ORIGINAL ARTICLE

THRESHOLD NETWORK GARCH MODEL

YUE PAN AND JIAZHU PAN[ID]

Department of Mathematics and Statistics, University of Strathclyde, Glasgow, UK

Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model and its variations have been widely adopted in the study of financial volatilities, while the extension of GARCH-type models to high-dimensional data is always difficult because of over-parameterization and computational complexity. In this article, we propose a multi-variate GARCH-type model that can simplify the parameterization by utilizing the network structure that can be appropriately specified for certain types of high-dimensional data. The asymmetry in the dynamics of volatilities is also considered as our model adopts a threshold structure. To enable our model to handle data with extremely high dimension, we investigate the near-epoch dependence (NED) of our model, and the asymptotic properties of our quasi-maximum-likelihood-estimator (QMLE) are derived from the limit theorems for NED random fields. Simulations are conducted to test our theoretical results. At last we fit our model to log-returns of four groups of stocks and the results indicate that bad news is not necessarily more influential on volatility if the network effects are considered.

1. INTRODUCTION

To pursue maximum return or to circumvent potential risk, investors constantly revise their portfolio according to any related information. Understanding how the volatility of financial assets responds to new information is crucial in risk management and a widely studied area in econometrics and statistics. In the literature, statistical models that describe the formation of financial risks have been developed and conducted in practice. The Autoregressive Conditional Heteroscedasticity (ARCH) model was proposed by Engle (1982) for estimating the variance of United Kingdom's inflation. In an ARCH(p) model, the volatilities of returns are affected by up to p lags of past observations. Bollerslev (1986) then proposed a generalized ARCH model (GARCH), to accommodate longer memory of past observations. It has become one of the most popular models in econometrics ever since, and numerous variations of GARCH model have been developed for modeling volatility with complicated structures. See Teräsvirta (2009) for a survey of different GARCH-type models.

When we study risks of multiple assets simultaneously, the conditional variances that represent individual risks are of interest as well as conditional covariances that represent risk-sharing relationships. On the other hand, risk of a particular individual could be affected by covariates of itself, and of those who closely related to it. This leads to the need of extending the GARCH-type models into the multi-variate case. For an N dimensional time series

*Correspondence to: Jiazhu Pan, Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, UK. Email:
jiazhu.pan@strath.ac.uk

$\{\mathbf{y}_t\}$, a canonical expression of multi-variate GARCH would be

$$
\begin{aligned}
\mathbf{y}_t &= H_t^{1/2}\mathbf{z}_t, \\
H_t &= g\left(\mathbf{y}_{t-1}, H_{t-1}\right),
\end{aligned}
\tag{1}
$$

where random vector $\mathbf{z}_t$ satisfies $\mathbb{E}(\mathbf{z}_t) = \mathbf{0}$ and $var(\mathbf{z}_t) = I_N$ and there could be various specifications of the function $g(\cdot)$ as it represents the structure of the conditional covariance matrix H_t. For more details on this subject, an excellent survey paper by Bauwens *et al.* (2006) on the family of multi-variate GARCH models is recommended.

However, in terms of parameter estimation, there are major challenges that could make multi-variate GARCH (MGARCH) models inapplicable in empirical analysis when it comes to dealing with high-dimensional data. For example, the number of parameters rises at the speed of $\mathcal{O}(N^4)$ in vectorized GARCH model (VEC-GARCH) proposed by Bollerslev *et al.* (1988). Over parameterized specification causes high computational complexity, and makes it problematic to derive conditions for positive definiteness of the conditional covariance matrix H_t. Plenty of efforts have been made in the literature, diagonal VEC-GARCH (DVEC-GARCH) model by Boller-slev *et al.* (1988) and Baba-Engle-Kraft-Kroner GARCH (BEKK-GARCH) model by Engle and Kroner (1995) are proposed with the aim of simplifying the conditions of positive definite by imposing structural restrictions on the conditional covariance matrix. The number of parameters could also be significantly reduced to $\mathcal{O}(N^2)$ in the Constant Conditional Correlation GARCH (CCC-GARCH) by Bollerslev (1990) and Dynamic Conditional Correlation GARCH (DCC-GARCH) by Engle (2002) and Tse and Tsui (2001). On the other hand, as an alternative way to overcome the over-parameterization problem, the idea of factor variables was imposed by Engle *et al.* (1990) on the multi-variate ARCH model as a dimension reduction technique. This idea is later introduced to MGARCH model by Bollerslev and Engle (1993), as well as succeeding work of Pan *et al.* (2010), Hu and Tsay (2014) and Li *et al.* (2016). At certain application scenarios when there is network structure behind the data we are interested in, multiple variables are connected and a multi-variate GARCH-type model could be fitted. These variations of MGARCH models solved the over-parameterization problem to some extent, but the number of parameters still expands along with the number of dimension nevertheless. With this flaw, MGARCH models could only be imposed on data with a small number of dimensions, such as stock indices of multiple markets or exchange rates of two currencies (see Karolyi, 1995 and Tse and Tsui, 2001).

Despite aforementioned difficulties due to dimensionality, for some specific types of multi-variate data where the connections between different components are actually observable, it is still possible to significantly simplify the model setup in the following aspects:

1. Instead of considering both volatilities and co-volatilities, we focus on studying the dynamics of volatilities only,
2. And instead of parameterizing every cross-individual effect, appropriate network structure can be embedded into the model.

In many cases such network structure can provide sufficient information about how the influence of pulses travel through edges between individual nodes. For instance, Nitzan and Libai (2011) found that customers connected with a defecting neighbor are 80% more likely to cancel their cellular service, and Goel and Goldstein (2014) concluded that the accuracy of individual behavior prediction can be significantly improved based on network data compared with conventional marketing practices.

Zhou *et al.* (2020) proposed a network GARCH model (see (2) for detailed specification of this model) that significantly reduces the parameterization complexity – the number of parameters remains fixed in their model no matter how large is the dimension N. However, they did not fully utilize such an advantage as their discussion on parameter estimation is limited to the case when N is fixed. Such setting narrows the variety of scenarios where their model could be applied since the size of network is often extremely large. In a study

on social network that consists of 2982 users, Zhu *et al.* (2017) proposed a network AR model and the corresponding least squares estimator is proved to remain valid when the sample size $T \to \infty$ and the dimension $N \to \infty$. Compared with their AR-type model, the unobserved volatility processes raise difficulties in extending such properties to GARCH-type models. We manage to address this problem by considering our network model as a spatial process on a two-dimensional lattice and adopting the asymptotic theorems for random fields proposed by Jenish and Prucha (2012) in the estimation of parameters. Since their limit theorems require NED, we will show such properties under certain restrictions on parameters and network structure. The idea of using limit theorems of spatial processes in the inference of high-dimensional time series has been considered by Xu *et al.* (2022) in the instrumental variable quantile regression estimation of their dynamic network quantile regression model. In this article, we will first introduce this idea into the estimation of high-dimensional GARCH-type models.

Aside from data with a diverging number of dimensions, we also aim to enable our model to handle data with asymmetry that was observed in empirical work such that positive and negative pulses affect volatilities differently in magnitude as well as in direction. While most GARCH-type models have an implicit assumption that the volatilities respond equally to the magnitude of positive and negative returns, Glosten *et al.* (1993) proposed a Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) model with a threshold structure, allowing the volatility to act asymmetrically in magnitude responding to positive and negative pulses. The threshold GARCH (TGARCH) by Zakoïan (1994) also accommodates the asymmetry, but in magnitude of influence on conditional standard deviation. The exponential GARCH (EGARCH) of Nelson (1991) takes log transformation on the conditional variances, lifting the limitation of non-negative coefficients in conventional GARCH-type models and making it possible for their model to explain asymmetry in the direction of how volatility change corresponding to positive and negative news.

To study the asymmetrical dynamics in the volatilities of high-dimensional financial data, we propose a threshold network GARCH (TNGARCH) in Section 2. Stationarity conditions of this model are derived in Section 3.1 with fixed N. In Section 3.2 we prove the L_2-NED of proposed model under certain restrictions. The asymptotic properties of QMLE are investigated in Section 4, in the case when $T \to \infty$ and $N \to \infty$ at a lower rate. Then we propose a Wald statistic in Section 5.1, to test the existence of threshold effect, and in Section 5.2 we introduce a test for high-dimensional white noise proposed by Li *et al.* (2019). In Section 6, our methodology is tested on simulated data that are generated based on four different kinds of network structure. We observed an asymmetry that is different from existing literature, in how much the volatility responds to good news and bad news at individual level by applying our model to high-dimensional time series of log returns in Section 7. At last in Section 8, conclusions and potential directions for future research are summarized.

2. MODEL SETUP

Consider an undirected and weightless network with N nodes. Define the adjacency matrix $A = (a_{ij})_{1 \le i,j \le N}$, where $a_{ij} = 1$ if there is a connection between node i and node j, otherwise $a_{ij} = 0$. Besides, self-connection is not allowed for any node i by letting $a_{ii} = 0$.

The *connection* can be defined differently with respect to practical scenarios, such as two social network accounts in a mutual followship, or two stocks who share at least one of top shareholders. As an interpretation of the network structure, A is symmetric since $a_{ij} = a_{ji}$, hence for any node i, the **out-degree** $d_i^{(\text{out})} = \sum_{j=1}^{N} a_{ij}$ is equal to the **in-degree** $d_i^{(\text{in})} = \sum_{j=1}^{N} a_{ji}$ and we use d_i to denote both for convenience.

For any node i in this network, let y_{it} be the observation at time t, and h_{it} be the unobservable conditional heteroscedasticity of y_{it}, i.e. $h_{it} := \text{var}(y_{it}|\mathcal{H}_{t-1})$ where $\mathcal{H}_{t-1}$ denotes the σ-algebra consisting of all available information up to $t-1$. A network GARCH(1, 1) specification of the conditional variance incorporates the network effect:

$$h_{it} = \omega + \alpha y_{i,t-1}^2 + \lambda \sum_{j=1}^{N} w_{ij} y_{j,t-1}^2 + \beta h_{i,t-1}, \quad i = 1, 2, \dots, N. \tag{2}$$

Model (2) indicates that the volatility h_{it} of each stock i, is influenced by not only its own previous change of price measured by $y_{i,t-1}^2$, but also the average (with weight $w_{ij} = \frac{a_{ij}}{d_i}$) of $y_{j,t-1}^2$ for all node j that are related to node i. To ensure the positiveness of conditional variance, it is constrained that $\omega > 0$ while $\alpha, \lambda, \beta \geq 0$.

