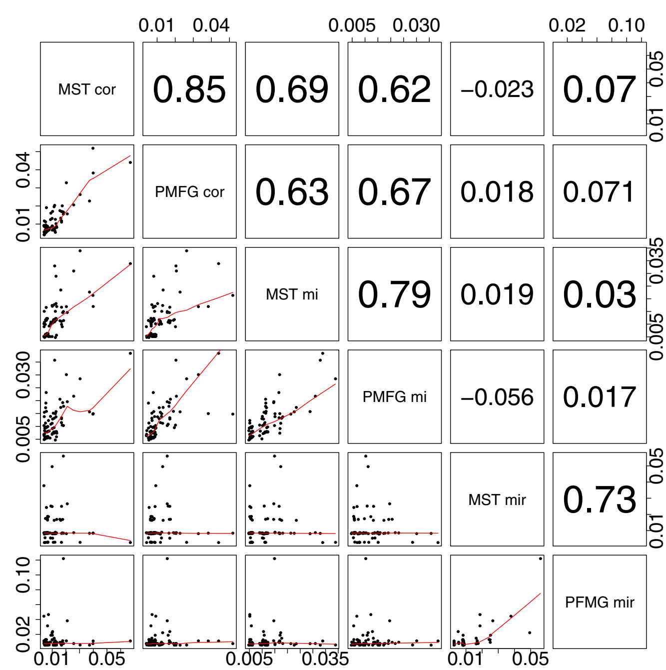

FIG. 4. (Color online) Correlations for Markov centrality.

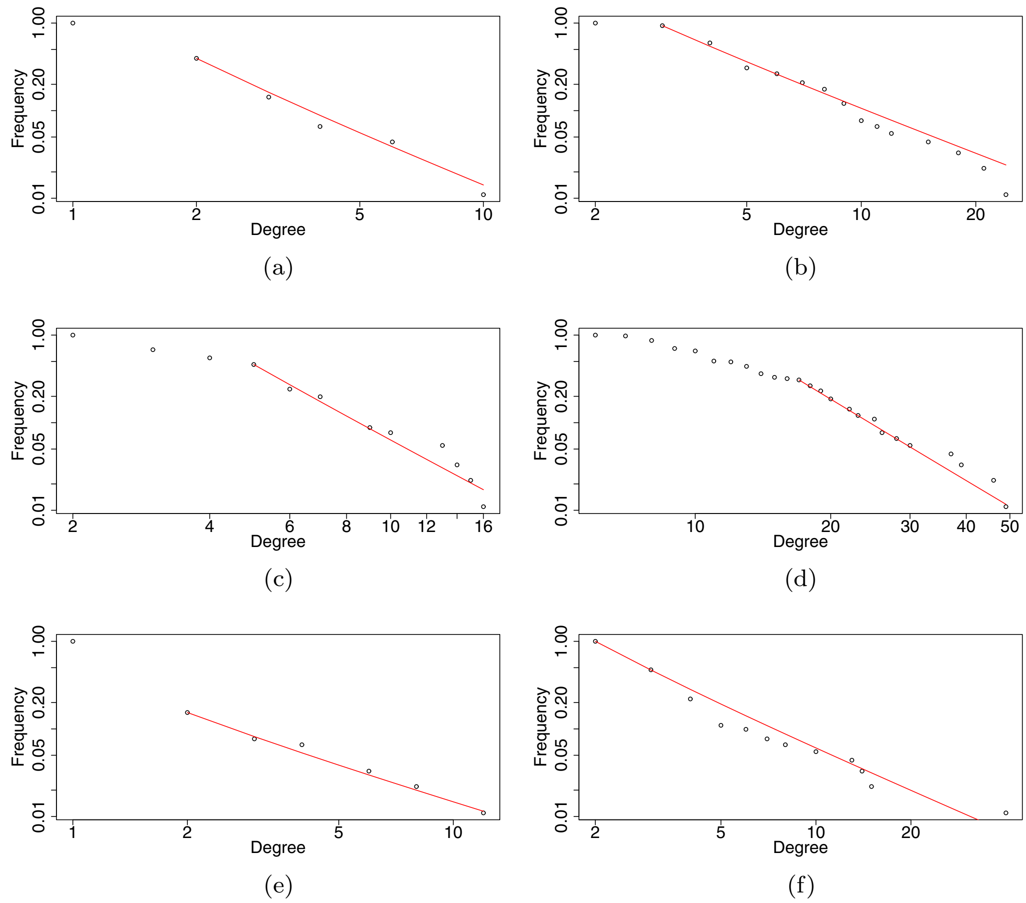

FIG. 5. (Color online) Degree distributions for (a) MST based on correlation; (b) PMFG based on correlation; (c) MST based on mutual information; (d) PMFG based on MI; (e) MST based on mutual information rate; and (f) PMFG based on MIR.

