

DATA 698: Capstone Methodology

Derek G Nokes & Paul Britton

2019-10-27

Methodology

In the following section, we detail our methodology for both establishing the stylized facts pertaining to return cross-correlations in the instrument universe under study, and determining whether or not the introduction of risk controls - primarily those intended to exploit the hierarchical structure of return cross-correlations - can significantly improve the risk-adjusted performance of a classical long-only momentum trading system. Section I describes the data used. Section II provides an overview of our approach for establishing that instrument return cross-correlations are non-random with significant hierarchical structure. In section III, we describe the base momentum trading system and introduce risk controls to be tested, namely a position-sizing approach intended to exploit the well-documented hierarchical structure of cross-correlations. Finally, in section IV, we outline our approach for establishing whether or not the risk controls introduced lead to statistically significant risk-adjusted performance improvements over the base momentum system.

I. Data

We study the characteristics of U.S. equity markets using an instrument universe comprised of all past and current constituents of the S&P 1500 index for the period 1994-11-30 to 2019-10-22. This universe includes large-, medium-, and small-capitalization single stocks across all sectors and industries. We use both price and instrument master data. Prices are adjusted for corporate actions - such as dividends and splits - and thus facilitate the calculation of total returns for each instrument. Instrument master data for each instrument includes a security identifier, security name, a business summary of the associated legal entity, and industry classification (GICS) information. Business summaries and GICS are not available for all instruments¹, particularly index constituents that exited the index long ago.

II. Cross-Correlations

The trading model presented in section III includes a set of filters to identify a subset of our full stock universe with both time series (absolute) and cross-sectional (relative) momentum to form a *momentum universe*. The characteristics of this subset of our full instrument universe is of particular interest because this is the set of stocks that will be used by our strategy to form a portfolio.

For each stock in the momentum universe, we compute the daily (logarithmic) return from the price, then *standardize* by removing the mean and dividing by the standard deviation. We define the daily logarithmic return $r_i(t)$ as

$$r_i(t) = \ln [p_i(t)] - \ln [p_i(t-1)]$$

where the price of instrument i at discrete time t is denoted by $p_i(t)$ ($i = 1, \dots, N$). We then define the rolling standardized return $\hat{z}_i(t)$ of stock i using a lookback of length T as

$$\hat{z}_i(t) = [r_i(t) - \bar{r}_i(T)] / \sigma_i(r_i(T))$$

¹Missing GICS can often be accurately inferred from business summaries using a simple classifier.

where the mean return $\bar{r}_i(T) = \frac{1}{T} \sum_{t=1}^T (\hat{r}(t))$ and the standard deviation of returns $\sigma_i(r_i(T))$ are computed over same time interval T . Given N instrument return series with T elements each, the elements of the $N \times N$ empirical correlation matrix \hat{C} are given by

$$\hat{c}_{i,j} = \frac{1}{T} \sum_{t=1}^T \hat{z}_i(t) \hat{z}_j(t)$$

Representing the standardized returns in matrix form, the empirical correlation matrix² is given by

$$\hat{C} = \frac{1}{T} \hat{Z} \hat{Z}'$$

with elements $\hat{c}(i, j) \in [-1, 1]$. \hat{Z} is the $T \times N$ matrix whose columns are the time series of standardized returns (one for each instrument). \hat{Z}' denotes the transpose of \hat{Z} .

We create a sequence of daily correlation matrices using a rolling time window of length T that advances one day at a time. We use these time-varying correlations to study the evolution of a momentum-filtered subset of the equity universe to examine the co-movement of momentum stocks.

In the next sub-section we introduce an approach to identify significant (i.e., non-random) structure in correlation matrices using Random Matrix Theory (RMT).

Spectral Analysis of Correlation Matrices

The spectral properties of observed return correlation matrices can be compared with those derived from both shuffled [1] and random [2][3][4][5][6] return time series. The following sub-section provides the theoretical basis for the next sub-section where we outline a set of tests for non-random structure in observed rolling correlation matrices.

The correlation matrix C is an $N \times N$ diagonalizable symmetric matrix that can be written in the form [1]

$$C = \frac{1}{T} E D E'$$

where D is a diagonal matrix of eigenvalues d and E is an orthogonal matrix of the corresponding eigenvectors.

