

Limit Theorems for Network Data without Metric Structure*

Wen Jiang[†] Yachen Wang[‡] Zeqi Wu[§] Xingbai Xu[¶]

Abstract

This paper develops limit theorems for random variables with network dependence, without requiring that individuals in the network to be located in a Euclidean or metric space. This distinguishes our approach from most existing limit theorems in network econometrics, which are based on weak dependence concepts such as strong mixing, near-epoch dependence, and ψ -dependence. By relaxing the assumption of an underlying metric space, our theorems can be applied to a broader range of network data, including financial and social networks. To derive the limit theorems, we generalize the concept of functional dependence (also known as physical dependence) from time series to random variables with network dependence. Using this framework, we establish several inequalities, a law of large numbers, and central limit theorems. Furthermore, we verify the conditions for these limit theorems based on primitive assumptions for spatial autoregressive models, which are widely used in network data analysis.

Keywords: functional dependence, network data, law of large numbers, central limit theorem, concentration inequality, spatial autoregressive model

*The authors are listed in alphabetical order, and contributed equally. Zeqi Wu and Xingbai Xu are corresponding authors.

[†]Jiang: Newhuadu Business School, Minjiang University. Email address: jiangwen@mju.edu.cn.

[‡]Wang: Paula and Gregory Chow Institute for Studies in Economics, Xiamen University. Email address: wangyachen@stu.xmu.edu.cn.

[§]Wu: Institute of Statistics and Big Data, Renmin University of China. Email address: wuzeqi@ruc.edu.cn.

[¶]Xu: MOE Key Lab of Econometrics and Fujian Key Lab of Statistics, Wang Yanan Institute for Studies in Economics (WISE), Department of Statistics and Data Science, School of Economics, Xiamen University. Email address: xuxingbai@xmu.edu.cn.

1. Introduction

In recent years, network data are becoming more and more popular and playing essential roles in various fields in economics and finance. For example, social network effects are essential for understanding behaviors and academic performance among students (Angrist and Lang, 2004, Graham, 2008, etc); financial risk spreads through networks of different countries or assets (Acemoglu et al., 2015, Zhu et al., 2019, etc); and spatial and network effects are critical in the spread of diseases like COVID-19 (Kraemer et al., 2020, Jia et al., 2020, Han et al., 2021, etc). As a result, network econometrics has become increasingly important in both empirical and theoretical research.

When studying the theoretical properties of estimators or test statistics for network econometric models, inequalities, laws of large numbers (LLNs), and central limit theorems (CLTs) for network data are indispensable. To derive such inequalities and limiting laws, weak dependence concepts are often necessary. In the next paragraph, we will review the related literature.

Martingale difference arrays (MDA) play a fundamental role in linear spatial autoregressive (SAR) models (Kelejian and Prucha, 2001, Lee, 2004, 2007, Lee et al., 2022, etc). Jenish and Prucha (2009) study some limit theorems for strong mixing and ϕ -mixing spatial random fields. Jenish and Prucha (2012) generalize the concept of near-epoch dependence (NED) from time series to spatial econometric settings. Wu et al. (2024) generalize the concept of functional dependence measure (FDM) from time series (Wu, 2005) and stationary random fields (El Machkouri et al., 2013) to nonstationary spatial processes, and establish some inequalities and limit theorems. Kuersteiner and Prucha (2013) employ a MDA CLT to derive a CLT for panel data models with cross-sectional dependence. In follow-up work, Kuersteiner and Prucha (2020) establish a CLT for dynamic panel data models with endogenous network formation based on a MDA structure of the errors. Kuersteiner (2019) proposes the concept of spatial mixingale for data with network structure and establishes some related limiting laws based on a model-dependent random metric. Leung (2019) defines a type of weak dependence for a sparse pairwise stable network generated by a strategic network formation model with homophilous agents, and obtains a weak LLN via the random-graph theory; Leung and Moon (2025) establish a CLT under similar settings. Kojevnikov et al. (2021) generalize the ψ -dependence proposed by

[Doukhan and Louhichi \(1999\)](#) to sparse networks and establish some related inequalities and limiting laws. These weak dependence concepts and their associated limit theorems have significantly advanced network econometrics. Further applications of these concepts can be found in papers such as [Xu and Lee \(2015\)](#), [Gao and Ding \(2025\)](#), [Kojevnikov \(2021\)](#), [Leung \(2022\)](#), [Xu and Lee \(2018\)](#), [Qu and Lee \(2015\)](#), [Liu et al. \(2025\)](#), among others.

To establish asymptotic properties for nonlinear network models or nonlinear estimators, researchers typically rely on some of the weak dependence concepts mentioned above, such as spatial NED and ψ -dependence. However, most of these weak dependence concepts assume that the individuals in the network are located in a metric space. For example, the spatial mixing, spatial NED and spatial FDM ([Jenish and Prucha, 2009, 2012](#), [Wu et al., 2024](#)) assume that individuals are located in a Euclidean space. The ψ -dependence ([Kojevnikov et al., 2021](#)) relies on the geodesic distance in sparse networks.

Relaxing the assumption that individuals are located in a Euclidean or metric space can be highly beneficial. Due to the limited theoretical tools available, empirical research often has to rely on assumptions that may not always be appropriate. For instance, [Xu et al. \(2022\)](#) use the spatial NED theory to examine quantile regression in a dynamic network model, applying it to study stocks traded in the NYSE and NASDAQ. However, it remains questionable whether stocks can be appropriately considered as points in a Euclidean space. Analogously, [Hoff et al. \(2002\)](#) assume that individuals in a social network are located in a Euclidean space. However, whether this assumption is reasonable is still debatable.¹ With the development of the internet and social media, individuals who are far apart in Euclidean distance may still have strong interactions, which challenges the weak dependence concept commonly employed in the literature. In theoretical research, traditional estimation methods, such as quasi-maximum likelihood estimation and the generalized method of moments, are applicable to both spatial and network data without requiring a metric structure. However, the asymptotic theories for robust estimation in the SAR model ([Liu et al., 2025](#)) rely on the NED theory from [Jenish and Prucha \(2012\)](#). Consequently, their methods are applicable only to spatial individuals in a Euclidean space. In conclusion, relaxing the metric space assumption would be advantageous. [Kuersteiner \(2019\)](#)

¹The work by [Hoff et al. \(2002\)](#) is further generalized to the studies of latent space models ([Athreya et al., 2018](#), [Smith et al., 2019](#)).

attempts to address this by introducing a model-dependent random metric, moving away from a fixed metric, but still maintains some form of underlying metric structure. Our work takes a further step by completely eliminating the need for any metric structure, offering a more flexible approach to studying network econometric models.

In this paper, we generalize the concept of functional dependence measure (FDM) in [Wu \(2005\)](#), [El Machkouri et al. \(2013\)](#), and [Wu et al. \(2024\)](#) to random variables with network structure. Unlike strong mixing, the FDM is usually much easier to calculate for a given network econometric model as it does not involve the calculation of the supremum over two sub- σ -fields. Our generalization of FDM is applicable to network data, regardless of whether a metric space exists for the individuals in the network. We obtain some moment inequalities, a concentration inequality, a weak LLN, and some CLTs for functionally dependent network data. We examine several examples carefully and provide some primitive conditions for the CLTs to hold.

We summarize the main contributions of this paper as follows. First, we do not impose any metric space assumption on the individuals in the network, which broadens the applicability of our limit theorems compared to existing weak dependence concepts in the literature. Second, we only require the average influence of each individual to be finite (see Remark 2.3), yet we allow for the possibility that individuals can have significant influence on others who are far apart (if there exists some metric structure in the network). This addresses a gap in existing weak dependence concepts, which typically assume that interactions between individuals diminish with distance. Third, our weak dependence concept applies to networks with small or moderate diameters (in terms of geodesic distance), a common characteristic of social and financial networks. In contrast, [Kojevnikov et al. \(2021\)](#) implicitly assume that the diameter of the network grows to infinity (see Remark 4.2).

The structure of this paper is as follows. In Section 2, we introduce the definition of FDM. In Section 3, we calculate FDM for some common examples in econometrics. In Section 4, we establish theoretical properties of FDM, including moment inequalities, a concentration inequality, a LLN and some CLTs. In Section 5, we study the FDM under various transformations commonly encountered in econometrics. Section 6 concludes this paper. The proofs for important theorems and lemmas are collected in Appendix B. The rest proofs are collected in the Supplement Appendix.

Notation. The set of positive integers is $\mathbb{N} \equiv \{1, 2, \dots\}$, \mathbb{Z} denotes the set of integers, and \mathbb{Z}^d is the d -dimensional integer lattice. For any column vector

$x = (x_1, x_2, \dots, x_d)' \in \mathbb{R}^d$, where \mathbb{R}^d is the d -dimensional Euclidean space, $\|x\| = (x'x)^{1/2}$ denotes its Euclidean norm, $\|x\|_\infty = \max_{1 \leq k \leq d} |x_k|$ represents its infinity norm, and $\|x\|_1 = \sum_{k=1}^d |x_k|$ is its 1-norm. For any matrix $A = (a_{ij})_{n \times m}$, its Frobenius norm is $\|A\| \equiv \sqrt{\sum_{ij} a_{ij}^2}$, its 1-norm (also called column sum norm) is defined as $\|A\|_1 = \max_{1 \leq j \leq m} \sum_{i=1}^n |a_{ij}|$, and its infinity norm (also called row sum norm) is defined as $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^m |a_{ij}|$. For any symmetric matrix A , $\min \text{eig}(A)$ denotes its minimum eigenvalue. For any two sequences a_n and b_n , $a_n \sim b_n$ if and only if $a_n/b_n \rightarrow 1$ as $n \rightarrow \infty$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Denote probability and expectation by \mathbb{P} and \mathbb{E} , respectively. For any random vector (matrix) X , its L^p norm is defined as $\|X\|_{L^p} \equiv [\mathbb{E}(\|X\|^p)]^{1/p}$ for any constant $p \geq 1$. The symbols “ $\xrightarrow{\mathbb{P}}$ ” and “ \xrightarrow{d} ” denote convergence in probability and convergence in distribution, respectively. For two positive non-random sequences a_n, b_n and random vector sequence X_n , $X_n = o_{\mathbb{P}}(a_n)$ means $\mathbb{P}(\|X_n\| > \epsilon a_n) \rightarrow 0$ as $n \rightarrow \infty$ for any $\epsilon > 0$ and $X_n = O_{\mathbb{P}}(a_n)$ means for any $\epsilon > 0$, there exists a constant $M > 0$ such that $\limsup_{n \rightarrow \infty} \mathbb{P}(\|X_n\| \geq Ma_n) < \epsilon$. For any sub- σ -field \mathcal{C} of \mathcal{F} , denote the conditional probability, conditional expectation, and conditional variance by $\mathbb{P}_{\mathcal{C}}(\cdot) \equiv \mathbb{P}(\cdot | \mathcal{C})$, $\mathbb{E}_{\mathcal{C}}(\cdot) \equiv \mathbb{E}(\cdot | \mathcal{C})$, and $\text{Var}_{\mathcal{C}}(\cdot) \equiv \text{Var}(\cdot | \mathcal{C})$, respectively. Besides, for a sub- σ -field \mathcal{G} of \mathcal{F} , we denote $\mathbb{E}_{\mathcal{C}}(\cdot | \mathcal{G}) \equiv \mathbb{E}(\cdot | \mathcal{G} \vee \mathcal{C})$, where $\mathcal{G} \vee \mathcal{C}$ denotes the σ -field generated by \mathcal{G} and \mathcal{C} . For a random vector (matrix) X , let $\|X\|_{L^p, \mathcal{C}} \equiv [\mathbb{E}_{\mathcal{C}}(\|X\|^p)]^{1/p}$.

2. Definition of Functional Dependence Measure

Consider a network with n nodes (also called individuals or units). For simplicity, name these n nodes as $1, 2, \dots, n$, and denote $[n] = \{1, 2, \dots, n\}$. Although we name the nodes as $1, 2, \dots, n$, we in fact consider a general setting where the order of the individuals can be arbitrary, i.e., individuals 1 and 2 might have no correlation, but individuals 1 and n can have strong correlation.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying probability space, and \mathcal{C}_n be a sub- σ -field of \mathcal{F} . Associated with each $i \in [n]$ is a random vector $e_{i,n} \in \mathbb{R}^{p_e}$, where $p_e \in \mathbb{N}$ is a constant. Suppose that $e_{i,n}$'s are conditionally independent given \mathcal{C}_n , but not necessary to be identically distributed. Denote $e_{[n]} = (e_{1,n}, e_{2,n}, \dots, e_{n,n})$. Suppose that $Y_{1,n}, \dots, Y_{n,n}$ are n random vectors generated by $e_{[n]}$:

$$Y_{j,n} = F_{j,n}(e_{[n]}). \quad (2.1)$$

The $\{Y_{j,n} : 1 \leq j \leq n, n \geq 1\}$ can be a random field or random variables with network structures. Although $Y_{j,n}$ can be vector-valued, for simplicity of presentation and without loss of generality, we will only discuss the real-valued case. See Remark 2.2 for details.

Suppose that conditional on \mathcal{C}_n , $e_{i,n}^*$ is an independent and identically distributed (i.i.d.) copy of $e_{i,n}$, and $e_{i,n}^*$ is independent to $e_{j,n}$ for all $j \neq i$. Denote $Y_{j,n,i}$ as the coupled version of $Y_{j,n}$ with $e_{i,n}$ replaced by its i.i.d. copy $e_{i,n}^*$, i.e., $Y_{j,n,i} \equiv F_{j,n}(e_{1,n}, \dots, e_{i-1,n}, e_{i,n}^*, e_{i+1,n}, \dots, e_{n,n})$. Now, we are ready to introduce the definition of functional dependence measure.

Definition 2.1 (Functional dependence measure). For $p \geq 1$, define the functional dependence measure (FDM), also called the physical dependence measure, as

$$\delta_{p,n}(j, i, \mathcal{C}_n) \equiv \|Y_{j,n} - Y_{j,n,i}\|_{L^p, \mathcal{C}_n}. \quad (2.2)$$

When \mathcal{C}_n is the trivial σ -field $\{\Omega, \emptyset\}$, we simplify the notation as $\delta_{p,n}(j, i) \equiv \delta_{p,n}(j, i, \mathcal{C}_n)$.

Remark 2.1. The $\delta_{p,n}(j, i, \mathcal{C}_n)$ in Eq.(2.2) measures the influence of $e_{i,n}$ on $Y_{j,n}$: conditional on \mathcal{C}_n , if $e_{i,n}$ is replaced by its i.i.d. version $e_{i,n}^*$, the magnitude of the change of $Y_{j,n}$ is $\delta_{p,n}(j, i, \mathcal{C}_n)$ under the norm $\|\cdot\|_{L^p, \mathcal{C}_n}$. We note that this definition of FDM may appear to be a metric. In essence, however, it is not, as it violates the triangle inequality. To illustrate, consider two nodes i and j with weak dependence. Under a metric-based concept of dependence, no node k could be strongly correlated with both. However, this scenario is indeed possible within the FDM framework.

Remark 2.2. When $Y_{j,n} = (Y_{j,n}^{(1)}, \dots, Y_{j,n}^{(p_Y)})'$ is a p_Y -dimensional random vector, since $\|y\| \leq \|y\|_1$, for all $i = 1, \dots, n$,

$$\|Y_{j,n} - Y_{j,n,i}\|_{L^p, \mathcal{C}_n} \leq \left\| \sum_{k=1}^{p_Y} |Y_{j,n}^{(k)} - Y_{j,n,i}^{(k)}| \right\|_{L^p, \mathcal{C}_n} \leq \sum_{k=1}^{p_Y} \|Y_{j,n}^{(k)} - Y_{j,n,i}^{(k)}\|_{L^p, \mathcal{C}_n},$$

where $Y_{j,n,i} = (Y_{j,n,i}^{(1)}, \dots, Y_{j,n,i}^{(p_Y)})'$ is the coupled version of $Y_{j,n}$ with $e_{i,n}$ replaced by its i.i.d. copy $e_{i,n}^*$. Thus, without loss of generality, it suffices to study the FDM of random variables.

Next, we compare our FDM to the weak dependence concepts in the literature.

(1) The index sets considered in Wu (2005), Wu et al. (2024), El Machkouri et al. (2013), Jenish and Prucha (2009, 2012), Kojevnikov et al. (2021) are either a subset

of \mathbb{Z}^d , \mathbb{R}^d or a general metric space. In contrast, our definition does not require the index i to belong to any metric space. This makes our approach applicable to network data, where individuals are not necessarily located in an underlying metric space.

(2) Compared to the strong mixing, the FDM is easier to calculate because the construction of the coupled version $Y_{j,n,i}$ is explicit, and the L^p -norm is straightforward to calculate. In contrast, the strong mixing coefficient requires the calculation of a supremum over two σ -fields, which can be quite challenging, especially for variables generated by spatial or network models. For more discussion on this point, see [Wu \(2005\)](#) and [Xu and Lee \(2024\)](#).

