A Time-Varying Network for Cryptocurrencies

Li Guo\textsuperscript{a,b}, Wolfgang Karl Härdle\textsuperscript{c,d,e,f,g}, and Yubo Tao\textsuperscript{e,hi}

\textsuperscript{a}School of Economics, Fudan University, Shanghai, China; \textsuperscript{b}Shanghai Institute of International Finance and Economics, Shanghai, China; \textsuperscript{c}BRC Blockchain Research Center, Humboldt-Universität zu Berlin, Berlin, Germany; \textsuperscript{d}Sim Kee Boon Institute, Singapore Management University, Singapore, Singapore; \textsuperscript{e}Asia Competitiveness Institute, National University of Singapore, Singapore, Singapore; \textsuperscript{f}Department of Information Management and Finance, National Yang Ming Chiao Tung University, Hsinchu, Taiwan; \textsuperscript{g}Department of Mathematics and Physics, Charles University, Prague, Czech Republic; \textsuperscript{h}Department of Economics, Faculty of Social Sciences, University of Macau, Taipa, Macao SAR, China; \textsuperscript{i}Asia-Pacific Academy of Economics and Management, University of Macau, Macao SAR, China

ABSTRACT
Cryptocurrencies return cross-predictability and technological similarity yield information on risk propagation and market segmentation. To investigate these effects, we build a time-varying network for cryptocurrencies, based on the evolution of return cross-predictability and technological similarities. We develop a dynamic covariate-assisted spectral clustering method to consistently estimate the latent community structure of cryptocurrencies network that accounts for both sets of information. We demonstrate that investors can achieve better risk diversification by investing in cryptocurrencies from different communities. A cross-sectional portfolio that implements an inter-crypto momentum trading strategy earns a 1.08% daily return. By dissecting the portfolio returns on behavioral factors, we confirm that our results are not driven by behavioral mechanisms.

1. Introduction
The invention of Bitcoin has spurred the creation of many cryptocurrencies (cryptos hereafter), commonly known as Altcoins. According to the statistics available at coinlore.com, as of July 31, 2022, a total of 10,442 cryptos (coins and tokens) are actively traded worldwide, with a total market capitalization of over 1.16 trillion dollars. Comparing with stocks and other traditional financial assets, the cryptos are based on blockchain technology instead of any tangible asset, not issued by a government or recognized financial institution, have no physical representation, and are infinitely divisible. More importantly, the crypto market is highly volatile (Chaim and Laurini 2018; Trimborn and Härdle 2018), which may increase estimation risk, and thus makes the use of portfolio theory (e.g., Markowitz 1952) problematic for portfolios including cryptocurrencies (see Platanakis, Sutcliffe, and Urquhart 2018; Platanakis and Urquhart 2019; Huang et al. 2022, etc.). Therefore, it is essential to understand the inter-relationships between the cryptos for achieving optimal investment decisions.

For traditional assets, a natural starting point is to compare their fundamental similarities. For example, we can rely on industrial classifications based on business operations, for example, Standard Industrial Classification (SIC) and the Global Industry Classification Standard (GICS), or product differentiation (Hoberg and Phillips 2016) to infer the similarity between stocks. Unlike stocks and many other traditional assets, the fundamentals of cryptos are difficult to determine as they neither deliver dividends nor release accounting information. In addition, many investors invest in cryptocurrencies because they believe in the blockchain technology embodied in these digital coins (Liu, Sheng, and Wang 2021). However, because of the open-source nature of blockchain technology, technological similarities could be uninformative since many Clonecoins are using the same technology but have utterly different price trajectories.

To resolve the dilemma, we introduce a novel method to classify cryptos by adding the return cross-predictability to technological similarities. The benefits are 2-fold. First, return cross-predictability provides timely information to understand the dynamics of market structure and the information propagation across financial assets (Menzly and Ozbas 2010). Second, the returns embody investors’ beliefs, which is crucial in classifying cryptos or other hard-to-value financial assets (Detzel et al. 2021). We naturally employ a directed time-varying network to model the evolution of return cross-predictability and contextualize the cryptos with the technological characteristics. Using network analysis, we develop a dynamic covariate-assisted spectral clustering (CASC) method, which accommodates important network features, such as degree heterogeneity, directionality, and time-varying membership, to systematically study the inter-relationships between cryptos. We also provide a theory and conduct extensive simulations to illustrate the classification consistency of our method.

We carry out three empirical tests to illustrate the economic contribution of the crypto communities. First, we demonstrate that the investors can achieve better risk diversification by...
investing in cross-community cryptos. Classical portfolio theory (Markowitz 1952; Roy 1952) suggests that investors can achieve the expected portfolio return with a lower risk (variance) by investing in less positively correlated assets. Therefore, we calculate and compare the within- and cross-community average pairwise return correlations using 7-day and 30-day future crypto returns. It shows that the return correlations between cryptos within the same community are, on average, significantly higher than those across communities. The pattern is robust to different evaluation time horizons and only valid for dynamic CASC-estimated communities.

Second, to formally test the information propagation effects within crypto communities, we form a cross-sectional portfolio that implements an inter-crypto momentum trading strategy. Prior literature has documented that, due to information propagation, assets that have fundamentals similarities will have high correlation in returns. Thus, past return of one asset predicts its future return, which is the essence of momentum trading strategy. We sort the cryptos into quartiles according to the trading signals and implement a momentum trading strategy through simultaneously buying the cryptos in the top quartile and selling the cryptos in the bottom quartile. It shows that the trading strategy generates a one-day-ahead average daily return of 1.08%, and the return does not revert in a week.

Finally, we explore the behavioral mechanisms, such as limit-to-arbitrage, investor attention, and macro uncertainty, as alternative explanations to momentum spillover effects. Specifically, we split the whole sample into high and low episodes of the behavioral factors by cutting off at their medians. It shows that the momentum portfolio returns exist regardless of these factors, indicating that behavioral mechanisms could not be the drivers of momentum spillover effects.

This article makes several fundamental contributions to the empirical understanding of the crypto market. To the best of our knowledge, our article is the first to establish a crypto network and study the crypto market segmentation problem using community detection. In addition, we are among the first to empirically justify the importance of crypto technologies in determining crypto prices. By creatively combining the return cross-predictability and technological similarities, we identify the community pattern between cryptos and document an inter-crypto momentum spillover effect, which is direct evidence for economic implications of the crypto communities. We present our conclusions in Section 5. All proofs and technical details are provided in the supplementary appendix. The R codes to implement the algorithms are available at QuantNet (search keywords “CASC”).
2. Construction of the Cryptocurrencies Network

2.1. Network of Information Propagation

Traditional asset pricing theory states that the cross-predictability of asset returns reflects the correlations in the fundamental determinants caused by common information of the future cash flows or discount rates (Menzly and Ozbas 2010). Extensive literature has applied network analysis to the equity market to gain insights about information propagation and spillover of behavioral bias across the firms (see, e.g., Cohen and Frazzini 2008; Aobdia, Caskey, and Ozel 2014; Acemoglu, Ozdaglar, and Tahbaz-Salehi 2015; Herskovic 2018; Chen, Hardle, and Okhrin 2019). Most recently, Sockin and Xiong (2020), Cong, Li, and Wang (2021), and Cong, Li, and Wang (2022) develop models for platform tokens and study their network effects caused by decentralization. Based on the above literature, we find it essential to adopt a network view on the crypto market to study how price information or shock transfers from one crypto to another.

Following Menzly and Ozbas (2010), we model the information propagation in the cryptocurrencies network using the return cross-predictability. We collect the historical daily prices, trading volumes, and contract attributes (e.g., algorithm, proof type, age, total coins, etc.) of the top 400 cryptos in market capitalization as of December 31, 2018 from an interactive platform (Cryptocompare.com) with free API access. The sampling period is from January 1, 2016 to December 31, 2018. After excluding cryptos under 1 year old or with incomplete price and contract information, we obtain a sample of 182 cryptos.

To form the linkages between cryptos, we regress a crypto’s current return on the other cryptos’ lagged returns in a 360-day (1-year) rolling window. We employ the adaptive Lasso (Zou 2006) to select the significant regressors and to estimate their coefficients simultaneously:

\[
\hat{b}_i^* = \arg \min \left\{ \left\| \begin{array}{c} r_{i,t+1}^s - \alpha_i - \sum_{j \neq i} b_{ij} r_{j,t}^s \\ \end{array} \right\|_2^2 + \lambda_i \sum_{j \neq i} \hat{w}_{ij} |b_{ij}| \right\} 
\]

where \( r_{i,t}^s \) is the standardized return for crypto \( j \), \( \hat{b}_i^* = (\hat{b}_{i,1}, \ldots, \hat{b}_{i,N})^T \) is the adaptive Lasso estimate, \( \lambda_i \) are nonnegative regularization parameters, and \( \hat{w}_{ij} \) are the weights corresponding to \( |b_{ij}| \) for \( j = 1, \ldots, N \) in the penalty term. Conventionally, one defines \( \hat{w}_{ij} = 1/(\hat{b}_{i,j}^{LS})^{\gamma} \) with some \( \gamma > 0 \), where \( \hat{b}_{i,j}^{LS} \) denotes the coefficient estimated by ordinary least squares. The Lasso technique yields an active set that has “parental” influence on the focal crypto. Thus, we obtain an adjacency matrix for each period, \( A_t, t = 1, \ldots, T \). Figure 1 presents a sub-network of 20 cryptos with linkages estimated by regression (1) on selected dates in January 2018.

Three critical features can be observed from the return-inferred network structure. First, the return cross-predictability is changing over time. For instance, the linkage between BTC and NMR on January 1 vanishes on January 5, and the predictive relationship between BTS and DGB on January 1 reverts on January 5. These observations indicate that information propagation is not always mono-directional or continues to exist between cryptos.

Second, the return cross-predictability induces a sparse network structure, that is, many isolated cryptos with no linkages in the network. The examples include BBR and STEEM on January 1; BBR, BLITS, BTC, and LSK on January 5; BBR on January 15; BBR, OMNI, and BTCD on January 31. The sparsity is an important feature that may create many challenges for conducting an in-depth analysis of the network structure, such as community detection, network formation modeling, network differential tests, and so forth. To resolve this issue, we need to explore new information sources to identify additional linkages between cryptos or apply advanced statistical methods designed for analyzing sparse networks. In fact, we use both strategies in the later analyses.

Finally, a larger market capitalization does not imply a leading position in return cross-predictability. For example, the largest three cryptos (BTC, LTC, and ETH) predict fewer crypto returns than CLAM on January 5 and LSK on January 15. These observations suggest that return cross-predictability is a noisy estimator for uncovering the authentic connections between cryptos, which motivates us to seek additional information sources to form crypto linkages so that the noises could cancel out with each other when we conduct a combinatorial analysis for the cryptos network.

