Overview of Network-based Methods for Analyzing Financial Markets

Pavel Tsankov

***Abstract*—Network based methods are suitable for the analysis of large number of financial time series and the better understanding of their interdependencies. Known approaches to reveal the underlying information about the complex structure of these interdependencies include network-wise and vertex-wise measures of the topology, as well as filtering techniques relying on minimum spanning trees, planar graphs, or spectral analysis. The aim of this study is to review relevant graph theoretical and statistical models and techniques for generating and examining the properties of financial networks, obtained by computing time series correlations or causality relationships. In particular, this study reviews literature discussing the time evolution of the observed phenomena from a network perspective, as well as applications in economy and finance, ranging from risk and diversification, through policy making and better understanding crisis impact, to forecasting. The information synthesized in this paper can be useful to gain further insights into this relatively new research area.**

***Index Terms*— causality, degree stability, financial networks, time-varying graphs, topology**

I. INTRODUCTION

GRAPH theory proposes useful mathematical tools applicable in various fields, including computer science, transportation and communication networks, biology and social sciences. In the field of finance and economics, there exist a growing body of literature that studies how graphs can be applied for modelling and forecasting. A network-based approach would be a natural tool for studying a large number of pairwise relationships between different financial assets, potentially depicting patterns and characteristics of an entire market or economy that cannot be observed if only considering a small number of time series of historical data. Graph representations are suitable for analyzing companies' stocks [1], [2] (e.g. all stocks of an index like S&P 500 or Dow Jones, highly capitalized European stocks, specific industries), but also the relationships between debt instruments [3], currencies [4] and cryptocurrencies [5], commodities [6], or also a mixture of different asset classes. The purpose can be modelling of information dissemination [7]; analyzing the influence of stock returns on another stocks [2], [8]; understanding the propagation of volatility and market shocks [9], [10], and also economic forecasting on the basis of existing data and the current market state. As discussed in the following, many well-known methods originating from the graph theory can be utilized in the context of financial networks. In this paper we review some of the most popular analytical approaches and methods considered in the financial networks literature from 2000 to 2020. The strength of relationships between financial assets is typically measured by computing correlations, lagged cross-correlations [11], [12], or causality. Sect. II provides a general framework for constructing a graph using this information. Sect. III discusses the literature considering Minimum Spanning Trees as a modelling tool. The vertex degree distributions of financial networks, and how they differ from random graphs, is discussed in Sect. IV. Sect. V outlines the applicability of Granger causality and other causality measures for identifying directional relationships, and in the Sect. VI some conclusions are discussed.

P. Tsankov is PhD student of the Faculty of Applied Mathematics and Informatics, FAMI, Technical University of Sofia, 1000 Sofia, Bulgaria (e-mail: p_tsankov@yahoo.fr).

II. NETWORK-BASED REPRESENTATION OF FINANCIAL MARKETS

A weighted graph $G = \Gamma(V_G, E_G)$ is a structure of vertices (also called nodes) from the vertex set denoted $V_G = \{v_1, v_2, \dots, v_n\}$, connected by edges from the edge set denoted $E_G = \{e(v_i, v_j, w) : v_i, v_j \in V_G\}$, where w is the edge weight. An edge $e(v_i, v_j, w)$ belongs to E_G if and only if the vertices v_i and v_j are connected. The graph G is directed (a digraph) if an edge has a beginning vertex and an end (target) vertex; otherwise the graph is undirected i.e. $e(v_i, v_j)$ and $e(v_j, v_i)$ is the same edge. Constructing a financial network does not require parallel directed edges, for instance if v_k and v_l are two stocks and edge weights correspond to correlations [1] between price returns, it is sufficient to create one single edge between v_k and v_l, in order to represent this information. Correlation based graphs are discussed in [1], [4], [11], [13]-[19]. On the other hand, unlike correlation, causality [2, 8-10, 20] is a directed measure of dependency, hence directed edges can be used to designate whether v_k influences v_l, or vice versa (for more details, see Sect. V). Most of the research known from the literature in this field relies on computing the correlation or measuring the causality between the log-return (1) of prices or realized volatility (2) time series, corresponding to different financial assets.

$$r_{i,t} = \log(x_{i,t}) - \log(x_{i,t-1}) \tag{1}$$

https://doi.org/10.47978/TUS.2021.71.01.001

$$\sigma_{i,t} = 100 \sqrt{\frac{1}{n}\sum_{k=t-n}^{t} r_{i,k}^2} \quad (2)$$

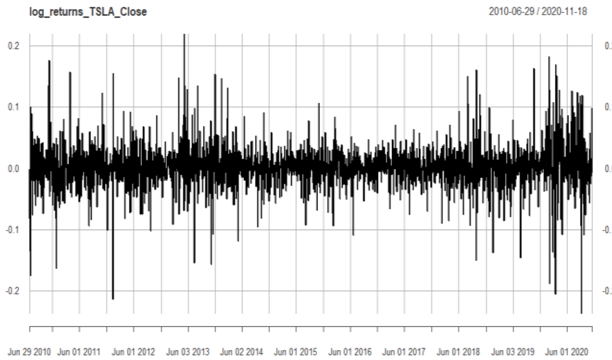

Fig. 1. Log-returns of TSLA (Tesla, Inc.) closing prices for the period between 29-Jun-2010 and 18-Nov-2020, calculated using Eq. (1).