To model the asymmetry in volatility, our TNGARCH model contains the threshold structure comparing with model (2). A TNGARCH(1,1) model is specified as follows:

$$
\begin{aligned}
y_{it} &= \varepsilon_{it}\sqrt{h_{it}}, \\
h_{it} &= \omega + \left(\alpha^{(1)}1_{\{y_{i,t-1}\geq 0\}} + \alpha^{(2)}1_{\{y_{i,t-1}<0\}}\right)y_{i,t-1}^2 + \lambda\sum_{j=1}^{N}w_{ij}y_{j,t-1}^2 + \beta h_{i,t-1}, \\
i &= 1, 2, \ldots, N,
\end{aligned}
\tag{3}
$$

where $1_{\{\cdot\}}$ is the indicator function. To assure the positiveness of h_{it}, the coefficients $\omega, \alpha^{(1)}, \alpha^{(2)}, \lambda$ and β are assumed to have the same constraints as in (2). $\{\varepsilon_{it}\}$ is a white noise process satisfying the following assumption:

Assumption 1. $\{\varepsilon_{it}\}$ is i.i.d. across i and t, with non-degenerate distribution, mean 0 and variance 1.

This assumption allows us to investigate, in the next section, the conditions for our model to have a unique strictly stationary solution, which serves as a precondition for further discussion on parameter estimation and statistical inference.

3. STATIONARITY AND NEAR-EPOCH DEPENDENCE

To derive the conditions under which model (3) is strictly stationary, we rewrite the conditional variance process in vector form

$$
\mathbf{h}_t = \omega \mathbf{1}_N + B_{t-1}\mathbf{h}_{t-1},
\tag{4}
$$

with notations as follows:

$$
\begin{aligned}
\mathbf{h}_t &= \left(h_{1t}, h_{2t}, \ldots, h_{Nt}\right)' \in \mathbb{R}^N, \\
\mathbf{1}_N &= (1, 1, \ldots, 1)' \in \mathbb{R}^N, \\
B_{t-1} &= \alpha^{(1)}R_{t-1}E_{t-1} + \alpha^{(2)}(I_N - R_{t-1})E_{t-1} + \lambda D^{-1}AE_{t-1} + \beta I_N, \\
R_{t-1} &= \text{diag}\left\{1_{\{y_{1,t-1}\geq 0\}}, 1_{\{y_{2,t-1}\geq 0\}}, \ldots, 1_{\{y_{N,t-1}\geq 0\}}\right\}, \\
E_{t-1} &= \text{diag}\left\{\varepsilon_{1,t-1}^2, \varepsilon_{2,t-1}^2, \ldots, \varepsilon_{N,t-1}^2\right\}, \\
D &= \text{diag}\left\{d_1, d_2, \ldots, d_N\right\}.
\end{aligned}
$$

In Section 3.1, the stationarity of (4) when N is a fixed number will be discussed. However, to estimate the parameters when $N \to \infty$, limit theorems based on stationarity and ergodicity of fixed-dimensional time series are not sufficient. Therefore, in Section 3.2 we will discuss the near-epoch dependence for a random field that supports the adoption of limit theorems for spatial processes in the subsequent sections.

3.1. Stationarity with N being Fixed

Since $y_{it} = \varepsilon_{it}\sqrt{h_{it}}$, $y_{it} \geq 0$ is equivalent to $\varepsilon_{it} \geq 0$. Hence

$$
R_{t-1} = \text{diag}\left\{1_{\{\varepsilon_{1,t-1}\geq 0\}}, 1_{\{\varepsilon_{2,t-1}\geq 0\}}, \ldots, 1_{\{\varepsilon_{N,t-1}\geq 0\}}\right\}.
$$

In this case, the random matrices $\{B_t\}$ are i.i.d. and model (4) is a generalized autoregressive equation by definition 1.4 in Bougerol and Picard (1992). It is easy to verify that $\mathbb{E}(\log^+ \|B_0\|_*) < \infty$. Therefore, the top Lyapunov exponent associated to $\{B_t\}$ is well-defined as follows:

$$\gamma := \inf \left\{ \mathbb{E} \left(\frac{1}{t+1} \log \|B_t B_{t-1} \dots B_0\|_* \right), \ t \in N \right\}, \tag{5}$$

where $\|\cdot\|_*$ is an operator norm of $N \times N$ matrices, corresponding to any norm on $\mathbb{R}^N$ through

$$\|M\|_* = \sup \left\{ \|M\mathbf{x}\| / \|\mathbf{x}\| ; \ \mathbf{x} \in \mathbb{R}^N, \ \mathbf{x} \neq 0 \right\}.$$

According to theorem 3.2 in Bougerol and Picard (1992), the series

$$\mathbf{h}_t = \omega \mathbf{1}_N + \omega \sum_{k=1}^{\infty} B_{t-1} \dots B_{t-k} \mathbf{1}_N, \tag{6}$$

is the unique strictly stationary and ergodic solution of model (4) if and only if the Lyapunov exponent $\gamma < 0$. Under this condition, process $\{\mathbf{y}_t\}$ is also strictly stationary and ergodic where $\mathbf{y}_t = \left(y_{1t}, y_{2t}, \dots, y_{Nt} \right)' \in \mathbb{R}^N$ since we could easily construct a continuous function $\Lambda : \mathbb{R}^N \to \mathbb{R}^N$ according to (3) such that $\mathbf{y}_t = \Lambda(\mathbf{h}_t)$. Besides, since $y_{it} = \varepsilon_{it} \sqrt{h_{it}}$, the almost sure convergence of (6) guarantees that $\mathbb{E}(h_{it}) < \infty$ for any i. Thus, $\mathbb{E}\|\mathbf{y}_t\|^2 = \sum_{i=1}^{N} \mathbb{E}(h_{it}) < \infty$ with $\|\cdot\|$ being an Euclidean norm.

By the subadditive ergodic theorem in Kingman (1973),

$$\gamma = \lim_{t \to \infty} \frac{1}{t+1} \log \|B_t B_{t-1} \dots B_0\|_*$$

almost surely. In this case, γ could be approximated through computer simulation technique given a specific distribution of ε_{it}. For the purpose of reducing computation complexity, we derive a sufficient condition that is simple and much easier to verify.

Theorem 1. Under Assumption 1, model (4) has a unique strictly stationary and ergodic solution in the form (6) if

$$\max \left\{ \alpha^{(1)}, \alpha^{(2)} \right\} + \beta + \lambda < 1. \tag{7}$$

3.2. Near-Epoch Dependence for Random Fields

Let $D := \{(i,t) : i \in \mathbb{N}_+, t \in \mathbb{Z}\}$ be a lattice on space $\mathbb{R}^2$, and $\rho((i,t),(j,\tau)) := \max\{|i-j|, |t-\tau|\}$ measures the distance between any two locations $(i,t), (j,\tau) \in D$. Assume we have observations from model (3) $\{y_{it}, 1 \leq i \leq N, 1 \leq t \leq T\}$, then these observations could be regarded as triangular array of random fields $\{y_{it} : (i,t) \in D_{NT}, NT \geq 1\}$ with $\{D_{NT}, NT \geq 1\}$ being a series of finite rectangular lattices $D_{NT} := \{(i,t) : 1 \leq i \leq N, 1 \leq t \leq T\}$. Then the growth of sample size is ensured by unbounded expansion of D_{NT} as $NT \to \infty$. Such expansion is represented as $|D_{NT}|_c \to \infty$, where $|\cdot|_c$ is the cardinality of D_{NT}. The discussions on the asymptotic behaviors of random fields concern only the expansion of sample region, therefore the theoretical results derived in this section will apply as long as $|D_{NT}|_c = NT \to \infty$.

Let $\|\cdot\|_p$ denote the L_p-norm, i.e. $\|X\|_p := (\mathbb{E}|X|^p)^{1/p}$ for an arbitrary random variable X. The definition of NED random fields is given as follows (see definition 1 in Jenish and Prucha, 2012):

Definition 1. A triangular array of random fields $\mathcal{Y} := \{y_{it} : (i,t) \in D_{NT}, NT \geq 1\}$ is said to be L_p-NED ($p \geq 1$) on $\mathcal{E} = \{\varepsilon_{it} : (i,t) \in D\}$ if $\sup_{(i,t) \in D} \|y_{it}\|_p < \infty$, and

$$\|y_{it} - E(y_{it}|\mathcal{F}_{it}(s))\|_p \leq d_{it} \psi(s),$$

where $\mathcal{F}_{it}(s) := \sigma\{\varepsilon_{j\tau} : \rho((i,t),(j,\tau)) \le s\}$, $\psi(s)$ is some non-negative sequence with $\lim_{s\to\infty} \psi(s) = 0$, and $\{d_{it} : (i,t) \in D_{NT}, NT \ge 1\}$ is an array of finite positive constants.

Remark. If $\psi(s) = \mathcal{O}(s^{-\mu})$ for some $\mu > 0$, then $\mathcal{Y}$ is said to be L_p-NED on $\mathcal{E}$ of size-μ; If $\psi(s) = \mathcal{O}(\rho^s)$ for some $0 < \rho < 1$, then $\mathcal{Y}$ is said to be L_p-NED on $\mathcal{E}$ geometrically; If $\sup_{(i,t)\in D} d_{it} < \infty$, then $\mathcal{Y}$ is said to be uniformly L_p-NED on random field $\mathcal{E}$. Note that geometric NED means NED of size-μ for all $\mu > 0$.

We need following assumptions before discussing the NED property of $\mathcal{Y}$. Assumption 2 is needed to prove that $\sup_{(i,t)\in D} \|h_{it}\|_2 < \infty$; Assumption 3 put restriction on the sparsity of the network: the power of connections between two nodes decays with their distance in case (a), or two nodes are only connected if they are sufficiently close in case (b). Similar restrictions on the network structure could also be seen in assumption 3 by Xu and Lee (2015) and assumption 3.2 by Xu *et al.* (2022).