For ASP the correlation between networks based on correlation and mutual information is equal to 42% for MST and 68% for PMFG. Same correlations for MC (and node degree, betweenness) are equal to 69% and 67%, respectively (thus it appears that a third of the relationships on the New York stock market is not linear). The degree distribution is different for these networks as well, as networks based on mutual information contain nodes with much larger degrees. Thus we can conclude that these networks are significantly statistically different. Any well-defined test of their equality would give us a p value of 0, and we have performed permutation tests with such results; thus we have ignored formal testing here due to the obvious nature of the results. The correlations between these two networks and networks based on MIR are equal to about 20% for ASP and around 0 for MC. The degree distribution of an MIR-based network is closer to this of a correlation-based one. Again we can obviously see that these networks are statistically different from the previous ones; thus a formal test is not needed.

We have confirmed that the differences between studied networks are statistically significant. Now we need to see if they contain valuable information. On a cluster level we look at whether the clusters have been aligned largely according to economic sectors. We want to preserve this characteristic as it is important and cannot be reproduced by simulating a virtual market [98]. This can be shown numerically as the ratio of arcs between stocks of the same sector to all nodes as presented in Table I. As can be seen without any topological restraints the full market has 11.58% links within sectors, but for our networks it is 49%–62% (correlation), 55%–67% (MI), and 24%–25% (MIR). We see that this important feature is preserved and even enhanced for networks based on MI, but is not well-preserved in networks based on MIR, but even these have a substantially higher ratio than unconstrained networks.

On a 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. We have calculated clustering coefficients (the ratio of the number of triangles observed to the number of possible triangles in the network) for the planar graphs, however, which are also

TABLE I. Network comparison for MSTs and PMFGs based on Pearson's correlation, mutual information, and mutual information rate.

Network	Clustering	Sector ratio
MST ρ	–	62.22%
PMFG ρ	17.60%	49.06%
MST I_S	–	66.67%
PMFG I_S	20.80%	55.81%
MST I_R	–	24.44%
PMFG I_R	5.60%	25.09%
Reference	50.00%	11.28%

presented in Table I. We can see that mutual information produces more clustering than correlation, but MIR creates significantly less clustered networks. The reference values in Table I are calculated for a network with no topological restraints.

On this basis we can conclude that mutual information is a better metric for the construction of financial networks than correlation. The resulting networks are statistically significantly, but not completely, different, and they show even stronger dependence on sector- than correlation-based ones. We also observe more clustering in those, and necessarily those networks contain information about nonlinear relationships the correlation-based networks ignore. We are therefore strongly suggest using this approach for evaluation of market structure. The conclusion is less clear with networks based on MIR. These are not similar to networks presented above and are also less clustered, and depend on sectors to a much lesser degree. We could therefore find them useless. But since MIR is a measure of market dynamics rather than market structure we are not necessarily expecting results similar to those based on correlation. Additionally these networks show strong dominance of the financial sector, which corroborates a similar result obtained using partial correlation [99]; thus we conclude that these networks contain useful information, but it may be necessary to see if other measures (e.g. partial mutual information or transfer entropy [100]) produce a more appropriate structures describing the dynamic side of the markets.

V. CONCLUSIONS

In this paper we have presented a methodology for creating hierarchical networks studying financial markets using mutual information and its dynamic extension: the mutual information rate. We have applied this methodology to the New York Stock Exchange. The resulting minimal spanning tress and planar maximally filtered graphs are significantly different from those obtained using Pearson's correlation as a similarity measure; therefore we conclude that the nonlinear dependencies not captured by Pearson's correlation coefficient but captured by mutual information are indeed relevant to the hierarchical structure of the financial markets. We find that mutual information is a good way of extending the correlation-based networks, while the mutual information rate presents a different side of the markets but produces networks with troubling characteristics. Further research should look into other measures which can find the dynamic structure of the markets while producing networks with better characteristics. The proposed methodology is sensitive to the choice of the number of bins into which the log returns are discretized and requires large sample sizes. Further research should look into the discretization step and the possibility of using differential entropy. Further studies based on other stock markets, market indexes, and currency exchange markets should also be performed to further analyze the usefulness of this approach.