Log-returns and volatilities can obviously be calculated from the price of any financial asset: Foreign exchange [15], [21]; American stocks [1], [7], [13]; Korean stocks [22]; European Government bonds are discussed in [3]; Cryptocurrencies [5], [23]; Oil prices [6]; Furthermore, a single network can represent a market or an economy over a long period of time, but it is also possible to create time-varying graphs for smaller consecutive overlapping or non-overlapping periods of time [9], [2]. This approach is appropriate when the objective is to analyze how the network changes in response to a financial crisis, and what is the time-evolving behavior of interdependencies between different assets.

III. Topology and Spanning Trees

Analysis and comparison of financial graphs can involve various techniques to examine connectivity, vertex degrees, central regions and important vertices, presence of clusters and homogeneity, surviving edges and vertices, as well as network robustness. The obtained networks can be highly complex, and a natural and straightforward approach is to analyze their spanning trees. The Minimum Spanning Trees (MSTs) [24] are a well-known graph theoretical tool, popular for constructing power line, or road networks, in computer science, and more recently in the context of financial networks. A MST is a tree of $n-1$ edges connecting all n vertices of a weighted graph, whose sum of the edge weights is as small as possible. There exist several algorithms for extracting MSTs, such as Kruskal's algorithm [25] and Prim's algorithm [26], which can be applied on a graph obtained as described in the previous section.

The analysis of pairwise relationships between assets, as discussed in Section II, would initially produce a fully connected weighted graph, hence, from this point of view, extracting MSTs can be considered as a filtering procedure [14], which highlights relevant relationships between assets [15]. Weak correlations or causal relationships are simply excluded from the resulting trees, as they correspond to edges with large weights. For instance, for the case of correlation-based trees, we can apply one of the following formulas in order to transform the correlation coefficient ρ into an edge weight:

$$w_{i,j} = \sqrt{2(1-\rho_{i,j})} \quad (3)$$

$$w_{i,j} = 1 - |\rho_{i,j}| \quad (4)$$

For both (3) and (4), i and j correspond to two vertices/assets.

Weights obtained using (3) can be interpreted as distances, as emphasized in [15] and [1], which means that strongly correlated time series would be represented by vertices which are close to each other, hence retained by the MST filtering procedure (e.g. Kruskal's algorithm). On the other hand, (4) would lead to preserving strongly anti-correlated pairs as well, as correlation ρ close to -1 also represents a meaningful relationship between the random variables (it is convenient to mention here the possibility of *short selling* in financial markets). For the case of causality, a similar approach can be used if the strength of the causal relationship is measured.

Existing real-world relationships between assets are more easily observable when using MSTs, and high-degree vertices in the center can be interpreted as important and influential assets [12]. Furthermore, clusters of closely related assets can be observed in the MSTs, as a result from the connectedness of vertices which are similar in terms of correlation/causality. For example, the analysis of the foreign exchange market performed in [4] shows that currencies of countries within geographically close regions are represented by nodes which are linked to each other and clustered together in the Minimum Spanning Trees. As can be expected, another finding, known from of [14], [4] and [17] is that the obtained trees are economically relevant, and major economic countries with stable economies tend to occupy central positions within the MST clusters, having higher vertex degrees.

These properties of MSTs are still relevant for the stock markets [13], [14], [1], [16], both for return-based and volatility-based networks, where highly capitalized companies are cluster centers. However, it is interesting whether the overall tree structure varies depending of the sampling frequency of the data. From a financial perspective, the period between data snapshots mainly depends on the type of business, for example in portfolio allocation and management the frequency is typically low (e.g. one data point per day or a longer period), human-driven trading and hedging of derivative instruments relies on intraday data, and in the case of market making and low latency trading millisecond/nanosecond timestamps are usual. In particular, for high frequencies, the market liquidity and intensity of quotes become important and decreasing cross-correlation for smaller time intervals (known as the Epps effect [27]) can be observed. An analysis of the stock market (NYSE) has been performed in [1], where the authors find that MST-based hierarchy of stocks changes for different sampling frequencies. The multi-cluster structure observed at lower frequencies tends to disappear, and the star-like structure (see also Sect. IV) of the trees emerges for smaller time-intervals. According to the MSTs obtained in [1], intra-sector correlation weakens faster than inter-sector correlation, when shrinking the sampling time window. The analysis given in [28] shows that the clustered structure of correlation-based MSTs can be preserved even at 5-minute sampling, by applying convenient filtering methods, such as removing the dominant "market mode" by zeroing the largest eigenvalue of the correlation matrix (see [28] for more details). Intraday lagged correlation patterns are discussed in detail [29] for the case of Bonferroni corrected networks [30], indicating to what extent the network topology could change between small time intervals during the trading day. Nonetheless, the network edges are largely persistent throughout the trading day, which is consistent with the findings of [28]. An analysis of MST structure depending on intraday volatility is performed in [22], for the case of Korean stock markets. The authors revealed that the normalized tree length (NTL), which measures the closeness among the components of network, decreases in periods of high volatility. Higher maximum vertex degree have also been observed during market stress periods, which is consistent with shorter tree lengths and network shrinkage, as this holds true for both intraday and low-frequency data.