Assumption 2. There exists $\kappa_4 := \mathbb{E}\varepsilon_{it}^4 < \infty$, such that

$$\kappa_4 \big(\max\{\alpha^{(1)}, \alpha^{(2)}\} + \beta + \lambda\big)^2 < 1.$$

Assumption 3. The row-normalized adjacency matrix W satisfies one of following conditions:

(a). $w_{ij} = \mathcal{O}(|i-j|^{-\frac{\mu+2}{2}})$ for some $\mu > 0$;
(b). $w_{ij} \ne 0$ if $|i-j| \le K$ for some constant $K \ge 1$, and $w_{ij} = 0$ otherwise.

Theorem 2. If condition (7) holds, under Assumptions 1, 2 and 3(a), $\{h_{it} : (i,t) \in D_{NT}, NT \ge 1\}$ is uniformly L_2-NED on $\{\varepsilon_{it} : (i,t) \in D\}$ of size-μ, where the NED size μ is the constant in Assumption 3(a). Moreover, if Assumption 3(b) holds instead of 3(a), $\{h_{it} : (i,t) \in D_{NT}, NT \ge 1\}$ is uniformly and geometrically L_2-NED on $\{\varepsilon_{it} : (i,t) \in D\}$.

Remark. Note that $\|y_{it}^2 - \mathbb{E}(y_{it}^2|\mathcal{F}_{it}(s))\|_2 = \|\varepsilon_{it}^2\|_2 \|h_{it} - \mathbb{E}(h_{it}^2|\mathcal{F}_{it}(s))\|_2$, then Assumption 2 facilitates the L_2-NED of y_{it}^2's given the L_2-NED of h_{it}'s. Besides, since $h_{it} \ge \omega > 0$, it is easy to verify that $\sqrt{h_{it}}$ is a Lipschitz transformation of h_{it} using mean value theorem. Then proposition 2 in Jenish and Prucha (2012) allows $\sqrt{h_{it}}$'s to inherit the NED properties from h_{it}'s, therefore we could also verify that y_{it}'s are also L_2-NED.

4. PARAMETER ESTIMATION

From model (3) we have observations $\{y_{it} : (i,t) \in D_{NT}, NT \ge 1\}$ with respect to true parameters $\theta_0 := (\omega_0, \alpha_0^{(1)}, \alpha_0^{(2)}, \lambda_0, \beta_0)' \in \mathbb{R}^5$. Based on the infinite past of observations, the quasi log-likelihood function is

$$L_{NT}(\theta) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} l_{it}(\theta),$$

$$l_{it}(\theta) = \log \sigma_{it}^2(\theta) + \frac{y_{it}^2}{\sigma_{it}^2(\theta)}, \tag{8}$$

where σ_{it}^2 is generated from model (3) as

$$\sigma_{it}^2 = \omega + \left\{ \alpha^{(1)} 1_{\{y_{i,t-1} \ge 0\}} + \alpha^{(2)} 1_{\{y_{i,t-1} < 0\}} \right\} y_{i,t-1}^2 + \lambda d_i^{-1} \sum_{j=1}^{N} a_{ij} y_{j,t-1}^2 + \beta \sigma_{i,t-1}^2,$$

and $\theta := (\omega, \alpha^{(1)}, \alpha^{(2)}, \lambda, \beta)' \in \mathbb{R}^5$ is the parameter vector.

Since the evaluation of the exact value of (8) is infeasible in practice, it is convenient to approximate (8) with

$$\tilde{L}_{NT}(\theta) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{l}_{it}(\theta),$$

$$\tilde{l}_{it}(\theta) = \log \tilde{\sigma}_{it}^2(\theta) + \frac{y_{it}^2}{\tilde{\sigma}_{it}^2(\theta)}, \tag{9}$$

where $\tilde{\sigma}_{it}^2$ is also generated from model (3) but with initial value $\tilde{\sigma}_{i0}^2 = 0$. And the QMLE of $\theta \in \Theta$ is given by

$$\hat{\theta}_{NT} := \underset{\theta \in \Theta}{\operatorname{argmin}}\, \tilde{L}_{NT}(\theta).$$

To prove the asymptotic properties of $\hat{\theta}_{NT}$, we need following assumptions aside from those required by Theorem 2:

Assumption 4. Θ is a compact subset of $\left\{\theta : \omega > 0, \alpha^{(1)} > 0, \alpha^{(2)} > 0, \lambda > 0, \beta > 0\right\}$ such that all $\theta \in \Theta$ satisfy (7) and Assumption 2, and the true parameter $\theta_0 \in \Theta$ is an interior point of Θ.

Assumption 5. $\sup_{(i,t)\in D} \sup_{\theta \in \Theta} \left\| \sigma_{it}^2(\theta) \right\|_p < \infty$ for some $p > 1$.

Assumption 6. $\mathbb{E}\varepsilon_{it}^{4r} < \infty$ for some $r > 2$, and following bounds exists:

$$\sup_{(i,t)\in D} \left\| \sigma_{it}^2(\theta_0) \right\|_{2r} < \infty;$$

$$\sup_{(i,t)\in D} \left\| \frac{\partial}{\partial \theta_k} \sigma_{i,t}^2(\theta_0) \right\|_{2r} < \infty;$$

$$\sup_{(i,t)\in D} \left\| \frac{\partial^2}{\partial \theta_j \partial_k \theta} \sigma_{it}^2(\theta_0) \right\|_2 < \infty,$$

where θ_k denotes the kth component of parameter vector θ.

Assumption 7. The NED-size μ in Theorem 2 satisfies $\frac{r-2}{2r-2}\mu > 2$ with r being the one in Assumption 6.

Assumptions 4 is also required by Zhou *et al.* (2020) to prove the asymptotic properties in the case when N being fixed. With both $T \to \infty$ and $N \to \infty$, additional assumptions as above are required to adopt the limit theorems of random fields. Specifically, Assumption 5 is required for $l_{it}(\theta)$ to satisfy the bound condition of law of large numbers (LLN) for random fields (assumption 2(a) in Jenish and Prucha, 2012); Assumption 6 facilitates the heredity of NED property from $\sigma_{it}^2(\theta_0)$ to the more complicated forms of first-order and second-order derivatives of $L_{NT}(\theta_0)$; Assumption 7 is a constraint on the decaying rate of NED coefficients, which is required by the central limit theorem (CLT) for random fields (assumption 4(c) in Jenish and Prucha, 2012). Of course, as we have remarked after Definition 1 that geometric NED means NED of size-μ for all $\mu > 0$, therefore Assumption 7 would be trivial under geometric NED.

Theorem 3. Under Assumptions required by Theorem 2, Assumption 4 and Assumption 5, the quasi-maximum likelihood estimator $\hat{\theta}_{NT}$ is consistent, i.e.

$$\hat{\theta}_{NT} \overset{p}{\longrightarrow} \theta_0,$$

as $NT \to \infty$; If Assumptions 6 and 7 also hold, and the smallest eigenvalue $\lambda_{\min}(\Sigma_{NT})$ of

$$\Sigma_{NT} := \frac{\kappa_4 - 1}{NT} \sum_{(i,t)\in D_{NT}} \mathbb{E}\left[\frac{1}{\sigma_{it}^4(\theta_0)} \frac{\partial}{\partial \theta} \sigma_{it}^2(\theta_0) \frac{\partial}{\partial \theta'} \sigma_{it}^2(\theta_0) \right],$$

satisfies that

$$\inf_{NT \geq 1} \lambda_{min}(\Sigma_{NT}) > 0, \tag{10}$$

then $\hat{\theta}_{NT}$ is asymptotically normal as $NT \to \infty$ and $N = o(T)$:

$$\sqrt{NT}\Sigma_{NT}^{1/2}(\hat{\theta}_{NT} - \theta_0) \overset{d}{\longrightarrow} \mathrm{N}(0, (\kappa_4 - 1)^2 I_5),$$

where I_5 is the 5×5 identity matrix.

Remark. Condition (10) can be implied if the smallest eigenvalues $\lambda_{\min}^{(i,t)}$ of $\mathbb{E}\left[\frac{1}{\sigma_{it}^4(\theta_0)}\frac{\partial}{\partial\theta}\sigma_{it}^2(\theta_0)\frac{\partial}{\partial\theta'}\sigma_{it}^2(\theta_0)\right]$ satisfy that $\inf_{NT\geq 1}\inf_{(i,t)\in D_{NT}}\lambda_{min}^{(i,t)} > 0$.

As we will show in the proof of Proposition 5.1, κ_4 and Σ_{NT} above could be approximated by

$$\hat{\kappa}_4 := \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{y_{it}^4}{\tilde{\sigma}_{it}^4(\hat{\theta}_{NT})}, \tag{11}$$

and

$$\hat{\Sigma}_{NT} := \frac{\hat{\kappa}_4 - 1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\frac{1}{\tilde{\sigma}_{it}^4(\hat{\theta}_{NT})}\frac{\partial\tilde{\sigma}_{it}^2(\hat{\theta}_{NT})}{\partial\theta}\frac{\partial\tilde{\sigma}_{it}^2(\hat{\theta}_{NT})}{\partial\theta'}\right], \tag{12}$$

respectively. The later could be calculated recursively as $\frac{\partial}{\partial\theta}\tilde{\sigma}_{it}^2(\hat{\theta}_{NT}) = \tilde{\mathbf{u}}_{i,t-1} + \hat{\beta}\frac{\partial}{\partial\theta}\tilde{\sigma}_{i,t-1}^2(\hat{\theta}_{NT})$ where

$$\tilde{\mathbf{u}}_{i,t-1} = \begin{pmatrix} 1 \\ y_{i,t-1}^2 1_{\{\hat{\varepsilon}_{i,t-1}\geq 0\}} \\ y_{i,t-1}^2 1_{\{\hat{\varepsilon}_{i,t-1}<0\}} \\ \sum_{j=1}^{N}w_{i,j}y_{j,t-1}^2 \\ \tilde{\sigma}_{i,t-1}^2(\hat{\theta}_{NT}) \end{pmatrix}.$$

5. TESTS ON THRESHOLD EFFECT AND RESIDUALS

5.1. A Wald Test for the Threshold Effect

Given a null hypothesis

$$H_0 : \Gamma\theta_0 = \eta, \tag{13}$$

where Γ is an $s \times 5$ matrix with rank s and η is an s-dimensional vector, we could define a Wald test statistic as follows:

$$W_{NT} := (\Gamma\hat{\theta}_{NT} - \eta)'\left\{\frac{\Gamma}{NT}(\hat{\kappa}_4 - 1)^2\hat{\Sigma}_{NT}^{-1}\Gamma'\right\}^{-1}(\Gamma\hat{\theta}_{NT} - \eta), \tag{14}$$

where $\hat{\kappa}_4$ and $\hat{\Sigma}_{NT}$ are defined in (11) and (12).