If the entries of \hat{Z} are random with variance σ^2 , then - in the limit where $T \rightarrow \infty$ and $N \rightarrow \infty$, such that $Q \equiv T/N$ is fixed [7] - the density of eigenvalues of C is given by [2]

$$\rho_C(\lambda) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda_- - \lambda)}}{\lambda}$$

where the maximum and minimum eigenvalues (λ_+ and λ_- respectively) are [2]

$$\lambda_{\pm} = \sigma^2 \left(1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}} \right)$$

$\rho(\lambda)$ is commonly referred to as the Marcenko-Pastur [8] density.

If N is small relative to T , the empirical correlation matrix \hat{C} approaches the true correlation matrix C .

We can see that as the number of observations T increases relative to the number of instruments N , Q increases and the shape of the eigenvalue density changes.

²Recall that the empirical correlation matrix \hat{C} of returns \hat{Z} is equal to the covariance matrix $\Sigma_{\hat{Z}}$ of \hat{Z} because we have standardized the return time series.

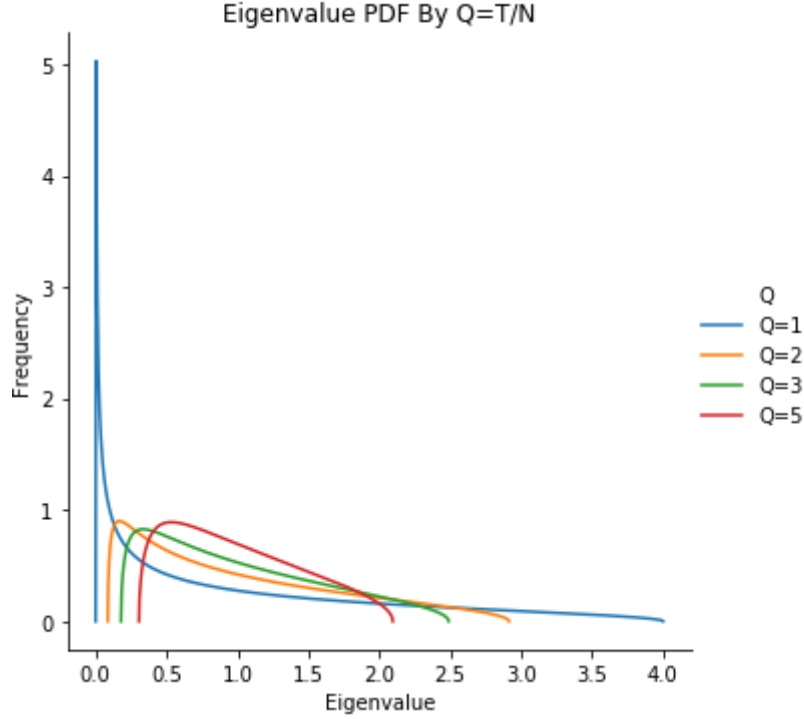


Figure 1: As Q increases, the expected range of eigenvalues decreases. When $Q = 1$ it is difficult to distinguish small eigenvalues from noise⁴.

To a large extent, the return correlation matrix is random and dominated by noise (i.e., most of the eigenvalues of return correlation matrices tend to be within the ranges predicted by RMT)[2][3][4].

We can explore the structure of correlation matrices using PCA [1] to map our high dimensional universe of momentum stocks into a lower dimensional statistical factor space.

Testing for Significant Structure in the Correlation Matrix

To test for the presence of significant non-random structure in the observed rolling correlation matrices, we follow [1] and perform two statistical tests. We generate two synthetic data sets to carry out these tests. For the first synthetic data set, we generate random Gaussian returns for the same number of instruments and time steps as our observed data, then compute rolling correlations using the same time window T as that used for the observed data. For the second synthetic data set, we independently shuffle each column of the observed data destroying the cross-correlation, but maintaining the distribution of instrument returns. We again compute the rolling correlations using time window T .

In the first test, we compare the distribution of all observed correlations from every rolling time window with the correlation distributions derived from each of the synthetic data sets. While the distributions for shuffled and simulated random data are very similar, the distribution of the observed data is markedly different with a significant non-zero mean correlation and a much higher dispersion of correlations around the mean.

In the second test, we compare the distribution of all eigenvalues from every time window derived from the observed returns with the distributions derived from shuffled and simulated random returns. We show that the eigenvalue distribution for observed correlations differs significantly from the equivalent distributions derived from both shuffled and random matrices.

The results of these tests imply significant structure in the return correlations of the stocks under study. In

the next sub-section, we outline a graph-based approach to uncover the hierarchical structure of observed return correlation matrices. In the results section of the paper we provide visualizations of these results.