(3) Compared to the spatial NED, FDM is more conveniently to calculate under any L^p -norm, while the L^p -NED property is often tractable only when $p = 2$, as it often relies on the fact that the conditional expectation is the best predictor under L^2 -distance.

(4) The FDM is easy to compute under a wide range of transformations, as discussed in Section 5. However, difficulties (e.g., undesirable moment conditions may be needed) have been found in preserving the L^p -NED property under transformations ([Davidson, 2021](#), p.388).

Definition 2.2. For any constants $p \geq 1$ and $q \geq 1$, $\{Y_{j,n}\}$ is said to be (L^p, q) -functionally dependent on $\{e_{i,n}\}$ given \mathcal{C}_n if

$$\Delta_{p,q}(\mathcal{C}_n) \equiv \frac{1}{n^q} \sum_{i=1}^n \left[\sum_{j=1}^n \delta_{p,n}(j, i, \mathcal{C}_n) \right]^q = o(1) \quad \text{a.s.}$$

as $n \rightarrow \infty$. When \mathcal{C}_n is the trivial σ -field $\{\Omega, \emptyset\}$, we simplify the notation as $\Delta_{p,q} \equiv \Delta_{p,q}(\mathcal{C}_n)$. When $\Delta_{p,q} = o(1)$, we call $\{Y_{j,n}\}$ is (L^p, q) -functionally dependent on $\{e_{i,n}\}$.

Remark 2.3. The term $\Delta_{p,q}(\mathcal{C}_n)$ plays a pivotal role throughout this paper. Since $\delta_{p,n}(j, i, \mathcal{C}_n)$ describes the impact of $e_{i,n}$ on $Y_{j,n}$, $\sum_{j=1}^n \delta_{p,n}(j, i, \mathcal{C}_n)$ is the total impact of $e_{i,n}$ on all $Y_{j,n}$'s, which can be regarded as the “influence power” of $e_{i,n}$ on $Y_{j,n}$'s. And $\frac{1}{n} \sum_{i=1}^n \left[\sum_{j=1}^n \delta_{p,n}(j, i, \mathcal{C}_n) \right]^q$ is the “average” q th power of the influence powers of all $e_{i,n}$'s. Thus, if the average q th power of the influence powers of all $e_{i,n}$'s on $Y_{j,n}$'s is almost surely bounded given \mathcal{C}_n for some $q > 1$, $\Delta_{p,q}(\mathcal{C}_n) = o(1)$ a.s. Critically, we allow some (but not all) individuals to have large influence powers, i.e., we allow $\sup_{i \in [n]} \sum_{j=1}^n \delta_{p,n}(j, i, \mathcal{C}_n)$ to increase to ∞ as $n \rightarrow \infty$. As illustrated in Section 3, for SAR models, $\sum_{j=1}^n \delta_{p,n}(j, i, \mathcal{C}_n)$ is proportional to the summation of the

absolute values of the elements in the i th column of the matrix $(I_n - \lambda W_n)^{-1}$, where W_n is the spatial weights matrix, i.e., $\sum_{j=1}^n \delta_{p,n}(j, i, \mathcal{C}_n) \propto \sum_{j=1}^n |[(I_n - \lambda W_n)^{-1}]_{ji}|$. Therefore, $\Delta_{p,q}(\mathcal{C}_n) = o(1)$ a.s. generalizes a standard assumption in many spatial econometric papers (Kelejian and Prucha, 2001, Lee, 2004, 2007, Yu et al., 2008, etc): $\sup_n \| (I_n - \lambda W_n)^{-1} \|_1 < \infty$, i.e., the column sum norm of $(I_n - \lambda W_n)^{-1}$ is uniformly bounded in n . However, $\Delta_{p,q}(\mathcal{C}_n) = o(1)$ a.s. excludes the case that all $Y_{j,n}$'s are mainly affected by the same very few $e_{i,n}$'s. Consider an extreme case that $Y_{j,n} = e_{1,n}$ for all $j = 1, \dots, n$. Then $\sum_{j=1}^n \delta_{p,n}(j, 1) \propto n$, and thus $\Delta_{p,q} = \frac{1}{n^q} \sum_{i=1}^n \left[\sum_{j=1}^n \delta_{p,n}(j, i) \right]^q \propto 1$, i.e., $\{Y_{j,n}\}$ is not (L^p, q) -functionally dependent on $\{e_{i,n}\}$ for any $p \geq 1$ and $q \geq 1$.

3. Some Examples

To illustrate the application of FDM, in this section, we calculate the FDM for some examples.

3.1. Linear Processes

This example generalizes the linear models studied in Wu (2005) and El Machkouri et al. (2013). Let $\{\epsilon_i : i = 1, 2, \dots, n\}$ be independent random variables; let $A_{ji,n}$ be constant real coefficients and define

$$Y_{j,n} \equiv \sum_{i=1}^n A_{ji,n} \epsilon_i, \quad j = 1, 2, \dots, n.$$

In this example, we consider \mathcal{C}_n as the trivial σ -field $\{\Omega, \emptyset\}$ and $\delta_{p,n}(j, i) = \delta_{p,n}(j, i, \mathcal{C}_n)$. Suppose $C_\epsilon \equiv \sup_i \|\epsilon_i\|_{L^p} < \infty$ for some $p > 1$. Then $\delta_{p,n}(j, i) \leq 2C_\epsilon \cdot |A_{ji,n}|$. With $\delta_{p,n}(j, i)$, it is easy to give a bound for $\Delta_{p,q}$: $\Delta_{p,q} \leq \frac{2C_\epsilon}{n^q} \sum_{i=1}^n \left[\sum_{j=1}^n |A_{ji,n}| \right]^q$.

3.2. SAR model

Let $F : \mathbb{R} \rightarrow \mathbb{R}$ be a Lipschitz function, i.e., $|F(x) - F(y)| \leq L|x - y|$ for some constant L and all $x, y \in \mathbb{R}$. A (possibly nonlinear) SAR model can be written as

$$Y_{j,n} = F(\lambda w_{j,n} Y_n + X'_{j,n} \beta + \epsilon_{j,n}) \tag{3.1}$$

for $j = 1, \dots, n$, where $w_{j,n}$ is the j th row of a non-stochastic and non-zero spatial/network weights matrix $W_n = (w_{ji,n})_{n \times n}$, $Y_n = (Y_{1,n}, Y_{2,n}, \dots, Y_{n,n})'$, $X_{j,n} \in \mathbb{R}^p$ is the exogenous regressor, $\epsilon_{j,n}$ is the disturbance term, and λ and β are model parameters. When $F(x) = x$, Eq.(3.1) is the standard (linear) SAR model; when $F(x) = \max(0, x)$, Eq.(3.1) becomes a SAR Tobit model. Now, we investigate the FDM of SAR model. The following two assumptions on $F(\cdot)$ and $\{\epsilon_{j,n}\}$ are needed.

Assumption 1. *F is a Lipschitz function with a Lipschitz constant $L > 0$, and $\zeta \equiv L|\lambda| \sup_n \|W_n\|_\infty < 1$.*

Remark 3.1. Assumption 1 is also considered by Wu et al. (2024), and it ensures the existence and uniqueness of the solution of the Eq.(3.1). It is similar to the stationary condition for autoregressive models in time series. When individuals are located in a Euclidean space, Assumption 1 implies that the correlation between $Y_{j,n}$'s is weak when their distance is large.

Assumption 2. *Denote $\mathcal{C}_n \equiv \bigvee_{j=1}^n \sigma(X_{j,n})$. $\epsilon_{j,n}$'s are conditionally independent given \mathcal{C}_n and $\|\epsilon\|_{L^p, \mathcal{C}_n} \equiv \sup_{j,n} \|\epsilon_{j,n}\|_{L^p, \mathcal{C}_n} < \infty$ a.s. for some $p > 1$.*

Let $e_{j,n} = X'_{j,n}\beta + \epsilon_{j,n}$ in the SAR model (3.1). Then $Y_{i,n}$'s are generated by the underlying random variables $e_{1,n}, \dots, e_{n,n}$. From Assumption 2, $e_{j,n}$'s are conditionally independent on \mathcal{C}_n . Denote $|W_n| \equiv (|w_{ij,n}|)_{n \times n}$ and

$$S_n^+ \equiv L(I_n - L|\lambda W_n|)^{-1} = (S_{ji,n}^+)_{n \times n}, \quad (3.2)$$

where I_n denotes the $n \times n$ identity matrix.

Proposition 3.1. *Under Assumptions 1-2, $\delta_{p,n}(j, i, \mathcal{C}_n) \leq 2\|\epsilon\|_{L^p, \mathcal{C}_n} S_{ji,n}^+$ a.s.*

Proposition 3.1 implies that to bound $\delta_{p,n}(j, i, \mathcal{C}_n)$, it suffices to bound $S_{ji,n}^+$. In the following, we discuss several examples of the weights matrix W_n , and calculate the corresponding bounds on $S_{ji,n}^+$.

Example 3.1 (Euclidean distance models). In the first example, we consider the case where the n individuals are located in a lattice of the d -dimensional Euclidean space \mathbb{R}^d . We identify the location of each individual with its name and denote the set of locations by $D_n \subset \mathbb{R}^d$. For any two individuals $j = (j_1, \dots, j_d)$ and $i = (i_1, \dots, i_d)$, $d_{ji} \equiv \max_{1 \leq k \leq d} |j_k - i_k|$ denotes their distance. For all $j \neq i \in D_n$, we assume

$d_{ji} \geq 1$. This assumption employs the increasing domain asymptotic and rules out the scenario of infilled asymptotic. Next, we consider two different structures of the spatial weights matrix W_n , which are presented in the following assumption.

Assumption 3. *The spatial weights matrix W_n satisfies one of the following conditions:*

- (i) *only individuals whose distances are less than some specific constant $\bar{d}_0 > 1$ may affect each other directly. In other words, whenever $d_{ji} \geq \bar{d}_0$, $w_{ji,n} = 0$;*
- (ii) *$w_{ji,n} \leq C_0 d_{ji}^{-\alpha}$ for some constants $C_0 > 0$ and $\alpha > d$.*

Note that Assumption 3(i) means that direct interactions only happen between individuals within the distance \bar{d}_0 . It rules out long-distance direct interactions, although indirect interactions are allowed. Under Assumption 3(ii), long-distance direct interactions are allowed; however, the strength of interactions should decrease as the distance increases.

Proposition 3.2. *Consider the settings in Example 3.1. Then (i) under Assumptions 1 and 3(i), we have $S_{ji,n}^+ \leq C_1 \zeta^{d_{ji}/\bar{d}_0}$ for some finite constant $C_1 > 0$, where $S_{ji,n}^+$ is defined by Eq.(3.2); and (ii) under Assumptions 1 and 3(ii), we have $S_{ji,n}^+ \leq C_2 d_{ji}^{-(\alpha-d)} (\log(2d_{ji}))^{\alpha-d}$ for some finite constant $C_2 > 0$.²*

Example 3.2 (Geodesic distance models). Suppose that we observe an undirected network $G_n = (D_n, E_n)$, where $D_n = [n]$ is the set of nodes and $E_n \subset \{(j, i) : j, i \in D_n, j \neq i\}$ denotes the set of links. Let $A_n = (A_{ji,n})_{n \times n}$ be the adjacency matrix of the graph G_n defined by

$$A_{ji,n} = \begin{cases} 1, & j \text{ and } i \text{ are adjacent in } G_n, \\ 0, & \text{otherwise.} \end{cases}$$

We consider the network weights matrix $W_n = (w_{ji,n})_{n \times n}$ given by the row normalization of the adjacency matrix A_n , i.e., $w_{ji,n} = \frac{A_{ji,n}}{\sum_{k=1}^n A_{jk,n}}$. For $j \neq i \in D_n$, let d_{ji} denote the geodesic distance between nodes j and i in G_n , i.e., the length of the shortest path between j and i in G_n , and define $d_{ii} = 0$ for $i \in D_n$.

Proposition 3.3. *For Example 3.2, under Assumption 1, $S_{ji,n}^+ \leq \frac{L}{1-\zeta} \zeta^{d_{ji}}$ for all $j, i \in D_n$, where ζ and L are defined in Assumption 1 and $S_{ji,n}^+$ is defined by Eq.(3.2).*

²The “2” in $\log(2d_{ji})$ is used to ensure that $\log(2d_{ji}) > 0$ for any $j \neq i$.

From Propositions 3.1 and 3.3, $\delta_{p,n}(j, i, \mathcal{C}_n) \leq \|\epsilon\|_{L^p, \mathcal{C}_n} \frac{2L}{1-\zeta} \zeta^{d_{ji}}$.

The above two examples are based on metric spaces for individuals. The following examples, however, do not rely on metric spaces. In these cases, bounding $S_{ji,n}^+$ analytically is challenging. Instead, in Section 4.3, we will compute $S_{ji,n}^+$ using Monte Carlo simulations and verify the conditions required for the CLT (Theorem 4.4).

Example 3.3 (Erdős-Rényi models). We consider the network weights matrix generated by the Erdős-Rényi (ER) model. Suppose that every two individuals have a link with probability $\frac{D}{n-1}$ independently, where $D \geq 1$ denotes the average degree of an individual. We can define a matrix $A \equiv (A_{ji})_{n \times n}$ by

$$A_{ji} = \begin{cases} 1, & j \text{ and } i \text{ are linked,} \\ 0, & \text{otherwise.} \end{cases} \quad (3.3)$$

Then the network weights matrix W_n is the row-normalized matrix A .

Example 3.4 (Triangle models). We consider the network weights matrix generated by the triangle model, which is popular to model the phenomenon of clustering in social networks. We first introduce the triangle structure. We say that individuals i, j, k forms a triangle if there are links between all pairs (i, j) , (j, k) and (i, k) . Suppose that every trio of individuals i, j, k forms a triangle with probability $T/\binom{n}{3}$, where T denotes the average number of triangles in the network. Based on this network formation mechanism, we define a link matrix $A^{(T)} \equiv (A_{ji}^{(T)})_{n \times n}$ as follows:

$$A_{ji}^{(T)} = \begin{cases} 1, & j \text{ and } i \text{ are linked,} \\ 0, & \text{otherwise.} \end{cases}$$

In addition, we consider a link matrix $A^{(\text{ER})}$ that generated from the Erdős-Rényi model with mean degree D . The link matrix of the triangle model is then defined as

$$A \equiv (A_{ji})_{n \times n} = (\max\{A_{ji}^{(T)}, A_{ji}^{(\text{ER})}\})_{n \times n}.$$

And the network weights matrix W_n is given by row-normalizing the matrix A .

Example 3.5 (Stochastic block models). We consider the network weights matrix generated by the stochastic block model (SBM). In this model, the units in the same

block have a higher probability to form links, while units from different blocks are less likely to connect. The SBM is popular in community detection for networks (Abbe, 2018, Wu et al., 2023). Suppose there are M blocks in total, and each individual is randomly assigned to a latent block $k \in \{1, 2, \dots, M\}$ with an equal probability of $\frac{1}{M}$. Denote the within-block mean degree by D_{wb} (i.e., each individual has D_{wb} links of the same block on average) and the between-block mean degree by D_{bb} (i.e., each individual has D_{bb} links from different blocks on average), respectively. The network weights matrix W_n is then given by row-normalizing the matrix $A \equiv (A_{ji})_{n \times n}$, where A_{ji} is defined as in Eq.(3.3).

4. Properties of Functional Dependence

In this section, we establish some inequalities, laws of large numbers (LLN), and central limit theorems (CLT) using FDM. These tools are indispensable to derive asymptotic theories for statistical and econometric models.

The basic idea of most proofs is to express $\sum_{j=1}^n (Y_{j,n} - \mathbb{E}_{\mathcal{C}_n} Y_{j,n})$ as the summation of a martingale difference array (MDA) and apply the theories of MDA. Thus, we first define a sequence of increasing σ -fields carefully. For $i \in [n]$, let $\mathcal{F}_{i,n} \equiv \sigma(e_{j,n} : j \leq i, j \in [n])$ denote the σ -field generated by $e_{1,n}, \dots, e_{i-1,n}, e_{i,n}$, and let $\mathcal{F}_{0,n} \equiv \{\Omega, \emptyset\}$.

4.1. Moment inequalities and a weak law of large numbers

In this subsection, we establish some moment inequalities in Theorem 4.1 for $\sum_{j=1}^n (Y_{j,n} - \mathbb{E}_{\mathcal{C}_n} Y_{j,n})$, which will directly lead to a LLN and will be useful for theoretical analysis in econometrics and statistics. We begin by stating an important lemma.

Lemma 4.1. Denote $P_i Y_{j,n} \equiv \mathbb{E}_{\mathcal{C}_n}(Y_{j,n} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{j,n} | \mathcal{F}_{i-1,n})$. Then, $\|P_i Y_{j,n}\|_{L^p, \mathcal{C}_n} \leq \delta_{p,n}(j, i, \mathcal{C}_n)$ a.s.