[1] B. B. Mandelbrot, J. Bus. **36**, 394 (1963).
[2] L. P. Kadanoff, Simulation **16**, 261 (1971).
[3] R. N. Mantegna, Physica A **179**, 232 (1991).
[4] A. Jain and R. Dubes, *Algorithms for Clustering Data* (Prentice Hall, Englewood Cliffs, NJ, 2000).
[5] L. Hubert and F. B. Baker, J. Educ. Stat. **1**, 87 (1976).
[6] A. Kraskov, H. Stogbauer, R. Andrzejak, and P. Grassberger, Europhys. Lett. **70**, 278 (2005).
[7] R. Mantegna, Eur. Phys. J. B **11**, 193 (1999).
[8] M. Tumminello, T. Aste, T. D. Matteo, and R. N. Mantegna, Eur. Phys. J. B **55**, 209 (2007).
[9] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters, Phys. Rev. Lett. **83**, 1467 (1999).
[10] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. Nunes-Amaral, and H. E. Stanley, Phys. Rev. Lett. **83**, 1471 (1999).
[11] G. Akemann, J. Baik, and P. D. Francesco, *The Oxford Handbook of Random Matrix Theory* (Oxford University Press, Oxford, 2011).
[12] G. Bonanno, N. Vandewalle, and R. N. Mantegna, Phys. Rev. E **62**, R7615(R) (2000).
[13] S. Maslov, Physica A **301**, 397 (2001).
[14] S. Drozdz, F. Grummer, F. Ruf, and J. Speth, Physica A **294**, 226 (2001).
[15] R. Coelho, C. Gilmore, B. Lucey, P. Richmond, and S. Hutzler, Physica A **376**, 455 (2007).
[16] C. G. Gilmore, B. M. Lucey, and M. Boscia, Physica A **387**, 6319 (2008).
[17] M. Eryigit and R. Eryigit, Physica A **388**, 3551 (2009).
[18] D.-M. Song, M. Tumminello, W.-X. Zhou, and R. N. Mantegna, Phys. Rev. E **84**, 026108 (2011).
[19] L. Sandoval and I. Franca, Physica A **391**, 187 (2012).
[20] M. McDonald, O. Suleman, S. Williams, S. Howison, and N. F. Johnson, Phys. Rev. E **72**, 046106 (2005).
[21] J.-P. Onnela, A. Chakraborti, K. Kaski, and J. Kertesz, Eur. Phys. J. B **30**, 285 (2002).
[22] J.-P. Onnela, A. Chakraborti, K. Kaski, and J. Kertesz, Physica A **324**, 247 (2003).
[23] M. Tumminello, T. Aste, T. D. Matteo, and R. N. Mantegna, Proc. Natl. Acad. Sci. USA **102**, 10421 (2005).
[24] M. Tumminello, C. Coronnello, F. Lillo, S. Micciche, and R. Mantegna, Int. J. Bifurcat. Chaos **17**, 2319 (2007).
[25] D. Hsieh, J. Financ. **46**, 1839 (1991).
[26] W. A. Brock, D. A. Hsieh, and B. LeBaron, *Nonlinear Dynamics, Chaos, and Instability. Statistical Theory and Economic Evidence* (MIT Press, Cambridge, MA, 1991).
[27] M. Qi, J. Bus. Econ. Stat. **17**, 419 (1999).
[28] D. McMillan, Int. Rev. Econ. Financ. **10**, 353 (2001).
[29] D. Sornette and J. Andersen, Int. J. Mod. Phys. C **13**, 171 (2002).
[30] K. Oh and K. Kim, Expert Syst. Appl. **22**, 249 (2002).
[31] P. H. Franses and D. V. Dijk, J. Forecasting **15**, 229 (1996).
[32] A. Abhyankar, L. Copeland, and W. Wong, Econ. J. **105**, 864 (1995).
[33] P. Chen, Stud. Nonlinear Dyn. E **1**, 87 (1996).
[34] A. Abhyankar, L. Copeland, and W. Wong, J. Bus. Econ. Stat. **15**, 1 (1997).
[35] J. A. Scheinkman and B. Lebaron, J. Bus. **62**, 311 (1989).
[36] P. A. Ammermann and D. M. Patterson, Pacific-Basin Finance J. **11**, 175 (2003).
[37] D. Hsieh, J. Bus. **62**, 339 (1989).