Correlation-Based Graphs

The spectral properties of return correlations highlight significant structure in the cross-correlations of momentum stocks. Statistical factor models provide a useful mapping from the high-dimensional momentum stock space to the much lower dimensional statistical factor space. Hierarchical clustering arguably provides a more interpretable approach for understanding co-movement among momentum stocks.

Graph theory is used to model pairwise relations between objects [9] and is widely employed in the analysis of complex systems [10][11]. Many, physical, biological, and social systems [12] – including financial markets [13] – are aptly described by networks (i.e., mathematical structures composed of nodes connected by edges)[10]. Common applications of graph theory include methods to extract statistically-reliable information from correlation-based systems [12]. More specifically, graph-based clustering techniques are used to uncover communities (clusters) of similar elements in a network [14][12][15][10]. Hierarchical clustering procedures in which communities are overlapping and organized in a nested structure, can be used to identify the fundamental frame of interactions within a system [16][17][12].

Given a correlation-based system of N elements - where all elements are connected (i.e., they form a *complete* graph) - the pairwise correlation coefficient between each set of elements can be interpreted as the strength of the link (i.e., edge weight) connecting the pairs of elements [12]. Very little information can be gleaned from the topology of such a complete graph without further filtering to selectively remove weakly-connected elements [12]. In principle, there are many different ways of filtering the correlation matrix in order to obtain noise-filtered information [12]. In this work, we focus on a hierarchical graph-based method equivalent to classical single linkage clustering [17].

Using log-return correlation we obtain a set of $N \times (N - 1)/2$ distances characterizing the similarity of any of the N instruments with respect to all of the other $N - 1$ instruments. This set of distances forms a complete graph with different edge strengths given by each correlation value. Every entry in the correlation matrix is associated with a metric distance⁵[19][18][13] representing a pair of instruments i and j according to

$$d_{i,j} = \sqrt{2(1 - c_{i,j})}$$

where $c_{i,j} \in [0, 2]$

The corresponding distance matrix D is used to perform hierarchical clustering on our momentum stock universe.

Hierarchical Clustering

From the fully-connected correlation-based graph outlined above, we extract a sub-graph commonly referred to as the Minimal Spanning Tree (MST). A MST is a subset of the edges of a connected, edge-weighted undirected graph that connects all of the n nodes (vertices) together - without any cycles - in such a way as to minimize total edge weight (i.e., distance)[9]. In other words, the MST is a tree comprised n nodes - one for each stock i - connected with $n - 1$ distance-weighted edges such that the total distance is as small as possible [9]. There are several algorithms for finding the MST [20][21]. For illustrative purposes, we describe the simplest one.

Informally, Prim's algorithm [21] for constructing the minimum spanning tree is as follows:

1. Initialize a tree with a single - arbitrarily chosen - node (vertex) from the graph.

⁵The derivation of this metric and its properties can be found in [18]

2. Grow the tree by one edge by finding the minimum-weight edge (i.e., shortest distance) among the edges that connect the tree to nodes (vertices) not yet in the tree, then add it to the tree.
3. Repeat step 2 until all nodes (vertices) are in the tree.

The topological properties of the MST provide an effective means of summarizing the most essential features of a correlation-based system. More importantly, these features can be exploited in the design of risk controls that can improve the risk-adjusted performance of a basic long-only momentum trading system.

Topological Properties

Some topological properties of a MST derived from a return correlation-based network are useful in identifying key stocks in our instrument universe. The (vertex) degree is the simplest quantity used to characterize the importance of nodes in a network [22]. It measures the number of connections (edges) adjacent to each node (vertex) [22]. The degree k_i of a node i is defined as the sum of the various elements $a_{i,j}$ of the adjacency matrix A ⁶[22]:

$$k_i = \sum_{j=1,n} a_{i,j}$$

The distribution of the node degrees $f(k)$ associated with a MST of a correlation-based network has been shown to follow a power law [23][24] of the following form:

$$f(k) \sim k^{-\alpha}$$

The scaling law associated with the MST of a correlation-based network constructed from simulated random or shuffled observed returns provides a useful set of reference points for identifying statistically significant central nodes (i.e., stocks that have an important position in the hierarchical structure). As with other metrics we study, we estimate the parameters of the scaling law for each rolling window T . A higher value of α represents greater diversity, while a lower value of α represents systems driven by fewer key nodes. We establish statistical significance by comparing the observed α against the distribution of α derived from random and shuffled returns.