2.2. Contextualization with Technology

As pointed out by Liu, Sheng, and Wang (2021), although cryptos do not distribute dividends or release accounting information, many investors invest in them simply because they believe blockchain technology is an important innovation, and cryptos are assets that represent a stake in the future of this technology. Specifically, there are two building blocks for blockchain technology: hashing algorithms and consensus mechanism (or proof types).

Algorithm, which is short for the hashing algorithm, plays a central role in determining the crypto’s security. For each crypto, there is a hash function in mining; for example, Bitcoin (BTC) uses double SHA-256 and Litecoin (LTC) uses Scrypt. As security is one of the most important features of cryptos, the hashing algorithm naturally—in terms of trust—determines the intrinsic value of a crypto. In the example above, the Scrypt system was used with cryptos to improve upon the SHA256 protocol. The SHA256 preceded the Scrypt system and was the basis for BTC. Specifically, Scrypt was employed as a solution to prevent specialized hardware from brute-force efforts to outmine others. Thus, Scrypt-based Altcoins require on average more computing effort per unit than the equivalent coin using SHA256. The relative difficulty of the algorithm confers a relative value.

Proof Types, or proof system/protocol, is an economic measure to deter denial of service attacks and other service abuses such as spam on a network by requiring some work from the service requester, usually the equivalent to processing time by a computer. For each crypto, at least one of the protocols will be chosen as a transaction verification method; for example, BTC and Ethereum (ETH) currently use the Proof-of-Work (PoW), whereas Diamond (DMD) and Blackcoin use the Proof-of-Stake (PoS). PoW-based cryptos such as BTC use mining—the solving of computationally intensive puzzles—to validate transactions.
Figure 1. Networks formed by return cross-predictability. This figure presents the return-based cryptos network on selected dates in January 2018. We selected 20 cryptos, including BTC, ETH, LTC, and other top cryptos by market capitalization as of December 31, 2017. We obtained the connections from the predictive regression \( r_{i,t+1} = \alpha_i + \sum_{j=1}^{N-1} b_{ij} r_{j,t} + \epsilon_{i,t} \), where \( r_{i,t} \) is the standardized daily return on crypto \( i \), and \( N \) is the total number of cryptos. Only the cryptos selected by adaptive Lasso will be linked to crypto \( i \). An arrow from crypto A to crypto B means crypto A’s return can predict the crypto B’s return.

and create new blocks. In PoS-based cryptos, the creator of the next block is chosen through various combinations of random selection and wealth (in terms of crypto) or age (i.e., the stake). In summary, the proof protocol determines the reliability, security, and effectiveness of the transactions.

Recent literature has tried to model the mining process and different proof types under a theoretical framework. Many papers (see, e.g., Biais et al. 2019; Easley, O’Hara, and Basu 2019; Cong, He, and Li 2021) have studied mining activities and transaction costs and their implications for crypto prices. Among them, Pagnotta (2022) is the first to build an equilibrium model embedding a security channel that leads to multiple equilibria. Abadi and Brunnermeier (2018), Budish (2018), and Hinzen, John, and Saleh (2022) studied the limitations and the pricing implications of the PoW mechanism, while Fanti, Kogan, and Viswanath (2019) and Saleh (2021) explored the PoS mechanism. Based on these models, we conjecture that the technological similarities could be vital for understanding the inter-relationships between cryptos.

For illustration, we present some sample data of representative cryptos in our analysis in Table 1. As shown in the sample, we collect the algorithm and proof type information
Table 1. Sample data of representative cryptos.

| Date       | CryptoID | Full name | Return | Algorithm | Proof type |
|------------|----------|-----------|--------|-----------|------------|
| 2016-01-02 | BTC      | Bitcoin   | −0.06% | SHA256    | PoW        |
| 2016-01-02 | ETH      | Ethereum  | 3.77%  | Ethash    | PoW        |
| 2016-01-02 | XRP      | Ripple    | −6.82% | Multiple dPoW/PoW |
| 2016-01-02 | BLK      | Blackcoin | 0.15% | Scrypt    | PoS        |
| ...        | ...      | ...       | ...    | ...       | ...        |
| 2018-12-31 | BTC      | Bitcoin   | −3.82% | SHA256    | PoW        |
| 2018-12-31 | ETH      | Ethereum  | −5.55% | Ethash    | PoW        |
| 2018-12-31 | XRP      | Ripple    | −5.01% | Multiple dPoW/PoW |
| 2018-12-31 | BLK      | Blackcoin | −5.80% | Scrypt    | PoS        |

Figure 2. Networks formed by technological similarities. This figure depicts the contract-based cryptos network. We connect two cryptos with double-sided arrows if they have the same algorithm or proof type.

(a) Algorithm

(b) Proof Types

(c) Combined Fundamentals

Figure 2. Networks formed by technological similarities. This figure depicts the contract-based cryptos network. We connect two cryptos with double-sided arrows if they have the same algorithm or proof type.

for each crypto in addition to the returns and contextualize the network induced by return cross-predictability with the technological similarities between cryptos. To further explain the contextualization, we plot the sub-network in Figure 2 using the same cryptos as in Figure 1, where the linkages are formed if the cryptos share at least one fundamental characteristic. For
example, LTC and DOGE are linked because they both adopt the Scrypt algorithm. Evidently, due to the limited choices of algorithms and proof types, the cryptos are more likely to connect with each other when using technological similarities to form linkages. Intuitively, the network will be denser if we naively link cryptos either by return cross-predictability or technological similarity, such as adding linkages in subfigure (c) of Figure 2 to each subfigure of Figure 1. In the following section, we will discuss how to combine the two sets of information in a more sophisticated manner.

In summary, we have built a time-varying crypto network that characterizes both return cross-predictability and technological similarity between cryptos. The price shocks (e.g., policy changes, macro events, etc.) will travel through return-based linkages, while technological shocks (e.g., innovation, security breaches, etc.) will travel through technology-based connections. In the following sections, we will investigate deeper on the inter-relationships between cryptos based on this network.

### 3. Communities in the Cryptocurrencies Network

Classical empirical studies on asset pricing have pointed out the importance of market segmentation in generating heterogeneous prices of risk and profitable trading strategies. However, most of them use ex-ante information, such as industry, listing exchange, and geography, to conjecture market segments (see, e.g., Fama and French 1993; Fama et al. 1993; Foerster and Karolyi 1999; Griffin 2002). Recently, Patton and Weller (2022) proposed a data-driven method (EM algorithm) to cluster stocks with similar risk prices and provide strong evidence of market segmentation across all portfolios, factor pricing models, and time periods. Based on these findings, we deduce market segmentation also plays an important role in understanding the risks and returns of cryptos and would like to cluster cryptos under the network framework.

A standard framework for studying clustering effect, that is, community detection, in a network setting is the stochastic blockmodel (SBM) proposed by Holland, Laskey, and Leinhardt (1983). Using this framework, we can represent the adjacency matrices stochastically via a block structure and adopt spectral clustering to identify latent communities. Binkiewicz, Vogelstein, and Rohe (2017) and Zhang et al. (2018) showed that, by introducing covariate assistance, they could improve the classification accuracy of the spectral clustering method under static network settings. This section presents an extension of the covariate-assisted spectral clustering (CASC) algorithm to deal with the time-varying directed networks. Theoretical justifications and simulations demonstrate the consistency of this method.

#### 3.1. Dynamic Contextual Stochastic Blockmodel

As return cross-predictability between any two cryptos could be mono-directional, we consider a time-varying network defined as a sequence of random directed graphs with $N$ nodes, $G_{t}, t = 1, \ldots, T$, on the vertex set $V_{N} = \{v_{1}, v_{2}, \ldots, v_{N}\}$, which does not change over time. For each period, we model the directed network structure with an adjacency matrix $A_{t}$ which is mostly asymmetric, that is,

$$A_{t}(i,j) = \begin{cases} \text{Bernoulli}(P_{t}(i,j)), & \text{if } i \neq j \\ 0, & \text{if } i = j \end{cases},$$

where $P_{t}(i,j) = \Pr(A_{t}(i,j) = 1)$. The adjacency matrix $A_{t}$ records the pattern of edges that are induced by crypto return cross-predictability in network at time $t$; if there is an edge starting from node $i \in \{1, \ldots, N\}$ and ending at node $j \in \{1, \ldots, N\}$ (i.e., $i \rightarrow j$), then $A_{t}(i,j) = 1$; otherwise, $A_{t}(i,j) = 0$. Thus, the $i$th row of $A_{t}$ records how crypto $i$ predicts other cryptos (sending edges), and the $j$th column of $A_{t}$ records how crypto $j$ is predicted by other cryptos (receiving edges).

To model the community structure, the probabilities of a connection $P_{t}(i,j)$ at period $t$ are partitioned into blocks. Let the block probability matrix in each period be $B_{t} \in \{0, 1\}^{K_{R} \times K_{C}}$ with rank $K = \min(K_{R}, K_{C})$. In particular, let us denote $z_{R,t}$ as the row community label of node $i$ at time $t$ and $z_{C,t}$ as the column community label of node $j$ at time $t$; respectively; then, if $z_{R,t} = k_{R}$ and $z_{C,t} = k_{C}$, then $P_{t}(i,j) = B_{t}(z_{R,t}, z_{C,t}) = B_{t}(k_{R}, k_{C})$. Hence, for any $t = 1, \ldots, T$, we can obtain the population adjacency matrix

$$A_{t} \overset{\text{def}}{=} E(A_{t}) = Z_{R,t}B_{t}Z_{C,t}^{T},$$

where $Z_{R,t} \in \{0, 1\}^{N \times K_{R}}$ and $Z_{C,t} \in \{0, 1\}^{N \times K_{C}}$ are the row-clustering matrices and column-clustering matrices, respectively. Notably, there is only one 1 in each row and at least one 1 in each column of each clustering matrix. The co-clustering of the adjacency matrix yields two partitions of the same set of nodes. The row communities contain nodes with similar sending patterns, and the column communities contain nodes with similar receiving patterns.

Qin and Rohe (2013) and Joseph and Yu (2016) showed that the regularization of adjacency matrix can substantially improve the spectral clustering performance, especially for sparse networks. Therefore, we define the regularized graph Laplacian $L_{t,t} \in \mathbb{R}^{N \times N}$ for the directed network as

$$L_{t,t} = (D_{R,t}^{-1/2}A_{t}D_{C,t}^{-1/2}),$$

where $D_{R,t}$ and $D_{C,t}$ are diagonal matrices with $D_{R,t}(i,i) = \sum_{j=1}^{N}A_{t}(i,j) + \tau_{R,t}$ and $D_{C,t}(i,i) = \sum_{j=1}^{N}A_{t}(i,j) + \tau_{C,t}$, where $\tau_{R,t}$ and $\tau_{C,t}$ are set to be the average row and column degrees at each period, respectively.