[38] R. Meese and A. Rose, Rev. Econ. Stud. **58**, 603 (1991).
[39] C. Brooks, Appl. Financ. Econ. **6**, 307 (1996).
[40] M. Qi and Y. Wu, J. Empir. Finance **10**, 623 (2003).
[41] T. Cover and J. Thomas, *Elements of Information Theory* (John Wiley & Sons, New York, 1991).
[42] F. Zhou, J. He, and W. Zhong, in *Proceedings of the 7th IEEE International Conference on Bioinformatics and Bioengineering* (IEEE, Boston, 2007), p. 1051.
[43] F. Zhou, J. He, W. Zhong, and Y. Pan, Int. J. Comput. Biol. Drug Des. **2**, 187 (2009).
[44] A. C. Muller, S. Nowozin, and C. H. Lampert, in *Pattern Recognition*, edited by A. Pinz, T. Pock, H. Bischof, and F. Leberl (Springer, Berlin, 2012), pp. 205–215.
[45] O. Sporns, D. R. Chialvo, M. Kaiser, and C. C. Hilgetag, Trends Cogn. Sci. **8**, 418 (2004).
[46] N. Brenner, O. Agam, W. Bialek, and R. R. de Ruyter van Steveninck, Phys. Rev. Lett. **81**, 4000 (1998).
[47] N. Brenner, O. Agam, W. Bialek, and R. de Ruyter van Steveninck, Phys. Rev. E **66**, 031907 (2002).
[48] J. Donges, Y. Zou, N. Marwan, and J. Kurths, Eur. Phys. J.-Spec. Top. **174**, 157 (2009).
[49] M. Palus, V. Komarek, T. Prochazka, Z. Hrncir, and K. Sterbova, IEEE Eng. Med. Biol. **20**, 65 (2001).
[50] A. M. Fraser and H. L. Swinney, Phys. Rev. A **33**, 1134 (1986).
[51] U. Parlitz, in *Nonlinear Modeling: Advanced Black-Box Techniques*, edited by J. A. K. Suykens and J. Vandewalle (Kluwer Academic Publishers, Boston, 1998), pp. 209–239.
[52] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, 2004).
[53] S. Haykin, *Communication Systems* (John Wiley & Sons, New York, 2001).
[54] F. Rossi, A. Lendasse, D. Francois, V. Wertz, and M. Verleysen, Chemometr. Intell. Lab. **80**, 215 (2006).
[55] L. Paninski, Neural Comput. **15**, 1191 (2003).
[56] C. Floros, S. Jaffry, and G. V. Lima, Stud. Econ. Finance **24**, 220 (2007).
[57] S.-F. Hsieh and S.-D. Shyu, J. Asia Bus. Stud. **4**, 49 (2009).
[58] T. A. A. Manap and S. H. Kassim, J. Risk Finance **12**, 356 (2011).
[59] R. Steuer, J. Kurths, C. Daub, J. Weise, and J. Selbig, Bioinformatics **18**, S231 (2002).
[60] A. Papana and D. Kugiumtzis, Int. J. Bifurcat. Chaos **19**, 4197 (2009).
[61] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).
[62] Y. Gao, I. Kontoyiannis, and E. Bienenstock, Entropy **10**, 71 (2008).
[63] A. Lempel and J. Ziv, IEEE T. Inform. Theory **22**, 75 (1976).
[64] J. Amigo, J. Szczepanski, E. Wajnryb, and M. Sanchez-Vives, Neural Comput. **16**, 717 (2004).
[65] M. Aboy, R. Hornero, D. Abasolo, and D. Alvarez, IEEE T. Bio-med. Eng. **53**, 2282 (2006).
[66] M. Christen, A. Kohn, T. Ott, and R. Stoop, J. Neurosci. Meth. **156**, 342 (2006).
[67] J. Amigo and M. Kennel, Chaos **16**, 043102 (2006).
[68] A. Lesne, J. L. Blanc, and L. Pezard, Phys. Rev. E **79**, 046208 (2009).
[69] N. Navet and S.-H. Chen, in *Natural Computing in Computational Finance*, Studies in Computational Intelligence, edited by T. Brabazon and M. O'Neill (Springer, Berlin Heidelberg, 2008), Vol. 100.