III. Momentum Trading Systems

Before we detail the trading system logic for each system variation, we define two cross-sectional momentum metrics. The classical cross-sectional momentum strategy uses the following definition of momentum [25]:

$$\text{RM}_i(t) = \ln [p_i(t)] - \ln [p_i(t - L_M)]$$

where L_M is the lookback window.

We define the risk-adjusted momentum as [26]

$$\text{RMVA}_i = \beta_i \times \frac{250}{L_M} \times R_i^2$$

where β_i is the slope and R_i^2 is the coefficient of determination both obtained from the exponential regression over a time window of length L_M

$$\ln [(P_i)] = \alpha + \beta \ln [X] + \epsilon$$

⁶Denoting n as the number of nodes in a graph, we can define the adjacency matrix A as a square ($n \times n$) tabular representation of the connections in a graph where elements in row i and column j are equal to 1 if there is an edge between node i and j , and equal to 0 otherwise. [22]

where $X = 1 \dots L_M$ and P_i is the vector of prices for stock i over the time interval L_M

We define an instrument as having positive time series momentum if its price is above its exponentially-weighted moving average (EMA).

Base Trading Model

We implement a very simple version of a common systematic equity momentum strategy [26]⁷. For each date in the simulation, we determine our starting instrument universe by looking up the constituents of the S&P 1500 for that date. All proposed variations of the momentum trading system outlined below filter the starting instrument universe each day to screen for stocks with particular characteristics, namely highly ranked cross-sectional momentum and a positive time series trend. Given available capital and a position-level risk budget, each system variation takes positions in as many stocks as possible, beginning with the highest ranked stock and successively working its way down the ranking list until not enough cash remains take a full additional position. When a position is put on, a trailing stop loss is initiated and maintained. If the price of the stock falls below the trailing stop loss level, the position is exited and cash is redeployed to the highest ranked momentum stock for which we do not yet have a position. If the stock is removed from the momentum universe by the screen, the position is exited. The following sub-sections detail the trading system logic.

Momentum Universe Screen

For any given date (t), the momentum universe screen ranks and sub-selects instruments from the full S&P 1500 index universe to form a universe of stocks with desirable momentum characteristics. This screen has four essential pieces: 1) The index filter determines (for a given lookback) the index constituents with enough history to evaluate for trading; 2) The cross-sectional momentum ranking orders stocks according to their relative momentum RM_i ; 3) The trend filter removes stocks that are not trending up, and; 4) The top decile filter reduces the universe to the relative best 10% of momentum stocks. This screen can be updated each date or at a lower frequency such as weekly or monthly.

Portfolio Allocation Logic

Once the momentum universe has been determined, the portfolio allocation logic determines the position size for each instrument and the total number of new positions that can be taken according to the available cash. Before any positions have been taken, the available cash is equal to the account size. On subsequent days, entries into new positions are only triggered by exits from open positions.

Entry Logic

The entry logic is triggered whenever we have enough cash to take a full position. Exiting positions generates changes to our cash level and thus can trigger the entry logic. Beginning with the top-ranked stock - as determined by the momentum screen - we take positions until we do not have enough cash to take another full position. A trailing stop is set and updated each day for all positions.

At t , if we have enough cash to take a position in stock i , we enter a *long* position of $u_i(t)$ units:

$$u_i(t) = \text{floor} \left[\frac{f \times A_i(t-1)}{\text{ATR}_i(t-1) \times M} \right]$$

where A is the account size, f is the fraction of account size risked per bet, $\text{ATR}_i(t-1)$ is the EMA of the true range for the previous time step, and M is the risk multiplier.

True range - a commonly-used measure of the daily price range of a financial instrument that accounts for gaps from the close of the previous period to the open of the current period - is defined as follows:

$$\text{TR}_i(t) = \max [p_{i,H}(t) - p_{i,L}(t), \text{abs}(p_{i,H}(t) - p_i(t-1)), \text{abs}(p_{i,L}(t) - p_i(t-1))]$$

⁷See the reference for a more comprehensive discussion of simple long-only momentum strategies

where $p_{i,H}(t)$ and $p_{i,L}(t)$ are the current daily high and low prices respectively, and $p_i(t-1)$ is the previous close price.

We set our initial stop loss level M units of ATR *below* the entry price level $p_i(t)$. For each subsequent time t we update our stop level $s_i(t)$ as follows:

$$s_i(t) = \max[p_i(t) - \text{ATR}_i(t-1) \times M, s_i(t-1)]$$

Exit Logic

We exit our long position in stock i if the price $p_i(t)$ moves below the stop loss level $s_i(t-1)$ or the stock is no longer in the momentum universe determined by the momentum screen.