Remark 4.1. The $P_i Y_{j,n}$ can be regarded as the change of the prediction in $Y_{j,n}$ when we have the new information $e_{i,n}$, conditional on $\mathcal{F}_{i-1,n}$ and \mathcal{C}_n . Notice that $\{P_i Y_{j,n}, \mathcal{F}_{i,n}\}$ is a MDA under both the conditional expectation $\mathbb{E}_{\mathcal{C}_n}$ and the unconditional expectation \mathbb{E} ,³ and $(Y_{j,n} - \mathbb{E}_{\mathcal{C}_n} Y_{j,n}) = \sum_{i=1}^n P_i Y_{j,n}$.

³ $\mathbb{E}_{\mathcal{C}_n}(P_i Y_{j,n} | \mathcal{F}_{i-1,n}) = \mathbb{E}_{\mathcal{C}_n}[\overline{\{\mathbb{E}_{\mathcal{C}_n}(Y_{j,n} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{j,n} | \mathcal{F}_{i-1,n})\}} | \mathcal{F}_{i-1,n}] = [\mathbb{E}_{\mathcal{C}_n}(Y_{j,n} | \mathcal{F}_{i-1,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{j,n} | \mathcal{F}_{i-1,n})] = 0$. Hence, $\mathbb{E}(P_i Y_{j,n} | \mathcal{F}_{i-1,n}) = \mathbb{E}\mathbb{E}_{\mathcal{C}_n}(P_i Y_{j,n} | \mathcal{F}_{i-1,n}) = 0$.

To obtain the moment inequalities in the following Theorem 4.1, we need two critical lemmas. The first one is Lemma 4.1, which claims that the conditional L^p -norm of the MDA term $P_i Y_{j,n}$ can be bounded by the FDM $\delta_{p,n}(j, i, \mathcal{C}_n)$. And the second one is Lemma A.1, which is the Burkholder's inequality in the martingale theory and is presented in the Appendix.

Theorem 4.1. *Let $C_p \equiv \sqrt{p-1}$ when $p \geq 2$ and $C_p \equiv \frac{1}{p-1}$ when $p \in (1, 2)$. Then for any constant $p > 1$, $\frac{1}{n} \left\| \sum_{j=1}^n (Y_{j,n} - \mathbb{E}_{\mathcal{C}_n} Y_{j,n}) \right\|_{L^p, \mathcal{C}_n} \leq C_p \{\Delta_{p, \min\{p, 2\}}(\mathcal{C}_n)\}^{1/\min\{p, 2\}}$ a.s. If $\Delta_{p, \min\{p, 2\}}(\mathcal{C}_n) = o(1)$ a.s., then $\frac{1}{n} \left\| \sum_{j=1}^n (Y_{j,n} - \mathbb{E}_{\mathcal{C}_n} Y_{j,n}) \right\|_{L^p, \mathcal{C}_n} = o(1)$ a.s.*

A direct result of Theorem 4.1 is a law of large numbers for functionally dependent network random variables, as presented in Theorem 4.2.

Theorem 4.2. *If $\Delta_{p, \min\{p, 2\}}(\mathcal{C}_n) = o(1)$ a.s. for some constant $p > 1$, then $\frac{1}{n} \sum_{j=1}^n (Y_{j,n} - \mathbb{E}_{\mathcal{C}_n} Y_{j,n}) \xrightarrow{\mathbb{P}} 0$.*

Remark 4.2. We compare our LLN to the one presented in Kojevnikov et al. (2021, Theorem 3.1). A key assumption in their LLN is Assumption 3.2, which implicitly assumes that the average dependence strength between any two individuals tends to zero as the geodesic distance tends to infinity, and the geodesic distances between most individuals are large. As they note in the discussion following Assumption 3.2, this assumption rules out networks with small diameters, which are common in social and financial networks. In contrast, our LLN accommodates networks with small diameters. As discussed in Remark 2.3, our assumption $\Delta_{p, \min\{p, 2\}}(\mathcal{C}_n) = o(1)$ only requires that the average influence powers of $e_{i,n}$'s on $Y_{j,n}$'s is finite, a condition that is feasible even for networks with small diameters.

4.2. Concentration inequality

Concentration inequalities, also called exponential inequalities in the literature, play an indispensable role in empirical process theory, semi-parametric and non-parametric econometrics, and high-dimensional statistics (Wainwright, 2019). Wu (2005) and Wu and Wu (2016) establish two concentration inequalities for functionally dependent stationary time series. In the literature, although there are some concentration inequalities for spatial data on irregular lattices (e.g., Xu and Lee (2018), Yuan and Spindler (2025), and Wu et al. (2024) establish some concentration inequalities for

NED and spatial functionally dependent random fields respectively), the concentration inequalities in network econometrics are rare. Hence, in this subsection, we follow the strategies in [Wu and Wu \(2016\)](#) to establish a concentration inequality for network data.

Theorem 4.3. *Let $Z_n \equiv \frac{1}{\sqrt{n}} \sum_{j=1}^n Y_{j,n}$. Assume $\mathbb{E}_{\mathcal{C}_n} Y_{j,n} = 0$ for all $j = 1, 2, \dots, n$ and $n \geq 1$. Assume that $\sup_{n \geq 1} n \Delta_{p,2}(\mathcal{C}_n) < \infty$ a.s. for all $p \geq 2$ and $\sup_{n \geq 1} \sqrt{n} \Delta_{p,2}^{1/2}(\mathcal{C}_n)$ increases slower than $O(p^\nu)$ for some constant $\nu \geq 0$ in the sense that*

$$\sup_{p \geq 2} \sup_{n \geq 1} p^{-\nu} \sqrt{n} \Delta_{p,2}^{1/2}(\mathcal{C}_n) \leq \gamma_0 \text{ a.s.} \quad (4.1)$$

for some finite constant $\gamma_0 > 0$. Let $\alpha = \frac{2}{1+2\nu}$. Then for $t \in [0, t_0]$,

$$m(t) \equiv \mathbb{E} [\exp(t |Z_n|^\alpha) | \mathcal{C}_n] \leq 1 + c_\alpha \left(1 - \frac{t}{t_0}\right)^{-1/2} \frac{t}{t_0} \text{ a.s.,}$$

where $t_0 = (e\alpha\gamma_0^\alpha)^{-1}$, c_α is a constant only depending on α . Consequently, by letting $t = \frac{t_0}{2}$, we have for all $x > 0$,

$$\mathbb{P}(|Z_n| \geq x | \mathcal{C}_n) \leq \exp(-tx^\alpha) m(t) \leq \left(1 + \frac{\sqrt{2}c_\alpha}{2}\right) \exp\left(-\frac{x^\alpha}{2e\alpha\gamma_0^\alpha}\right) \quad (4.2)$$

a.s.

We now illustrate when the condition (4.1) holds. Consider the SAR model from Section 3.1. From Proposition 3.1, we have $\delta_{p,n}(j, i) \leq 2 \|\epsilon\|_{L^p} S_{ji,n}^+$.⁴ As a result, we obtain

$$\Delta_{p,2}^{1/2} = \sqrt{\frac{1}{n^2} \sum_{i=1}^n \left[\sum_{j=1}^n \delta_{p,n}(j, i) \right]^2} \leq \frac{2 \|\epsilon\|_{L^p}}{\sqrt{n}} \sqrt{\frac{1}{n} \sum_{i=1}^n \left[\sum_{j=1}^n S_{ji,n}^+ \right]^2}.$$

If $\sup_{n \geq 1} \frac{1}{n} \sum_{i=1}^n \left[\sum_{j=1}^n S_{ji,n}^+ \right]^2 < C^2 < \infty$ for some constant $C > 0$, then

$$\sup_{p \geq 2} \sup_{n \geq 1} p^{-\nu} \sqrt{n} \Delta_{p,2}^{1/2} \leq 2C \sup_{p \geq 2} p^{-\nu} \|\epsilon\|_{L^p}.$$

⁴For simplicity, here we assume that \mathcal{C}_n is the trivial σ -field $\{\emptyset, \Omega\}$.

Thus, the condition (4.1) holds with $\nu = 1$ if $\epsilon_{i,n}$'s are uniformly sub-exponential, with $\nu = \frac{1}{2}$ if $\epsilon_{i,n}$'s are uniformly sub-Gaussian, and with $\nu = 0$ if $\epsilon_{i,n}$'s are uniformly bounded. Compared to the Hoeffding's inequality (Wainwright, 2019, Proposition 2.5), we see that when $\nu = 0$, the decay rate on the right-hand-side of (4.2) with respect to x is the same as in the independent case. When $\nu > 0$, the decay rate is slower.

4.3. Central limit theorems

In this subsection, we establish two central limit theorems (CLTs) based on the FDM. Denote $S_n \equiv \sum_{j=1}^n Y_{j,n}$, $\sigma_n^2 \equiv \text{Var}_{\mathcal{C}_n}(S_n)$, and $Z_{i,n} \equiv \sum_{j=1}^n P_i Y_{j,n}$ for $i \in [n]$ where $P_i Y_{j,n} = \mathbb{E}_{\mathcal{C}_n}(Y_{j,n} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{j,n} | \mathcal{F}_{i-1,n})$. Since $\{Z_{i,n} : i \in [n]\}$ is a MDA, we will apply a CLT for MDA (Lemma A.2) to establish our CLTs.

4.3.1. CLT for finite index set

Theorem 4.4. *Let $p > 2$ be a constant. Suppose that (1) $\sigma_n^{-2} = O_{\mathbb{P}}(n^{-1})$, (2)*

$$\sup_{n,i} \sum_{k=1}^n \delta_{p,n}(k, i, \mathcal{C}_n) < \infty \text{ a.s.}, \quad (4.3)$$

and (3)

$$\frac{1}{n^{\min\{2,p/2\}}} \sum_{i=1}^n \left\{ \sum_{j=i}^n \sum_{k=1}^n \min \{ \delta_{p,n}(k, i, \mathcal{C}_n), \delta_{p,n}(k, j, \mathcal{C}_n) \} \right\}^{\min\{2,p/2\}} = o(1) \text{ a.s.} \quad (4.4)$$

as $n \rightarrow \infty$. Then

$$\frac{S_n - \mathbb{E}_{\mathcal{C}_n} S_n}{\sigma_n} \xrightarrow{d} N(0, 1).$$

A sufficient, and slightly stronger, condition for Eq.(4.4) is that

$$\frac{1}{n^{\min\{2,p/2\}}} \sum_{i=1}^n \left\{ \sum_{j=1}^n \sum_{k=1}^n \min \{ \delta_{p,n}(k, i, \mathcal{C}_n), \delta_{p,n}(k, j, \mathcal{C}_n) \} \right\}^{\min\{2,p/2\}} = o(1) \text{ a.s.}$$

as $n \rightarrow \infty$, which is invariant to the order of individuals $1, \dots, n$. This condition reflects the fact that our CLT is designed for network data, where the conditions for

the CLT should not depend on the order of individuals within the network. In the following lemma, we derive a more primitive condition for Eq.(4.4).

Lemma 4.2. *For any $i \in [n]$, denote $j(i)$ as the index of the j th largest value of $\{\delta_{p,n}(i, j, \mathcal{C}_n) : 1 \leq j \leq n\}$. If Eq.(4.3) holds and there exist two sequences $\kappa_n = o(n^{(\min\{2,p/2\}-1)/\min\{2,p/2\}})$ and $\tau_n = o(n^{-1/\min\{2,p/2\}})$ such that*

$$\sup_{1 \leq i \leq n} \sum_{j(i) \geq \kappa_n} \delta_{p,n}(i, j(i), \mathcal{C}_n) \leq \tau_n \text{ a.s.}, \quad (4.5)$$

then Eq.(4.4) holds.

We provide a sufficient condition for Eq.(4.5). If

$$\delta_{p,n}(i, j(i), \mathcal{C}_n) \leq C\{j(i)\}^{-\alpha}, \quad \forall i \in [n], \quad (4.6)$$

where $\alpha > \frac{\min\{2,p/2\}}{\min\{2,p/2\}-1} > 1$ and $C > 0$ are some constants, then choosing $\kappa_n = \{n/\log n\}^{1/\alpha}$ yields that

$$\begin{aligned} \sup_{1 \leq i \leq n} \sum_{j(i) \geq \kappa_n} \delta_{p,n}(i, j(i), \mathcal{C}_n) &\leq C \sup_{1 \leq i \leq n} \sum_{j(i) \geq \kappa_n} \{j(i)\}^{-\alpha} \leq C \int_{\kappa_n-1}^{\infty} x^{-\alpha} dx \\ &= \frac{C}{\alpha-1} (\kappa_n - 1)^{-(\alpha-1)} = o(n^{-1/\min\{2,p/2\}}), \end{aligned}$$

i.e., Eq.(4.5) holds.

4.3.2. Comparison with the CLTs in the literature

Next, we will compare our CLT to some in the literature.

- (1) The CLTs developed in Wu (2005), Wu et al. (2024), El Machkouri et al. (2013), Jenish and Prucha (2009, 2012), Kojevnikov et al. (2021) require that the individuals are located in some metric space, which is not required in our CLT. This makes our approach applicable to network data, where individuals are not necessarily located in an underlying metric space.
- (2) The key condition for CLTs in the literature (e.g., Jenish and Prucha, 2009, 2012, Kojevnikov et al., 2021, Wu et al., 2024) typically takes the form $\sum_{s=1}^{\infty} \nu_s \theta_s < \infty$, where s denotes the distance, ν_s represents the number of individuals at a distance approximately equal to s from a specific individual (e.g., in both spatial mixing and

NED, $\nu_s = O(s^{d-1})$ when the individuals are located in \mathbb{R}^d , and θ_s is the weak dependence measure at distance s (e.g., the strong mixing coefficient, NED coefficient, ψ -dependence coefficient). For the condition $\sum_{s=1}^{\infty} \nu_s \theta_s < \infty$ to hold, it is typically required that $\theta_s = O(s^{-1-\delta} \nu_s^{-1}) \rightarrow 0$ as $s \rightarrow \infty$ for some constant $\delta > 0$, meaning that the weak dependence coefficient θ_s must decrease rapidly enough as the distance s increases. Consequently, this condition does not allow ν_s to increase too quickly, thereby excluding networks with small diameters.⁵ In contrast, our condition (4.6) only requires that for any node i , the magnitude of the impact it receives from other nodes (i.e., the FDM $\delta_{p,n}(i, j(i), \mathcal{C}_n)$) decreases as a power function of $j(i)$ when the impacts are ordered from largest to smallest. This allows our CLT to accommodate networks with small or moderate diameters, a common characteristic of social and financial networks.

(3) The CLT in [Kuersteiner \(2019\)](#) is for spatial mixingale processes. [Kuersteiner \(2019\)](#) relaxes the assumption of a fixed metric of distance and introduces a model-dependent random metric. [Kuersteiner \(2019\)](#) assumes the network to be sparse that “rules out a buildup of a mass of nodes with very similar features is captured by a summability condition of the probabilities that two nodes are close in an appropriate sense”. In contrast, our approach does not require any form of underlying metric and allows for networks with small diameters.

(4) Going beyond the conventional notion of weak dependence (e.g., mixing and NED), [Leung and Moon \(2025\)](#) propose the “stabilization” conditions, which requires some form of weak dependence for node degrees. This motivation is similar to our work. However, they consider the setting of strategic network formation, while our concept is applicable to various networks, including social network, spatial network, and financial network, among others.

(5) [Lee et al. \(2022\)](#) study a CLT for a linear-quadratic form of independent random variables with the existence of dominant units (also called popular units in their paper). Their CLT is mainly used for linear SAR type models, but ours is applicable to nonlinear ones.

4.3.3. Verifying the CLT conditions for some examples

We illustrate that conditions (4.3)-(4.4) hold under certain conditions for the examples considered in Section 3.

⁵See, e.g., the first paragraph in [Kojevnikov et al. \(2021, Section 3.1\)](#).

For the linear process considered in Section 3.1, conditions (4.3)-(4.4) become

$$\sup_{n,i} \sum_{k=1}^n |A_{ki,n}| < \infty \text{ and } \frac{1}{n^{\min\{2,p/2\}}} \sum_{i=1}^n \left\{ \sum_{j=i}^n \sum_{k=1}^n \min \{ |A_{ki,n}|, |A_{kj,n}| \} \right\}^{\min\{2,p/2\}} = o(1).$$

For SAR models, by Proposition 3.1, conditions (4.3)-(4.4) become

$$\sup_{n,i} \sum_{k=1}^n S_{ki,n}^+ < \infty \quad (4.7)$$

and

$$\frac{1}{n^{\min\{2,p/2\}}} \sum_{i=1}^n \left\{ \sum_{j=i}^n \sum_{k=1}^n \min \{ S_{ki,n}^+, S_{kj,n}^+ \} \right\}^{\min\{2,p/2\}} = o(1), \quad (4.8)$$

where $S_{ji,n}^+$ is defined in Eq.(3.2). We verify Eqs.(4.7)-(4.8) for Examples 3.1-3.5 considered in Section 3.2.