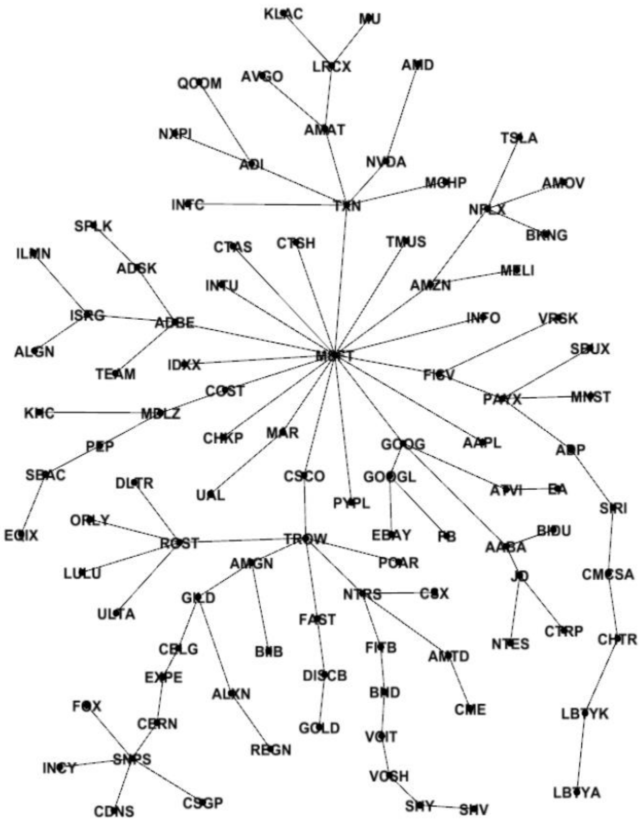

Fig. 2. Example from [23] of a correlation-based Minimum Spanning Tree (MST) obtained from the log-returns of 100 highly capitalized NASDAQ stocks, for the period between 16-Dec-2017 and 30-Apr-2019. Vertex labels correspond to stock symbols (tickers).

A natural approach to analyze and compare the obtained networks, either trees or not, is to study the distribution of their vertex degrees, which is discussed in more detail in Sect. IV. Apart from in-degree/out-degree centrality [2], [31] and network diameter, there exist several centrality measures popular in network analysis, applicable as well in the context of financial graphs. The closeness centrality [27], [32], which is suitable for connected directed graphs (see for example Sect. V for the case of causality graphs), is the average path length of the shortest paths between a vertex and all other vertices in the graph. The harmonic centrality (5), which is a slight modification of closeness centrality ([27] and the references therein), is the harmonic mean of distances between every pair of distinct vertices. These are natural measures of the level of connectedness of a market, as for example, higher harmonic centrality corresponds to a more densely connected market [9], [2]. If we denote with $d(v_i, v_j)$ the distance between vertices v_i and v_j, the harmonic centrality is defined as follows:

$$hc(v_i) = \sum_{i \neq j} \frac{1}{d(v_i, v_j)} \quad (5)$$

The normalized betweenness centrality [8], [33] and [22] is useful to determine important vertices within the network, which can be considered as information/shock transmission mediums. It is computed by counting the number of shortest paths between a pair of vertices (v_j, v_k) that pass through a given vertex v_i, denoted $p_{jk}(v_i)$, and the total number of shortest paths p_{jk} between (v_j, v_k):

$$bc(v_i) = \frac{1}{N^2 - 3N + 2} \sum_j \sum_{k>j} \frac{p_{jk}(v_i)}{p_{jk}}, \; i \neq j \neq k \quad (6)$$

According to [27], "…the intuition behind betweenness is that if a large fraction of shortest paths passes through v_x, then v_x is an important junction point of the network". Another measure, applicable for evaluating relative node importance, is the PageRank centrality [34], based on eigenvalues of the graph's adjacency matrix [27]. Other potentially interesting measures for the importance of a vertex are the Katz centrality [35] and the Bonacich centrality [36].

All these network measures and statistics can be useful for investigating the behaviour of time-varying graphs, and to compare networks obtained over time windows of different length. Advantages of comparing graphs corresponding to different time periods includes: the possibility of implementing and improved filtering in the context of diversification and portfolio management [19], [14]; evaluation of risk spillovers using volatility-based networks [37], [9], and [32]; better understanding of market shocks global impact and precisely measuring the changes of the network topology during periods of market turbulence [38], [39], and [7]. Vertex-wise centrality measures can be applied on a sequence of networks estimated on a rolling window of historical data, in order to produce time series that describe the time evolution of the influence of specific assets (e.g. company stocks, currencies), or the average importance of entire economic sectors as well. Furthermore, time series can be used to describe the properties of an entire network using *network-wise* measures, for instance how the diameter, average/maximum degrees [22] and density change pre, during and post-crisis, and to what extent the interactions between network core and periphery vary in time, which facilitates to figure out how the available information may be used for forecasting. From this perspective, it might be also interesting to study the edge survival ratios [40], with the purpose to examine the network stability over long periods (or also, at the occurrence of market shocks), and the robustness of the model and results [37], [9] and [2]. Moreover, the above is applicable for the comparison of different markets (e.g. European, US and APAC stock markets) and the time-evolution patterns of different asset classes.