Enhanced Trading Model

Our enhanced momentum strategy is identical to the base strategy with one exception - we adapt the position-sizing to exploit the hierarchical structure of the correlation matrix. In particular, we use a recently-developed approach referred to as Hierarchical Risk Parity (HRP)[27].

Sizing positions using Markowitz’s classical mean-variance (MV) portfolio optimization is notoriously unstable. As a result, practitioners have developed a number of alternative approaches, including various forms of volatility-based weightings. Our base momentum trading model uses one such commonly-used volatility-weighting scheme⁸.

When allocating capital based on the classical MV approach, we are effectively using a complete (fully-connected) correlation-based graph where every node is a potential substitute to all others in our allocation decisions [27]. Inverting the correlation matrix is equivalent to evaluating the rates of substitution across the complete graph [27]. Small estimation errors over some of these edges often results in broad and highly impactful changes in allocations [27].

In section II we illustrated that as the number of instruments N in a portfolio increases relative to the time window T used to estimate the correlation matrix, measurement error increases. When N/T is less than one - as is common with large portfolios such as broad-based indices - this error causes significant practical problems. As the components of the portfolio get more correlated, this instability magnifies. By applying a suitable filtration to the fully-connected correlation-based graph (i.e., by dropping insignificant edges), we can significantly improve the stability in our allocations. The MST provides one potential filtering approach with two particularly desirable properties [27], namely: 1) It has only $N - 1$ edges to connect N nodes, so the position weights only rebalance among peers at various hierarchical levels, and; 2) Consistent with how many asset managers build their portfolios, the weights are distributed top-down. In other words, instruments that move together because they belong to the same asset class are treated as non-diversifying substitutes leaving only instruments that exhibit lower co-movement to be considered as diversifying.

The HRP algorithm works in three stages. The first stage groups instruments with similar returns into clusters based on the distance matrix D defined above. The second stage reorganizes the rows and columns of the covariance matrix associated with D so that the largest values lie along the diagonal. The final stage determines the position sizes through recursive bisection of the reordered covariance matrix⁹.

IV. Testing the Effectiveness of Strategy Enhancements

We evaluate the proposed position-sizing improvement by directly comparing the simulated risk-adjusted performance of the base strategy with that of the enhanced strategy. We examine the difference in the time series characteristics of the simulated cumulative total returns for each strategy. We supplement this

⁸There is a broad academic literature - and much practitioner-focused research - about risk-based allocation approaches commonly referred to as ‘risk parity’. Our base model uses the simplest of such approaches.

⁹See [27] for more detail.

comparison of *backtests* with statistical tests intended to uncover the sensitivity of our backtests to the starting instrument universe. We randomly sample S stocks from the starting universe and run a backtest for each variation of the strategy on the subset universe comprised of the S stocks. We repeat this process B times. This procedure generates B total return paths for each strategy variation. For each strategy variation, we create a distribution for each performance metric. We then use these distributions to determine whether enhancements are statistically significant.

We divide the time series data associated with our starting instrument universe into three partitions, forming training, testing, and validation sets. We perform our in-sample parameter selection for each strategy variation of the trading model applying a brute-force grid search on the training set to get a coarse understanding of the parameter space¹⁰. We then refine the grid search for key parameters. Once trading model parameters have been selected, we then evaluate the trading model using the testing set. During the model development process we use only the training and testing sets. Final results are generated truly out-of-sample using the validation set.