Lemma 4.3. *For Example 3.1, if Assumptions 1 and 3 hold with $\alpha > 3d$, then*

$$\sup_{i,n} \sum_{j=1}^n \sum_{k=1}^n \min \{ S_{ki,n}^+, S_{kj,n}^+ \} < \infty. \quad (4.9)$$

As a result, conditions (4.3)-(4.4) hold when $p > 2$.

Example 3.2: To establish Eqs.(4.7)-(4.8), we need an additional condition regarding the denseness of the network. Let $\mathcal{N}_n^\partial(i; \ell)$ ($i \in D_n, \ell \geq 0$) denote the set of the nodes that are exactly at the distance ℓ from node i . That is,

$$\mathcal{N}_n^\partial(i; \ell) = \{j \in D_n : d_{ij} = \ell\}.$$

We have the following result.

Lemma 4.4. *For Example 3.2, if Assumption 1 holds and $\sup_{i,n} |\mathcal{N}_n^\partial(i; \ell)| \leq C\eta^\ell$ for some constants $C > 0$ and $0 < \eta < 1/\sqrt{\zeta}$ where ζ is defined in Assumption 1, then it holds that*

$$\sup_{i,n} \sum_{j=1}^n \sum_{k=1}^n \min \{ S_{ki,n}^+, S_{kj,n}^+ \} < \infty. \quad (4.10)$$

As a result, conditions (4.3)-(4.4) hold when $p > 2$.

		Eq.(4.7)			Eq.(4.8)		
		$n = 100$	$n = 400$	$n = 900$	$n = 100$	$n = 400$	$n = 900$
$\lambda = 0.2$	$D = 3$	1.726	1.825	1.881	0.038	0.010	0.005
	$D = 5$	1.606	1.699	1.754	0.064	0.020	0.010
	$D = 10$	1.461	1.524	1.550	0.149	0.073	0.048
$\lambda = 0.3$	$D = 3$	2.163	2.378	2.454	0.084	0.025	0.012
	$D = 5$	1.994	2.141	2.243	0.187	0.079	0.046
	$D = 10$	1.796	1.898	1.938	0.517	0.377	0.324
$\lambda = 0.4$	$D = 3$	2.796	3.059	3.194	0.210	0.072	0.038
	$D = 5$	2.538	2.750	2.868	0.561	0.333	0.244
	$D = 10$	2.215	2.371	2.458	1.666	1.702	1.786
$\lambda = 0.8$	$D = 3$	10.944	12.296	13.357	38.673	48.659	56.889
	$D = 5$	9.788	11.043	11.562	107.801	225.678	356.614
	$D = 10$	8.178	9.161	9.582	207.483	557.337	1024.002

Table 1: Simulation results of Eqs.(4.7)-(4.8) for ER models. We repeat each simulation setting 100 times.

Examples 3.3-3.5: For the spatial weights matrices generated by ER, triangle, and SBM models (Examples 3.3-3.5), it is challenging to give primitive conditions. Hence, we run some simulations to verify the conditions (4.7)-(4.8) with $p = 4$ and $L = 1$ (as defined in Assumption 1). Since the spatial weights matrices generated by these models are random, we compute the mean of the left-hand-side quantities in (4.7) and (4.8), respectively. The simulation results are shown in Tables 1-3.⁶ From the simulation results, we can see that condition (4.7) generally hold for any fixed mean degree D and λ . And condition (4.8) is more plausible to hold with a lower degree D and smaller λ . In other words, the condition (4.8) will fail if the network is too dense and/or the dependence factor λ is too large.

⁶For the SBM models, we generate A as follows. We set $M = \sqrt{n}/2$ blocks, randomly assigning each node a latent label $k \in \{1, 2, \dots, M\}$ with an equal probability of $1/M$. Subsequently, the within-block mean degree D is designed to be 3, 5, and 10 (i.e., each node has D friends of the same group on average), and the between-block mean degree is designed to be 2 (i.e., each node has 2 friends from different groups on average). Lastly, we obtain the spatial weights matrix W by row-normalizing A .

		Eq.(4.7)			Eq.(4.8)		
		$n = 100$	$n = 400$	$n = 900$	$n = 100$	$n = 400$	$n = 900$
$\lambda = 0.2$	$D = 3$	1.719	1.856	1.915	0.037	0.010	0.004
	$D = 5$	1.640	1.730	1.782	0.063	0.020	0.010
	$D = 10$	1.476	1.552	1.576	0.147	0.071	0.047
$\lambda = 0.3$	$D = 3$	2.190	2.399	2.530	0.079	0.023	0.010
	$D = 5$	2.074	2.213	2.317	0.175	0.073	0.042
	$D = 10$	1.802	1.904	1.974	0.511	0.365	0.312
$\lambda = 0.4$	$D = 3$	2.817	3.127	3.357	0.187	0.063	0.032
	$D = 5$	2.593	2.863	2.990	0.534	0.298	0.214
	$D = 10$	2.253	2.400	2.488	1.623	1.648	1.727
$\lambda = 0.8$	$D = 3$	11.290	13.286	13.807	30.042	34.813	37.950
	$D = 5$	10.090	11.521	12.373	98.917	201.064	308.821
	$D = 10$	8.444	9.356	9.902	202.649	542.418	993.015

Table 2: Simulation results of Eqs.(4.7)-(4.8) for triangular models. We repeat each simulation setting 100 times.

In Appendix C, we also verify conditions (4.3)-(4.4) for the weights matrices constructed from real financial networks.

4.3.4. CLTs for multivariate Y

Using the Cramér-Wold device, we can generalize Theorem 4.4 to multivariate case. Now, $Y_{j,n}$'s are random vectors taking values in \mathbb{R}^{p_Y} ($p_Y \geq 1$). Denote $S_n \equiv \sum_{j=1}^n Y_{j,n}$, $\Sigma_n \equiv \text{Var}_{\mathcal{C}_n}(S_n)$, $Z_{i,n} \equiv \sum_{j=1}^n P_i Y_{j,n}$, and I_{p_Y} as a $p_Y \times p_Y$ identity matrix.

Theorem 4.5. Suppose that (1) $\{\min \text{eig}(\Sigma_n)\}^{-1} = O_{\mathbb{P}}(n^{-1})$, (2) $\sup_{n,i} \sum_{k=1}^n \delta_{p,n}(k, i, \mathcal{C}_n) < \infty$ a.s., and (3)

$$\frac{1}{n^{\min\{2,p/2\}}} \sum_{i=1}^n \left\{ \sum_{j=i}^n \sum_{k=1}^n \min \{\delta_{p,n}(k, i, \mathcal{C}_n), \delta_{p,n}(k, j, \mathcal{C}_n)\} \right\}^{\min\{2,p/2\}} = o(1) \quad (4.11)$$

		Eq.(4.7)			Eq.(4.8)		
		$n = 100$	$n = 400$	$n = 900$	$n = 100$	$n = 400$	$n = 900$
$\lambda = 0.2$	$D = 3$	1.589	1.679	1.742	0.064	0.019	0.009
	$D = 5$	1.504	1.588	1.627	0.087	0.030	0.016
	$D = 10$	1.397	1.464	1.499	0.128	0.049	0.029
$\lambda = 0.3$	$D = 3$	1.991	2.129	2.248	0.182	0.072	0.042
	$D = 5$	1.849	1.987	2.062	0.263	0.127	0.083
	$D = 10$	1.677	1.791	1.845	0.377	0.198	0.146
$\lambda = 0.4$	$D = 3$	2.531	2.750	2.856	0.546	0.295	0.208
	$D = 5$	2.301	2.520	2.622	0.805	0.531	0.433
	$D = 10$	2.038	2.225	2.318	1.044	0.743	0.669
$\lambda = 0.8$	$D = 3$	9.703	10.882	11.579	103.269	204.683	310.930
	$D = 5$	8.726	9.581	10.311	130.542	276.419	452.188
	$D = 10$	7.275	8.141	8.928	126.650	265.136	439.604

Table 3: Simulation results of Eqs.(4.7)-(4.8) for SBM models. We repeat each simulation setting 100 times.

a.s. as $n \rightarrow \infty$. Then

$$\Sigma_n^{-1/2}(S_n - \mathbb{E}_{\mathcal{C}_n} S_n) \xrightarrow{d} N(0, I_{p_Y}).$$

5. Functional Dependence Measure Under Transformations

To study the asymptotic properties of an estimator, we usually need to deal with various functions of random variables, e.g., $Y_{i,n}^2$ or $\Phi(Y_{i,n})$, where $\Phi(\cdot)$ is a distribution function of some random variable. So, in order that our theory has more applications, we study the FDM under different transformations, and investigate whether the important inequalities and the CLTs still hold. As in Section 2, denote $Y_{j,n} = F_{j,n}(e_{[n]}) \in \mathbb{R}$, $Y_{j,n,i} \equiv F_{j,n}(e_{1,n}, \dots, e_{i-1,n}, e_{i,n}^*, e_{i+1,n}, \dots, e_{n,n})$, and $\delta_{p,n}(j, i, \mathcal{C}_n) \equiv \|Y_{j,n} - Y_{j,n,i}\|_{L^p, \mathcal{C}_n}$ as the FDM of $\{Y_{j,n}\}$ conditional on \mathcal{C}_n . We add

“(Y)” in $\delta_{p,n}^{(Y)}(j, i, \mathcal{C}_n)$ to differentiate it from the FDM of other random variables.

We first consider Lipschitz-type functions $H_{j,n} : \mathbb{R} \rightarrow \mathbb{R}$. Suppose for all $(y, y^\bullet) \in \mathbb{R} \times \mathbb{R}$ and all j and $n \geq 1$:

$$|H_{j,n}(y) - H_{j,n}(y^\bullet)| \leq B_{j,n}(y, y^\bullet) |y - y^\bullet|. \quad (5.1)$$

In Propositions 5.1-5.3, we denote $Z_{j,n} \equiv H_{j,n}(Y_{j,n})$.

Proposition 5.1. *Suppose $H_{j,n}(\cdot)$ satisfies Eq.(5.1) with $\sup_{n,j} \sup_{y,y^\bullet} B_{j,n}(y, y^\bullet) \leq C < \infty$ for some constant C . Then*

$$\delta_{p,n}^{(Z)}(j, i, \mathcal{C}_n) \leq C \delta_{p,n}^{(Y)}(j, i, \mathcal{C}_n) \text{ a.s.}$$

Obviously, when the FDM of $\{Y_{j,n} : j \in [n], n \in \mathbb{N}\}$ satisfies the conditions in Theorem 4.1, 4.3, or 4.4, so will the FDM of $\{Z_{j,n} : j \in [n], n \in \mathbb{N}\}$.

Next, we consider unbounded $B_{j,n}(y, y^\bullet)$, e.g., $H_{j,n}(y) = y^2$.

Proposition 5.2. *Suppose $H_{j,n}(\cdot)$ satisfies Eq.(5.1) with $B_{j,n}(y, y^\bullet) \leq C_1(|y|^a + |y^\bullet|^a + 1)$ for some finite constants $C_1 > 0$ and $a \geq 1$, constants $p, q, r \geq 1$ satisfying $p^{-1} = q^{-1} + r^{-1}$, and $\|Y\|_{L^{ar}, \mathcal{C}_n} \equiv \sup_{n,j} \|Y_{j,n}\|_{L^{ar}, \mathcal{C}_n} < \infty$ a.s. Then*

$$\delta_{p,n}^{(Z)}(j, i, \mathcal{C}_n) \leq C_1(2 \|Y\|_{L^{ar}, \mathcal{C}_n}^a + 1) \delta_{q,n}^{(Y)}(j, i, \mathcal{C}_n) \text{ a.s.}$$

Hence, if $\frac{1}{n^{\tilde{q}}} \sum_{i=1}^n \left[\sum_{j=1}^n \delta_{q,n}^{(Y)}(j, i, \mathcal{C}_n) \right]^{\tilde{q}} = o(1)$ a.s. and $\|Y\|_{L^{ar}, \mathcal{C}_n} < \infty$ a.s., then

$$\frac{1}{n^{\tilde{q}}} \sum_{i=1}^n \left[\sum_{j=1}^n \delta_{p,n}^{(Z)}(j, i, \mathcal{C}_n) \right]^{\tilde{q}} = o(1) \text{ a.s.}$$

So Theorem 4.1 is applicable. Similarly, condition (4.3) in Theorem 4.4 can be verified for $\{Z_{j,n}\}$. In Proposition 5.2, there is a trade-off between p , q and r . If we want a larger p , then we need a larger q or larger r . If we further impose some conditions on $\delta_{p,n}^{(Y)}(j, i, \mathcal{C}_n)$, then we can avoid the trade-off between p , q and r , which is presented in the next proposition.

Proposition 5.3. *Suppose $H_{j,n}(\cdot)$ satisfies condition (5.1) with $B_{j,n}(y, y^\bullet) \leq C_1(|y|^a + |y^\bullet|^a + 1)$ for some finite constants $C_1 > 0$ and $a \geq 1$. If $\|Y\|_{L^q, \mathcal{C}_n} \equiv \sup_{n,j} \|Y_{j,n}\|_{L^q, \mathcal{C}_n} <$*

∞ a.s. for some $q > \max\left(\frac{ap}{p-1}, ap + p\right)$. Then

$$\delta_{p,n}^{(Z)}(j, i, \mathcal{C}_n) \leq C_2(\mathcal{C}_n) [\delta_{p,n}^{(Y)}(j, i, \mathcal{C}_n)]^{(q-ap-p)/(pq-ap-p)} \text{ a.s.,}$$

where $C_2(\mathcal{C}_n) < \infty$ a.s.

When $\frac{1}{n^q} \sum_{j=1}^n \left\{ \sum_{i=1}^n [\delta_{p,n}^{(Y)}(i, j, \mathcal{C}_n)]^{(q-ap-p)/(pq-ap-p)} \right\}^q = o(1)$ a.s., we have

$$\frac{1}{n^q} \sum_{j=1}^n \left\{ \sum_{i=1}^n \delta_{p,n}^{(Z)}(i, j, \mathcal{C}_n) \right\}^q = o(1) \text{ a.s.}$$

by Proposition 5.3. So, Theorem 4.1 is applicable. And we can verify condition (4.3) in Theorem 4.4 using similar arguments.

In addition to Lipschitz transformation, another important nonlinear transformation is $\mathbf{1}(y > 0)$, which is useful in discrete choice and censor data model. See, e.g., Xu and Lee (2015, 2018).

Proposition 5.4. Denote $Z_{j,n} \equiv \mathbf{1}(Y_{j,n} > 0)$. Denote the density function of $Y_{j,n}$ conditional on \mathcal{C}_n by $f_{j,n}(y | \mathcal{C}_n)$. If $\sup_{n \geq 1, 1 \leq j \leq n} \sup_y f_{j,n}(y | \mathcal{C}_n) < C_1 < \infty$ a.s. for some constant $C_1 > 0$, then

$$\delta_{p,n}^{(Z)}(j, i, \mathcal{C}_n) \leq C [\delta_{p,n}^{(Y)}(j, i, \mathcal{C}_n)]^{1/(p+1)},$$

for some constant C not depending on i, j , nor n .

Moreover, we are interested in the sum or the product of two random variables. Let $\{Y_{j,n}\}$ and $\{Z_{j,n}\}$ be two sets of random variables, and they are both functions of some conditionally independent random vectors $e_{i,n}$ given \mathcal{C}_n , where $i \in [n]$. Denote the FDMs of $\{Y_{j,n} + Z_{j,n}\}$ and $\{Y_{j,n}Z_{j,n}\}$ by $\delta_{p,n}^{(Y+Z)}(j, i, \mathcal{C}_n)$ and $\delta_{p,n}^{(YZ)}(j, i, \mathcal{C}_n)$, respectively. For summation, the conclusion is a direct result of Minkowski's inequality, and we summarize it below.

Proposition 5.5. $\delta_{p,n}^{(Y+Z)}(j, i, \mathcal{C}_n) \leq \delta_{p,n}^{(Y)}(j, i, \mathcal{C}_n) + \delta_{p,n}^{(Z)}(j, i, \mathcal{C}_n)$ a.s.

The FDM of product is more complicated. For simplicity, we assume the original random fields are real-valued, since if Y is functionally dependent, then all the elements of Y are still functionally dependent and vice versa. Like Propositions 5.2 and 5.3, we have two versions of conclusions, as are given below.