By the asymptotic normality of $\hat{\theta}_{NT}$, W_{NT} could also be proved to follow a canonical asymptotic distribution as in Proposition 5.1.

Proposition 5.1. Under the same assumptions required by Theorem 3, as $T \to \infty$ and $N = o(T)$, the Wald test statistic defined in (14) asymptotically follows a χ^2 distribution with degree of freedom s, i.e.

$$W_{NT} \xrightarrow{d} \chi_s^2.$$

5.2. A White Noise Test on the Residuals

There has been a large literature investigating high-dimensional time series models, including Xu and Lee (2015), Zhu *et al.* (2017) and Xu *et al.* (2022) among others, but none of them has used diagnostic tools to check the model adequacy. In this section, we will introduce a high-dimensional white noise test developed by Li *et al.* (2019) that can be applied to the diagnostic of high-dimensional models including ours.

Assume we have residuals $\{\mathbf{r}_t : 1 \leq t \leq T\}$, where $\mathbf{r}_t := (r_{1t}, \ldots, r_{Nt})'$. We want to test whether $\{\mathbf{r}_t : 1 \leq t \leq T\}$ are high-dimensional white noises, i.e. there exists a matrix P such that

$$H_0 : \mathbf{r}_t = P\mathbf{z}_t, \tag{15}$$

where $\mathbf{z}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$. The test statistic is the sum of squared singular values of first q lagged sample autocovariance matrices:

$$G_q := \sum_{\tau=1}^{q} tr\left(\hat{S}_\tau \hat{S}_\tau'\right), \tag{16}$$

where $\hat{S}_\tau = \frac{1}{T}\sum_{t=1}^{T} \mathbf{r}_t \mathbf{r}_{t-\tau}'$ with $\mathbf{r}_t = \mathbf{r}_{t+T}$ when $t \leq 0$.

If A is unknown, the sample covariance matrix of $\mathbf{r}_t$ is $\hat{S}_0 = \frac{1}{T}\sum_{t=1}^{T} \mathbf{r}_t \mathbf{r}_t'$. According to (2.8) in Li *et al.* (2019), we reject (15) if

$$\frac{G_q - \frac{N^2 q}{T}\hat{s}_1^2}{\sqrt{\frac{2N^2 q}{T^2}\left(\hat{s}_2 - \frac{N}{T}\hat{s}_1^2\right)^2}} > Z_\alpha,$$

where $\hat{s}_1 = \frac{1}{N}tr(\hat{S}_0)$, $\hat{s}_2 = \frac{1}{N}tr(\hat{S}_0^2)$ and Z_α is the upper-α quantile of standard normal distribution.

Note that $\{\mathbf{r}_t : 1 \leq t \leq T\}$ being white-noise means that the residuals are uncorrelated over t. However, it doesn't indicate that the residuals are uncorrelated over both i and t. The later indicates a stronger adequacy of high-dimensional model. We could assume that $P = I_N$ in the null hypothesis, and by (2.5) in Li *et al.* (2019), we reject $H_0 : \mathbf{r}_t = \mathbf{z}_t$ if

$$\frac{G_q - \frac{N^2 q}{T}}{\sqrt{\frac{2N^2 q}{T^2} + \frac{4N^3 q^2 (\kappa_4 - 3)}{T^3} + \frac{8N^3 q^2}{T^3}}} > Z_\alpha.$$

6. SIMULATION STUDY

6.1. Network Simulation

The symmetric matrix A in model 4 represents an undirected network structure, the pattern of which varies over different application scenarios. In this simulation study, we tend to use four different mechanisms of simulating corresponding network. The network structure in Example 1 adapts to Assumption 3(b), which is required by geometric NED as we have shown in Theorem 2. Simulation mechanisms introduced in Examples 2–4 are for testing the robustness of our estimation, against network structures that may violate Assumption 3.

Example 1. For each node $i \in \{1, 2, \ldots, N\}$, it is connected to node j only if j is inside i's D-neighborhood. That is, in the adjacency matrix, $a_{ij} = 1$ if $0 < |i - j| \leq D$ and $a_{ij} = 0$ otherwise. Figure 1(a) is a visualization of such a network with $N = 100$ and $D = 10$.

Example 2 (Network structure with random distribution). For each node $i \in \{1, 2, \ldots, N\}$, we generate D_i from uniform distribution $U(0, 5)$, and then draw $[D_i]$ samples randomly from $\{1, 2, \ldots, N\}$ to form a set S_i ($[x]$ denotes the integer part of x). $A = (a_{ij})$ could be generated by letting $a_{ij} = 1$ if $j \in S_i$ and $a_{ij} = 0$ otherwise. In a network simulated with such mechanism, as it is indicated in Figure 1(b), there is no significantly influential node (i.e. node with extremely large in-degree).

Example 3 (Network structure with power-law distribution). According to Clauset *et al.* (2009), for each node i in such a network, D_i is generated the same way as in Example 2. Instead of uniformly selecting $[D_i]$ samples from $\{1, 2, \ldots, N\}$, these samples are collected w.r.t. probability $p_i = s_i / \sum_{i=1}^{N} s_i$ where s_i is generated from a discrete power-law distribution $\mathbb{P}\{s_i = x\} \propto x^{-a}$ with scaling parameter $a = 2.5$. As shown in Figure 1(c), a few nodes have much larger in-degrees while most of them have less than 2. Compared to Example 2, network structure with power-law distribution exhibits larger gaps between the influences of different nodes. This type of network
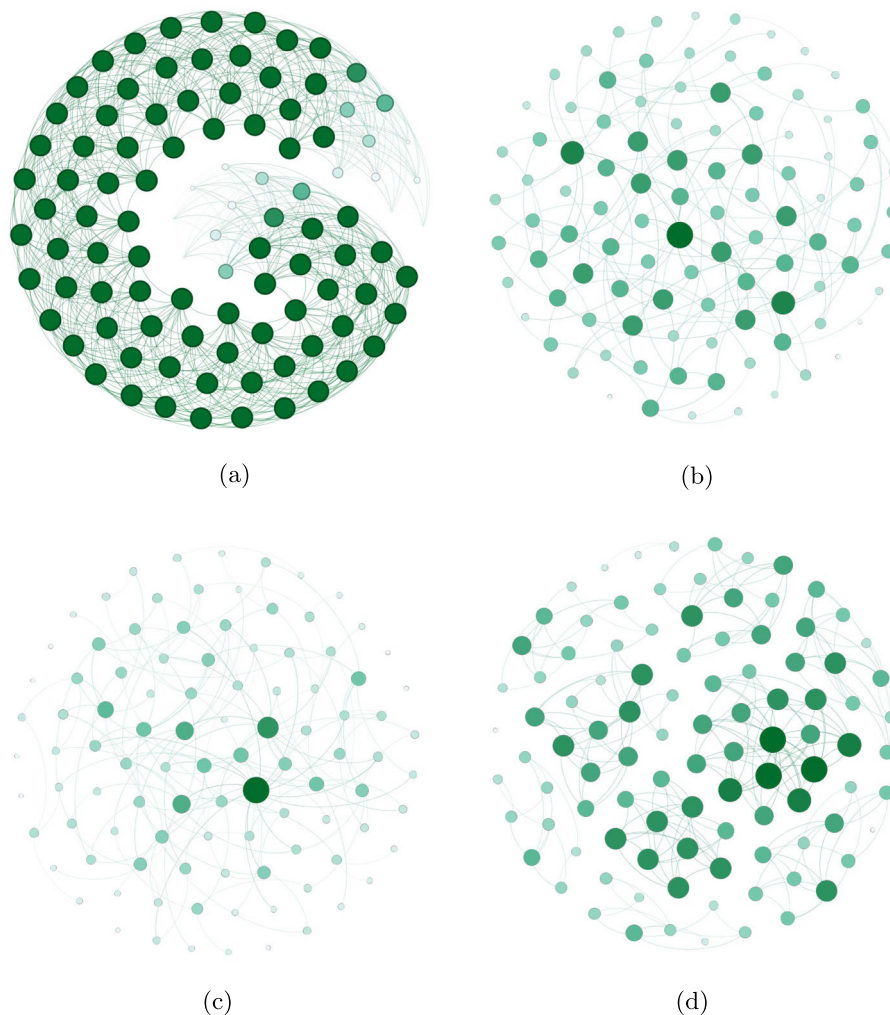

Figure 1. Visualized network structures with $N = 100$. (a) Example 1 ($D = 10$) (b) Example 2; (c) Example 3; (d) Example 4 ($K = 10$)

is suitable for modeling social media such as Twitter and Instagram, where celebrities have huge influence while the ordinary majority has little.

Example 4 (Network structure with stochastic blocks). As it was proposed in Nowicki and Snijders (2001), in a network with stochastic block structure, all nodes are divided into blocks and nodes from the same block are more likely to be connected comparing to those from different blocks. To simulate such structure, these N nodes are randomly divided into K groups by assigning labels $\{1, 2, \ldots, K\}$ to every nodes with equal probability. For any two nodes i and j from the same group, let $\mathbb{P}(a_{ij} = 1) = 0.5$ while for those two from different groups, $\mathbb{P}(a_{ij} = 1) = 0.001/N$. Hence, it is very unlikely for nodes to be connected across groups. Our simulated network successfully mimics this characteristic as Figure 1(d) shows clear boundaries between groups. Block network also has its advantage in practical perspective. For instance, the price of one stock is highly relevant to those in the same industry sector.