To contextualize the network with the technological similarities between cryptos, we introduce the covariate matrix $X \in \{0, 1\}^{N \times R}$, where each row of $X$ denotes a crypto, and each column denotes a type of technology. Thus, $X(i,j) = 1$ if crypto $i$ adopts technology $j$ and is 0 otherwise. Following Binkiewicz, Vogelstein, and Rohe (2017), we assume the covariate $X$ also reflects the row- and column-cluster’s information, that is, let $M_{R,t} \in \{0, 1\}^{K_{R} \times R}$ and $M_{C,t} \in \{0, 1\}^{K_{C} \times R}$ denote the covariate expectation matrix for row- and column-clusters in each period $t$, respectively, then the population covariate matrix is

$$X = E(X) = Z_{R,t}M_{R,t} = Z_{C,t}M_{C,t}.$$
Then, we combine return cross-predictability and technological similarities between cryptos by constructing a similarity matrix from regularized graph Laplacian $L_{r,t}$ and covariate matrix $X$, that is, for each $t = 1, \ldots, T$,

$$S_t = L_{r,t} + \alpha_t C_w^t = D_{R,t}^{-1/2} A_t D_{S,t}^{-1/2} + \alpha_t X W_t X^T,$$

(6)

where $\alpha_t \in [0, \infty)$ is the tuning parameter that controls for the informational balance between return cross-predictability and technological similarities, and $W_t$ is a time-varying weight matrix that controls for the degree of technological similarities.

The setup in Equation (6) addresses several extensions of existing methods. First, $W_t$ creates a time-varying interaction between different covariates. For instance, we may think of different refined algorithms that have the same origin. Such inheritance relationships will potentially lead to an interaction between the cryptos. In addition, as shown in Figure 3, some algorithms may become more popular over time, while the others may be near extinction. Thus, this interaction would also change over time. Second, we can easily select covariates by setting certain elements of $W_t$ to 0. This is necessary as it helps us to model the evolution of technologies. At some point in time, some technology may disappear due to upgrades or cracking. Therefore, $W_t$ offers us the flexibility to exclude covariates. Finally, due to the open-source nature of the blockchain, crypto developers can easily copy and paste the source code and launch a new coin without any costs. Consequently, we observe a high degree of homogeneity in the cryptos market. However, this homogeneity does not necessarily result in a cross-predictability of prices: some cryptos are negatively correlated. In this case, we may set $W_t(i, i)$ to be negative, and $C_w^t$ will eventually bring the cryptos with different technologies closer in the similarity matrix.

Following Karrer and Newman (2011), we employ degree correction to allow for heterogeneous degrees. Specifically, we introduce two sets of degree parameters $\psi^R = (\psi^R_1, \psi^R_2, \ldots, \psi^R_N)$ and $\psi^C = (\psi^C_1, \psi^C_2, \ldots, \psi^C_N)$ so that the edge probability are given by

$$P_t(i, j) = \psi^R_s \psi^C_s B_t(i, j)$$

under the restriction that

$$\sum_{i \in \mathcal{G}_R^k} \psi^R_s = 1, \forall k \in \{1, 2, \ldots, K_R\},$$

$$\sum_{i \in \mathcal{G}_C^k} \psi^C_s = 1, \forall k \in \{1, 2, \ldots, K_C\}.$$ 

Then, the population adjacency matrix for dynamic degree-corrected contextual stochastic blockmodel (DCCBM) is

$$A_t = E(A_t) = \psi^R_s Z_{R,t} B_t Z_{C,t}^\top \psi^C_s,$$

(7)

where $\psi^s = \text{diag}(\psi^s), s \in \{R, C\}$, and the population regularized graph Laplacian is

$$L_{r,t} = D_{R,t}^{-1/2} A_t D_{C,t}^{-1/2}.$$  

(8)

Therefore, the population similarity matrix is

$$S_t = L_{r,t} + \alpha_t C_w^t,$$

(9)

where $C_w^t = X W_t X^T$ and $W_t$ is the population version of the weight matrix $W_t$.

By construction, we know that $D_{R,t}(i, i) = \sum_{j=1}^N 1_{(i \rightarrow j)} + \tau_{R,t}$, which controls for the number of the parents of node $j$, and that $D_{C,t}(j, j) = \sum_{i=1}^N 1_{(i \rightarrow j)} + \tau_{C,t}$, which controls the number of the offspring of node $j$. To analyze the asymmetric adjacency matrix $A_t$ caused by directional information, Rohe, Qin, and Yu (2016) proposed using the singular value decomposition instead of eigen-decomposition for the regularized graph Laplacian. The intuition behind this methodology is to use both the eigenvectors of $L_{r,t}^T L_{r,t}$ and $L_{r,t} L_{r,t}^T$, which contains

---

**Figure 3.** The evolution of technology adoption rates. This figure presents the adoption rates (or the popularity) of the major technologies in cryptos market from January 1, 2016 to December 31, 2018.
information about “the number of common parents” and “the number of common offsprings”; that is, for each $t = 1, \ldots, T$:

$$(L^T_{t,t}L_{t,t})_{ab} = \sum_{i=1}^{N} L_{t,t}(i,a)L_{t,t}(i,b)$$

$$= \frac{1}{\sqrt{D_{C_{t}}(a,a)D_{C_{t}}(b,b)}} \sum_{i=1}^{N} \frac{1(t)}{D_{R_{t}}(i,i)},$$

$$(L_{t,t}L^T_{t,t})_{ab} = \sum_{i=1}^{N} L_{t,t}(a,i)L_{t,t}(b,i)$$

$$= \frac{1}{\sqrt{D_{R_{t}}(a,a)D_{R_{t}}(b,b)}} \sum_{i=1}^{N} \frac{1(t)}{D_{C_{t}}(i,i)}.$$

Therefore, the column communities consist of cryptos that receive similar price and technology shocks (common parents), and the row communities contain cryptos transmitting shocks to similar cryptos (common offspring).

### 3.2. Dynamic Covariate-Assisted Spectral Clustering

When setting up a dynamic covariate-assisted spectral clustering algorithm, we need to (i) determine the pre-specified interactive weights $W_t$ between covariates and (ii) estimate the similarity matrix with time-varying network information. For the first issue, we use the adoption rates to present the popularity of the crypto technology. Specifically, for any two technologies (algorithms or proof types) $a$ and $b$ at time $t$, we have

$$W_t(a, b) = \frac{N_{a,t}}{N_t} \times \frac{N_{b,t}}{N_t},$$

where $N$ is the total number of cryptos in the market, $N_a$ and $N_b$ are the numbers of cryptos that adopt technology $a$ and $b$, respectively. Intuitively, the weight is greater when the technologies are more popular among the cryptos, indicating a higher strength of technological similarity.

Regarding the second issue, we make an important assumption to ensure that neither the connection probabilities nor the community memberships change drastically over the horizons. Specifically, we assume:

**Assumption 1.** For $1 \leq k \leq K' \leq K$, let $\zeta_t = t/T$ denote the time instance, there exists a function $f(\cdot; k, k')$ such that $B_t(k, k') = f(\zeta_t; k, k')$ and $f(\cdot; k, k') \in \Sigma(\beta, L)$, where $\Sigma(\beta, L)$ is a Hölder class of functions $f(\cdot)$ on $[0, 1]$ such that $f(\cdot)$ are $\ell$ times differentiable and

$$|f^{(\ell)}(x) - f^{(\ell)}(x')| \leq L|x - x'|^{\beta - \ell},$$

with $\ell$ being the largest integer smaller than $\beta$.

**Assumption 1** ensures that $S_t$ is continuous through assuming that the block probability matrix $B_t$ transits smoothly from one time instance to another as a Hölder continuous function. Han, Xu, and Airoldi (2015) and Pensky (2019) showed that continuity assumptions allow one to take advantage of the well-studied methodologies in nonparametric regression estimation and significantly improve precision of the inference.

Based on **Assumption 1**, we follow Pensky and Zhang (2019) by constructing the estimator of $S_t$ with a discrete kernel to bring in historical linkage information. Specifically, we first pick an integer $r \geq 0$, obtain a set of integers $F_r = \{-r, \ldots, 0\}$ and assume that $\|W_{r,\ell}(i)\| \leq W_{\text{max}}$, where $W_{\text{max}}$ is independent of $r$ and $i$, and satisfies

$$\frac{1}{|F_r|} \sum_{i \in F_r} W_{r,\ell}(i) = \begin{cases} 1, & \text{if } k = 0, \\ 0, & \text{if } k = 1, 2, \ldots, \ell - 1. \end{cases} \quad (12)$$

By construction, the $W_{r,\ell}$ is a discretized version of the order $\ell$ optimal continuous boundary kernel that weighs only the historical observations (Gasser and Müller 1979; Altman 1990). As shown in Figure 4, the kernel assigns more recent similarity matrices higher scores. To choose an optimal bandwidth $r$, Pensky and Zhang (2019) proposed an adaptive estimation procedure using Lepski, Mammen, and Spokoiny (1997). We employ

![Figure 4](image-url)

**Figure 4.** Examples of discrete kernel. The first figure plots the value of discrete kernel $W_{r,\ell}(i)$ with bandwidth $r = 4$, smoothing parameter $\ell = 4$, and $i \in F_r$. The second figure plots $W_{r,\ell}(i)$ with bandwidth $r = 6$, smoothing parameter $\ell = 4$, and $i \in F_r$. 
the same method (see Appendix A, supplementary materials) and construct the estimator for the smoothed similarity matrices:

$$\hat{S}_{t,r} = \frac{1}{|T_t|} \sum_{i \in T_t} W_{r,t}(i) S_{t,i}. \quad (13)$$

**Remark 1.** Different from Pensky and Zhang (2019), we only apply a boundary kernel on the historical similarity matrices for community detection. The changes are to avoid the “look-ahead bias” (Kan and Kirikos 1996) in asset pricing inference in subsequent sections.

Once we obtain $\hat{S}_{t,r}$, we create a singular value decomposition of $\hat{S}_{t,r} = \hat{U}_t \hat{\Sigma}_t \hat{V}_t^\top$ for each $t = 1, 2, \ldots, T$. As Lei and Rinaldo (2015) discussed, the matrix $\hat{U}_t$ may now have more than $K$ distinct rows due to the degree correction, whereas the rows of $\hat{U}_t$ still only point to at most $K$ directions. Therefore, we apply the spherical clustering algorithm to find a community structure among the rows of the normalized matrix $\hat{U}_t^+$ with $\hat{U}_t^+(i, *) = \hat{U}_t(i, *) / \| \hat{U}_t(i, *) \|$. More specifically, we consider the following spherical $k$-medians spectral clustering:

$$\| \hat{Z}_{r,t}^+ \hat{Y}_t - \hat{U}_t^+ \|_F^2 \leq (1 + \epsilon) \min_{\hat{Z}_{r,t} \in \mathcal{M}_{N R \times K}} \| \hat{Z}_{r,t} \hat{Y}_t - \hat{U}_t^+ \|_F^2 \quad (14)$$

where $\hat{Y}_t$ is some rotation matrix. In the last step, we extend $\hat{Z}_{r,t}$ to obtain $\hat{Z}_{r,t}$ by adding $N - N_{K_t}^+$ canonical unit row vectors at the end. $\hat{Z}_{r,t}$ is the estimate of $Z_{r,t}$ from this method. Same techniques are applied to $\hat{V}_t$ to obtain column communities. We summarize the detailed steps in Algorithm 1.