Proposition 5.6. Suppose constants $p, q_1, q_2, r_1, r_2 > 1$ satisfy $p^{-1} = q_1^{-1} + r_1^{-1} = q_2^{-1} + r_2^{-1}$, $\|Y\|_{L^{r_2}, \mathcal{C}_n} \equiv \sup_{n,j} \|Y_{j,n}\|_{L^{r_2}, \mathcal{C}_n} < \infty$ a.s. and $\|Z\|_{L^{r_1}, \mathcal{C}_n} \equiv \sup_{n,j} \|Z_{j,n}\|_{L^{r_1}, \mathcal{C}_n} < \infty$ a.s. Then

$$\delta_{p,n}^{(YZ)}(j, i, \mathcal{C}_n) \leq \|Z\|_{L^{r_1}, \mathcal{C}_n} \delta_{q_1,n}^{(Y)}(j, i, \mathcal{C}_n) + \|Y\|_{L^{r_2}, \mathcal{C}_n} \delta_{q_2,n}^{(Z)}(j, i, \mathcal{C}_n) \text{ a.s.}$$

Let us investigate whether $\{Y_{j,n}Z_{j,n}\}$ satisfies the conditions in Theorem 4.1. Denote $\Delta_{p,2}^{(YZ)}(\mathcal{C}_n) \equiv \frac{1}{n^2} \sum_{i=1}^n [\sum_{j=1}^n \delta_{p,n}^{(YZ)}(j, i, \mathcal{C}_n)]^2$, and $\Delta_{q_1,2}^{(Y)}(\mathcal{C}_n)$ and $\Delta_{q_2,2}^{(Z)}(\mathcal{C}_n)$ are defined similarly. Then by Proposition 5.6,

$$\begin{aligned} \Delta_{p,2}^{(YZ)}(\mathcal{C}_n) &\leq \frac{1}{n^2} \sum_{i=1}^n \left[\|Z\|_{L^{r_1}, \mathcal{C}_n} \sum_{j=1}^n \delta_{q_1,n}^{(Y)}(j, i, \mathcal{C}_n) + \|Y\|_{L^{r_2}, \mathcal{C}_n} \sum_{j=1}^n \delta_{q_2,n}^{(Z)}(j, i, \mathcal{C}_n) \right]^2 \\ &\leq \frac{1}{n^2} \sum_{i=1}^n \left\{ 2 \|Z\|_{L^{r_1}, \mathcal{C}_n}^2 \left[\sum_{j=1}^n \delta_{q_1,n}^{(Y)}(i, j, \mathcal{C}_n) \right]^2 + 2 \|Y\|_{L^{r_2}, \mathcal{C}_n}^2 \left[\sum_{j=1}^n \delta_{q_2,n}^{(Z)}(i, j, \mathcal{C}_n) \right]^2 \right\} \\ &= 2 \|Z\|_{L^{r_1}, \mathcal{C}_n}^2 \Delta_{q_1,2}^{(Y)}(\mathcal{C}_n) + 2 \|Y\|_{L^{r_2}, \mathcal{C}_n}^2 \Delta_{q_2,2}^{(Z)}(\mathcal{C}_n) \rightarrow 0 \text{ a.s.}, \end{aligned}$$

where the second inequality follows from the fact that $(a+b)^2 \leq 2a^2+2b^2$ for arbitrary $a, b \in \mathbb{R}$. Consequently, when $\max \{\Delta_{q_1,2}^{(Y)}(\mathcal{C}_n), \Delta_{q_2,2}^{(Z)}(\mathcal{C}_n)\} = o(1)$ a.s., $\{Y_{j,n}Z_{j,n}\}$ also satisfies the conditions in Theorem 4.1.

Similar to Proposition 5.2, there is a trade-off between p, q_1, r_1 and p, q_2, r_2 . Since $q_1 > p$ and $q_2 > p$, when three or more random fields are multiplied, larger values for q_1 and q_2 are required, which may not always be feasible. However, by introducing additional conditions, it may be possible to avoid such trade-offs, as demonstrated below.

Proposition 5.7. Suppose $\|Y\|_{L^q, \mathcal{C}_n} \equiv \sup_{n,j} \|Y_{j,n}\|_{L^q, \mathcal{C}_n} < \infty$ a.s., $\|Z\|_{L^q, \mathcal{C}_n} \equiv \sup_{n,j} \|Z_{j,n}\|_{L^q, \mathcal{C}_n} < \infty$ a.s. for some $q > \max(p/(p-1), 2p)$ and $p > 1$. Then

$$\delta_{p,n}^{(YZ)}(j, i, \mathcal{C}_n) \leq C_1(\mathcal{C}_n) [\delta_{p,n}^{(Y)}(j, i, \mathcal{C}_n)]^{\frac{q-2p}{pq-2p}} + C_2(\mathcal{C}_n) [\delta_{p,n}^{(Z)}(j, i, \mathcal{C}_n)]^{\frac{q-2p}{pq-2p}}$$

a.s., where $C_1(\mathcal{C}_n) < \infty$ a.s. and $C_2(\mathcal{C}_n) < \infty$ a.s.

With Proposition 5.7 and $(a+b)^r \leq \max\{1, 2^{r-1}\}(a^r + b^r)$ for any $a, b, r \geq 0$, we

can show that when

$$\frac{1}{n^2} \sum_{i=1}^n \left[\sum_{j=1}^n (\delta_{p,n}^{(Y)}(i, j, \mathcal{C}_n))^{(q-2p)/(pq-2p)} \right]^2 = o(1) \text{ a.s.}$$

and

$$\frac{1}{n^2} \sum_{i=1}^n \left[\sum_{j=1}^n (\delta_{p,n}^{(Z)}(i, j, \mathcal{C}_n))^{(q-2p)/(pq-2p)} \right]^2 = o(1) \text{ a.s.,}$$

we have

$$\Delta_{p,2}^{(YZ)}(\mathcal{C}_n) = \frac{1}{n^2} \sum_{i=1}^n \left[\sum_{j=1}^n \delta_{p,n}^{(YZ)}(i, j) \right]^2 = o(1) \text{ a.s.}$$

Then, Theorem 4.1 is applicable. We can verify condition (4.3) in Theorem 4.4 under some similar conditions.

In practice, we can use either Proposition 5.6 or 5.7 to study the FDM of $\{Y_{j,n}Z_{j,n}\}$. Proposition 5.6 requires a stronger moment condition, while Proposition 5.7 requires that most of $\delta_{p,n}^{(Y)}(i, j, \mathcal{C}_n)$ are smaller, as $\frac{(q-2p)}{pq-2p} < 1$ implies that $\left[\delta_{p,n}^{(Y)}(i, j, \mathcal{C}_n) \right]^{(q-2p)/(pq-2p)} > \delta_{p,n}^{(Y)}(i, j, \mathcal{C}_n)$.⁷

6. Conclusion

This paper develops limit theorems for random variables with network links, without requiring individuals to be located in a Euclidean or metric space. This approach stands apart from most existing limit theorems in network econometrics, which typically rely on weak dependence assumptions such as strong mixing, near-epoch dependence, or ψ -dependence. To derive the limit theorems, we generalize the concept of functional dependence measure proposed by Wu (2005). Using this concept, we derive several inequalities, as well as law of large numbers and central limit theorems. We also demonstrate the applicability of these theorems by verifying their conditions for spatial autoregressive models, which are widely used in network data analysis. Finally, we study the functional dependence measure under various transformations that are needed for econometrics.

We outline two potential extensions of our work: (i) relaxing the condition (4.3)

⁷Note that most of $\delta_{p,n}^{(Y)}(i, j, \mathcal{C}_n)$ are smaller than 1 asymptotically.

in Theorem 4.4 to allow for the existence of powerful nodes; and (ii) generalizing our theories so that it can be applied to panel data with network links.

Appendices

A. Some Useful Lemmas

Lemma A.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and \mathcal{C} is a sub- σ -field of \mathcal{F} . Let X_1, X_2, \dots, X_n be a zero-mean martingale difference array under the conditional expectation $\mathbb{E}_{\mathcal{C}}$. Let $p > 1$ be a constant. Let $C_p \equiv \sqrt{p-1}$ when $p \geq 2$ and $C_p \equiv (p-1)^{-1}$ when $p \in (1, 2)$. Then for any $p > 1$, it holds that

$$\left\| \sum_{i=1}^n X_i \right\|_{L^p, \mathcal{C}} \leq C_p \left(\sum_{i=1}^n \|X_i\|_{L^p, \mathcal{C}}^{\min\{p, 2\}} \right)^{1/\min\{p, 2\}} \quad a.s. \quad (\text{A.1})$$

Lemma A.2. (Hall and Heyde, 1980, Theorem 3.2). Given a triangular MDA $\{X_{nk}, \mathcal{F}_{nk} : 1 \leq k \leq k_n, n \geq 1\}$, if (a) $\sum_{k=1}^{k_n} X_{nk}^2 \xrightarrow{\mathbb{P}} 1$, (b) $\max_{1 \leq k \leq k_n} |X_{nk}| \xrightarrow{\mathbb{P}} 0$, and (c) $\mathbb{E} \max_{1 \leq k \leq k_n} X_{nk}^2 < K$ for all $n \geq 1$, where $K > 0$ is a constant, then $S_n = \sum_{k=1}^{k_n} X_{nk} \xrightarrow{d} N(0, 1)$.

Lemma A.3. Consider the system (2.1). Suppose index set \mathcal{I} is a non-empty subset of $[n]$. Then for any $1 \leq j \leq n$ and $i \in I$,

$$\|\mathbb{E}_{\mathcal{C}_n}(Y_{j,n} | \sigma(e_{k,n} : k \in \mathcal{I})) - \mathbb{E}_{\mathcal{C}_n}(Y_{j,n} | \sigma(e_{k,n} : k \in \mathcal{I} \setminus \{i\}))\|_{L^p, \mathcal{C}_n} \leq \delta_{p,n}(j, i, \mathcal{C}_n).$$

B. Proofs for Section 4

Proof of Lemma 4.1: This directly follows from Lemma A.3. \square

Proof of Theorem 4.1: Notice that $(Y_{j,n} - \mathbb{E}_{\mathcal{C}_n} Y_{j,n}) = \sum_{i=1}^n P_i Y_{j,n}$. As a result,

$$\sum_{j=1}^n (Y_{j,n} - \mathbb{E}_{\mathcal{C}_n} Y_{j,n}) = \sum_{j=1}^n \sum_{i=1}^n P_i Y_{j,n} = \sum_{i=1}^n \left(\sum_{j=1}^n P_i Y_{j,n} \right). \quad (\text{B.1})$$

Because $\left\{\sum_{j=1}^n P_i Y_{j,n} : i \in [n]\right\}$ is a MDA and measurable with respect to \mathcal{C}_n ,

$$\begin{aligned}
& \left\| \sum_{j=1}^n (Y_{j,n} - \mathbb{E}_{\mathcal{C}_n} Y_{j,n}) \right\|_{L^p, \mathcal{C}_n} = \left\| \sum_{i=1}^n \left(\sum_{j=1}^n P_i Y_{j,n} \right) \right\|_{L^p, \mathcal{C}_n} \\
& \leq C_p \left(\sum_{i=1}^n \left\| \sum_{j=1}^n P_i Y_{j,n} \right\|_{L^p, \mathcal{C}_n}^{\min\{p,2\}} \right)^{\frac{1}{\min\{p,2\}}} \leq C_p \left[\sum_{i=1}^n \left(\sum_{j=1}^n \delta_{p,n}(j, i, \mathcal{C}_n) \right)^{\min\{p,2\}} \right]^{\frac{1}{\min\{p,2\}}} \\
& \leq C_p \left\{ \Delta_{p, \min\{p,2\}}(\mathcal{C}_n) \right\}^{\frac{1}{\min\{p,2\}}} n
\end{aligned} \tag{B.2}$$

a.s., where the first equality follows from Eq.(B.1), the first inequality follows from Lemma A.1, the second one follows from conditional Minkowski's inequality and Lemma 4.1, and the last one follows from the definition of $\Delta_{p, \min\{p,2\}}(\mathcal{C}_n)$. \square

Proof of Theorem 4.3: The idea of the proof is borrowed from that in Wu and Wu (2016). From Theorem 4.1, for any $p \geq 2$,

$$\left\| \sum_{j=1}^n Y_{j,n} \right\|_{L^p, \mathcal{C}_n} \leq \sqrt{p-1} \sqrt{\Delta_{p,2}(\mathcal{C}_n)} n \text{ a.s.}$$

Consequently, $\|Z_n\|_{L^p, \mathcal{C}_n} \leq \sqrt{p-1} \sqrt{n} \Delta_{p,2}^{1/2}(\mathcal{C}_n)$ a.s. for $p \geq 2$. Recall a Taylor's formula from Wu and Wu (2016): $(1-s)^{-1/2} = 1 + \sum_{k=1}^{\infty} a_k s^k$, where $|s| < 1$ and $a_k = (2k)! / (2^{2k} (k!)^2)$, $a_0 = 1$. By Stirling's formula, $a_k \sim (\pi k)^{-1/2}$ as $k \rightarrow \infty$. Hence, $k! \sim \sqrt{2} (k/e)^k a_k^{-1}$ and $\frac{a_k}{a_{k-1}} \rightarrow 1$, and there exist constants $c_1, c_2 > 0$ such that $k! \geq c_1 (k/e)^k a_k^{-1}$ and $a_k \leq c_2 a_{k-1}$ hold for all $k \geq 1$. By Eq.(4.1), when $\alpha k \geq 2$, we have $\sup_{n \geq 1} \sqrt{n} \Delta_{\alpha k,2}^{1/2}(\mathcal{C}_n) \leq \gamma_0 (\alpha k)^\nu$. As a result, when $\alpha k \geq 2$,

$$\begin{aligned}
& \frac{t^k \|Z_n\|_{L^{\alpha k}, \mathcal{C}_n}^{\alpha k}}{k!} \leq \frac{t^k (\alpha k - 1)^{\alpha k/2} \left[\sqrt{n} \Delta_{\alpha k,2}^{1/2}(\mathcal{C}_n) \right]^{\alpha k}}{c_1 (k/e)^k a_k^{-1}} \leq \frac{t^k (\alpha k - 1)^{\alpha k/2} \gamma_0^{\alpha k} (\alpha k)^{\alpha k \nu}}{c_1 (k/e)^k a_k^{-1}}, \\
& = \frac{a_k t^k (\alpha k - 1)^{\alpha k/2}}{c_1 t_0^k (\alpha k)^{\alpha k/2}} \leq \frac{a_k}{c_1 \sqrt{e}} \frac{t^k}{t_0^k} \text{ a.s.},
\end{aligned}$$

where the equality follows from $t_0 = (e \alpha \gamma_0^\alpha)^{-1}$ and $\nu = \frac{1}{\alpha} - \frac{1}{2}$, and the last inequality is due to the fact that $(x-1)^{x/2} / x^{x/2} \leq e^{-1/2}$ for all $x \geq 2$. When $0 < \alpha k < 2$ and

$k \geq 1$, we have $\|Z_n\|_{L^{\alpha k}, \mathcal{C}_n} \leq \|Z_n\|_{L^2, \mathcal{C}_n} \leq \sqrt{n} \Delta_{2,2}^{1/2}(\mathcal{C}_n) \leq 2^\nu \gamma_0$ a.s. and

$$\frac{t^k \|Z_n\|_{L^{\alpha k}, \mathcal{C}_n}^{\alpha k}}{k!} \leq \frac{t^k 2^{\nu \alpha k} \gamma_0^{\alpha k}}{c_1 (k/e)^k a_k^{-1}} = \frac{a_k t^k}{c_1 t_0^k} \frac{2^{\nu \alpha k}}{(\alpha k)^k} \leq \frac{2^{2/\alpha-1}}{\min\{\alpha, \alpha^{2/\alpha}\}} \frac{a_k t^k}{c_1 t_0^k} \text{ a.s.},$$

where the equality is because $t_0 = (e\alpha \gamma_0^\alpha)^{-1}$ and the last inequality follows from $2^{\nu \alpha k} \leq 2^{2\nu}$ (as $\nu = \frac{1}{\alpha} - \frac{1}{2} \geq 0$), and $(\alpha k)^k \geq \alpha 1(\alpha \geq 1) + \alpha^{2/\alpha} 1(\alpha < 1) \geq \min\{\alpha, \alpha^{2/\alpha}\}$. Using $e^x = 1 + \sum_{k=1}^{\infty} \frac{x^k}{k!}$ and the above two displayed inequalities, we obtain

$$\begin{aligned} m(t) &= 1 + \sum_{k=1}^{\infty} \frac{t^k \mathbb{E}[|Z_n|^{\alpha k} | \mathcal{C}_n]}{k!} = 1 + \sum_{1 \leq k < 2/\alpha} \frac{t^k \|Z_n\|_{L^{\alpha k}, \mathcal{C}_n}^{\alpha k}}{k!} + \sum_{k \geq 2/\alpha} \frac{t^k \|Z_n\|_{L^{\alpha k}, \mathcal{C}_n}^{\alpha k}}{k!} \\ &\leq 1 + \sum_{1 \leq k < 2/\alpha} \frac{2^{2/\alpha-1}}{\min\{\alpha, \alpha^{2/\alpha}\}} \frac{a_k t^k}{c_1 t_0^k} + \sum_{k \geq 2/\alpha} \frac{a_k}{c_1 \sqrt{e}} \frac{t^k}{t_0^k} \leq 1 + c'_\alpha \sum_{k=1}^{\infty} a_k \frac{t^k}{t_0^k} \\ &\leq 1 + c'_\alpha \sum_{k=1}^{\infty} c_2 a_{k-1} \frac{t^k}{t_0^k} = 1 + c_\alpha \frac{t}{t_0} \sum_{k=0}^{\infty} a_k \frac{t^k}{t_0^k} = 1 + c_\alpha \frac{t/t_0}{(1-t/t_0)^{1/2}} \text{ a.s.}, \end{aligned}$$