IV. VERTEX DEGREES AND DEGREE STABILITY

In a network where vertices represent different companies' stocks, a high vertex degree i.e. a densely connected vertex indicates that the corresponding company is highly influential to other companies [12]. This interpretation can easily be generalized to other financial assets. Whether a specific value of the degree can be considered high, obviously depends on the network, for instance, in [17], which analyses the currency market, vertices with degrees higher than two are considered to be central, but in the case of stock correlation networks of hundreds of nodes, the maximum degrees can reach values between 10 and more than 100 [1], [13], [14], [16], [18], [19], depending on the network size. Unlike the randomly growing graphs introduced by P. Erdos and A. Rényi [41], for which the degree distribution has a Poissonian form, many real-world networks growing via preferential attachment exhibit power-law distributed vertex degrees [42], according to the following formula: $P_D(k) \rightarrow ck^{-\alpha}$ where α is the scaling parameter, and c is a normalization constant. A network following power-law in terms of vertex degrees is called a *scale-free network*, and this effect has been observed for various financial graphs, incl. stock correlation networks. In practice, this means that while the network grows, its diameter remains almost constant (the so called *small world* property). The typical values of α vary between 2 and 3 for correlation-based MSTs [43], [16], and can decrease to values lower than 1.8 during crisis periods [40]. Smaller value of α correspond to shrinking network diameters, and inversely, larger diameters / values of α are observed in calm market conditions.

The authors of [16] compared fixed-size graph generated using two random simulation models (uncorrelated multivariate Gaussian returns time series and one-factor model representing market and stock returns), and a stock correlation graph derived from the daily closing prices of 1071 stocks continuously traded at the NYSE over a 12-years period. They extracted the minimum spanning trees (see also the previous section) from three networks and investigated their topologies and node degree distributions. In the case of the multivariate Gaussian model, the observed distribution of degrees tends to a Poissonian form, and the obtained tree has only a few nodes with a degree greater than a few units, which also means that long files of vertices are observed. The one-factor model produces a star-like tree structure with power-law distributed degrees and one single central vertex. However, both random simulation models considered in [16], cannot capture the complexity of the real-world tree, where the degree distribution is confirmed to follow a power law, but there are several high-degree nodes in the center and more complex hierarchy is observed. Indeed, if we consider degree as a measure of the importance of a vertex, real-world networks exhibit a more structured hierarchy of the stocks' importance, which cannot be reproduced using the considered simulation models.

The time-varying degree of vertices gives us information about whether the network topology is stable over time and how often structural changes can be observed. An analysis of the degree stability is carried out in [13], where the authors constructed return-based and volatility based minimum spanning trees (MSTs) from 93 highly capitalized stocks traded over a 12 years' period. They noted that the volatility based networks are more stable if Spearman rank-order correlation coefficient is used instead of Pearson correlation coefficient, since the probability density functions of volatility time series are highly skewed, unlike the case of returns. For both return and volatility MSTs they considered a rolling time window, computed series of time-varying degrees for all vertices (i.e. the importance of different stocks), and measured autocorrelation functions (ACFs) of the obtained series. The results of [13] indicate that volatility networks and MSTs are less stable over time compared to the case of returns, since the ACF values are significantly lower for degree series calculated on the basis of the volatility MSTs. Furthermore, the degrees of stocks have a slowly changing dynamics, with long range memory of approximately 3 calendar years for return based MSTs and 6 years for volatility MSTs. It should be noted that this analysis can be extended to larger networks (for example, a large number of stocks), different markets and assets, and different periods. The Kolmogorov-Smirnov (KS) test can also be performed to compare the degree distributions of different MSTs, hence to quantify the difference between their structures, as suggested in [28].

In addition to studying the empirical distribution of the degrees for various financial graphs, the concept of vertex degree can also be used in the graph construction process, in order to control the topology and the formation of clusters and cliques. An algorithm, called *Proportional Degree* (PD) is introduced in [18] in such a way that the vertex degrees in the obtained graphs are proportional to the importance of stocks. If the similarity between two assets is the Pearson correlation or the Mutual information between their corresponding log-return (volatility) series, the sum of similarities between a particular asset and all other assets in the dataset, i.e. the sum of the edge weighs between this asset and its first degree neighbors (in the case of a fully connected undirected graph as described in Section II) can be represented as an asset importance. The PD algorithm is an iterative procedure which starts with an empty graph and allows the creation of edges between vertices only if their current degree is less than a boundary value proportional to their estimated importance. In other words, the networks generated using this algorithm would not contain densely connected vertices with small importance compared to all other vertices. As consequence, the properties of the obtained graphs will differ from fully connected graphs, MSTs [13], [1], asset graphs [44], planar graphs [45], [46], and [47], etc... In the particular case of networks where vertices correspond to companies' stocks, the quality criterion of the results could be to observe stocks belonging to the same industrial sector to be clustered together in the network. According to the analysis conducted in [18], the Proportional Degree algorithm produces more homogeneous maximal 3-cliques and 4-cliques from the perspective of sectoral classification, compared to planar graphs obtained using the planar maximally filtered graphs (PMFG) method [45], [47]. Furthermore, the networks obtained using the PD algorithm are more robust with respect to removing random edges compared to PMFG-based networks, and have better partitioning properties.

V. Lead-lag relationships, causality and directed graphs

The analytical approaches for studying topological properties and degree distributions discussed above are generally applicable for different types of graphs, including planar graphs, directed graphs, MSTs, and they do not depend on the exact model for creating edges in function of time series information. Correlation-based networks are widely used, but they do not take into account relationships such that an asset influences, but is not influenced by, another one. From a network perspective, it is necessary to use directed edges, as a directed edge between a pair of assets (e.g. currencies) U and V will only be created if a lead-lag / causal relationship is detected. For instance, if the cross-correlation between two series is highest at $lag \neq 0$, the next step would be to identify whether the time-lagged values of the series V are those that have influence on U, or the opposite, i.e. what is the sign of the lag, and finally, to create an edge [48], [49], and [50]. Unlike cross-correlation and partial correlation [51] computations, which provide a direct estimate for the strength of linear lead-lag relationship between stochastic processes [21], a causality test gives information on the predictive power of one process to another. Formally, let $\{r_{U,t}\}_{0\leq t\leq T}$ and $\{r_{V,t}\}_{0\leq t\leq T}$ be the log-returns of U and V. Let us denote by $\Omega_{U,t}$, $\Omega_{V,t}$ the information sets of all returns known at t for U and V respectively. The null hypothesis (7) means that knowledge of the past time-lagged information relevant to V does not improve the prediction ability for the future values of U.