### 3.3. Classification Consistency

In the subsequent analysis, we illustrate that the dynamic CASC achieves classification consistency uniformly over time. We first make some assumptions on the graph that generates the dynamic network. The major assumption required here is assortativity, which ensures that the nodes within the same community are more likely to share an edge than nodes in different communities. Following Yan and Sarkar (2021), we assume:

**Assumption 2.** $\min_k B_t(k, k) > \max_{k \neq l} B_t(k, l), \forall t = 1, \ldots, T$.

Intuitively, the more frequent the community membership changes, the less stable the network will be. Consequently, it becomes harder to use the information from the historical and future network structures to detect the communities in the present network structure. In Assumption 3, we restrict the maximum number of nodes that switch memberships ($s$) to some finite number. Based on this assumption, the proportion of nodes that switch their memberships shrinks to 0 as the size of the network grows to infinity. Additionally, we can easily bind the dynamic behavior of clustering matrices ($Z_{r,t+i} - Z_{r,t}$ or $Z_{c,t+i} - Z_{c,t}$) by noting that there are at most $rs$ nonzero rows in the difference matrix.

**Assumption 3.** At most, $s < \infty$ number of nodes can switch their memberships between any consecutive time instances.

### Algorithm 1: CASC in the Dynamic DCCBM

**Input:** Adjacency matrices $A_t$ for $t = 1, \ldots, T$; Covariates matrix $X$; Number of row communities $K_R$ and number of column communities $K_C$; Approximation parameter $\epsilon$.

**Output:** Membership matrices of rows and columns $Z_{r,t}$ and $Z_{c,t}$ for $t = 1, \ldots, T$.

1. Calculate regularized graph Laplacian $L_{t,t}$.
2. Estimate $S_t$ by $\hat{S}_{t,r}$, as in (13).
3. Compute the singular value decomposition of $\hat{S}_{t,r} = \hat{U}_t \hat{\Sigma}_t \hat{V}_t^\top$ for $t = 1, \ldots, T$.
4. Extract the first $K$ columns of $U_t$ and $V_t$ that correspond to the $K$ largest singular values in $\Sigma_t$, where $K = \min(K_R, K_C)$. Denote the resulting matrices $\hat{U}_t^K \in \mathbb{R}^{NR \times K}$ and $\hat{V}_t^K \in \mathbb{R}^{KC \times K}$.
5. Let $N_{K_t}^+$ be the number of nonzero rows of $\hat{U}_t^+$; then, obtain $\hat{U}_t^{+} \in \mathbb{R}^{NR \times K}$ consisting of normalized nonzero rows of $\hat{U}_t^+$; that is, $\hat{U}_t^{+}(i, *) = \hat{U}_t(i, *) / \| \hat{U}_t(i, *) \|$ for $i$ such that $\| \hat{U}_t(i, *) \| > 0$.
6. Similarly, let $N_{K_t}^+$ be the number of nonzero rows of $\hat{V}_t^+$; then, obtain $\hat{V}_t^{+} \in \mathbb{R}^{KC \times K}$ consisting of normalized nonzero rows of $\hat{V}_t^+$; that is, $\hat{V}_t^{+}(i, *) = \hat{V}_t(i, *) / \| \hat{V}_t(i, *) \|$ for $i$ such that $\| \hat{V}_t(i, *) \| > 0$.
7. Apply the $(1 + \epsilon)$-approximate $k$-medians algorithm to cluster the rows (columns) of $\hat{S}_{t,r}$ into $K_R$ ($K_C$) communities by treating each row of $\hat{U}_t^{+}$ ($\hat{V}_t^{+}$) as a point in $\mathbb{R}^K$ to obtain $\hat{Z}_{r,t}$ ($\hat{Z}_{c,t}$).
8. Extend $\hat{Z}_{r,t}^+$ and $\hat{Z}_{c,t}^+$ to obtain $\hat{Z}_{r,t}$ and $\hat{Z}_{c,t}$, respectively, by arbitrarily adding $N - N_{K_t}^+$ and $N - N_{K_t}^+$ canonical unit row vectors at the end, such as $\hat{Z}_{r,t}(i) = (1, 0, \ldots, 0)$ and $\hat{Z}_{c,t}(i) = (1, 0, \ldots, 0)$ for $i$ such that $\| \hat{U}_t(i, *) \| = 0$ and $\| \hat{V}_t(i, *) \| = 0$.
9. Output $\hat{Z}_{r,t}$ and $\hat{Z}_{c,t}$.

Finally, to guarantee the performance of our clustering method, we impose some conditions to regularize the behavior of the covariate matrix and the eigenvalues of the similarity matrices.

**Assumption 4.** Define $\delta = \inf_i \{ \min_j \{ \min_i \mathcal{D}_{R,i}(i, j), \min_i \mathcal{D}_{C,i}(i, j) \} \}$ as the minimum average degree. Then, for any $\epsilon \in (0, 1)$, assume:

(i) $\delta > 3 \ln(16NT/\epsilon)$;

(ii) $\alpha_{\max} = \sup_i a_t \leq (NR)^{-1} a_t$, where $a_t = \sqrt{3\ln(16NT/\epsilon)/2}$.

**Remark 2.** Assumption 4 embodies several important conditions for classification consistency. Condition (i) indicates the graph should be at least semi-dense. Condition (ii) ensures the population eigen-gap to be greater than the maximum of the absolute difference between the sum of covariate variances within a block and the mean of the sums across all blocks, which
means covariate similarities will not dominate the clustering results.

Following Rohe, Qin, and Yu (2016), we define the "R-misclustered" and "C-misclustered" vertices as

\[ M_i^R = \left\{ j : \| C_{ii}^R - C_{jj}^R \| > \| C_{ii}^C - C_{jj}^C \| , \text{for any } j \neq i \right\}, \]

\[ p \in \{R, C\}, \]

where \( C_{ii}^R \) and \( C_{jj}^C \) for \( p \in \{R, C\} \) are the community centroids of the \( i \)-th node at time \( t \) generated using the \( k \)-medians clustering on the left/right singular vectors and the population left/right singular vectors, respectively.

**Theorem 1.** Let \( \lambda_1, \lambda_2, \ldots, \lambda_K > 0 \) be the \( K = \min\{K_R, K_C\} \) largest singular values of \( S_t \). Let \( Z_{R,t} \in \mathcal{M}_{N,K_R}, Z_{C,t} \in \mathcal{M}_{N,K_C} \), and \( P_{\max} = \max\{\max_i(Z_{R,t}^T Z_{R,t})_{ii}, \max_i(Z_{C,t}^T Z_{C,t})_{ii}\} \) denote the size of the largest block over the horizons. Then, under Assumptions 1–4 and \( K_R \leq K_C \), the misclustering rate satisfies

\[ \sup_t \left\{ \frac{\mathbb{E}[M^R_t]}{N} \right\} = \frac{c_2(\epsilon)KW^2_{\max}}{m^2N\lambda_{\max}^2} \left( (6 + c_1) \frac{b}{\delta^{1/2}} + \frac{2K_C}{\delta} (2P_{\max}r + 2P_{\max}) + \frac{P_{\max}}{\delta \cdot \ell!} \left( \frac{r}{T} \right) \right)^2, \]

\[ \sup_t \left\{ \frac{\mathbb{E}[M^C_t]}{N} \right\} = \frac{c_3(\epsilon)KW^2_{\max}}{m^2N\lambda_{\max}^2} \left( (6 + c_1) \frac{b}{\delta^{1/2}} + \frac{2K_C}{\delta} (2P_{\max}r + 2P_{\max}) + \frac{P_{\max}}{\delta \cdot \ell!} \left( \frac{r}{T} \right) \right)^2, \]

with a probability of at least \( 1 - \epsilon \), where \( \lambda_{\max} = \max_i(\lambda_{K_i,i}) \), \( c_2(\epsilon) = 2^6(2 + \epsilon)^2 \), \( c_3(\epsilon) = 2^7(2 + \epsilon)^2, b = \sqrt{3\ln(16N/\epsilon)} \). \( m_r, N, r \), and \( \ell \) are defined by equations (S5), (S18), and (S20) in the Appendix, supplementary materials, respectively.

**Remark 3.** The uniform upper bound is analogous to the static upper bounds in Binkiewicz, Vogelstein, and Rohe (2017) and Zhang et al. (2018). Abbe, Fan, and Wang (2022) pointed out that these bounds are not tight enough to inform the information gain in combining the network and node attributes. Dessimone et al. (2018), Lu and Sen (2020), and Abbe, Fan, and Wang (2022) tried to derive the threshold holographic for exact recovery under static network settings. It is definitely worth another theoretical paper to derive the uniform information bound for dynamic networks, but it is beyond the scope of the current article. For empirical analysis, we only need consistency to guarantee the performance of our method. In subsequent sections, we will illustrate the improvement of classification accuracy by combining two sets of information from simulations and empirical tests.

If group memberships of the nodes remain unchanged over time \( s = 0 \) and one chooses to cluster \( S_t \) period by period \( (W_{\max} = 1 \) and \( r = 0 \) ), then the misclustering rates will be analogous to the clustering error in Theorem 5.2 of Zhang et al. (2018). Specifically, one can formulate the following corollary:

**Corollary 1.** Under the assumptions of Theorem 1, suppose the membership is constant and one uses a period-by-period spectral clustering method for estimation. Then, with a probability of at least \( 1 - \epsilon \), the misclustering rates

\[ \sup_t \left\{ \frac{\mathbb{E}[M^R_t]}{N} \right\} \leq \frac{C_0(\epsilon)K \ln(16NT/\epsilon)}{m^2N\lambda_{\max}^2 \delta}, \]

\[ \sup_t \left\{ \frac{\mathbb{E}[M^C_t]}{N} \right\} \leq \frac{C_1(\epsilon)K \ln(16NT/\epsilon)}{m^2N\gamma^2 \lambda_{\max}^2 \delta}. \]

**Remark 4.** Zhang et al. (2018) derived the period-wise upper bounds for misclustering rates that is proportional to \( \ln(16N/\epsilon) \). In contrast, our upper bounds are proportional to \( \ln(16NT/\epsilon) \). The difference is a result of Boole’s inequality and matrix Bernstein inequality (see Proof of Lemma S2 in the supplementary materials.) Based on this corollary, the remarks following Theorem 5.2 in Zhang et al. (2018) are all applicable to our results (e.g., when \( \delta \) grows faster than \( \ln(NT) \), we have misclustering rates go to zeros as \( N, T \to \infty \)).