In the next section, the simulation study is carried out on datasets that are generated according to the process (3) in conjunction with three types of adjacency matrices in Examples 1–4.

6.2. Simulation Results

Setting the true parameters θ_0 as $(0.1, 0.1, 0.2, 0.2, 0.2)'$, we generate data according to process (3) with different sample sizes T and number of dimensions N. In our setting, T increases from 50 to 4000, while N also increases at relatively slower rates of $\mathcal{O}(\sqrt{T})$ and $\mathcal{O}(T/\log(T))$ respectively, as it is showed in the following table:

T	50	100	200	500	800	1000	1500	2000	2500	3000	4000
$N \approx \sqrt{T}$	7	10	14	22	28	31	38	44	50	54	63
$N \approx T/\log(T)$	12	21	37	80	119	144	205	263	319	374	482

For each combination of (T, N), $M = 1000$ datasets will be simulated independently, according to (3). Based on the mth ($m = 1, 2, \ldots, M$) dataset, the estimation of θ_0 will be carried out and the estimation result is denoted as $\hat{\theta}_m = (\hat{\theta}_{km})' = (\hat{\omega}_m, \hat{\alpha}_m^{(1)}, \hat{\alpha}_m^{(2)}, \hat{\lambda}_m, \hat{\beta}_m)'$. For $k = \{1, 2, 3, 4, 5\}$, the following two measurements are used to evaluate the performance of simulation results:

1. **root-mean-square error**: $\text{RMSE}_k = \sqrt{M^{-1} \sum_{m=1}^{M} (\hat{\theta}_{km} - \theta_{k0})^2}$,
2. **coverage probability**: $\text{CP}_k = M^{-1} \sum_{m=1}^{M} 1_{\{\theta_{k0} \in CI_{km}\}}$.

CI_{km} is the 95% confidence interval defined as

$$\text{CI}_{km} = \left(\hat{\theta}_{km} - z_{0.975} \widehat{\text{SE}}_{km}, \hat{\theta}_{km} + z_{0.975} \widehat{\text{SE}}_{km} \right),$$

where the estimated SE $\widehat{\text{SE}}_{km}$ could be calculated as the square root of kth diagonal element of $(NT)^{-1}(\hat{\kappa}_4 - 1)\widehat{\Sigma}^{-1}$ and $z_{0.975}$ is the 0.975th quantile of standard normal distribution. To eliminate the effect of starting points, a different initial guess of θ is used for each m.

As it is demonstrated in line graphs (c) and (d) in Figures 2–5, the consistency of the estimator is obvious since RMSE drops toward zero when T and N increases. Additionally, $\widehat{\text{SE}}$ is proved to provide reliable estimation of true SE since the coverage probability (CP) converges to its theoretical value of 95% in graphs (a) and (b) in Figures 2–5. In conclusion, the asymptotic properties of our estimator in Theorem 3 are well supported by our simulation results, even for network structures in Examples 2–4 that may violate Assumption 3.
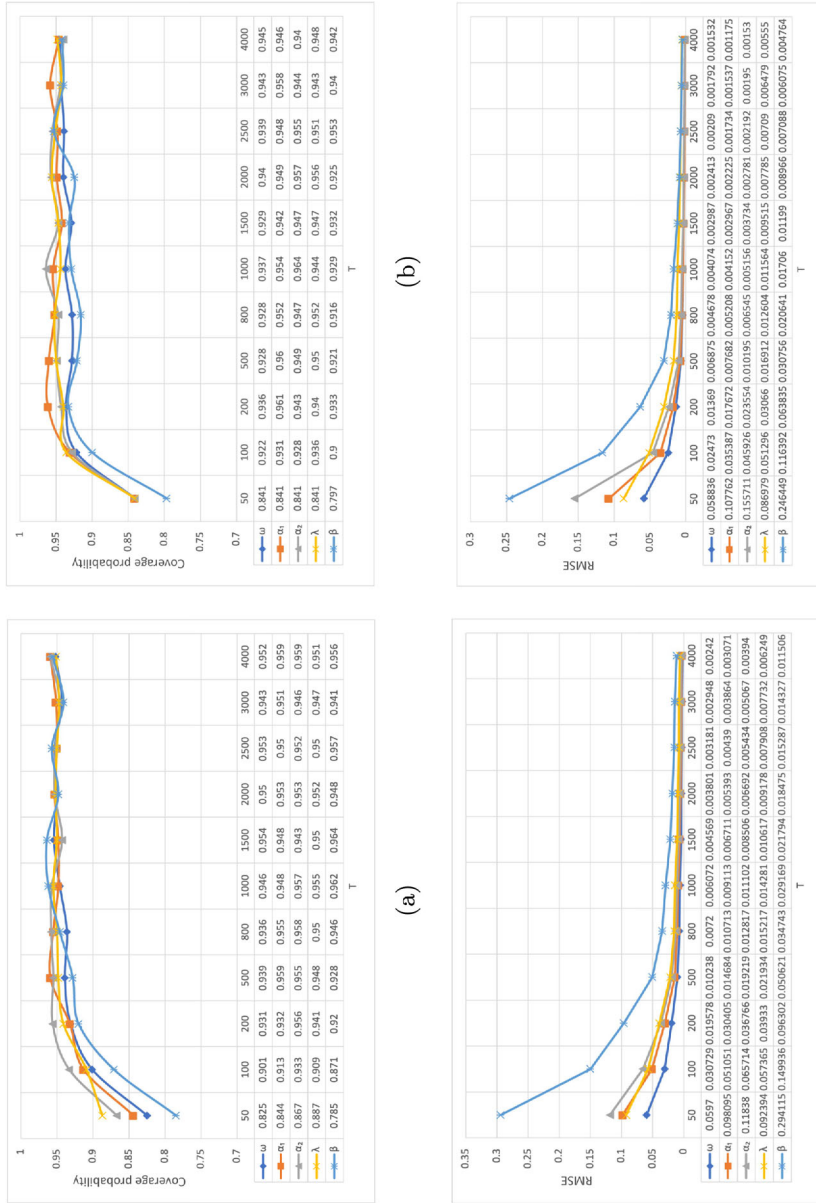
Figure 2. Results of simulation (Example 1). (a) $N \approx \sqrt{T}$; (b) $N \approx T/\log(T)$; (c) $N \approx \sqrt{T}$; (d) $N \approx T/\log(T)$
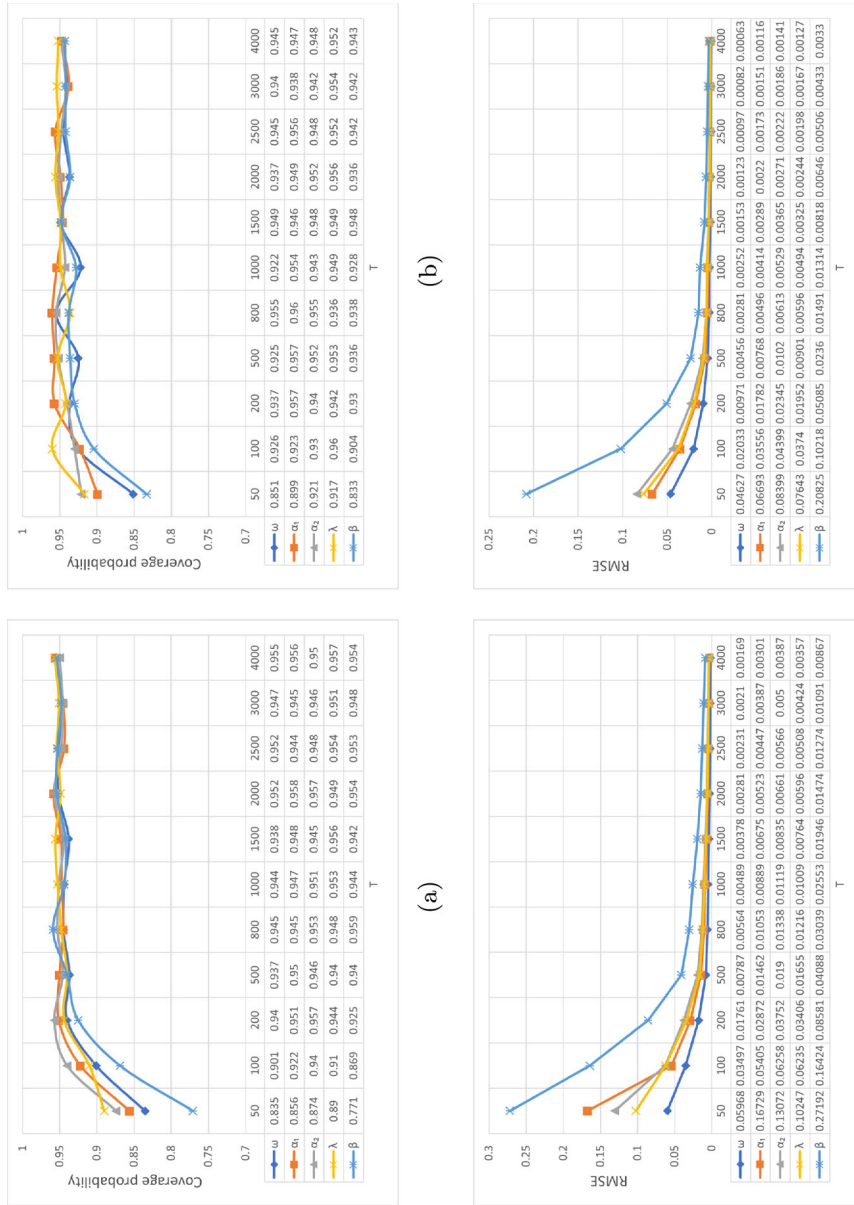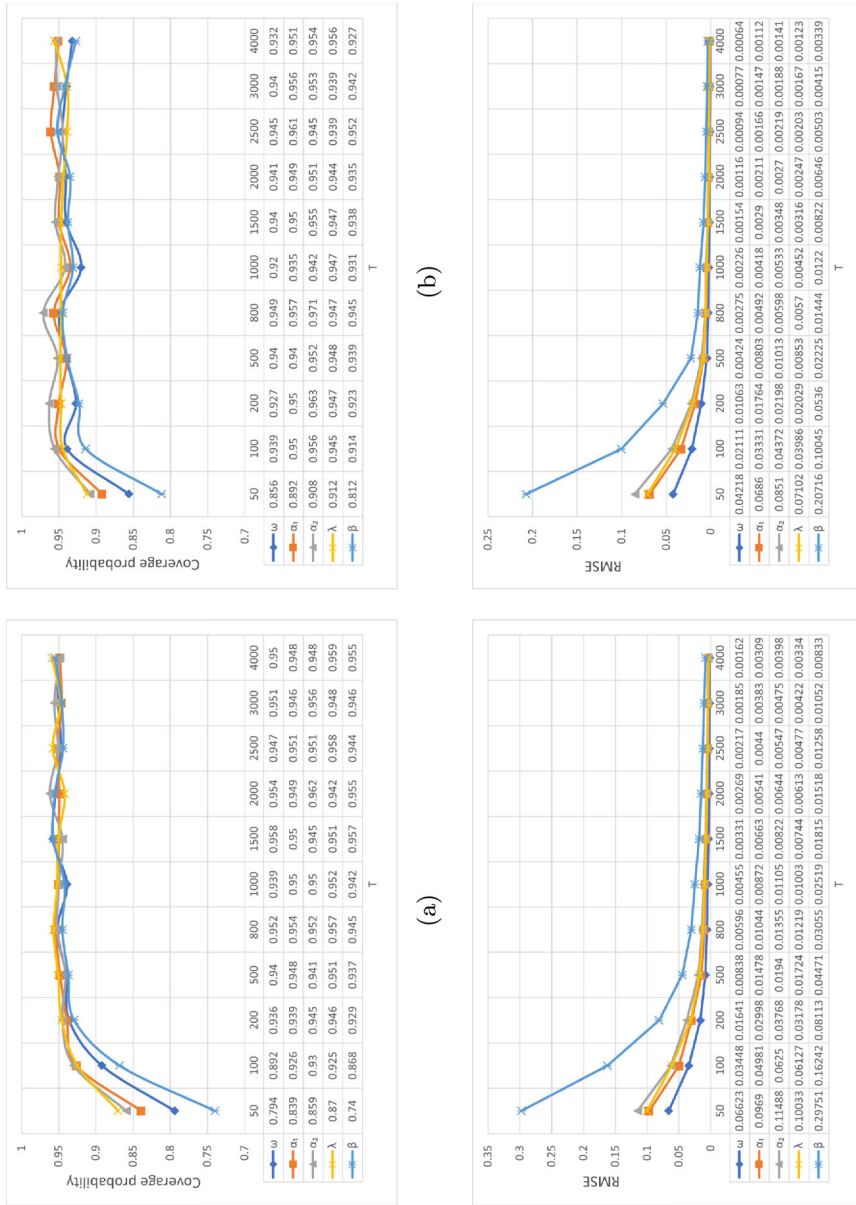
Figure 3. Results of simulation (Example 2). (a) $N \approx \sqrt{T}$; (b) $N \approx T/\log(T)$; (c) $N \approx \sqrt{T}$; (d) $N \approx T/\log(T)$