where $c'_\alpha, c_\alpha \geq 0$ are constants depending only on α . By conditional Markov's inequality, letting $t = t_0/2$,

$$\begin{aligned} \mathbb{P}(|Z_n| \geq x | \mathcal{C}_n) &= \mathbb{P}(\exp(t|Z_n|^\alpha) \geq \exp(tx^\alpha) | \mathcal{C}_n) \\ &\leq \exp(-tx^\alpha) m(t) \leq \left(1 + \frac{\sqrt{2}c_\alpha}{2}\right) \exp\left(-\frac{x^\alpha}{2e\alpha\gamma_0^\alpha}\right) \end{aligned}$$

a.s. \square

Proof of Theorem 4.4: Recall: $P_i Y_{j,n} \equiv \mathbb{E}_{\mathcal{C}_n}(Y_{j,n} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{j,n} | \mathcal{F}_{i-1,n})$, and $Z_{i,n} \equiv \sum_{j=1}^n P_i Y_{j,n}$. The proof of this theorem is based on a CLT for MDA (Lemma A.2). Notice that

$$\frac{S_n - \mathbb{E}_{\mathcal{C}_n} S_n}{\sigma_n} = \sigma_n^{-1} \sum_{j=1}^n (Y_{j,n} - \mathbb{E}_{\mathcal{C}_n} Y_{j,n}) = \sigma_n^{-1} \sum_{j=1}^n \sum_{i=1}^n P_i Y_{j,n} = \sum_{i=1}^n \sigma_n^{-1} Z_{i,n} = \sum_{i=1}^n X_{i,n},$$

where $X_{i,n} \equiv \sigma_n^{-1} Z_{i,n}$. And $\{X_{i,n}, \mathcal{F}_{i,n} : i = 1, \dots, n, n \geq 1\}$ is a MDA under the unconditional probability measure \mathbb{P} . To apply Lemma A.2, we need to show that (a) $\sum_{i=1}^n X_{i,n}^2 \xrightarrow{\mathbb{P}} 1$, (b) $\max_{i=1, \dots, n} |X_{i,n}| \xrightarrow{\mathbb{P}} 0$, and (c) $\mathbb{E}(\max_{i=1, \dots, n} X_{i,n}^2) < K$ for all $n \geq 1$, where $K > 0$ is a constant.

Condition (a): Since $\{X_{i,n}, i = 1, \dots, n\}$ is also a MDA under the conditional probability measure $\mathbb{P}_{\mathcal{C}_n}$ and σ_n is measurable with respect to \mathcal{C}_n , we have

$$\sigma_n^2 = \text{Var}_{\mathcal{C}_n}(S_n) = \text{Var}_{\mathcal{C}_n} \left[\sum_{i=1}^n \left(\sum_{j=1}^n P_i Y_{j,n} \right) \right] = \sigma_n^2 \sum_{i=1}^n \text{Var}_{\mathcal{C}_n}(X_{i,n}),$$

and thus

$$\sum_{i=1}^n \mathbb{E}_{\mathcal{C}_n} X_{i,n}^2 = \sum_{i=1}^n \text{Var}_{\mathcal{C}_n}(X_{i,n}) = 1 \quad (\text{B.3})$$

a.s. Then, condition (a) is equivalent to $\sigma_n^{-2} \sum_{i=1}^n (Z_{i,n}^2 - \mathbb{E}_{\mathcal{C}_n} Z_{i,n}^2) \xrightarrow{\mathbb{P}} 0$. From Condition (1) in this theorem, $\sigma_n^{-2} = O_{\mathbb{P}}(\frac{1}{n})$. Then it suffices to show that

$$\frac{1}{n} \sum_{i=1}^n (Z_{i,n}^2 - \mathbb{E}_{\mathcal{C}_n} Z_{i,n}^2) \xrightarrow{\mathbb{P}} 0$$

as $n \rightarrow \infty$. We intend to apply Theorem 4.1. Let

$$\mathcal{G}_{i,j} = \begin{cases} \sigma(e_{1,n}, \dots, e_{j,n}^*, \dots, e_{i,n}) & \text{if } j \leq i \\ \mathcal{F}_{i,n} & \text{if } j > i \end{cases}$$

and $Z_{i,n,j} = \sum_{k=1}^n [\mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i-1,j})]$. Here $Z_{i,n,j}$ is the coupled version of $Z_{i,n}$ with $e_{j,n}$ being replaced by its i.i.d. copy $e_{j,n}^*$ conditional on \mathcal{C}_n . The FDM $\delta_{p/2}^*(i, j, \mathcal{C}_n) \equiv \|Z_{i,n}^2 - Z_{i,n,j}^2\|_{L^{p/2}, \mathcal{C}_n}$ measures the impact of $e_{j,n}$ on $Z_{i,n}^2$ under the $L^{p/2}$ -norm. An immediate observation is $\delta_{p/2}^*(i, j, \mathcal{C}_n) = 0$ a.s. for all $j > i$ since $Z_{i,n} = Z_{i,n,j}$ for all $j > i$. Then, to apply Theorem 4.1, it suffices to show

$$\frac{1}{n^{\min\{p/2, 2\}}} \sum_{i=1}^n \left[\sum_{j=i}^n \delta_{p/2}^*(j, i, \mathcal{C}_n) \right]^{\min\{p/2, 2\}} = o(1) \quad a.s. \quad (\text{B.4})$$

as $n \rightarrow \infty$. To show Eq.(B.4), we control the FDM $\delta_{p/2}^*(i, j, \mathcal{C}_n)$ for $i \geq j$. Note that we have

$$\begin{aligned} \delta_{p/2}^*(i, j, \mathcal{C}_n) &= \|Z_{i,n}^2 - Z_{i,n,j}^2\|_{L^{p/2}, \mathcal{C}_n} = \|(Z_{i,n} - Z_{i,n,j})(Z_{i,n} + Z_{i,n,j})\|_{L^{p/2}, \mathcal{C}_n} \\ &\leq \|Z_{i,n} + Z_{i,n,j}\|_{L^p, \mathcal{C}_n} \|Z_{i,n} - Z_{i,n,j}\|_{L^p, \mathcal{C}_n} \end{aligned}$$

$$\begin{aligned}
&\leq \left(\|Z_{i,n,j}\|_{L^p, \mathcal{C}_n} + \|Z_{i,n}\|_{L^p, \mathcal{C}_n} \right) \|Z_{i,n} - Z_{i,n,j}\|_{L^p, \mathcal{C}_n} \\
&= 2 \|Z_{i,n}\|_{L^p, \mathcal{C}_n} \|Z_{i,n} - Z_{i,n,j}\|_{L^p, \mathcal{C}_n} = 2 \left\| \sum_{j=1}^n P_i Y_{j,n} \right\|_{L^p, \mathcal{C}_n} \|Z_{i,n} - Z_{i,n,j}\|_{L^p, \mathcal{C}_n} \\
&\leq 2 \left(\sum_{j=1}^n \|P_i Y_{j,n}\|_{L^p, \mathcal{C}_n} \right) \|Z_{i,n} - Z_{i,n,j}\|_{L^p, \mathcal{C}_n} \leq 2 \left(\sum_{j=1}^n \delta_{p,n}(j, i, \mathcal{C}_n) \right) \|Z_{i,n} - Z_{i,n,j}\|_{L^p, \mathcal{C}_n} \\
&\leq 2C \|Z_{i,n} - Z_{i,n,j}\|_{L^p, \mathcal{C}_n} = 2C \left\| \sum_{k=1}^n \left[\mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i-1,n}) \right. \right. \\
&\quad \left. \left. - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j}) + \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i-1,j}) \right] \right\|_{L^p, \mathcal{C}_n} \leq 2C \sum_{k=1}^n \left\| \mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i,n}) \right. \\
&\quad \left. - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i-1,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j}) + \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i-1,j}) \right\|_{L^p, \mathcal{C}_n}, \tag{B.5}
\end{aligned}$$

where the first inequality follows from Hölder's inequality, the fourth one follows from Lemma 4.1, and the fifth one follows from Condition (2) in this theorem (C denotes the upper bound in Condition (2)), the third equality is from the fact that $Z_{i,n,j}$ and $Z_{i,n}$ are identically distributed conditional on \mathcal{C}_n , and the fourth and the fifth equalities are from the definitions of $Z_{i,n}$ and $Z_{i,n,j}$.

On the one hand, by Minkowski's inequality,

$$\begin{aligned}
&\|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i-1,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j}) + \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i-1,j})\|_{L^p, \mathcal{C}_n} \\
&\leq \|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i-1,n})\|_{L^p, \mathcal{C}_n} \\
&\quad + \|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i-1,j})\|_{L^p, \mathcal{C}_n} \leq 2\delta_{p,n}(k, i, \mathcal{C}_n) \tag{B.6}
\end{aligned}$$

a.s., where the second inequality follows from Lemma A.3.

On the other hand,

$$\begin{aligned}
&\|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i-1,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j}) + \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i-1,j})\|_{L^p, \mathcal{C}_n} \\
&\leq \|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j})\|_{L^p, \mathcal{C}_n} \\
&\quad + \|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i-1,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i-1,j})\|_{L^p, \mathcal{C}_n}. \tag{B.7}
\end{aligned}$$

We note that

$$\|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j})\|_{L^p, \mathcal{C}_n}$$

$$\begin{aligned}
&= \|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{F}_{i,n}) + \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j})\|_{L^p, \mathcal{C}_n} \\
&\leq \|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n} - Y_{k,n,j} | \mathcal{F}_{i,n})\|_{L^p, \mathcal{C}_n} + \|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j})\|_{L^p, \mathcal{C}_n} \\
&\leq \|Y_{k,n} - Y_{k,n,j}\|_{L^p, \mathcal{C}_n} + \|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j})\|_{L^p, \mathcal{C}_n} \\
&= \delta_{p,n}(k, j, \mathcal{C}_n) + \|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j})\|_{L^p, \mathcal{C}_n} \\
&\leq \delta_{p,n}(k, j, \mathcal{C}_n) + \delta_{p,n}(k, j, \mathcal{C}_n) = 2\delta_{p,n}(k, j, \mathcal{C}_n),
\end{aligned} \tag{B.8}$$

where the second inequality follows from the conditional Jansen's inequality. The last inequality in (B.8) holds because (i) $\|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j})\|_{L^p, \mathcal{C}_n} = 0$ for $j > i$ and (ii)

$$\begin{aligned}
&\|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j})\|_{L^p, \mathcal{C}_n} \\
&= \|\mathbb{E}_{\mathcal{C}_n}(F_{k,n}(e_{1,n}, \dots, e_{j,n}^*, \dots, e_{n,n}) | \sigma(e_{1,n}, \dots, e_{i,n})) - \\
&\quad \mathbb{E}_{\mathcal{C}_n}(F_{k,n}(e_{1,n}, \dots, e_{j,n}^*, \dots, e_{n,n}) | \sigma(e_{1,n}, \dots, e_{j,n}^*, \dots, e_{i,n}))\|_{L^p, \mathcal{C}_n} \\
&= \|\mathbb{E}_{\mathcal{C}_n}(F_{k,n}(e_{1,n}, \dots, e_{j,n}^*, \dots, e_{n,n}) | \sigma(e_{1,n}, \dots, e_{j-1,n}, e_{j+1,n}, \dots, e_{i,n})) - \\
&\quad \mathbb{E}_{\mathcal{C}_n}(F_{k,n}(e_{1,n}, \dots, e_{j,n}^*, \dots, e_{n,n}) | \sigma(e_{1,n}, \dots, e_{j,n}^*, \dots, e_{i,n}))\|_{L^p, \mathcal{C}_n} \\
&\leq \delta_{p,n}(k, j, \mathcal{C}_n)
\end{aligned} \tag{B.9}$$

for $j \leq i$, where the first equality is by the definitions of $Y_{k,n,j}$, $\mathcal{F}_{i,n}$, and $\mathcal{G}_{i,j}$, the second equality is from the independence between $e_{j,n}$ and $(e_{1,n}, \dots, e_{j,n}^*, \dots, e_{n,n})$ conditional on \mathcal{C}_n , and the last inequality follows from Lemma A.3 and the fact that $e_{j,n}^*$ is an i.i.d. (conditional on \mathcal{C}_n) copy of $e_{j,n}$.

Replacing the i in Eq.(B.8) by $i - 1$, we obtain

$$\|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i-1,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i-1,j})\|_{L^p, \mathcal{C}_n} \leq 2\delta_{p,n}(k, j, \mathcal{C}_n). \tag{B.10}$$

Thus, plugging Eqs.(B.8) and (B.10) into the right-hand-side of Eq.(B.7), we have

$$\begin{aligned}
&\|\mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n} | \mathcal{F}_{i-1,n}) - \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i,j}) + \mathbb{E}_{\mathcal{C}_n}(Y_{k,n,j} | \mathcal{G}_{i-1,j})\|_{L^p, \mathcal{C}_n} \\
&\leq 4\delta_{p,n}(k, j, \mathcal{C}_n).
\end{aligned} \tag{B.11}$$

Combining Eqs.(B.5), (B.6) and (B.11), we have

$$\begin{aligned}
\delta_{p/2}^*(i, j, \mathcal{C}_n) &\leq 2C \sum_{k=1}^n \min\{2\delta_{p,n}(k, i, \mathcal{C}_n), 4\delta_{p,n}(k, j, \mathcal{C}_n)\} \\
&\leq 8C \sum_{k=1}^n \min\{\delta_{p,n}(k, i, \mathcal{C}_n), \delta_{p,n}(k, j, \mathcal{C}_n)\}.
\end{aligned}$$

Thus, it follows from condition (3) in this theorem that Eq.(B.4) holds. We conclude that Condition (a) is satisfied.

Condition (b): Note that

$$\begin{aligned}
\left\{ \mathbb{E}_{\mathcal{C}_n} \left(\max_{i=1, \dots, n} |X_{i,n}| \right) \right\}^p &\leq \mathbb{E}_{\mathcal{C}_n} \left[\left(\max_{i=1, \dots, n} |X_{i,n}| \right)^p \right] \leq \mathbb{E}_{\mathcal{C}_n} \sum_{i=1}^n |X_{i,n}|^p \\
&= \sigma_n^{-p} \sum_{i=1}^n \left\| \sum_{j=1}^n P_i Y_{j,n} \right\|_{L^p, \mathcal{C}_n}^p \leq \sigma_n^{-p} \sum_{i=1}^n \left[\sum_{j=1}^n \delta_{p,n}(j, i, \mathcal{C}_n) \right]^p \leq \sigma_n^{-p} \sum_{i=1}^n C^p = o_{\mathbb{P}}(1),
\end{aligned}$$

where the first inequality follows from conditional Jensen's inequality, the third one is by Lemma 4.1, the fourth one is by Condition (2) in this theorem, and the last step follows from conditions (1) in this theorem. Then Condition (b) follows from conditional Markov's inequality and Chernozhukov et al. (2018, Lemma 6.1 (a)).