### 3.4 Monte Carlo Simulations

To illustrate the theoretical results intuitively, we carry out several simulation studies using our algorithm and existing clustering methods under different model setups. Our benchmark algorithms for the directed networks are the degree-corrected DI-SIM (DI-SIM-DC) by Rohe, Qin, and Yu (2016) and the covariate-assisted DI-SIM for static networks (CA-DI-SIM-Stc) by Zhang et al. (2018). Since both methods are intended for static networks, we apply them to the adjacency matrix period by period and report the average misclustering rate over time.

To start with, we briefly summarize the data generating process below in order to explain how we generate simulated data. The adjacency matrix \( A_t \) is generated according to the Bernoulli distribution:

\[ A_t(i,j) = \begin{cases} 
\text{Bernoulli}(P_t(i,j)), & \text{if } i \neq j \\
0, & \text{if } i = j
\end{cases} \]

where \( P_t(i,j) = B_t(z_{R,it}, z_{C,it}) \) with \( z_{R,it} \) being the community label of node \( i \) and \( z_{C,it} \) being the community label of node \( j \). The covariates that incorporate clustering information is generated by

\[ X = E(X) = Z_{R,t}M_{R,t} = Z_{C,t}M_{C,t}, \]

where \( M_{R,t} \in [0,1]^{K_R \times R} \) and \( M_{C,t} \in [0,1]^{K_C \times R} \) denote the covariate expectation matrix for row- and column-clusters in each period \( t \), respectively. Since we use the same set of covariates for parent and offspring nodes, if we assume \( X \) are Bernoulli covariates, then the rth covariate of node \( i \) is

\[ X(i,r) = \text{Bernoulli}(M_{R,t}(z_{R,it}, r)) \cap \text{Bernoulli}(M_{C,t}(z_{C,it}, r)). \]

Throughout the simulations, we assume that the number of communities \( K_R = K_C = 4 \) is known, and we set the block probability matrix \( B_t \) as

\[ B_t = \begin{bmatrix}
0.6 & 0.3 & 0.2 & 0.1 \\
0.3 & 0.5 & 0.2 & 0.1 \\
0.2 & 0.4 & 0.3 & 0.1 \\
0.1 & 0.1 & 0.1 & 0.3
\end{bmatrix} \times \frac{T + 2t}{2T} , \text{ with } 1 \leq t \leq T. \]
Then, we generate the initial clustering matrices $Z_{R,1}$ and $Z_{C,1}$ by randomly choosing one entry in each row and assign it to 1. For $t = 2, \ldots, T$, we fix the last $N - s$ rows of $Z_{R,t-1}$ and $Z_{C,t-1}$ and reassign 1s in the first $s$ rows of $Z_{R,t-1}$ and $Z_{C,t-1}$, respectively, to mimic the community membership changes. Lastly, we generate $R = 4$ Bernoulli covariates such that, in the $k$th block, the probability of the $k$th covariate being 1 is 0.8 and the probability of the other covariates being 1 is 0.2 as in Binkiewicz, Vogelstein, and Rohe (2017).

We conduct the spectral clustering by setting the order of the polynomials for kernel construction at $L = 4$ for all simulations and replicate all experiments 1000 times. The misclustering rate we report is the temporal average of the misclustering rates; that is, $T^{-1} \sum_{t=1}^{T} |M_{R,t}| / N$ and $T^{-1} \sum_{t=1}^{T} |M_{C,t}| / N$.

We first examine the clustering performance with a growing network size. The number of vertices in the network varies from 20 to 200 with step size 20. The time span is $T = 10$ and the number of membership changes is $s = 10$. We summarize the results in Figures 5(a) and (b). Evidently, as the size of the network increases, the misclustering rates of the CA-DI-SIM-Dyn decrease sharply and dominate both CA-DI-SIM-Stc and DI-SIM-DC throughout the simulation. The CA-DI-SIM-Stc dominates DI-SIM-DC in classification accuracy, which is consistent with the argument that some similarity is sufficient for the covariates to assist in the discovery of the graph structure (Binkiewicz, Vogelstein, and Rohe 2017). The dominating performance of CA-DI-SIM-Dyn over CA-DI-SIM-Stc proves that historical linkage information is crucial for improving classification accuracy. The misclustering rate of CA-DI-SIM-Dyn-Id ($W_t = I_S$) is slightly larger than that of CA-DI-SIM-Dyn, but it is still significantly better than other methods, indicating that the choice of weight matrix has limited impact on the clustering efficiency.

Next, we check the relative performance for a growing maximum number of community membership changes. Here, we fix the total number of vertices at 200 and vary the community membership changes for each period, $s$, from 20 to 100 with step size 20. The total number of horizons is $T = 10$. We summarize the results in Figures 5(c) and (d).

![Figure 5](image-url)

**Figure 5.** Misclustering rates. This figure reports misclustering rates of different spectral clustering algorithms for networks with a growing number of vertices or with a growing number of membership changes. The CA-DI-SIM-Dyn represents Algorithm 1. The CA-DI-SIM-Dyn-Id represents Algorithm 1 with $W_t = I_S$. The DI-SIM-DC is the degree-corrected DI-SIM in Rohe, Qin, and Yu (2016) and CA-DI-SIM-Stc is the static covariate-assisted DI-SIM method in Zhang et al. (2018). In subfigures (a) and (b), the number of nodes ($N$) grows from 20 to 200 with a step size of 20, and the number of membership changes is fixed at $s = 10$. In subfigures (c) and (d), the number of nodes is fixed at 200, and the number of membership changes ($s$) grows from 20 to 100 with step size of 20. The number of periods is $T = 10$, and the number of blocks is $K = 4$, with the block probability matrix given in Equation (16). The smoothness parameter for kernel construction is $L = 4$. All simulations are conducted 1000 times.
the results in Figure 5(c) and (d). Evidently, the misclustering rate of our method increases with the number of community membership changes. The benchmark methods are insensitive to the change of community memberships because they estimate communities using a one-period adjacency matrix. Despite the result, our method still achieves the lowest misclustering rate among all methods when the community memberships are relatively stable ($s < N/2$).

To perceptually illustrate the supreme accuracy of the dynamic CASC over the benchmark methods, we visualize the block structure of the dynamic network in Figure 6 by drawing the average adjacency matrices over a certain period and sort the rows and columns according to their community labels. Specifically, we simulate a dynamic network of 200 nodes with four blocks across 10 periods according to the block probability matrix $B_t$ in Equation (16) by setting $s = 10$ and Bernoulli covariates. Then, we calculate the average adjacency matrix over six periods, $1/6 \sum_{t=1}^{6} A_t$, and sort the rows and columns jointly according to their community labels. We choose six periods because the optimal bandwidth $r$ is chosen as 5 in the simulation. Thus, using six periods will lose no information in the kernel. Taking the average adjacency matrix as a $200 \times 200$ grid, we paint the square with a darker color when it has a larger value.

Evidently, the graph with true communities exhibits a clear block structure where each block is associated with reasonable connection frequencies abide by the block probability matrix. All three methods successfully recover the dense blocks (blocks in the northwest corner), whereas only the dynamic CASC can recover the less dense blocks (blocks in the southeast corner). The simulation results serve as remarkable evidence to confirm the advantages of the dynamic CASC in dealing with dense networks.

### 3.5. Empirical Communities

Applying the dynamic CASC algorithm to the cryptos network, we classify the crypto network into four column communities. We present three representative cryptos in each estimated community.
community in Table 2. It shows that as of December 31, 2017, the largest cryptos (BTC, ETH, and LTC) in terms of market capitalization are not necessarily categorized into the same community. For example, although the return patterns of BTC and ETH are closely related, their fundamental attributes are somewhat different: BTC employs SHA256, while ETH uses KECCAK-256. For comparison, we also report the representative cryptos of each community estimated by DI-SIM (Rohe, Qin, and Yu 2016). It shows that the representative cryptos of each community are pretty different when choosing different classification algorithms, indicating that the DI-SIM-estimated communities could have distinctive features compared to dynamic CASC-estimated communities.

To evaluate the classification efficiency under empirical settings, we define two important measures, the average within- and cross-community degrees:

**Within-Community Degrees**

\[
\text{Within} = \frac{1}{T} \sum_{t=1}^{T} \sum_{i \in C} \sum_{j \notin C} \sum_{k \in C} \{A_t(i,k) + A_t(k,i)\},
\]

**Cross-Community Degrees**

\[
\text{Cross} = \frac{1}{2NC(N - NC)} \sum_{i \in C} \sum_{j \notin C} \sum_{k \notin C} \{A_t(i,k) + A_t(k,i)\},
\]

where \(N_C\) is the number of cryptos in community \(C\), and \(N\) is the total number of cryptos. Intuitively, if the clustering method correctly classifies all cryptos, then the within-community degrees should be larger than the cross-community degrees, that is, the difference between the two measures should be positive. Table 3 summarizes the average within- and cross-community degrees of network linkages induced by the return cross-predictability, the technological similarities, and the joint similarities based on the dynamic CASC and the DI-SIM, respectively.

Consistent with our simulations, the dynamic CASC method shows superior classification efficiency than that of DI-SIM. It delivers more positive and significant differences in average degrees for all three types of network linkages. More interestingly, even though the DI-SIM only uses the return information to classify crypto, its performance for classifying return linkages is worse than that of the dynamic CASC, which needs to balance the return and technology information. This striking result proves that technological similarity assists extracting community information from return dynamics by emphasizing the return cross-predictability cause by common technological risk, and thus improves the classification.

### 4. Economic Implications

In this section, we conduct three empirical tests to assess the economic contributions of crypto communities to the crypto market. First, we examine whether the investors can diversify risks better by investing cryptos across different communities than those within the same community. Second, we explore the asset pricing implication of information propagation effects on within-community cryptos by constructing a cross-sectional portfolio that implements an inter-crypto momentum trading
strategy. Finally, we dissect the portfolio return on several behavioral factors as falsification tests to support our informational mechanism.