Figure 4. Results of simulation (Example 3). (a) $N \approx \sqrt{T}$; (b) $N \approx T / \log(T)$; (c) $N \approx \sqrt{T}$; (d) $N \approx T / \log(T)$
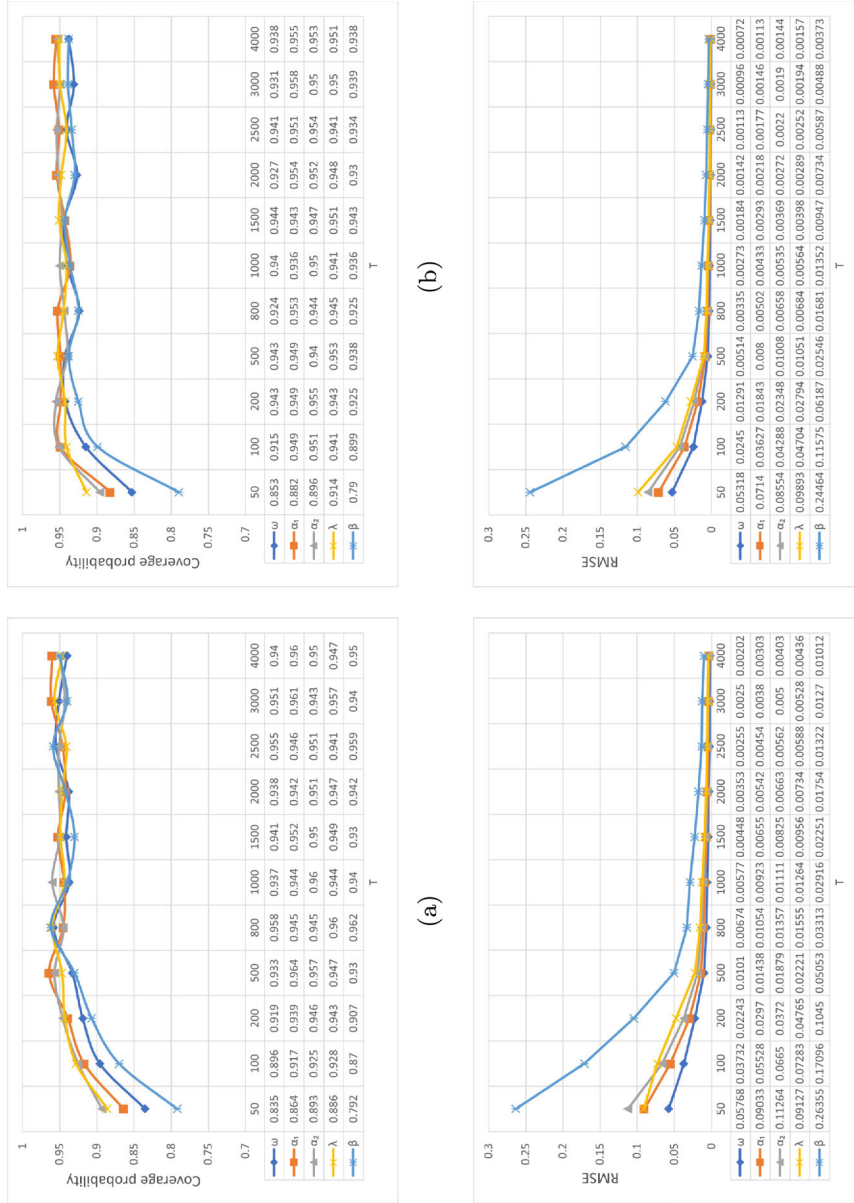
Figure 5. Results of simulation (Example 4). (a) $N \approx \sqrt{T}$; (b) $N \approx T/\log(T)$; (c) $N \approx \sqrt{T}$; (d) $N \approx T/\log(T)$
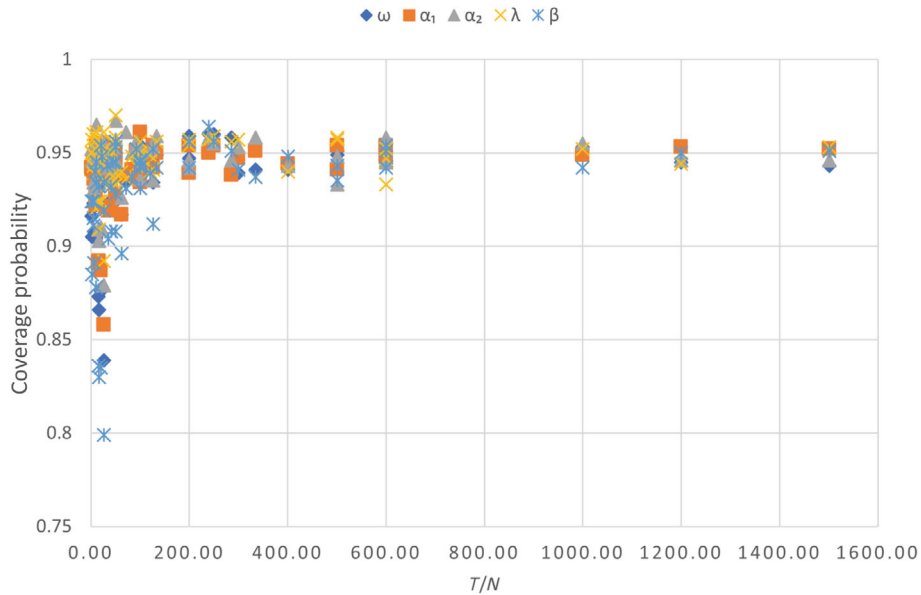
Figure 6. Coverage probabilities varies with the ratio of T and N

Remark. It is worth noticing that the CPs show a lower efficiency of convergence in general when $N = \mathcal{O}(T/\log(T))$, comparing with the case when $N = \mathcal{O}(\sqrt{T})$. Such phenomenon raises an assumption that the performance of the estimator $\widehat{SE}$ is highly related to the ratio of T and N. We repeat the simulation for 61 different combinations of (T, N) and the scatter graph Figure 6 indicates that such assumption could be true, and the convergence of $\widehat{SE}$ only requires $T/N \to \infty$, which includes what we have in Theorem 3, where $T \to \infty$ and $N \to \infty$ at a lower rate.

7. EMPIRICAL DATA ANALYSIS

In addition to simulation studies, we want to test our model using real data from Chinese Shanghai Stock Exchange (SSE) and Shenzhen Stock Exchange (SZSE). The dataset consists of daily log returns of 286 stocks, which are observed in two consecutive years of 2019 and 2020 ($T = 487$ except for closing days). These stocks come from four industry sectors as follows:

- 75 stocks from automotive industry sector;
- 73 stocks from financial industry sector;
- 68 stocks from information industry sector;
- 70 stocks from pharmaceutical industry sector.

And our model is tested within each sector, in which the number of stocks is approximately $T/\log(T) \approx 79$. Hence the estimates and inferences could be trusted according to the simulation study.