Condition (c): Condition (c) is satisfied because

$$\mathbb{E} \left(\max_{i=1, \dots, n} X_{i,n}^2 \right) \leq \mathbb{E} \left(\sum_{i=1}^n X_{i,n}^2 \right) = \mathbb{E} \left[\mathbb{E}_{\mathcal{C}_n} \left(\sum_{i=1}^n X_{i,n}^2 \right) \right] = \mathbb{E} \left[\sum_{i=1}^n \mathbb{E}_{\mathcal{C}_n} (X_{i,n}^2) \right] = 1,$$

where the last inequality follows from Eq.(B.3). Thus the conclusion follows. \square

Proof of Lemma 4.2: It suffices to show that

$$\sup_{1 \leq i \leq n} \sum_{j=1}^n \sum_{k=1}^n \min \{ \delta_{p,n}(k, i, \mathcal{C}_n), \delta_{p,n}(k, j, \mathcal{C}_n) \} = o(n^{(\min\{2,p/2\}-1)/\min\{2,p/2\}}) \text{ a.s.} \quad (\text{B.12})$$

as $n \rightarrow \infty$, since Eq.(B.12) will imply that

$$\frac{1}{n^{\min\{2,p/2\}}} \sum_{i=1}^n \left\{ \sum_{j=1}^n \sum_{k=1}^n \min \{ \delta_{p,n}(k, i, \mathcal{C}_n), \delta_{p,n}(k, j, \mathcal{C}_n) \} \right\}^{\min\{2,p/2\}} = o(1) \text{ a.s.}$$

In fact, Eq.(B.12) follows from

$$\begin{aligned}
& \sup_{1 \leq i \leq n} \sum_{j=1}^n \sum_{k=1}^n \min \{\delta_{p,n}(k, i, \mathcal{C}_n), \delta_{p,n}(k, j, \mathcal{C}_n)\} \\
&= \sup_{1 \leq i \leq n} \sum_{k=1}^n \sum_{j(k)=1}^n \min \{\delta_{p,n}(k, i, \mathcal{C}_n), \delta_{p,n}(k, j(k), \mathcal{C}_n)\} \\
&\leq \sup_{1 \leq i \leq n} \sum_{k=1}^n \sum_{j(k) < \kappa_n} \min \{\delta_{p,n}(k, i, \mathcal{C}_n), \delta_{p,n}(k, j(k), \mathcal{C}_n)\} \\
&\quad + \sup_{1 \leq i \leq n} \sum_{k=1}^n \sum_{j(k) \geq \kappa_n} \min \{\delta_{p,n}(k, i, \mathcal{C}_n), \delta_{p,n}(k, j(k), \mathcal{C}_n)\} \\
&\leq \sup_{1 \leq i \leq n} \sum_{k=1}^n \sum_{j(k) < \kappa_n} \delta_{p,n}(k, i, \mathcal{C}_n) + \sup_{1 \leq i \leq n} \sum_{k=1}^n \sum_{j(k) \geq \kappa_n} \delta_{p,n}(k, j(k), \mathcal{C}_n) \\
&\leq C\kappa_n + n\tau_n = o(n^{(\min\{2,p/2\}-1)/\min\{2,p/2\}}) \text{ a.s.},
\end{aligned}$$

where the last inequality follows from $\sup_{1 \leq i \leq n} \sum_{k=1}^n \delta_{p,n}(k, i, \mathcal{C}_n) < C < \infty$ a.s. for some constant $C > 0$ and $\sup_{1 \leq k \leq n} \sum_{j \geq \kappa_n} \delta_{p,n}(k, j(k), \mathcal{C}_n) \leq \tau_n$ a.s. \square

C. Verifying the CLT using weights matrices from real data

We verify the conditions in Eqs.(4.7)-(4.8) for four financial networks, showing that the conditions in our central limit theorem can be easily satisfied in real applications. In the following four examples, we all consider listed companies in the A-share market, but adopt different network construction approaches. Data on stocks and fund holdings were collected from the China Stock Market & Accounting Research (CSMAR) database, while fund classification data was sourced from the Wind database.

(1) Fund common ownership network

In the A-share market, the fund common ownership network plays a significant role in understanding the co-movement of stock returns (Barberis et al., 2005, Lou, 2012, Anton and Polk, 2014). In the first real data example, we regard stocks as nodes and establish an edge between two stocks whenever they are commonly held by at least one fund. Moreover, the adjacency matrix $A = (a_{ij})_{n \times n}$ is weighted. We

do not simply let $a_{ij} = 1$ when there is an edge between stock i and j . Instead, we calculate the degree of the pair (i, j) 's common ownership defined as FCAP by [Anton and Polk \(2014\)](#):

$$FCAP_{i,j} = \sum_{f=1}^F \frac{S_i^f P_i + S_j^f P_j}{S_i P_i + S_j P_j},$$

where S_i^f is the number of the shares of stock i held by fund f at the end of the proceeding quarter at price P_i with total shares outstanding of S_i , similarly for stock j , and F is the total number of funds holding stocks i and j commonly in the sample. We let $a_{ij} = FCAP_{i,j}$ if stock i and j are connected. Conventionally, the diagonal elements a_{ii} are zero. Then, the spatial weights matrix W can be obtained by row-normalizing the adjacency matrix A .

We collected data on the listed companies in China's A-share market for the years 2006, 2010, 2014, 2018, and 2022.⁸ The number of listed companies (sample size n) and the mean degree (D) for these years are reported in Table 4. The fund samples included equity funds, equity mixed funds, and flexible allocation funds, with the requirement that the proportion of equity holdings to total net assets is not less than 50%. Structured funds, ETF funds, and index funds were excluded. We can see from Table 5 that condition (4.7) generally holds for any fixed λ . In Table 6, the results for the years 2014, 2018, and 2022, decrease as n increases for any fixed λ , satisfying condition (4.8). It is worth noting that the results for the years 2006, 2010, and 2014 increase with n , which may be due to the sharp rise in the mean degree for these three years. The mean degree in 2010 even reached approximately twice that of 2006. A denser network leads to an increase in this statistic (as mentioned in the simulation studies). If the mean degree remains at a relatively stable level, even with a slight increase, as observed in 2014 and 2018, the statistic would decrease with n , still satisfying condition (4.8).

(2) Return correlation network

In this example, we construct financial networks based on stock return correlations ([Bekaert et al., 2009](#), [Pollet and Wilson, 2010](#), [Lou and Polk, 2022](#)), with edges created between each stock and its top D most correlated stocks ($D = 50, 100$, and 200). Note that the adjacency matrix is also weighted in this example. The weight is the Pearson coefficient between stocks i and j .

⁸Fund holding data is released every quarter. For these five different years, we uniformly select the data from the third quarter for our analysis.

Year	2006	2010	2014	2018	2022
n	1203	1768	2356	3305	4625
D	6.4771	11.8382	17.2929	18.9773	8.6301

Table 4: Sample sizes (n) and mean degrees (D) for the fund common ownership networks in different years.

Eq.(4.7)				
	$n = 1203$	$n = 1768$	$n = 2356$	$n = 3305$
$\lambda = 0.2$	2.1166	3.3045	2.6533	3.5181
$\lambda = 0.3$	2.9547	5.0473	3.9119	5.4597
$\lambda = 0.4$	4.1109	7.4631	5.6647	8.1850
$\lambda = 0.8$	25.5273	45.4683	33.6427	52.7505
				33.2012

Table 5: Verifying Eq.(4.7) for the fund common ownership networks.

Eq.(4.8)				
	$n = 1203$	$n = 1768$	$n = 2356$	$n = 3305$
$\lambda = 0.2$	0.0076	0.0150	0.0242	0.0150
$\lambda = 0.3$	0.0272	0.0658	0.1283	0.0756
$\lambda = 0.4$	0.0879	0.2417	0.5428	0.3044
$\lambda = 0.8$	11.8800	41.8507	135.0150	65.8056
				28.2899

Table 6: Verifying Eq.(4.8) for the fund common ownership networks.

The stocks we select is the same as the first example. The number of stocks (n) corresponding to the five years 2006, 2010, 2014, 2018, and 2022 can be seen in Table 4. Results are reported in Tables 7-8. From these results, we can see that condition (4.7) generally holds for any fixed mean degree D and λ . Condition (4.8) is more plausible to hold with a smaller λ and fail when the dependence factor λ is large (e.g., $\lambda = 0.8$).

(3) Industry network

Industry networks are frequently used, such as in Ahern and Harford (2014), Aobdia et al. (2014), and He et al. (2024), to examine risk propagation within industry-based structures. In this example, the construction of the adjacency ma-

		Eq.(4.7)				
		$n = 1203$	$n = 1768$	$n = 2356$	$n = 3305$	$n = 4625$
$\lambda = 0.2$	$D = 50$	6.8711	12.4716	8.2058	10.2532	23.8892
	$D = 100$	4.3308	7.2482	5.1293	6.8979	13.3177
	$D = 200$	2.8914	4.3921	3.3721	4.5419	7.7220
$\lambda = 0.3$	$D = 50$	10.9222	20.4824	13.1912	17.6514	39.6805
	$D = 100$	6.6305	11.6260	8.0077	11.4696	22.0397
	$D = 200$	4.1966	6.7680	5.0263	7.2144	12.4094
$\lambda = 0.4$	$D = 50$	16.2225	31.0442	19.7145	28.2705	60.2453
	$D = 100$	9.6363	17.4031	11.7931	17.8975	33.3646
	$D = 200$	5.9005	9.9012	7.2048	10.8945	18.5819
$\lambda = 0.8$	$D = 50$	87.5653	176.4463	106.9930	211.5733	344.3621
	$D = 100$	49.8095	96.8896	69.3964	121.7419	190.0621
	$D = 200$	28.5783	52.8307	42.6707	67.1354	104.1565

Table 7: Verifying Eq.(4.7) for the return correlation networks.

trix $A = (a_{ij})_{n \times n}$ is straightforward. We let $a_{ij} = 1$ if companies i and j belong to the same industry, and $a_{ij} = 0$ otherwise. Conventionally, the diagonal elements a_{ii} are zero. Then, the spatial weights matrix W can be obtained by row-normalizing the adjacency matrix A .

Note that this network is block-diagonal. Each block is a complete graph (i.e., any two companies in the same industry are connected). The mean degrees for the years 2006, 2010, 2014, 2018, and 2022 are 38.6799, 60.0964, 79.9318, 122.7469, and 203.3153, respectively, showing a rapid increase over these five years. To investigate the values of Eqs.(4.7)-(4.8) across different sample sizes, we need to control for the same level of mean degree. We conduct random sampling of industries in the year 2022, ensuring that the mean degree of the company network in the sampled industries falls within the range of [190, 210]. Table 9 presents the number of industries and listed companies we have sampled and the corresponding mean degree. Results are collected in Tables 10-11. Controlling the level of mean degrees, conditions (4.7)-(4.8) generally hold for any fixed λ .

(4) City network

		Eq.(4.8)				
		$n = 1203$	$n = 1768$	$n = 2356$	$n = 3305$	$n = 4625$
$\lambda = 0.2$	$D = 50$	0.4118	0.3644	0.2726	0.3077	0.1953
	$D = 100$	0.9174	0.7736	0.6850	0.7508	0.4828
	$D = 200$	2.0947	1.7826	1.6666	1.8175	1.2004
$\lambda = 0.3$	$D = 50$	1.7429	1.6997	1.2717	1.3379	0.9717
	$D = 100$	3.5416	3.2718	2.8959	2.9701	2.1851
	$D = 200$	7.4130	6.8135	6.4070	6.6079	4.9735
$\lambda = 0.4$	$D = 50$	5.9654	6.2886	4.6570	4.5995	3.8287
	$D = 100$	11.0072	11.0875	9.7352	9.4287	7.8859
	$D = 200$	21.3768	21.1529	19.8840	19.5249	16.6010
$\lambda = 0.8$	$D = 50$	552.7762	896.5796	630.1783	450.7894	733.4698
	$D = 100$	808.2361	1242.7139	984.1271	718.3961	1116.2232
	$D = 200$	1278.2769	1818.9832	1567.4807	1225.7036	1811.6045

Table 8: Verifying Eq.(4.8) for the return correlation networks.

Number of industries	10	30	50	70	84
n	777	1777	2889	3684	4625
D	199.9408	200.9781	206.8750	203.6352	203.3153

Table 9: Number of industries, number of listed companies (n) and mean degrees (D) for the sampled industry networks.

		Eq.(4.7)				
		$n = 777$	$n = 1777$	$n = 2889$	$n = 3684$	$n = 4625$
$\lambda = 0.2$		1.2500	1.2500	1.2500	1.2500	1.2500
$\lambda = 0.3$		1.4286	1.4286	1.4286	1.4286	1.4286
$\lambda = 0.4$		1.6667	1.6667	1.6667	1.6667	1.6667
$\lambda = 0.8$		5.0000	5.0000	5.0000	5.0000	5.0000

Table 10: Verifying Eq.(4.7) for the sampled industry networks.

Eq.(4.8)					
	$n = 777$	$n = 1777$	$n = 2889$	$n = 3684$	$n = 4625$
$\lambda = 0.2$	1.4706	0.6852	0.5961	0.4329	0.3225
$\lambda = 0.3$	4.2361	1.9759	1.7280	1.2539	0.9334
$\lambda = 0.4$	10.1440	4.7346	4.1518	3.0114	2.2406
$\lambda = 0.8$	358.8146	167.6468	147.6931	107.0428	79.5906

Table 11: Verifying Eq.(4.8) for the sampled industry networks.

Number of cities	20	120	240	300	379
n	298	1545	2771	3525	4625
D	131.8993	137.2634	141.0220	140.1214	141.5834

Table 12: Number of cities, number of listed companies (n) and mean degrees (D) for the sampled city networks.

Eq.(4.7)					
	$n = 298$	$n = 1545$	$n = 2771$	$n = 3525$	$n = 4625$
$\lambda = 0.2$	1.2500	1.2500	1.2500	1.2500	1.2500
$\lambda = 0.3$	1.4286	1.4286	1.4286	1.4286	1.4286
$\lambda = 0.4$	1.6667	1.6667	1.6667	1.6667	1.6667
$\lambda = 0.8$	5.0000	5.0000	5.0000	5.0000	5.0000

Table 13: Verifying Eq.(4.7) for the sampled city networks.

Geographic proximity has been a common method for modeling inter-company relations (Holmes, 2005, Defever, 2006, Alcácer and Delgado, 2016). In this example, we let $a_{ij} = 1$ if the headquarters of companies i and j locate in the same city, and $a_{ij} = 0$ otherwise. Conventionally, the diagonal elements a_{ii} are zero. Then, the spatial weights matrix W can be obtained by row-normalizing the adjacency matrix A .

Similar with the industry network, the city network is block-diagonal with each block being a complete graph and the mean degrees increase rapidly over the five years. Thus, we also conduct random sampling of cities in the year 2022 ($D = 141.5834$), ensuring that the mean degree of the company network in the sampled cities falls

Eq.(4.8)					
	$n = 298$	$n = 1545$	$n = 2771$	$n = 3525$	$n = 4625$
$\lambda = 0.2$	1.8940	0.6301	0.3816	0.2793	0.2123
$\lambda = 0.3$	5.4156	1.8216	1.1041	0.8075	0.6137
$\lambda = 0.4$	12.9189	4.3705	2.6505	1.9373	1.4723
$\lambda = 0.8$	454.0162	155.1216	94.1475	68.7506	52.2488

Table 14: Verifying Eq.(4.8) for the sampled city networks.

within the range of [130, 150]. Table 12 presents the number of cities and listed companies we have sampled and the corresponding mean degree. Results are collected in Tables 13-14. Conditions (4.7)-(4.8) generally hold for any fixed λ after controlling the level of mean degrees.

References

- ABBE, EMMANUEL (2018): “Community detection and stochastic block models: recent developments,” *Journal of Machine Learning Research*, 18 (177), 1–86.
- ACEMOGLU, DARON, ASUMAN OZDAGLAR, AND ALIREZA TAHBAZ-SALEHI (2015): “Systemic risk and stability in financial networks,” *American Economic Review*, 105 (2), 564–608.
- AHERN, KENNETH R AND JARRAD HARFORD (2014): “The importance of industry links in merger waves,” *The Journal of Finance*, 69 (2), 527–576.
- ALCÁCER, JUAN AND MERCEDES DELGADO (2016): “Spatial organization of firms and location choices through the value chain,” *Management Science*, 62 (11), 3213–3234.
- ANGRIST, JOSHUA D AND KEVIN LANG (2004): “Does school integration generate peer effects? Evidence from Boston’s Metco Program,” *American Economic Review*, 94 (5), 1613–1634.
- ANTON, MIGUEL AND CHRISTOPHER POLK (2014): “Connected stocks,” *The Journal of Finance*, 69 (3), 1099–1127.

- AOBDIA, DANIEL, JUDSON CASKEY, AND N BUGRA OZEL (2014): “Inter-industry network structure and the cross-predictability of earnings and stock returns,” *Review of Accounting Studies*, 19 (3), 1191–1224.
- ATHREYA, AVANTI, DONNIELL E. FISHKIND, MINH TANG, CAREY E. PRIEBE, YOUNGSE R PARK, JOSHUA T. VOGELSTEIN, KEITH LEVIN, VINCE LYZINSKI, YICHEN QIN, AND DANIEL L. SUSSMAN (2018): “Statistical Inference on Random Dot Product Graphs: A Survey,” *Journal of Machine Learning Research*, 18 (226), 1–92.
- BARBERIS, NICHOLAS, ANDREI SHLEIFER, AND JEFFREY WURGLER (2005): “Comovement,” *Journal of Financial Economics*, 75 (2), 283–317.
- BEKAERT, GEERT, ROBERT J HODRICK, AND XIAOYAN ZHANG (2009): “International stock return comovements,” *The Journal of Finance*, 64 (6), 2591–2626.
- BERNSTEIN, DENNIS S. (2009): *Matrix Mathematics: Theory, Facts, and Formulas (Second Edition)*, Princeton University Press, Princeton, NJ.
- BURKHOLDER, DONALD L (1988): “Sharp inequalities for martingales and stochastic integrals,” *Astérisque*, 157 (158), 75–94.
- CHERNOZHUKOV, VICTOR, DENIS CHETVERIKOV, MERT DEMIRER, ESTHER DU-FLO, CHRISTIAN HANSEN, WHITNEY NEWHEY, AND JAMES ROBINS (2018): “Double/Debiased Machine Learning for Treatment and Structural Parameters,” *The Econometrics Journal*, 21 (1), C1–C68.
- CHOW, YUAN SHIH AND HENRY TEICHER (2003): *Probability Theory: Independence, Interchangeability, Martingales*, New York: Springer-Verlag, third ed.
- DAVIDSON, JAMES (2021): *Stochastic Limit Theory: An Introduction for Econometricians*, Advanced Texts in Econometrics, Oxford: Oxford University Press, second edition ed.
- DEFEVER, FABRICE (2006): “Functional fragmentation and the location of multinational firms in the enlarged Europe,” *Regional Science and Urban Economics*, 36 (5), 658–677.