$$H_0 : \mathbb{E}(r_{U,t} \mid \Omega_{U,t-1}) = \mathbb{E}(r_{U,t} \mid \Omega_{U,t-1}, \Omega_{V,t-1}) \qquad (7)$$

On the other hand, if the null hypothesis is rejected, i.e. taking into account the information for V available between 0 and $t-1$ changes the expectation of $r_{U,t}$

$$H_1 : \mathbb{E}(r_{U,t} \mid \Omega_{U,t-1}) \neq \mathbb{E}(r_{U,t} \mid \Omega_{U,t-1}, \Omega_{V,t-1}) \qquad (8)$$

we say that V causes U ($V \rightarrow U$). It is sufficient for $r_{V,t-1}$ to have a statistically significant impact on $r_{U,t}$ (i.e. at $lag = 1$) to assume that a causal relationship exists, but notice that $r_{U,t}$ can be influenced as well by older values $r_{V,t-k}, k > 1$. *Granger causality* [20], along with the *Vector Autoregressive Model* (VAR) [52], [39] are the baseline used for the inference of causal networks in finance [2], [34], [8], [33], [9], and [32]. However, it is important to note that the relationships between time series in finance are very complex and it is useful to consider non-linear models for testing causality.

Unlike cross-correlation and Granger causality, the *Mutual information* [53], is a measure based on the Kullback–Leibler divergence, able to capture both linear and non-linear dependencies. *Transfer entropy* [54], can be seen as an improvement of Mutual Information, which uses appropriate transition probabilities computations in order to better "distinguish information that is actually exchanged from shared information due to common history and input signals". *Rényi Transfer Entropy* is also a model-free measure capable to detect non-linear interactions, discussed in [55] in the context of detecting spillovers of rare events (market drawdowns) at the network-wide level. *Causation Entropy* and an algorithm for *causal network inference* (called oCSE) has been developed in [56], as an improvement of the Conditional Mutual Information and the Transfer Entropy. According to the authors, some remaining problems need to be addressed, in order to make this measure applicable in a wider practical context. A discussion of alternative approaches to detect causality and measures with good properties can be found in [57], [58], and [38].

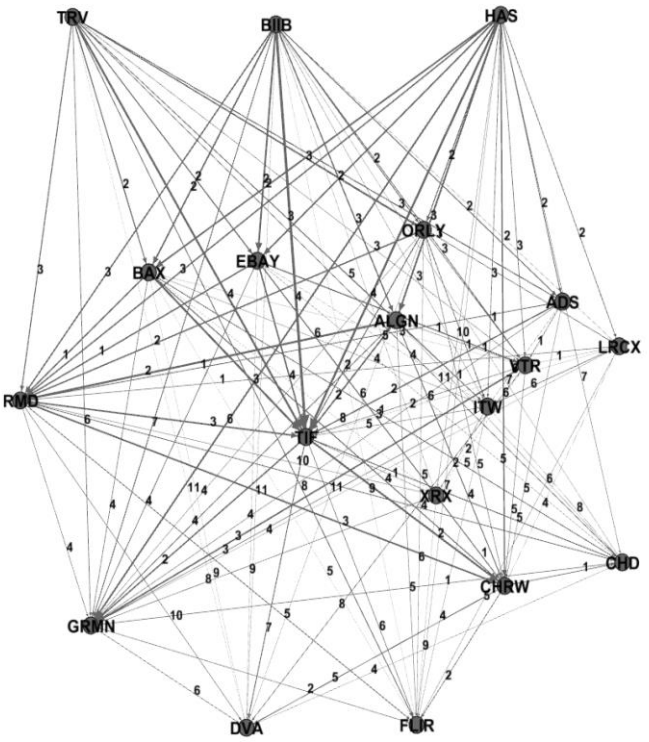

Fig. 3. Example from [59] of a Directed Acyclic Graph (DAG) with retained only the vertices with $\deg(v) \geq 20$, extracted from a log-return based causality graph of S&P500 stocks for the period 01-Oct-2019 to 31-Dec-2019. Vertex labels correspond to stock tickers; edge labels correspond to estimated lags of the causal relationships.

Causality networks, similarly to correlation-based networks, can be useful for identifying the vertices which play a role of information diffusion mediums [33]. In the context of transmission of market shocks and measuring systemic risks, causality inference allows to explain the volatility of an asset by taking into account not only its own past volatilities, but the past volatility series of other assets, as well (see [37] and [60]). This is applicable not only on an individual vertex level, but also for network clusters and industrial sectors in the case of company stocks, allowing to find risk spillover propagators. Power-law tail-exponents of the in and out degrees, as discussed in the previous section, can be further analyzed from the perspective of time-varying causality graphs. Jump growth of node degrees of global stock market indices networks have been observed after the stock market crash in Asian-Pacific region (Jan 2008), after announcements of significant losses (HSBC in Jan 2007, Deutsche Bank in Feb 2016), during the period of the subprime mortgage crisis and debt crises (Greece in Apr 2010, EU in Sep 2011), and after the Brexit vote (Jun 2016) [8]. According to [9] and [32], this information can improve the forecasting of returns and volatility, and moreover, the predictability of extreme risk in the economy and to help constructing early indicators. The time evolution and stability of volatility spillover networks was studied in [2] and [9] by calculating the surviving ratios of edges. The topological properties and core-periphery interactions are susceptible to change during market distress and crisis, which were analyzed [8]-[10], [31], [60], and [61].