As an initial impression of data from each category the time plots of daily average log returns are presented in Figure 7. We also have the shareholder information of each stock, based on which two stocks are considered as connected when they share at least one common shareholder among their top ten shareholders. By this principle, four adjacency matrices are constructed and visualized as Figure 8 for four different industry sectors. Although it is quite intuitive to tell from Figure 8 the sparsity of these four networks, we tend to use the network density (ND) as a quantified measurement, which is defined by the ratio of the number of existing edges to the number of potential connections:

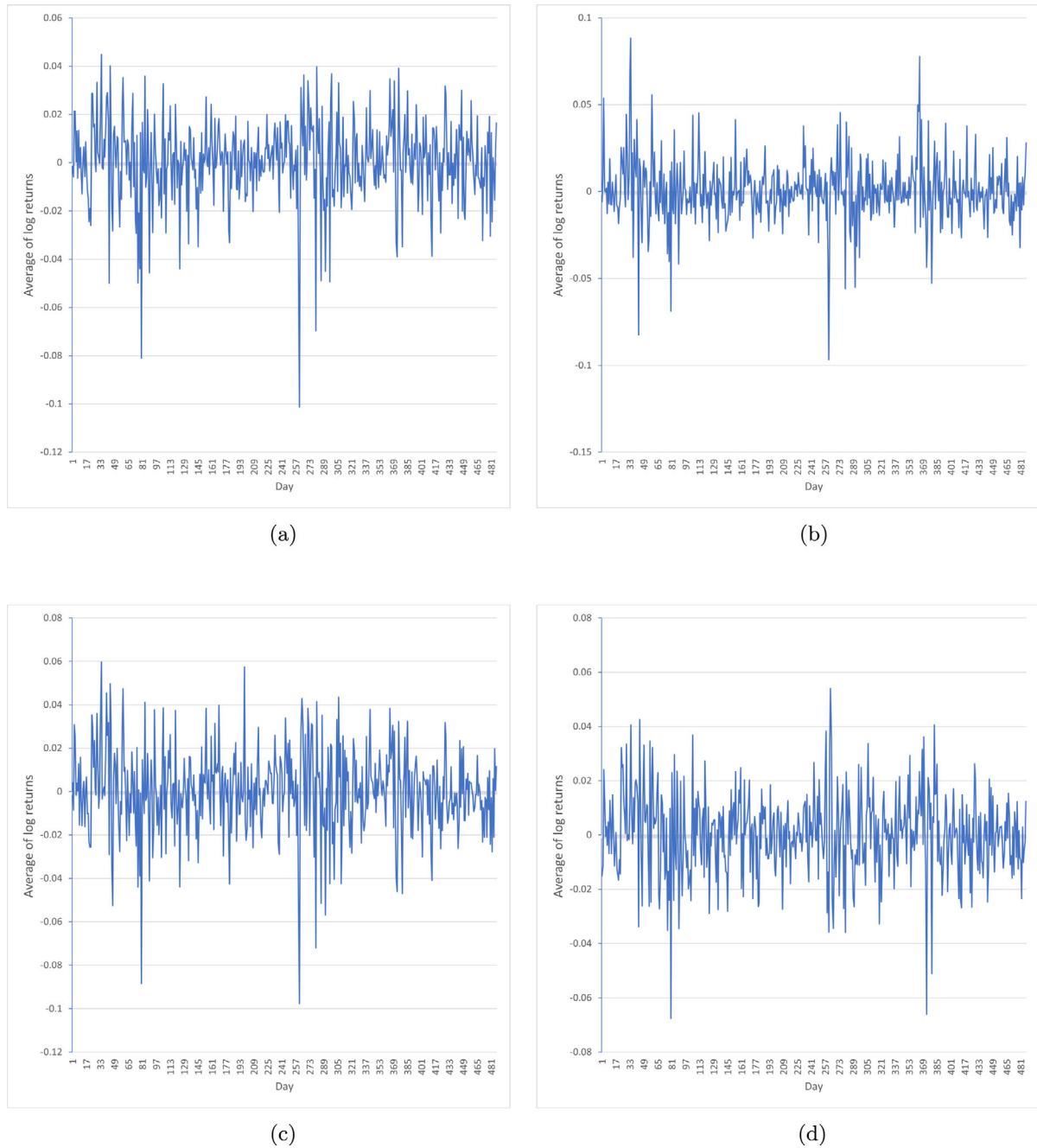$$\text{ND} := 100\% \times \frac{\sum_{i=1}^{N} d_i}{N(N-1)}.$$

Figure 7. Average log returns of stocks from different industry sectors. (a) Automotive industry; (b) Financial industry; (c) Information industry; (d) Pharmaceutical industry
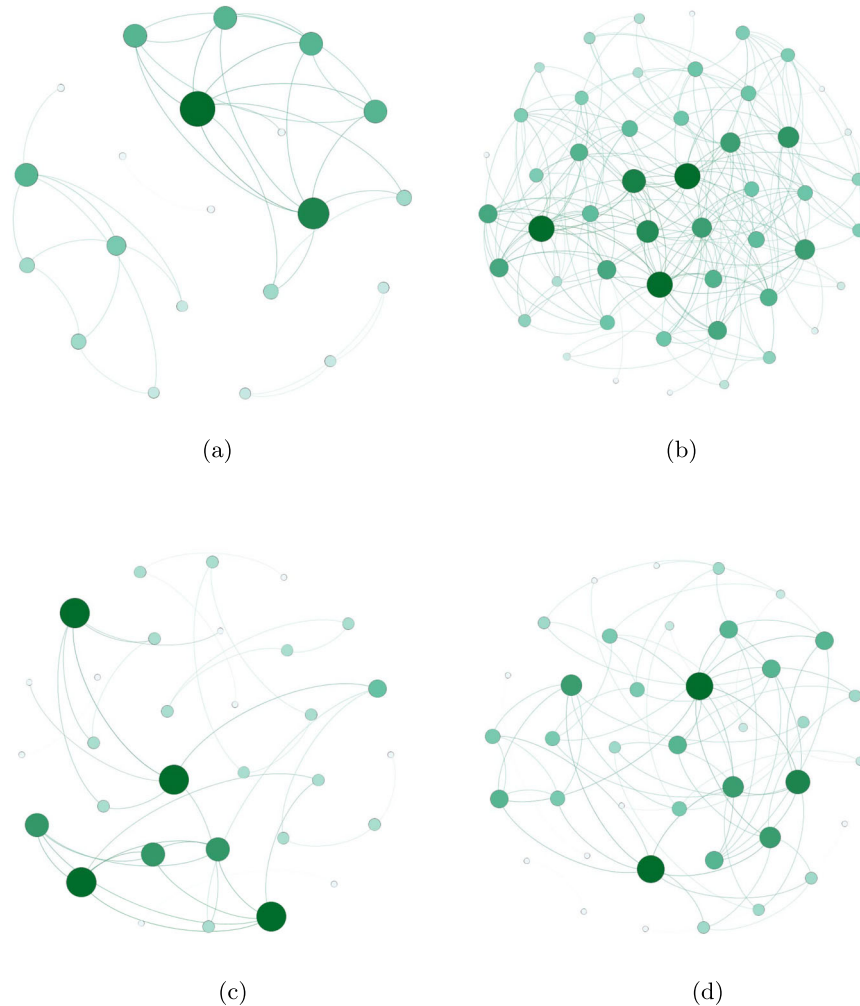
Figure 8. Visualization of networks for stocks from different industry sectors. (a) Automotive industry (ND = 1.26%); (b) Financial industry (ND = 8.11%); (c) Information industry (ND = 1.58%); (d) Pharmaceutical industry (ND = 2.82%)

The results of parameter estimation is summarized in Table I. It is worth noting that the estimated network effect λ for automotive industry sector is not statistically significant while the other estimates for other coefficients or estimates from other sectors are all significant at 5% level. As indicated in Figure 8(a), this could be caused by the sparsity of the network structure as the data from automotive industry has the lowest network density comparing to others. Positive estimates of λ indicate positive correlation between the return of a stock and the returns of its neighbors. Comparing with other parameters, the estimates of β are much larger for all four categories. Strong memory of volatility has been observed in many econometric studies on daily data, and such persistence would be stronger with data sampled at higher frequency according to Nelson (1991).

We now conduct a Wald test on the existence of threshold effect based on the estimated parameters. By letting $\Gamma := (0, 1, -1, 0, 0)$ and $\eta := 0$ in (13), we can make a null hypothesis as follows:

$$H_0 : \alpha_0^{(1)} = \alpha_0^{(2)}.$$

© 2024 The Authors.
Journal of Time Series Analysis published by John Wiley & Sons Ltd.

Table I. Estimation results based on daily log-returns (2019&2020) of stocks from four industries.

Parameter	Estimation	SE	p-Value	Parameter	Estimation	SE	p-Value
	Automotive industry				Financial industry		
ω	0.000099	5.83e−07	<0.05	ω	0.000043	3.12e−06	<0.05
$\alpha^{(1)}$	0.199408	1.08e−02	<0.05	$\alpha^{(1)}$	0.247765	1.41e−02	<0.05
$\alpha^{(2)}$	0.136423	1.01e−02	<0.05	$\alpha^{(2)}$	0.202237	1.47e−02	<0.05
λ	0.004591	4.71e−03	0.16465	λ	0.010469	5.35e−03	<0.05
β	0.727756	1.17e−02	<0.05	β	0.737272	1.09e−02	<0.05
	Information industry				Pharmaceutical industry		
ω	0.000105	6.39e−06	<0.05	ω	0.000063	4.15e−06	<0.05
$\alpha^{(1)}$	0.172737	9.34e−03	<0.05	$\alpha^{(1)}$	0.180950	1.05e−02	<0.05
$\alpha^{(2)}$	0.122312	8.86e−03	<0.05	$\alpha^{(2)}$	0.131722	1.06e−02	<0.05
λ	0.009475	4.03e−03	<0.05	λ	0.012929	4.06e−03	<0.05
β	0.745699	1.11e−02	<0.05	β	0.753305	1.11e−02	<0.05

Table II. p-Values of Wald test on $H_0 : \alpha_0^{(1)} = \alpha_0^{(2)}$

Automotive industry	Financial industry	Information industry	Pharmaceutical industry
1.09e−10	2.16e−07	3.8e−06	3.17e−06

Table III. Results of high-dimensional white noise test on $H_0 : \mathbf{r}_t = P\mathbf{z}_t$ with $q = 3$ and $\alpha = 0.01$

	Automotive industry	Financial industry	Information industry	Pharmaceutical industry
P is unknown	Not rejected	Not rejected	Not rejected	Not rejected
$P = I_N$	Rejected	Rejected	Rejected	Rejected

As it is indicated in Table II, we could reject the null hypothesis with strong confidence and conclude that there exists extremely significant threshold effect within each industry sector.