### 4.1. Risk Diversification

Diversification is one of the major components of investment decision making under risk and uncertainty. Its underlying philosophy is to minimize both the probability of portfolio loss and its severity through multilateral insurance, in which each asset is insured by the remaining assets. Seminal papers by Markowitz (1952) and Roy (1952) mathematically formalized the idea of diversification of investment and showed that diversification can reduce risk without changing the expected portfolio returns. According to Rubinstein (2002), the key ingredient of risk diversification is the return covariance between risky assets that to be selected into the portfolio. The less positively correlated the assets, the lower the probability that the assets suffer losses simultaneously in the same proportion, and the better the protection offered by this multilateral insurance, namely diversification (Koumou 2020).

To testify that crypto communities facilitate risk diversification, we compare the out-of-sample average pairwise return correlations of cryptos within and across the communities. Specifically, we first estimate the crypto communities on each trading day and compute the pairwise Pearson correlation coefficients using the next 7- (1-week) and 30-day (1-month) crypto returns. Then, we calculate the average within-community correlations for each community by averaging the correlation coefficients of the crypto pairs whose members both belong to the focal community. We also obtain the average cross-community correlations for each community by averaging the correlation coefficients of the crypto pairs where only one of the members belongs to the focal community. Ideally, the average within-community correlations should be significantly higher than the average cross-community correlations for all communities because the cryptos within the same community share common sources of price and technological shocks. For comparison, we also present the results with the communities being estimated by DI-SIM.

Table 4 summarizes the results, and three findings are obtained. First, the return correlations between cryptos within the same community are, on average, significantly higher than those across communities. For example, in Panel A, the average within- and cross-community 30-day-ahead return correlations for Community 1 are 34.5 % and 31.7%, respectively. The correlation difference is 2.8%, which is significant at the 1% level with a Newey–West adjusted t-statistic of 9.0. Economically, the substantial reduction in return correlations suggests that investors can find attractive upside and diversification possibilities through allocating portfolio weights on cryptos from different communities.

Second, the dynamic CASC-estimated communities exhibit better risk diversification performance than the DI-SIM-estimated communities. In particular, all the dynamic CASC-estimated communities possess significantly higher within-community return correlations than cross-community return correlations. For example, only one DI-SIM community (Community 3) shows significantly positive differences between the within- and cross-community return correlations. The consistent correlation patterns across the dynamic CASC-estimated communities makes it more convenient for investors to form a diversified crypto portfolio following the cross-community principle. This is also empirical evidence that supports the validity of the dynamic CASC algorithm for community detection in time-varying networks.

Finally, the differences between the within- and cross-community return correlations are robust to the length of the evaluation window. Specifically, the changes between the 7- and 30-day return correlation differences for dynamic CASC-estimated communities are around 0.1%. From practitioners’ perspective, a stable return pattern correlation allows the investors to make flexible investment decisions on diversifying short- or long-term risks.

In summary, the dynamic CASC-estimated communities have meaningful implications for investors’ portfolio choices and risk diversification through delivering lower cross-community return correlations robustly across all communities and over different horizons. We acknowledge that Markowitz (1952) mean-variance portfolio optimization framework is highly sensitive to estimation errors in the input parameters (see, Kan and Zhou 2007, for a detailed discussion), and the high volatility feature of cryptos (Chaim and Laurini 2018) makes

| Community 1 | 0.366 | 0.337 | 0.029 | 7.55 | 0.345 | 0.317 | 0.028 | 9.00 |
| Community 2 | 0.369 | 0.337 | 0.032 | 8.10 | 0.359 | 0.317 | 0.033 | 9.87 |
| Community 3 | 0.381 | 0.341 | 0.040 | 9.46 | 0.358 | 0.318 | 0.040 | 10.68 |
| Community 4 | 0.381 | 0.340 | 0.041 | 10.99 | 0.362 | 0.319 | 0.042 | 11.64 |

**Table 4.** Within- and cross-community average return correlations.

| 7-day return correlations | 30-day return correlations |
|---------------------------|-----------------------------|
| **Within** | **Cross** | **Diff.** | **t-stat.** | **Within** | **Cross** | **Diff.** | **t-stat.** |
| Community 1 | 0.336 | 0.328 | 0.008 | 1.24 | 0.322 | 0.309 | 0.014 | 0.85 |
| Community 2 | 0.295 | 0.307 | -0.012 | -1.02 | 0.268 | 0.284 | -0.016 | -0.43 |
| Community 3 | 0.562 | 0.388 | 0.173 | 10.15 | 0.550 | 0.369 | 0.181 | 4.34 |
| Community 4 | 0.296 | 0.310 | -0.014 | -1.78 | 0.266 | 0.284 | -0.018 | -1.48 |

NOTE: This table reports the out-of-sample average return correlations of the cryptos within and across communities. Panel A presents the results for communities estimated by dynamic CASC and Panel B reports the results for communities estimated by DI-SIM. We first calculate the pairwise return correlations between cryptos in each period using the 7- and 30-day future returns. Then, we compute the within-community (cross-community) average return correlations using the correlation coefficients between the cryptos from the same community (different communities). The t-statistics reported are based on the Newey–West robust standard errors with four lags. The sample period spans from January 1, 2016 to December 31, 2018.
the use of portfolio theory problematic for portfolios including cryptocurrencies. However, our risk diversification analysis is not to dispute or to defend the Markowitz’s theory. Instead, we simply suggest that investors may achieve better diversification under any diversification framework (e.g., 1/N, Markowitz mean-variance, advanced Black–Litterman model with VBCs, etc.) that requires low correlations between risky assets if they select cryptos from different communities to construct a portfolio because the return correlation is lower for cryptos across communities than those within.

4.2. Momentum Spillover

Since information diffuses gradually across investors (Hong and Stein 1999; Hong, Lim, and Stein 2000), the price of an asset may react sluggishly to the arrival of relevant news/information about other assets. This suggests that assets having fundamental similarities will have momentum spillovers, wherein past return of one asset predicts the returns of assets linked to it. Numerous papers have verified such spillovers in the equity market by using a variety of proxies for inter-firm linkages (see, e.g., Moskowitz and Grinblatt 1999; Cohen and Frazzini 2008; Menzly and Ozbas 2010; Lee et al. 2019; Ali and Hirshleifer 2020; Parsons, Sabbatucci, and Titman 2020). Similarly, the estimated communities in the cryptos market are designed to optimally capture the information propagation and technological similarities between cryptos. Therefore, we expect the momentum spillover effects to exist within each community in the cryptos market.

We construct a long-short portfolio that implements an inter-crypto momentum trading strategy to formally test the momentum spillover effects and justify the information propagation mechanism between similar cryptos in the network. The portfolio is constructed in three steps. First, for each crypto i on day t, we compute the average return of all the other cryptos within the same community (\( R_{it} \)) to form a trading signal:

\[
R_{it} = \frac{1}{|C_{-i}|} \sum_{j \in C_{-i}} r_{jt},
\]

where \( C_{-i} \) denotes the set of other cryptos that are in the same community as crypto i. Second, we sort cryptos into four quartiles according to the trading signal and label the cryptos in the top/bottom quartile as the winner/loser cryptos. Finally, we buy the winner cryptos and sell the loser cryptos with equal positions at the end of day t and rebalance the trading positions at the end of next trading day.

We examine the portfolio return up to 7 days after the signal formation period. The hypothesis is that the information propagation mechanism should predict a drifting return pattern instead of a return reversal. Ideally, the winner/loser cryptos will generate higher/lower returns as the positive/negative shocks transmit through the network linkages between similar cryptos.

Table 5 reports the results. Consistent with our hypothesis, the long-short portfolio generates a one-day-ahead average daily return of 1.08% (t-stat. 12.24), statistically significant at the 1% level. The results exhibit a strong monotonic relationship between past connected crypto returns and future crypto returns. More importantly, the positive portfolio return does not revert over the horizons. The portfolio analysis confirms that our network model and community structure effectively capture the information propagation between cryptos and provide economically significant investment opportunities.

4.3. Behavioral Interpretation

The momentum spillover effect evidences strongly that return cross-predictability and technological similarities identify fundamental relationships between cryptos. However, the information propagation mechanism is not the only possibility that could generate the momentum spillover effect. Investors may be slow to carry information across cryptos due to a behavioral bias or constraint. Therefore, we explore three classical behavioral mechanisms commonly used to explain cross-sectional return anomalies (i.e., limit-to-arbitrage, investor attention, and economic policy uncertainty) to see if they can provide significant explanatory power for the momentum spillover effects in the cryptos market.

The first behavioral explanation is the limit-to-arbitrage. Cryptos market exhibits periods of large, recurrent arbitrage opportunities across exchanges (Makarov and Schoar 2020). According to Shleifer and Vishny (1997), sophisticated investors would quickly eliminate any return predictability arising from anomalies in a frictionless market without impediments to arbitrage. Hou and Moskowitz (2005) documented that the asset prices will delay to respond to information due to the impact of market frictions. Therefore, when the cryptos market is more frictional, the information travels slower in the cryptos network, and delayed cryptos will not exhibit price momentum. Analogous to market friction measures in the equity market (Brunnermeier, Nagel, and Pedersen 2008; Brunnermeier and Pedersen 2009), we employ the cryptos market volatility index (VCRIX) constructed by Kim, Trimborn, and Härdle.
which is consistent with Hou and Moskowitz (2005). In summary, positive, albeit insignificant, when market frictions are lower, the average portfolio returns are generally more significant differences across the high and low episodes of market frictions. The average portfolio returns are generally more significant differences across the high and low episodes of market frictions. Then, we calculate the average portfolio returns with in each subperiods of high and low investor attention. Table 6 by splitting the sample into two subperiods (high and low) by cutting at the median of the market friction indicators. Then, we calculate the average portfolio returns within each subperiods.

Table 6. Market frictions and the inter-crypto momentum portfolio returns.

|       | VCR IX | TED |
|-------|--------|-----|
|       | Low    | High | High–low |
|       |        |      |         |
| Loser | −0.03  | 0.10 | 0.13     |
| 2     | 0.43   | 0.36 | −0.06    |
| 3     | 0.81   | 0.73 | −0.08    |
| Winner| 1.27   | 1.08 | −0.19    |
| Winner–loser | 1.30 | 0.98 | −0.32 |
| t-statistic | 8.98 | 8.86 | −1.59 |

NOTE: This table reports the inter-crypto momentum portfolio returns (in %) in subperiods of high and low market frictions. t-statistics are computed based on Newey–West robust standard errors with four lags. The sample period spans from January 1, 2016 to December 31, 2018.

Table 7. Investor attention and the inter-crypto momentum portfolio returns.

|       | Bitcoin attention | Cryptos attention |
|-------|-------------------|------------------|
|       | Low               | High             | High–low |
| Loser | −0.02             | 0.20             | 0.22     |
| 2     | 0.24              | 0.66             | 0.42     |
| 3     | 0.48              | 1.26             | 0.78     |
| Winner| 1.02              | 1.36             | 0.33     |
| Winner–loser | 1.04 | 1.15 | 0.11 |
| t-statistic | 9.61 | 7.54 | 0.55 |

NOTE: This table reports the inter-crypto momentum portfolio returns (in %) in subperiods of high and low investor attention. t-statistics are computed based on Newey–West robust standard errors with four lags. The sample period spans from January 1, 2016 to December 31, 2018.