VI. Conclusion

Graph theory, combined with probability and statistics, can provide useful tools for large-scale analysis of financial data. This study discusses models and methods for constructing correlation or causality based financial networks, and also explores their fundamental properties. Furthermore, the literature reviewed in this paper considers some sophisticated methods for filtering and analyzes some networks depending on the context of application and the type of input data used (e.g. asset class, sampling frequency, volume and time horizon). Taking into account this variety of methods can lead to the identification of potentially useful new tools and clues in what direction to expand the research. Given the wide range of practical contexts and applicable analytical methods, as a future work we consider to perform more comparisons between existing methods and models, using different data sets. In addition, we believe a good direction would be to investigate further the clustering behavior within financial networks which evolve over time.

References

[1] G. Bonanno, G. Caldarelli, F. Lillo, S. Micciche, N. Vandewalle, and R. N. Mantegna, "Networks of equities in financial markets," *The European Physical Journal B*, vol. 38, no. 2, pp. 363–371, 2004. https://doi.org/10.1140/epjb/e2004-00129-6

[2] T. V`yrost, Š. Lyócsa, and E. Baumöhl, "Granger causality stock market networks: Temporal proximity and preferential attachment," *Physica A: Statistical Mechanics and its Applications*, vol. 427, pp. 262–276, 2015.

[3] G. JUODŽIUKYNIEN˙ E, "Financial contagion among new member states of the european union: Granger causality approach," *DEVELOPMENT IN ECONOMICS: THEORY AND PRACTICE*, p. 26.

[4] T. Mizuno, H. Takayasu, and M. Takayasu, "Correlation networks among currencies," *Physica A: Statistical Mechanics and its Applications*, vol. 364, pp. 336–342, 2006. https://doi.org/10.1016/j.physa.2005.08.079

[5] G. Milunovich, "Cryptocurrencies, mainstream asset classes and risk factors: A study of connectedness," *Australian Economic Review*, vol. 51, no. 4, pp. 551–563, 2018. https://doi.org/10.1111/1467-8462.12303

[6] M. Jiang, X. Gao, H. An, H. Li, and B. Sun, "Reconstructing complex network for characterizing the time-varying causality evolution behavior of multivariate time series," *Scientific reports*, vol. 7, no. 1, pp. 1–12, 2017. https://doi.org/10.1038/s41598-017-10759-3

[7] L. S. Junior, "Dynamics in two networks based on stocks of the us stock market," *arXiv preprint* arXiv:1408.1728, 2014.

[8] Q. Zheng and L. Song, "Dynamic contagion of systemic risks on global main equity markets based on granger causality networks," *Discrete Dynamics in Nature and Society*, vol. 2018, 2018. https://doi.org/10.1155/2018/9461870

[9] E. Baumöhl, E. Koˇcenda, Š. Lyócsa, and T. V`yrost, "Networks of volatility spillovers among stock markets," *Physica A: Statistical Mechanics and its Applications*, vol. 490, pp. 1555–1574, 2018. https://doi.org/10.1016/j.physa.2017.08.123

[10] F. Corsi, F. Lillo, D. Pirino, and L. Trapin, "Measuring the propagation of financial distress with granger-causality tail risk networks," *Journal of Financial Stability*, vol. 38, pp. 18–36, 2018. https://doi.org/10.1016/j.jfs.2018.06.003

[11] B. Podobnik, D. Wang, D. Horvatic, I. Grosse, and H. E. Stanley, "Timelag cross-correlations in collective phenomena," *EPL (Europhysics Letters)*, vol. 90, no. 6, p. 68001, 2010. https://doi.org/10.1209/0295-5075/90/68001

[12] G. Bonanno, F. Lillo, and R. N. Mantegna, "High-frequency crosscorrelation in a set of stocks," 2001. https://doi.org/10.1080/713665554

[13] S. Miccichè, G. Bonanno, F. Lillo, and R. N. Mantegna, "Degree stability of a minimum spanning tree of price return and volatility," *Physica A: Statistical Mechanics and its Applications*, vol. 324, no. 1-2, pp. 66–73, 2003. https://doi.org/10.1016/S0378-4371(03)00002-5

[14] A. Esfahanipour and S. Zamanzadeh, "A stock market filtering model based on minimum spanning tree in financial networks," *AUT Journal of Modeling and Simulation*, vol. 45, no. 1, pp. 67–75, 2015.

[15] E. Limas, "An application of minimal spanning trees and hierarchical trees to the study of latin american exchange rates," *Journal of Dynamics & Games*, vol. 6, no. 2, p. 131, 2019. https://doi.org/10.3934/jdg.2019010

[16] G. Bonanno, G. Caldarelli, F. Lillo, and R. N. Mantegna, "Topology of correlation-based minimal spanning trees in real and model markets," *Physical Review E*, vol. 68, no. 4, p. 046130, 2003. https://doi.org/10.1103/PhysRevE.68.046130

[17] M. Rešovsk`y, D. Horváth, V. Gazda, and M. Siniˇcáková, "Minimum spanning tree application in the currency market," *Biatec*, vol. 21, no. 7, pp. 21–23, 2013.