[70] P. Fiedor, arXiv:1310.5540 (2013).
[71] S. Zozor, P. Ravier, and O. Buttelli, Physica A **345**, 285 (2005).
[72] J.-L. Blanc, N. Schmidt, L. Bonnier, L. Pezard, and A. Lesne, in *Deuxiéme Conférence Française de Neurosciences Computationnelles* (Elsevier, Marseille, 2008).
[73] Z. Dian-Zhong, Acta Phys. Sin. **56**, 3152 (2007).
[74] W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1971).
[75] R. Mantegna, in *Proceedings of the ANDM 97 International Conference*, edited by J. Kadtke (AIP Press, San Diego, 1997).
[76] J. L. Blanc, L. Pezard, and A. Lesne, Phys. Rev. E **84**, 036214 (2011).
[77] L. Dobrushin, Usp. Mat. Nauk **14**, 3 (1959).
[78] R. Gray and J. Kieffer, IEEE T. Inform. Theory **26**, 412 (1980).
[79] S. Verdu and T. Han, IEEE T. Inform. Theory **40**, 1147 (1994).
[80] C. Stevens and A. Zador, in *Proceedings of the 1995 Conference on Advances in Neural Information Processing Systems 8*, edited by D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo (MIT Press, Cambridge, MA, 1995) pp. 75–81.
[81] S. P. Strong, R. Koberle, R. R. de Ruyter van Steveninck, and W. Bialek, Phys. Rev. Lett. **80**, 197 (1998).
[82] P. Reinagel, Current Biol. **10**, R542 (2000).
[83] I. Nemenman, W. Bialek, and R. de Ruyter van Steveninck, Phys. Rev. E **69**, 056111 (2004).
[84] D. Warland, P. Reinagel, and M. Meister, J. Neurophysiol. **78**, 2336 (1997).
[85] T. Schurmann and P. Grassberger, Chaos **6**, 414 (1996).
[86] J. A. Bonachela, H. Hinrichsen, and M. A. Muñoz, J. Phys. A: Math. Theor. **41**, 202001 (2008).
[87] J. Ziv and A. Lempel, IEEE Trans. Inform. Theory **23**, 337 (1977).
[88] I. Kontoyiannis, in *IEEE International Symposium on Information Theory* (MIT, Cambridge, MA, 1998).
[89] M. Kennel, J. Shlens, H. Abarbanel, and E. Chichilnisky, Neural Comput. **17**, 1531 (2005).
[90] R. Steuer, L. Molgedey, W. Ebeling, and M. Jiménez-Montaño, Eur. Phys. J. B **19**, 265 (2001).
[91] F. Kaspar and H. Schuster, Physica A **36**, 842 (1987).
[92] S. Ihara, *Information Theory for Continuous Systems* (World Scientific, Singapore, 1993).
[93] M. Pinsker, *Information and Information Stability Random Process* (Holden Day, San Francisco, 1964).
[94] M. Li, J. Badger, X. Chen, S. Kwong, P. Kearney, and H. Zhang, Bioinformatics **17**, 149 (2001).
[95] C. Papadimitriou and K. Steiglitz, *Combinatorial Optimization* (Prentice-Hall, Englewood Cliffs, NJ, 1982).
[96] M. Tumminello, F. Lillo, and R. N. Mantegna, J. Econ. Behav. Organ. **75**, 40 (2010).
[97] P. Fiedor, in *Proceedings of the IEEE Computational Intelligence for Financial Engineering & Economics 2014*, edited by A. Serguieva, D. Maringer, V. Palade, and R. Jorge Almeida (IEEE, London, 2014), pp. 247–254.
[98] G. Bonanno, G. Caldarelli, F. Lillo, and R. N. Mantegna, Phys. Rev. E **68**, 046130 (2003).
[99] D. Kenett, M. Tumminello, A. Madi, G. Gur-Gershgoren, R. Mantegna, and E. Ben-Jacob, PloS ONE **5**, e15032 (2010).
[100] K. Hlavackova-Schindler, M. Palus, M. Vejmelka, and J. Bhattacharya, Phys. Rep. **441**, 1 (2007).