(2021) and TED spread (TED) as the market-friction indicators. Specifically, the volatility index affects the market liquidity, and TED determines the investors’ funding constraints. We split the sample into two subperiods (high and low) by cutting at the median of the market friction indicators. Then, we calculate and report the average portfolio returns within each subperiods in Table 6.

Evidently, the average portfolio returns do not exhibit significant differences across the high and low episodes of market frictions. The average portfolio returns are generally more positive, albeit insignificant, when market frictions are lower, which is consistent with Hou and Moskowitz (2005). In summary, the momentum spillover effect exists regardless of market frictions, indicating that the limit-to-arbitrage mechanism does not explain our crypto anomalies.

The second behavioral explanation is investor attention. Peng and Xiong (2006) pointed out that investors have limited attention and are subject to overconfidence. Therefore, they need to process information to infer the value of cryptos and overestimate the precision of the acquired information due to overconfidence. Guo et al. (2022) show that investor attention could spill over along the network linkages. Together, the limited attention, attention spillover, and overconfidence lead to an overreaction-induced momentum effect in the crypto network. Following Liu and Tsyvinski (2021), we proxy the market-wide investor attention by using the daily Google search volume for the words “crypto” and “cryptocurrency,” subtracting its 1-month historical average. Considering the leading position of Bitcoin in the crypto market (Griffin and Shams 2020), we also proxy the market-wide investor attention with the investor attention paid to Bitcoin, that is, the Google search volume for the word “Bitcoin.” Then, we repeat the test in Table 6 by splitting the sample into high and low attention subperiods.

The results are presented in Table 7. Clearly, the inter-crypto momentum portfolio returns do not exhibit significant differences in high and low episodes of investor attention. For example, the portfolio generates an average daily return of 1.15% during low Bitcoin attention episodes while retaining a 1.04% average daily return for the high ones. The difference is merely 0.11% and statistically insignificant. The portfolio return patterns under high and low crypto attention episodes are similar to that of Bitcoin attention. These results imply that investor attention is not the mechanism that drives the momentum spillover effects in the crypto market.

Finally, observing that government policy and crypto prices have a strong synchronization (Demir et al. 2018; Cheng and Yen 2020), it is natural to wonder whether the portfolio returns are related to underlying policy uncertainty. By representing uncertainty with investor belief dispersion (disagreement), Atmaz and Basak (2018) theoretically showed that higher disagreement leads to higher average bias and more overvaluation, thereby implying a negative disagreement–return relationship, as empirically documented in Brogaard and Detzel (2015). Based on Daniel, Hirshleifer, and Subrahmanyam (1998), when there are misinformed investors who underreact to firm-specific news, the price momentum will occur. Therefore, more uncertainty leads to price momentum. To examine this hypothesis, we employ the U.S. and Global economic policy uncertainty indices (EPU and GEPU) introduced by Baker, Bloom, and Davis (2016) to examine the portfolio performance in high and low uncertainty subperiods, respectively.

Table 8 summarizes the results. It is evident that the portfolio returns are statistically indifferent across high and low uncertainty subperiods. Specifically, in high EPU episodes, the portfolio return is 1.15% bps, whereas it is 1.01% in low EPU episodes. The difference between the portfolio returns is 0.14% with a Newey–West adjusted t-statistic of 0.77. When replacing the
EPU with the GEPU indicator, the difference in portfolio returns remains the same magnitude and statistically insignificant. Both results indicate that policy uncertainty or investor disagreement has no impact on the momentum spillover effects in the cryptos market.

In summary, by dissecting the portfolio returns with limit-to-arbitrage, investor attention, and economic policy uncertainty factors, we falsify the potential behavioral mechanisms that would generate momentum spillover effects in the cryptos market. The joint evidence in the current section and Section 4.2 strongly supports that the momentum spillover effects result from information propagation between fundamentally similar cryptos.

5. Conclusions

In this article, we examine the market segmentation problem and uncover the information propagation mechanism in the cryptos market by constructing a time-varying network of cryptos that combines return cross-predictability and technological similarities. To analyze this network, we propose a dynamic CASC algorithm for community detection and prove its classification consistency. We demonstrate the performance of our methodology on simulated networks and show that it generally performs better than other state-of-the-art methods. Empirically, we apply dynamic CASC to the cryptos network and obtain four latent communities. Based on the estimated communities, we show that investors can better diversify risk and obtain four latent communities. Based on the estimated communities, we show that investors can better diversify risk and obtain four latent communities. Based on the estimated communities, we show that investors can better diversify risk and obtain four latent communities. Based on the estimated communities, we show that investors can better diversify risk and obtain four latent communities.

This article mainly focuses on the uniform consistency of the classification method. Recent literature has developed inspiring framework for statistical inference on membership profiles as well as on estimation of low-rank matrices (e.g., Chen et al. 2019; Fan et al. 2022). It would be an interesting and crucial research direction to investigate how to accommodate dynamic networks and time-varying membership in these frameworks. We will leave them to future research.

### Appendix: List of Notations

| Notation | Definition |
|----------|------------|
| \( A_t, \hat{A}_t \) | Adjacency matrix at time \( t \). \( A_t(i, j) = 1 \) if \( i \rightarrow j \), and 0 otherwise. \( \hat{A}_t = E(\hat{A}_t) \). |
| \( Z_{Rt}, Z_{Ct} \) | Row-clustering matrices (\( N \)-by-\( K_R \)) and column-clustering matrices (\( N \)-by-\( K_C \)). |
| \( B_t \) | Block probability matrix at time \( t \), \( K_R \)-by-\( K_C \) dimensional. |
| \( \tau_{Rt}, \tau_{Ct} \) | Regularizer for graph Laplacian. Our default choices are average row and column degrees. |
| \( G_{RC}^t, \psi_C^t \) | Subgraph composed by the nodes within the \( k \)th row/column block at time \( t \). |
| \( \psi_R, \psi_C \) | Degree correction parameters that sum up to 1 within each block, that is, \( \sum_{i \in G_{k}^t} \psi_R^i = 1, \text{ for } k = 1, \ldots, K_R \) and \( \sum_{i \in D_{k}^t} \psi_C^i = 1, \text{ for } k = 1, \ldots, K_C \). |
| \( D_{R,t}, D_{C,t} \) | Diagonal row and column regularization matrices, where the diagonal elements are: \( D_{R,t}(i, i) = \frac{N}{\sum_{j=1}^{N} A_t(i, j) + \tau_{R,t}} \). |
| \( \Delta_{R,t}, \Delta_{C,t} \) | Population version of \( D_{R,t} \) and \( D_{C,t} \): \( \Delta_{R,t}(i, i) = \frac{N}{\sum_{j=1}^{N} A_t(i, j) + \tau_{R,t}} \). |
| \( L_{R,t}, L_{C,t} \) | Regularized graph Laplacian matrix, \( L_{R,t} = D_{R,t}^{-1/2} A_t D_{C,t}^{-1/2} \) and \( L_{C,t} = D_{C,t}^{-1/2} A_t D_{R,t}^{-1/2} \). |
| \( \mathbf{X}, \mathbf{X}' \) | Covariate matrix, \( N \)-by-\( d \) dimensional. \( \mathbf{X} = E(\mathbf{X}) \). |
| \( W_t, \mathbf{Y}_t \) | Weight matrix indicating the degree of covariate similarity, \( R \)-by-\( R \) dimensional. \( \mathbf{Y}_t = E(\mathbf{W}_t) \). |
| \( C_{R}^X, C_{C}^X \) | Covariate similarity matrix, \( C_{R}^X = AW_tX'X \) and \( C_{C}^X = X'W_tX \). |
| \( \alpha_t \) | Tuning parameter that controls for informational balance between adjacency matrices and covariate similarity matrices. \( \alpha_t \in [0, +\infty), \forall t \). |
| \( S_t, S^t \) | Similarity matrix, \( S_t = L_{R,t} + \alpha C_{R}^X \) and \( S^t = L_{C,t} + \alpha C_{C}^X \). |
| \( r, r_t \) | Bandwidth of the kernel function, and \( F_t = \{-r, \ldots, 0\} \). |
| \( \ell_t \) | Smoothing parameter of the kernel function. |
| \( W_t, \mathcal{W}_t \) | \( \ell_t \)th order kernel weights under the bandwidth \( r_t \). |
| \( S_{R,t} \) | Kernel smoothed similarity matrices: \( S_{R,t} = \frac{F_t^{-1}}{\sum_{i \in R,t} W_t(i)S_{R,t} + } \). |
| \( \hat{S}_{R,t} \) | Estimator for the kernel smoothed similarity matrices: \( \hat{S}_{R,t} = \frac{F_t^{-1}}{\sum_{i \in R,t} W_t(i)S_{R,t} + } \). |

### Supplementary Materials

The supplementary material contains details about tuning parameter choice, proofs of Theorem 1 in Section 3.
Acknowledgments

The authors sincerely thank the editor, Professor Jianqing Fan, an associate editor, and two anonymous referees for their helpful and insightful comments that have significantly improved the article. We are also grateful to Ying Chen, Dashaun Huang, Oliver Linton, Peter C. B. Phillips, Shuyang Sheng, Vladimir Spokoiny, Liangjun Su, Jun Yu, Yichong Zhang, and all participants of the 2017 "Crypto-Currencies in a Digital Economy" workshop at Humboldt-Universität zu Berlin, the 2018 China Meeting of the Econometric Society, the 2019 SH3 Conference on Econometrics, the 2019 STAT of ML Conference, and the 2019 SoFiE 12th Annual Conference for their helpful discussions and comments.

Disclosure Statement

The authors report there are no competing interests to declare.

Funding

Li Guo acknowledges financial support by the National Science Foundation of China (Project No. 72003040). Wolfgang K. Härdle acknowledges the financial support by the European Union’s Horizon 2020 research and innovation program “FIN-TECH: A Financial supervision and Technology compliance training programme” under the grant agreement No. 825215, the European Cooperation in Science & Technology (COST) Action grant CA19130, the Deutsche Forschungsgemeinschaft’s grant no. 19-28231X/CAS: XDA 23020303. Yubo Tao acknowledges the funding support by the Startup Research Grant (SRG2022-00016-FSS) from the University of Macau.