[18] S. S. Hosseini, N. Wormald, and T. Tian, "A weight-based information filtration algorithm for stock-correlation networks," *Physica A: Statistical Mechanics and its Applications*, p. 125489, 2020. https://doi.org/10.1016/j.physa.2020.125489

[19] W. Sun, C. Tian, and G. Yang, "Network analysis of the stock market," 2015.

[20] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: journal of the Econometric Society*, pp. 424–438, 1969. https://doi.org/10.2307/1912791

[21] L. Basnarkov, V. Stojkoski, Z. Utkovski, and L. Kocarev, "Lead–lag relationships in foreign exchange markets," *Physica A: Statistical Mechanics and its Applications*, vol. 539, p. 122986, 2020. https://doi.org/10.1016/j.physa.2019.122986

[22] J. Lee, J. Youn, andW. Chang, "Intraday volatility and network topological properties in the korean stock market," *Physica A: Statistical mechanics and its Applications*, vol. 391, no. 4, pp. 1354–1360, 2012. https://doi.org/10.1016/j.physa.2011.09.016

[23] M. Durcheva and P. Tsankov, "Analysis of similarities between stock and cryptocurrency series by using graphs and spanning trees," in *AIP Conference Proceedings*, vol. 2172, no. 1, p. 090004. AIP Publishing LLC, 2019. https://doi.org/10.1063/1.5133581

[24] R. L. Graham and P. Hell, "On the history of the minimum spanning tree problem," *Annals of the History of Computing*, vol. 7, no. 1, pp. 43–57, 1985. https://doi.org/10.1109/MAHC.1985.10011

[25] J. B. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem," *Proceedings of the American Mathematical society*, vol. 7, no. 1, pp. 48–50, 1956. https://doi.org/10.1090/S0002-9939-1956-0078686-7

[26] R. C. Prim, "Shortest connection networks and some generalizations," *The Bell System Technical Journal*, vol. 36, no. 6, pp. 1389–1401, 1957. https://doi.org/10.1002/j.1538-7305.1957.tb01515.x

[27] P. Boldi and S. Vigna, "Axioms for centrality," *Internet Mathematics*, vol. 10, no. 3-4, pp. 222–262, 2014. https://doi.org/10.1080/15427951.2013.865686

[28] C. Borghesi, M. Marsili, and S. Micciche, "Emergence of time-horizon invariant correlation structure in financial returns by subtraction of the market mode," *Physical Review E*, vol. 76, no. 2, p. 026104, 2007. https://doi.org/10.1103/PhysRevE.76.026104

[29] C. Curme, M. Tumminello, R. N. Mantegna, H. E. Stanley, and D. Y. Kenett, "How lead-lag correlations affect the intraday pattern of collective stock dynamics," Available at SSRN 2648490, 2019.

[30] M. Tumminello, S. Micciche, F. Lillo, J. Piilo, and R. N. Mantegna, "Statistically validated networks in bipartite complex systems," *PloS*

one, vol. 6, no. 3, p. e17994, 2011. https://doi.org/10.1371/journal.pone.0017994

[31] S. K. Stavroglou, A. A. Pantelous, K. Soramaki, and K. Zuev, "Causality networks of financial assets," *Journal of Network Theory in Finance*, vol. 3, no. 2, pp. 17–67, 2017. https://doi.org/10.21314/JNTF.2017.029

[32] J.W. Song, B. Ko, P. Cho, andW. Chang, "Time-varying causal network of the korean financial system based on firm-specific risk premiums," *Physica A: Statistical Mechanics and its Applications*, vol. 458, pp. 287–302, 2016. https://doi.org/10.1016/j.physa.2016.03.093

[33] X. Gao, S. Huang, X. Sun, X. Hao, and F. An, "Modelling cointegration and granger causality network to detect long-term equilibrium and diffusion paths in the financial system," *Royal Society open science*, vol. 5, no. 3, p. 172092, 2018. https://doi.org/10.1098/rsos.172092

[34] Y. Tang, J. J. Xiong, Y. Luo, and Y.-C. Zhang, "How do the global stock markets influence one another? evidence from finance big data and granger causality directed network," *International Journal of Electronic Commerce*, vol. 23, no. 1, pp. 85–109, 2019. https://doi.org/10.1080/10864415.2018.1512283

[35] L. Katz, "A new status index derived from sociometric analysis," *Psychometrika*, vol. 18, no. 1, pp. 39–43, 1953. https://doi.org/10.1007/BF02289026

[36] P. Bonacich, "Power and centrality: A family of measures," *American journal of sociology*, vol. 92, no. 5, pp. 1170–1182, 1987. https://doi.org/10.1086/228631

[37] G.-J. Wang, C. Xie, K. He, and H. E. Stanley, "Extreme risk spillover network: application to financial institutions," *Quantitative Finance*, vol. 17, no. 9, pp. 1417–1433, 2017. https://doi.org/10.1080/14697688.2016.1272762

[38] A. Papana, C. Kyrtsou, D. Kugiumtzis, and C. Diks, "Financial networks based on granger causality: A case study," *Physica A: Statistical Mechanics and its Applications*, vol. 482, pp. 65–73, 2017. https://doi.org/10.1016/j.physa.2017.04.046

[39] F. X. Diebold and K. Yılmaz, "On the network topology of variance decompositions: Measuring the connectedness of financial firms," *Journal of Econometrics*, vol. 182, no. 1, pp. 119–134, 2014. https://doi.org/10.1016/j.jeconom.2014.04.012

[40] J.-P. Onnela, A. Chakraborti, K. Kaski, J. Kertesz, and A. Kanto, "Dynamics of market correlations: Taxonomy and portfolio analysis," *Physical Review E*, vol. 68, no. 5, p. 056110, 2003. https://doi.org/10.1103/PhysRevE.68.056110

[41] P. Erdos and A. Rényi, "On the evolution of random graphs," Publ. Math. Inst. Hung. Acad. Sci, vol. 5, no. 1, pp. 17–60, 1960.