Using the diagnostic tool introduced in Section 5.2, we could check the model adequacy by inspecting the correlations between residual vectors $\mathbf{r}_t = \left[\frac{y_{1t}}{\bar{\sigma}_{1t}(\hat{\theta}_{NT})}, \ldots, \frac{y_{Nt}}{\bar{\sigma}_{Nt}(\hat{\theta}_{NT})} \right]'$. We will test null hypothesis $H_0 : \mathbf{r}_t = P\mathbf{z}_t$ with P being unknown and $P = I_N$ respectively, the results are summarized in Table III. In all sectors, we cannot reject the hypothesis that the residual vectors are high-dimensional white noises with $\mathbb{E}\mathbf{r}_t = 0$ and $\mathrm{Var}(\mathbf{r}_t) = PP'$ over t. However, the stronger hypothesis $H_0 : \mathbf{r}_t = \mathbf{z}_t$ is rejected, as there exist correlations between residuals $\left\{ \frac{y_{it}}{\bar{\sigma}_{it}(\hat{\theta}_{NT})} \right\}$ with different i. We might be able to eliminate such deficiency in the adequacy of our model by heterogeneous parameterization with coefficients as ω_i. $\alpha_i^{(1)}$, $\alpha_i^{(2)}$, λ_i and β_i, or by considering a dynamic network structure. However, the purpose of the introduction of network structure is to reduce the number of parameters of high-dimensional time series. Besides, deriving limit theorems for models with heterogeneous parameters or dynamic network could be theoretically challenging.

On the other hand, our results on asymmetric effect of positive and negative news are quite different compared to what was derived from univariate data in the literature. For instance, in a study by Engle and Ng (1993) on daily returns of Japanese stock index TOPIX, it was found that negative news would have larger impact on future volatility. Such phenomenon is reasonable in stock market since investors would lose confidence to a certain asset when it performs badly, hence they would adjust their portfolio and add more uncertainty to the future. However, it is not necessarily the case if we take into consideration the whole picture instead of looking at one individual and ignoring possible impact of its neighbors in the same system. In our estimation results, $\alpha^{(1)}$ are uniformly larger than $\alpha^{(2)}$, indicating a larger impact of good news on volatility. A more precise conclusion would be that the volatility of one individual is more sensitive to its own good news, which actually does not contradict the

conclusion of Engle and Ng ([1993](#)), since in the univariate case, how much proportion of the 'bad news' effect is actually contributed by bad performance in systematic perspective remains unknown. Our results show that good news has larger 'local influence' as it is indicated by $\alpha^{(1)}$, while there is a possibility that bad news, despite of having less 'local influence', spreads faster and has larger 'global influence' on the neighbors through network connection. Such potential leads to a future extension of our model that the threshold effect could be further applied on the coefficient λ, allowing good news and bad news to have asymmetric network effect.

8. CONCLUSION

In this article, we propose a TNGARCH model by taking consideration of network effect, as well as the threshold effect of shocks on volatilities. Our model can be applied to describe asymmetric properties for volatilities of high dimensional time series without increasing the parameterization complexity. Strict stationarity when N is fixed, as well as near-epoch dependence when $N \to \infty$ are discussed. Then the parameters are estimated by quasi-maximum likelihood estimation, the consistency and the asymptotic normality of the proposed estimator are proved as well. The results of simulation study support the theoretical properties of QMLE. At last, our model is fitted to real stock data containing 710 stocks of four industries from SSE and SZSE. Empirical results reveal that although volatility is more sensitive to bad news in univariate case, with network structure being considered, there is a possibility that majority of the revision of individual volatility is due to the impact of bad news of its neighbors, hence the 'local influence' of bad news is not necessarily larger than that of good news in such case.

There is room for extension of our methodology, which could lead to interesting topics for future research. In Theorem [3](#) we have derived asymptotic properties when $T \to \infty$ and $N \to \infty$ at a lower rate, our estimation method enables us to make reliable inference on parameters even when the data has hundreds of dimensions according to the simulation study. The limitation is also obvious, for example, as shown in Figures [2–5](#), to get a decent approximation of standard errors, we need to collect 4000 samples even though the number of dimensions is about 482, which could be even higher in real-world situations: user data collected from a social network often consists of millions of individual accounts whereas it may be impossible to collect sufficient number of samples over time even at daily frequency. Therefore our model would be applicable in a much larger scale if its statistical properties could be derived when N increases at the same rate, or even higher rate compared to T. As far as we know, there is no published work in the literature to solve this problem theoretically for GARCH-type models. Another limitation of our model is that the way we consider the network effect is simplified in two aspects: The network structure is deterministic rather than stochastic over time, embedding a random network in our model would make more sense, it would nevertheless raise the complexity of the model, and may cause problems in the estimation of parameters (see Chandrasekhar and Lewis, [2011](#)); Moreover, there is only one type of individual-to-individual relation considered since the network structure is constructed solely based on common shareholders. Zhu *et al.* ([2023](#)) constructed their factor-augmented network using several types of relations, including individual-to-individual relations and factor-to-individual relations. Bringing more information into consideration would possibly improve the adequacy of the model and we will leave it for future research.

DATA AVAILABILITY STATEMENT

Our dataset consists of daily log returns of 286 stocks, which are observed in two consecutive years of 2019 and 2020 (except for holidays) from Chinese Shanghai Stock Exchange (SSE) and Shenzhen Stock Exchange (SZSE). The data are available at request. The data that support the findings of this study are available at request.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

REFERENCES

Bauwens L, Laurent S, Rombouts JV. 2006. Multivariate GARCH models: a survey. *Journal of Applied Econometrics* **21**(1):79–109.

Bollerslev T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**(3):307–327.

Bollerslev T. 1990. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The Review of Economics and Statistics* **72**(3):498–505.

Bollerslev T, Engle RF. 1993. Common persistence in conditional variances. *Econometrica* **61**(1):167–186.

Bollerslev T, Engle RF, Wooldridge JM. 1988. A capital asset pricing model with time-varying covariances. *Journal of Political Economy* **96**(1):116–131.

Bougerol P, Picard N. 1992. Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* **52**(1):115–127.

Chandrasekhar A, Lewis R. 2011. Econometrics of sampled networks. Unpublished Manuscript, MIT.[422].

Clauset A, Shalizi CR, Newman MEJ. 2009. Power-law distributions in empirical data. *SIAM Review* **51**(4):661–703.

Engle RF. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**(4):987–1007.

Engle R. 2002. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* **20**(3):339–350.

Engle RF, Kroner KF. 1995. Multivariate simultaneous generalized ARCH. *Econometric Theory* **11**(1):122–150.

Engle RF, Ng VK. 1993. Measuring and testing the impact of news on volatility. *The Journal of Finance* **48**(5):1749–1778.

Engle RF, Ng VK, Rothschild M. 1990. Asset pricing with a factor-ARCH covariance structure: empirical estimates for treasury bills. *Journal of Econometrics* **45**(1):213–237.

Glosten LR, Jagannathan R, Runkle DE. 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance* **48**(5):1779–1801.

Goel S, Goldstein DG. 2014. Predicting individual behavior with social networks. *Marketing Science* **33**(1):82–93.

Hu Y-P, Tsay RS. 2014. Principal volatility component analysis. *Journal of Business & Economic Statistics* **32**(2):153–164.

Jenish N, Prucha IR. 2012. On spatial processes and asymptotic inference under near-epoch dependence. *Journal of Econometrics* **170**(1):178–190.

Karolyi GA. 1995. A multivariate GARCH model of international transmissions of stock returns and volatility: the case of the United States and Canada. *Journal of Business & Economic Statistics* **13**(1):11–25.

Kingman JFC. 1973. Subadditive ergodic theory. *Annals of Probability* **1**(6):883–899.

Li W, Gao J, Li K, Yao Q. 2016. Modeling multivariate volatilities via latent common factors. *Journal of Business & Economic Statistics* **34**(4):564–573.

Li Z, Lam C, Yao J, Yao Q. 2019. On testing for high-dimensional white noise. *The Annals of Statistics* **47**(6):3382–3412.

Nelson DB. 1991. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* **59**(2):347–370.

Nitzan I, Libai B. 2011. Social effects on customer retention. *Journal of Marketing* **75**(6):24–38.

Nowicki K, Snijders TAB. 2001. Estimation and prediction for stochastic block structures. *Journal of the American Statistical Association* **96**(455):1077–1087.

Pan J, Polonik W, Yao Q. 2010. *Estimating factor models for multivariate volatilities: an innovation expansion method*. In *Proceedings of COMPSTAT'2010*, Physica-Verlag, Heidelberg; 305–314.

Teräsvirta T. 2009. *An introduction to univariate GARCH models*. In *Handbook of Financial Time Series*, Springer, Berlin; 17–42.

Tse Y, Tsui AK. 2001. A multivariate GARCH model with time-varying correlations. *Mathematics Preprint Archive* **2001**(6):1184–1217.

Xu X, Lee L. 2015. Maximum likelihood estimation of a spatial autoregressive tobit model. *Journal of Econometrics* **188**(1):264–280.

Xu X, Wang W, Shin Y, Zheng C. 2022. Dynamic network quantile regression model. *Journal of Business & Economic Statistics* **42**(2):407–421.

Zakoïan J-M. 1994. Threshold heteroskedastic models. *Journal of Economic Dynamics and Control* **18**(5):931–955.

Zhou J, Li D, Pan R, Wang H. 2020. Network GARCH model. *Statistica Sinica* **30**:1–18.

Zhu X, Pan R, Li G, Liu Y, Wang H. 2017. Network vector autoregression. *Annals of Statistics* **45**(3):1096–1123.

Zhu Z, Zhang N, Zhu K. 2023. Big portfolio selection by graph-based conditional moments method. *ArXiv 2301:11697*.