ORCID

Wolfgang Karl Härdle http://orcid.org/0000-0001-5600-3014
Yubo Tao http://orcid.org/0000-0002-1013-9984

References

Abadi, J., and Brunnermeier, M. (2018), “Blockchain Economics,” NBER working paper, 25407. [440]
Abbe, E., Fan, J., and Wang, K. (2022), “An $\ell_p$ Theory of PCA and Spectral Clustering,” Annals of Statistics, forthcoming. [438,446]
Acemoglu, D., Oxidaglar, A., and Tahbaz-Salehi, A. (2015), “Systemic Risk and Stability in Financial Networks,” American Economic Review, 105, 564–608. [439]
Ali, U., and Hirshleifer, D. (2020), “Shared Analyst Coverage: Unifying Momentum Spillover Effects,” Journal of Financial Economics, 136, 649–675. [438,451]
Altman, N. S. (1990), “Kernel Smoothing of Data with Correlated Errors,” Journal of the American Statistical Association, 85, 749–759. [444]
Amini, A. A., Chen, A., Bickel, P. I., and Levina, E. (2013), “Pseudo-Likelihood Methods for Community Detection in Large Sparse Networks,” Annals of Statistics, 41, 2097–2122. [438]
Amini, A. A., and Levina, E. (2018), “On Semidefinite Relaxations for the Block Model,” Annals of Statistics, 46, 149–179. [438]
Aodbia, D., Caskey, I., and Ozel, N. B. (2014), “Inter-Industry Network Structure and the Cross-Predictability of Earnings and Stock Returns,” Review of Accounting Studies, 19, 1191–1224. [439]
Atmaz, A., and Basak, S. (2018), “Belief Dispersion in the Stock Market,” Journal of Finance, 73, 1225–1279. [452]
Baker, S. R., Bloom, N., and Davis, S. J. (2016), “Measuring Economic Policy Uncertainty,” Quarterly Journal of Economics, 131, 1593–1636. [452]
Biais, B., Bisiere, C., Bouvard, M., and Casamatta, C. (2019), “The Blockchain Folk Theorem,” Review of Financial Studies, 32, 1662–1715. [440]
Bickel, P., Choi, D., Chang, X., and Zhang, H. (2013), “Asymmetric Normality of Maximum Likelihood and its Variational Approximation for Stochastic Blockmodels,” Annals of Statistics, 41, 1922–1943. [438]
Bickel, P. J., and Chen, A. (2009), “A Nonparametric View of Network Models and Newman–Girvan and other Modularities,” Proceedings of the National Academy of Sciences, 106, 21068–21073. [438]
Binkiewicz, N., Vogelstein, J. T., and Rohe, K. (2017), “Correlated-Assisted Spectral Clustering,” Biometrika, 104, 361–377. [438,442,446,447]
Brogaard, J., and Detzel, A. (2015), “The Asset Pricing Implications of Government Economic Policy Uncertainty,” Management Science, 61, 3–18. [452]
Brunnermeier, M. K., Nagel, S., and Pedersen, L. H. (2008), “Carry Trades and Currency Crashes,” NBER Macroeconomics Annual, 23, 313–348. [451]
Brunnermeier, M. K., and Pedersen, L. H. (2009), “Market Liquidity and Funding Liquidity,” Review of Financial Studies, 22, 2201–2238. [451]
Budish, E. (2018), “The Economic Limits of Bitcoin and the Blockchain,” NBER working paper, 24717. [440]
Cai, T. T., and Li, X. (2015), “Robust and Computationally Feasible Community Detection in the Presence of Arbitrary Outlier Nodes,” Annals of Statistics, 43, 1072–1059. [438]
Celisse, A., Daudin, J.-J., and Pierre, L. (2012), “Consistency of Maximum-Likelihood and Variational Estimators in the Stochastic Block Model,” Electronic Journal of Statistics, 6, 1847–1899. [438]
Chaim, P., and Laurini, M. P. (2018), “Volatility and Return Jumps in Bitcoin,” Economics Letters, 173, 158–163. [437,450]
Chen, C. Y.-H., Härdle, W. K., and Okhrin, Y. (2019), “Tail Event Driven Networks of SIFIs,” Journal of Econometrics, 208, 282–298. [439]
Chen, Y., Fan, J., Ma, C., and Yan, Y. (2019), “Inference and Uncertainty Quantification for Noisy Matrix Completion,” Proceedings of the National Academy of Sciences, 116, 22931–22937. [453]
Chen, Y., and Xu, J. (2016), “Statistical-Computational Tradeoffs in Planted Problems and Submatrix Localization with a Growing Number of Clusters and Submatrices,” Journal of Machine Learning Research, 17, 882–938. [438]
Cheng, H.-P., and Yen, K.-C. (2020), “The Relationship between the Economic Policy Uncertainty and the Cryptocurrency Market,” Finance Research Letters, 35, 101308. [452]
Cohen, L., and Frazzini, A. (2008), “Economic Links and Predictable Returns,” Journal of Finance, 63, 1977–2011. [438,439,451]
Cong, L. W., He, Z., and Li, J. (2021), “Decentralized Mining in Centralized Pools,” Review of Financial Studies, 34, 1191–1235. [440]
Cong, L. W., Li, Y., and Wang, N. (2021), “Tokenomics: Dynamic Adoption and Valuation,” Review of Financial Studies, 34, 1105–1155. [438,439]
——— (2022), “Token-based Platform Finance,” Journal of Financial Economics, 144, 972–991. [439]
Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (1998), “Investor Psychology and Security Market under-and Overreactions,” Journal of Finance, 53, 1839–1885. [452]
Decelle, A., Krzakala, F., Moore, C., and Zdeborová, L. (2011), “Asymptotic Analysis of the Stochastic Block Model for Modular Networks and its Algorithmic Applications,” Physical Review E, 84, 066106. [438]
Demir, E., Gorgoz, G., Lau, C. K. M., and Vigne, S. A. (2018), “Does Economic Policy Uncertainty Predict the Bitcoin Returns? An Empirical Investigation,” Finance Research Letters, 26, 145–149. [452]
Deshpande, Y., Sen, S., Montanari, A., and Mossel, E. (2018), “Contextual Stochastic Block Models,” in Advances in Neural Information Processing Systems, Vol. 31, pp. 8590–8602. [438,446]
Detzel, A., Liu, H., Strauss, J., Zhou, G., and Zhu, Y. (2021), “Learning and Predictability via Technical Analysis: Evidence from Bitcoin and Stocks with Hard-to-Value Fundamentals,” Financial Management, 50, 107–137. [437]
Easley, D., O’Hara, M., and Basu, S. (2019), “From Mining to Markets: The Evolution of Bitcoin Transaction Fees,” Journal of Financial Economics, 134, 91–109. [440]
Esmaeili, M., Saad, H. M., and Nosratinia, A. (2021), “Semidefinite Programming for Community Detection with Side Information,” IEEE Transactions on Network Science and Engineering, 8, 1957–1973. [438]
Fama, E. F., and French, K. R. (1993), “Common Risk Factors in Stock and Bond Returns,” Journal of Financial Economics, 33, 3–56. [442]

454 L. GUO, W. K. HÄRDLE, AND Y. TAO
Qin, T., and Rohe, K. (2013), "Regularized Spectral Clustering under the Degree-Corrected Stochastic Blockmodel," in Advances in Neural Information Processing Systems (Vol. 26), pp. 3120–3128. [442]
Rohe, K., Chatterjee, S., and Yu, B. (2011), "Spectral Clustering and the High-Dimensional Stochastic Blockmodel," Annals of Statistics, 39, 1878–1915. [438]
Rohe, K., Qin, T., and Yu, B. (2016), "Co-clustering Directed Graphs to Discover Asymmetries and Directional Communities," Proceedings of the National Academy of Sciences, 113, 12679–12684. [438,443,446,447,449]
Roy, A. D. (1952), "Safety First and the Holding of Assets," Econometrica, 20, 431–449. [438,450]
Rubinstein, M. (2002), "Markowitz’s "Portfolio Selection": A Fifty-Year Retrospective," Journal of Finance, 57, 1041–1045. [450]
Saleh, F. (2021), "Blockchain without Waste: Proof-of-Stake," Review of Financial Studies, 34, 1156–1190. [440]
Shleifer, A., and Vishny, R. W. (1997), "The Limits of Arbitrage," Journal of Finance, 52, 35–55. [451]
Sockin, M., and Xiong, W. (2020), "A Model of Cryptocurrencies," NBER working paper, 26816. [438,439]
Trimborn, S., and Härdle, W. K. (2018), "CRIX an Index for Cryptocurrencies," Journal of Empirical Finance, 49, 107–122. [437]
Wang, Y. R., and Bickel, P. J. (2017), "Likelihood-based Model Selection for Stochastic Block Models," Annals of Statistics, 45, 500–528. [438]
Wong, H., and Feng, Y. (2022), "Community Detection with Nodal Information: Likelihood and its Variational Approximation," Stat, 11, e428. [438]
Yan, B., and Sarkar, P. (2021), "Covariate Regularized Community Detection in Sparse Graphs," Journal of the American Statistical Association, 116, 734–745. [438,445]
Yermack, D. (2017), "Corporate Governance and Blockchains," Review of Finance, 21, 7–31. [438]
Zhang, Y., Levina, E., and Zhu, J. (2016), "Community Detection in Networks with Node Features," Electronic Journal of Statistics, 10, 3153–3178. [438]
Zhang, Y., Poux-Berthe, M., Wells, C., Koc-Michalska, K., and Rohe, K. (2018), "Discovering Political Topics in Facebook Discussion Threads with Graph Contextualization," Annals of Applied Statistics, 12, 1096–1123. [438,442,446,447]
Zhao, J., Liu, X., Wang, H., and Leng, C. (2022), "Dimension Reduction for Covariates in Network Data," Biometrika, 109, 85–102. [438]
Zhao, Y., Levina, E., and Zhu, J. (2012), "Consistency of Community Detection in Networks under Degree-Corrected Stochastic Block Models," Annals of Statistics, 40, 2266–2292. [438]
Zhou, Z., and Amini, A. A. (2019), "Analysis of Spectral Clustering Algorithms for Community Detection: The General Bipartite Setting," Journal of Machine Learning Research, 20, 1774–1820. [438]
Zhou, Z., and Amini, A. A. (2020), "Optimal Bipartite Network Clustering," Journal of Machine Learning Research, 21, 1–68. [438]
Zhu, X., Cai, Z., and Ma, Y. (2021), "Network Functional Varying Coefficient Model," Journal of the American Statistical Association, forthcoming. [438]
Zou, H. (2006), "The Adaptive Lasso and its Oracle Properties," Journal of the American Statistical Association, 101, 1418–1429. [439]
Zou, T., Lan, W., Li, R., and Tsai, C.-L. (2022), "Inference on Covariance-Mean Regression," Journal of Econometrics, 230, 318–338. [438]
Zou, T., Lan, W., Wang, H., and Tsai, C.-L. (2017), "Covariance Regression Analysis," Journal of the American Statistical Association, 112, 266–281. [438]