[42] L. Lovász, Large networks and graph limits, vol. 60. *American Mathematical Soc.*, 2012. https://doi.org/10.1090/coll/060

[43] N. Vandewalle, F. Brisbois, X. Tordoir et al., "Non-random topology of stock markets," *Quantitative Finance*, vol. 1, no. 3, pp. 372–374, 2001. https://doi.org/10.1088/1469-7688/1/3/308

[44] J.-P. Onnela, A. Chakraborti, K. Kaski, J. Kertesz, and A. Kanto, "Asset trees and asset graphs in financial markets," *Physica Scripta*, vol. 2003, no. T106, p. 48, 2003. https://doi.org/10.1238/Physica.Topical.106a00048

[45] M. Tumminello, T. Aste, T. Di Matteo, and R. N. Mantegna, "A tool for filtering information in complex systems," *Proceedings of the National Academy of Sciences*, vol. 102, no. 30, pp. 10 421–10 426, 2005. https://doi.org/10.1073/pnas.0500298102

[46] G.-J.Wang, C. Xie, and S. Chen, "Multiscale correlation networks analysis of the us stock market: a wavelet analysis," *Journal of Economic Interaction and Coordination*, vol. 12, no. 3, pp. 561–594, 2017. https://doi.org/10.1007/s11403-016-0176-x

[47] G.-J.Wang, C. Xie, and H. E. Stanley, "Correlation structure and evolution of world stock markets: Evidence from pearson and partial correlationbased networks," *Computational Economics*, vol. 51, no. 3, pp. 607–635, 2018. https://doi.org/10.1007/s10614-016-9627-7

[48] M. Eryi˘git and R. Eryi˘git, "Network structure of cross-correlations among the world market indices," *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 17, pp. 3551–3562, 2009. https://doi.org/10.1016/j.physa.2009.04.028

[49] B. Podobnik, D. Horvatic, A. M. Petersen, and H. E. Stanley, "Crosscorrelations between volume change and price change," *Proceedings of the National Academy of Sciences*, vol. 106, no. 52, pp. 22 079–22 084, 2009. https://doi.org/10.1073/pnas.0911983106

[50] W.-Q. Duan and H. E. Stanley, "Cross-correlation and the predictability of financial return series," *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 2, pp. 290–296, 2011. https://doi.org/10.1016/j.physa.2010.09.013

[51] D. Y. Kenett, M. Tumminello, A. Madi, G. Gur-Gershgoren, R. N. Mantegna, and E. Ben-Jacob, "Dominating clasp of the financial sector revealed by partial correlation analysis of the stock market," *PloS one*, vol. 5, no. 12, p. e15032, 2010. https://doi.org/10.1371/journal.pone.0015032

[52] J. Geweke, "Measurement of linear dependence and feedback between multiple time series," *Journal of the American statistical association*, vol. 77, no. 378, pp. 304–313, 1982. https://doi.org/10.1080/01621459.1982.10477803

[53] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

[54] T. Schreiber, "Measuring information transfer," *Physical review letters*, vol. 85, no. 2, p. 461, 2000. https://doi.org/10.1103/PhysRevLett.85.461

[55] J. Korbel, X. Jiang, and B. Zheng, "Transfer entropy between communities in complex financial networks," *Entropy*, vol. 21, no. 11, p. 1124, 2019. https://doi.org/10.3390/e21111124

[56] J. Sun, D. Taylor, and E. M. Bollt, "Causal network inference by optimal causation entropy," *SIAM Journal on Applied Dynamical Systems*, vol. 14, no. 1, pp. 73–106, 2015. https://doi.org/10.1137/140956166

[57] A. Péguin-Feissolle, S. Strikholm, and T. Teräsvirta, "Testing the granger noncausality hypothesis in stationary nonlinear models of unknown functional form," *Communications in Statistics-Simulation and Computation*, vol. 42, no. 5, pp. 1063–1087, 2013. https://doi.org/10.1080/03610918.2012.661500

[58] D. Kugiumtzis, "Direct-coupling information measure from nonuniform embedding," *Physical Review E*, vol. 87, no. 6, p. 062918, 2013. https://doi.org/10.1103/PhysRevE.87.062918

[59] M. Durcheva and P. Tsankov, "Granger causality networks of S&P 500 stocks," 2020, to be published.

[60] P. Mazzarisi, S. Zaoli, C. Campajola, and F. Lillo, "Tail granger causalities and where to find them: Extreme risk spillovers vs spurious linkages," Available at SSRN 3591958, 2020. https://doi.org/10.2139/ssrn.3591958

[61] C. Hendahewa and V. Pavlovic, "Analysis of causality in stock market data," in 2012 *11th International Conference on Machine Learning and Applications*, vol. 1, pp. 288–293. IEEE, 2012. https://doi.org/10.1109/ICMLA.2012.56