References

- [1] D. J. Fenn, M. A. Porter, S. Williams, M. McDonald, N. F. Johnson, and N. S. Jones, “Temporal evolution of financial-market correlations,” *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, vol. 84, no. 2 Pt 2, p. 026109, Aug. 2011.
- [2] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters, “Noise dressing of financial correlation matrices,” *Phys. Rev. Lett.*, vol. 83, no. 7, pp. 1467–1470, Aug. 1999.
- [3] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. Nunes Amaral, and H. E. Stanley, “Universal and nonuniversal properties of cross correlations in financial time series,” *Phys. Rev. Lett.*, vol. 83, no. 7, pp. 1471–1474, Aug. 1999.
- [4] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, and H. E. Stanley, “Random matrix approach to cross correlations in financial data,” *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 65, no. 6 Pt 2, p. 066126, Jun. 2002.
- [5] M. J. Bommarito and A. Duran, “Spectral analysis of time-dependent market-adjusted return correlation matrix,” *Physica A: Statistical Mechanics and its Applications*, vol. 503, pp. 273–282, Aug. 2018.
- [6] D. Y. Kenett, Y. Shapira, and E. Ben-Jacob, “RMT assessments of the market latent information embedded in the stocks’ raw, normalized, and partial correlations,” *J. Probab. Stat.*, vol. 2009, Mar. 2010.
- [7] A. M. Sengupta and P. P. Mitra, “Distributions of singular values for some random matrices,” *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics*, vol. 60, no. 3, pp. 3389–3392, Sep. 1999.
- [8] L. A. P. V. A. Marchenko, “Distribution of eigenvalues for some sets of random matrices,” *Mathematics of the USSR-Sbornik*, vol. 1, no. 4, pp. 457–483, 1967.
- [9] D. B. West and Others, *Introduction to graph theory*, vol. 2. Prentice-Hall, Englewood Cliffs, NJ, 1996.
- [10] M. Newman, *Networks: An introduction*. New York, NY, USA: Oxford University Press, Inc., 2010.
- [11] 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 of the United States of America*, vol. 102, no. 30, pp. 10421–10426, Jul. 2005.
- [12] M. Tumminello, C. Coronnello, F. Lillo, S. Miccichè, and R. N. Mantegna, “Spanning trees and bootstrap reliability estimation in correlation-based networks,” *International Journal of Bifurcation and Chaos*, vol. 17, no. 7, pp. 2319–2329, Jul. 2007.

¹⁰The selection of robust trading model parameters is a complex process. Typically, bootstrapping inputs, determining the trading model performance for each bootstrapped path for each coordinate in the parameter space, then averaging the results for each coordinate in the parameter space, vastly improves the continuity of the space for visualization. Given the computationally intensive nature of such a task, this type of process can really only be achieved through the use of parallel processing.

- [13] R. N. Mantegna, “Hierarchical structure in financial markets,” *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 11, pp. 193–197, Sep. 1999.
- [14] M. Tumminello, F. Lillo, and R. N. Mantegna, “Hierarchically nested factor model from multivariate data,” *Europhys. Lett.*, vol. 78, no. 3, p. 30006, Apr. 2007.
- [15] M. Tumminello, F. Lillo, and R. N. Mantegna, “Correlation, hierarchies, and networks in financial markets,” *Journal of Economic Behavior & Organization*, vol. 75, pp. 40–58, 2010.
- [16] H. A. Simon, “The architecture of complexity,” *Proc. Am. Philos. Soc.*, vol. 106, no. 6, pp. 467–482, 1962.
- [17] M. R. Anderberg, *Cluster analysis for applications: Probability and mathematical statistics: A series of monographs and textbooks*. Academic Press, 1973.
- [18] R. N. Mantegna and H. Eugene Stanley, *Introduction to econophysics: Correlations and complexity in finance*. Cambridge University Press, 1999.
- [19] J. C. Gower, “Some distance properties of latent root and vector methods used in multivariate analysis,” *Biometrika*, vol. 53, nos. 3-4, pp. 325–338, Dec. 1966.
- [20] J. B. Kruskal, “On the shortest spanning subtree of a graph and the traveling salesman problem,” *Proc. Am. Math. Soc.*, vol. 7, no. 1, pp. 48–50, 1956.
- [21] R. C. Prim, “Shortest connection networks and some generalizations,” *The Bell System Technical Journal*, vol. 36, no. 6, pp. 1389–1401, Nov. 1957.
- [22] G. Caldarelli and A. Chessa, *Data science and complex networks*. Oxford University Press, 2016.
- [23] G. Caldarelli, S. Battiston, D. Garlaschelli, and M. Catanzaro, “Emergence of complexity in financial networks,” in *Complex networks*, E. Ben-Naim, H. Frauenfelder, and Z. Toroczkai, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 399–423.
- [24] J.-P. Onnela, A. Chakraborti, K. Kaski, J. Kertész, and A. Kanto, “Asset trees and asset graphs in financial markets,” *Physica Scripta*, vol. 2003, no. T106, p. 48, Jul. 2006.
- [25] N. Jegadeesh and S. Titman, “Returns to buying winners and selling losers: Implications for stock market efficiency,” *The Journal of Finance*, vol. 48, p. 65, 1993.
- [26] A. F. Clenow, *Stocks on the move: Beating the market with hedge fund momentum strategies*. CreateSpace Independent Publishing Platform, 2015.
- [27] M. L. de Prado, “Building diversified portfolios that outperform out of sample,” *The Journal of Portfolio Management*, 2016.