- DOUKHAN, PAUL AND SANA LOUHICHI (1999): “A new weak dependence condition and applications to moment inequalities,” *Stochastic Processes and their Applications*, 84 (2), 313–342.
- EL MACHKOURI, MOHAMED, DALIBOR VOLNÝ, AND WEI BIAO WU (2013): “A central limit theorem for stationary random fields,” *Stochastic Processes and their Applications*, 123 (1), 1–14.
- GAO, MENGSI AND PENG DING (2025): “Causal Inference in Network Experiments: Regression-based Analysis and Design-Based Properties,” *Journal of Econometrics*, 252, 106119.
- GRAHAM, BRYAN S (2008): “Identifying social interactions through conditional variance restrictions,” *Econometrica*, 76 (3), 643–660.
- HALL, PETER AND C. C. HEYDE (1980): *Martingale Limit Theory and Its Application*, Probability and Mathematical Statistics, New York: Academic Press.
- HAN, XIAOYI, YILAN XU, LINLIN FAN, YI HUANG, MINHONG XU, AND SONG GAO (2021): “Quantifying COVID-19 importation risk in a dynamic network of domestic cities and international countries,” *Proceedings of the National Academy of Sciences*, 118 (31), e2100201118.
- HE, JIE, LANTIAN LIANG, HUI WANG, AND HAN XIA (2024): “Networking behind the scenes: Institutional cross-industry holdings and corporate loan markets,” *Management Science*, 70 (7), 4932–4952.
- HOFF, PETER D., ADRIAN E. RAFTERY, AND MARK S. HANDCOCK (2002): “Latent Space Approaches to Social Network Analysis,” *Journal of the American Statistical Association*, 97 (460), 1090–1098.
- HOLMES, THOMAS J (2005): “The location of sales offices and the attraction of cities,” *Journal of Political Economy*, 113 (3), 551–581.
- JENISH, NAZGUL AND INGMAR R PRUCHA (2009): “Central limit theorems and uniform laws of large numbers for arrays of random fields,” *Journal of Econometrics*, 150 (1), 86–98.

- (2012): “On spatial processes and asymptotic inference under near-epoch dependence,” *Journal of Econometrics*, 170 (1), 178–190.
- JIA, JAYSON S, XIN LU, YUN YUAN, GE XU, JIANMIN JIA, AND NICHOLAS A CHRISTAKIS (2020): “Population flow drives spatio-temporal distribution of COVID-19 in China,” *Nature*, 582 (7812), 389–394.
- KELEJIAN, HARRY H. AND INGMAR R. PRUCHA (2001): “On the asymptotic distribution of the Moran I test statistic with applications,” *Journal of Econometrics*, 104 (2), 219–257.
- KOJEVNIKOV, DENIS (2021): “The Bootstrap for Network Dependent Processes,” arXiv:2101.12312.
- KOJEVNIKOV, DENIS, VADIM MARMER, AND KYUNGCHUL SONG (2021): “Limit theorems for network dependent random variables,” *Journal of Econometrics*, 222 (2), 882–908.
- KRAEMER, MORITZ UG, CHIA-HUNG YANG, BERNARDO GUTIERREZ, CHIEH-HSI WU, BRENNAN KLEIN, DAVID M PIGOTT, OPEN COVID-19 DATA WORKING GROUP, LOUIS DU PLESSIS, NUNO R FARIA, RUORAN LI, ET AL. (2020): “The effect of human mobility and control measures on the COVID-19 epidemic in China,” *Science*, 368 (6490), 493–497.
- KUERSTEINER, GUIDO M (2019): “Limit theorems for data with network structure,” arXiv:1908.02375.
- KUERSTEINER, GUIDO M. AND INGMAR R. PRUCHA (2013): “Limit theory for panel data models with cross sectional dependence and sequential exogeneity,” *Journal of Econometrics*, 174 (2), 107–126.
- (2020): “Dynamic Spatial Panel Models: Networks, Common Shocks, and Sequential Exogeneity,” *Econometrica*, 88 (5), 2109–2146.
- LEE, LUNG-FEI (2004): “Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models,” *Econometrica*, 72 (6), 1899–1925.
- (2007): “GMM and 2SLS estimation of mixed regressive, spatial autoregressive models,” *Journal of Econometrics*, 137 (2), 489–514.

- LEE, LUNG-FEI, CHAO YANG, AND JIHAI YU (2022): “QML and efficient GMM estimation of spatial autoregressive models with dominant (popular) units,” *Journal of Business & Economic Statistics*, 1–13.
- LEUNG, MICHAEL AND HYUNGSIK ROGER MOON (2025): “Normal Approximation in Large Network Models,” arXiv:1904.11060.
- LEUNG, MICHAEL P. (2019): “A weak law for moments of pairwise stable networks,” *Journal of Econometrics*, 210 (2), 310–326.
- (2022): “Causal Inference Under Approximate Neighborhood Interference,” *Econometrica*, 90 (1), 267–293.
- LIU, TUO, XINGBAI XU, LUNG-FEI LEE, AND YINGDAN MEI (2025): “Robust Estimation for the Spatial Autoregressive Model,” *Working Paper*.
- LOU, DONG (2012): “A flow-based explanation for return predictability,” *The Review of Financial Studies*, 25 (12), 3457–3489.
- LOU, DONG AND CHRISTOPHER POLK (2022): “Comomentum: Inferring arbitrage activity from return correlations,” *The Review of Financial Studies*, 35 (7), 3272–3302.
- POLLET, JOSHUA M AND MUNGO WILSON (2010): “Average correlation and stock market returns,” *Journal of Financial Economics*, 96 (3), 364–380.
- QU, XI AND LUNG-FEI LEE (2015): “Estimating a spatial autoregressive model with an endogenous spatial weight matrix,” *Journal of Econometrics*, 184 (2), 209–232.
- RIO, EMMANUEL (2009): “Moment inequalities for sums of dependent random variables under projective conditions,” *Journal of Theoretical Probability*, 22 (1), 146–163.
- SMITH, ANNA L., DENA M. ASTA, AND CATHERINE A. CALDER (2019): “The Geometry of Continuous Latent Space Models for Network Data,” *Statistical Science*, 34 (3), 428–453.
- WAINWRIGHT, MARTIN J. (2019): *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*, Cambridge University Press, 1 ed.

- WU, SHIAO, ZHE LI, AND XUENING ZHU (2023): “A distributed community detection algorithm for large scale networks under stochastic block models,” *Computational Statistics & Data Analysis*, 187, 107794.
- WU, WEI BIAO (2005): “Nonlinear system theory: Another look at dependence,” *Proceedings of the National Academy of Sciences*, 102 (40), 14150–14154.
- WU, WEI BIAO AND YING NIAN WU (2016): “Performance bounds for parameter estimates of high-dimensional linear models with correlated errors,” *Electronic Journal of Statistics*, 10 (1), 352 – 379.
- WU, ZEQI, WEN JIANG, AND XINGBAI XU (2024): “Applications of Functional Dependence to Spatial Econometrics,” *Econometric Theory*, 1–36.
- XU, XINGBAI AND LUNG-FEI LEE (2015): “Maximum likelihood estimation of a spatial autoregressive Tobit model,” *Journal of Econometrics*, 188 (1), 264–280.
- (2018): “Sieve maximum likelihood estimation of the spatial autoregressive Tobit model,” *Journal of Econometrics*, 203 (1), 96–112.
- (2024): “Conditions for α -Mixing Linear Random Fields: with Applications to Spatial Econometric Models,” *China Journal of Econometrics*, 4 (1), 26–57.
- XU, XIU, WEINING WANG, YONGCHEOL SHIN, AND CHAOWEN ZHENG (2022): “Dynamic Network Quantile Regression Model,” *Journal of Business & Economic Statistics*, 1–15.
- YU, JIHAI, ROBERT DE JONG, AND LUNG-FEI LEE (2008): “Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large,” *Journal of Econometrics*, 146 (1), 118–134.
- YUAN, ZIHAO AND MARTIN SPINDLER (2025): “Bernstein-Type Inequalities and Nonparametric Estimation under near-epoch Dependence,” *Journal of Econometrics*, 251, 106054.
- ZHU, XUENING, WEINING WANG, HANSHENG WANG, AND WOLFGANG KARL HÄRDLE (2019): “Network quantile autoregression,” *Journal of Econometrics*, 212 (1), 345–358.

Supplement Appendix

This Supplement Appendix contains the proofs of Lemmas 4.3-4.4 and A.1-A.3, as well as the proofs for Section 3 and 5.

D. Proofs for Lemmas A.1-A.3

Proof of Lemma A.1. When $p > 2$, the conclusion follows from Theorem 2.1 in Rio (2009). When $p = 2$, since $\mathbb{E}_{\mathcal{C}}(X_i, X_j) = 0$ for all $i \neq j$, $\|\sum_{i=1}^n X_i\|_{L^2, \mathcal{C}} = \left(\sum_{i=1}^n \|X_i\|_{L^p, \mathcal{C}}^2\right)^{1/2}$.

When $1 < p < 2$, by Theorem 3.1 in Burkholder (1988), we have

$$\left\| \sum_{i=1}^n X_i \right\|_{L^p, \mathcal{C}} \leq C_p \left\| \left(\sum_{i=1}^n X_i^2 \right)^{1/2} \right\|_{L^p, \mathcal{C}}$$

a.s. Notice that for any non-negative real numbers a_1, \dots, a_n and $q > r > 1$, we have $(\sum_{i=1}^n a_i^r)^{1/r} \geq (\sum_{i=1}^n a_i^q)^{1/q}$ (Proposition 9.1.5 in Bernstein (2009), p.599). Thus, from the above two inequalities, we have

$$\begin{aligned} \left\| \sum_{i=1}^n X_i \right\|_{L^p, \mathcal{C}} &\leq C_p \left\| \left(\sum_{i=1}^n X_i^2 \right)^{1/2} \right\|_{L^p, \mathcal{C}} \leq C_p \left\| \left(\sum_{i=1}^n |X_i|^p \right)^{1/p} \right\|_{L^p, \mathcal{C}} \\ &= C_p \left(\mathbb{E}_{\mathcal{C}} \sum_{i=1}^n |X_i|^p \right)^{1/p} = C_p \left(\sum_{i=1}^n \|X_i\|_{L^p, \mathcal{C}}^p \right)^{1/p} \end{aligned}$$

a.s. \square

Proof of Lemma A.2. See Theorem 3.2 in Hall and Heyde (1980) and the Remarks in Hall and Heyde (1980, p.59). \square

Proof of Lemma A.3. Recall that $Y_{j,n,i}$ is the coupled version of $Y_{j,n}$ with $e_{i,n}$ replaced by its i.i.d. copy $e_{i,n}^*$. Then,

$$\begin{aligned} &\|\mathbb{E}_{\mathcal{C}_n}(Y_{j,n}|\sigma(e_{k,n} : k \in \mathcal{I})) - \mathbb{E}_{\mathcal{C}_n}(Y_{j,n}|\sigma(e_{k,n} : k \in \mathcal{I} \setminus \{i\}))\|_{L^p, \mathcal{C}_n} \\ &= \|\mathbb{E}_{\mathcal{C}_n}(Y_{j,n}|\sigma(e_{k,n} : k \in \mathcal{I})) - \mathbb{E}_{\mathcal{C}_n}(Y_{j,n,i}|\sigma(e_{k,n} : k \in \mathcal{I} \setminus \{i\}))\|_{L^p, \mathcal{C}_n} \\ &= \|\mathbb{E}_{\mathcal{C}_n}(Y_{j,n}|\sigma(e_{k,n} : k \in \mathcal{I})) - \mathbb{E}_{\mathcal{C}_n}(Y_{j,n,i}|\sigma(e_{k,n} : k \in \mathcal{I}))\|_{L^p, \mathcal{C}_n} \end{aligned}$$

$$= \|\mathbb{E}_{\mathcal{C}_n} (Y_{j,n} - Y_{j,n,i} | \sigma(e_{k,n} : k \in \mathcal{I}))\|_{L^p, \mathcal{C}_n} \leq \|Y_{j,n} - Y_{j,n,i}\|_{L^p, \mathcal{C}_n} = \delta_{p,n}(j, i, \mathcal{C}_n),$$

where the first equality follows from the fact that $e_{i,n}^*$ is an i.i.d. (conditional on \mathcal{C}_n) copy of $e_{i,n}$, the second equality follows from Corollary 2 in Chow and Teicher (2003, p.231), and the inequality follows from conditional Jensen's inequality. \square

E. Proofs for Section 3

Proof of Proposition 3.1. In this proof, for any vector or matrix $A = (a_{ij})_{n \times m}$, we denote $|A| \equiv (\|a_{ij}\|)_{n \times m}$. Direct calculations show that $|A + B| \leq^* |A| + |B|$ and $|AB| \leq^* |A||B|$, where $A = (a_{ij})_{m \times n} \leq^* B = (b_{ij})_{m \times n}$ means $\forall i, j : a_{ij} \leq b_{ij}$. To shorten formulas, denote $e_{i,n} \equiv X'_{i,n}\beta + \epsilon_{i,n}$ and $e_n \equiv X_n\beta + \epsilon_n$. Denote the solution of (3.1) as $Y_n(e_n)$. Then $Y_n(e_n) = F(\lambda W_n Y_n(e_n) + e_n)$. For any two realizations of e_n , denoted by $e_n^{(1)}$ and $e_n^{(2)}$, consider $Y_n^{(1)} = Y_n(e_n^{(1)})$ and $Y_n^{(2)} = Y_n(e_n^{(2)})$. So, for any $1 \leq i \leq n$,

$$\begin{aligned} |Y_{j,n}^{(1)} - Y_{j,n}^{(2)}| &= \left| F\left(\lambda w_{j,n} Y_n^{(1)} + e_{j,n}^{(1)}\right) - F\left(\lambda w_{j,n} Y_n^{(2)} + e_{j,n}^{(2)}\right) \right| \\ &\leq L |\lambda| \sum_{i=1}^n |w_{ji,n}| \left| Y_{i,n}^{(1)} - Y_{i,n}^{(2)} \right| + L \left| e_{j,n}^{(1)} - e_{j,n}^{(2)} \right|. \end{aligned}$$

The above inequality can be written in a matrix form:

$$(I_n - L |\lambda| |W_n|) \left| Y_n^{(1)} - Y_n^{(2)} \right| \leq^* L \left| e_n^{(1)} - e_n^{(2)} \right|. \quad (\text{E.1})$$

Since $S_n^+ = L(I_n - L |\lambda W_n|)^{-1} = L \sum_{l=0}^{\infty} (L |\lambda W_n|)^l$, all entries of S_n^+ are nonnegative. As a result, we can multiply $\frac{1}{L} S_n^+$ on both sides of (E.1): $\left| Y_n^{(1)} - Y_n^{(2)} \right| \leq^* S_n^+ \left| e_n^{(1)} - e_n^{(2)} \right|$. So, for any $1 \leq i \leq n$,

$$\left| Y_{j,n}^{(1)} - Y_{j,n}^{(2)} \right| \leq \sum_{i=1}^n S_{ji,n}^+ \left| e_{i,n}^{(1)} - e_{i,n}^{(2)} \right|. \quad (\text{E.2})$$

Denote $Y_{j,n,i}$ as the coupled version of $Y_{j,n}$ with $e_{i,n}$ replaced by its i.i.d. copy $e_{i,n}^* \equiv X'_{i,n}\beta + \epsilon_{i,n}^*$ (conditional on \mathcal{C}_n). By Eq.(E.2),

$$\delta_{p,n}(j, i, \mathcal{C}_n) = \|Y_{j,n} - Y_{j,n,i}\|_{L^p, \mathcal{C}_n} \leq S_{ji,n}^+ \|e_{i,n} - e_{i,n}^*\|_{L^p, \mathcal{C}_n}$$

$$= S_{ji,n}^+ \|X'_{i,n}\beta + \epsilon_{i,n} - X'_{i,n}\beta - \epsilon_{i,n}^*\|_{L^p, \mathcal{C}_n} = S_{ji,n}^+ \|\epsilon_{i,n} - \epsilon_{i,n}^*\|_{L^p, \mathcal{C}_n} \leq 2 \|\epsilon\|_{L^p, \mathcal{C}_n} S_{ji,n}^+.$$

□

Proof of Proposition 3.2. This proof is based on the proof of Proposition 4.3 in Wu et al. (2024). Under Assumption 3(i), by Eq.(S.29) in the Supplementary Material of Wu et al. (2024), we have

$$\sup_{n,j} \sum_{i \in \{1, \dots, n\}: d_{ji} \geq s} S_{ji,n}^+ \leq C_1 \zeta^{s/\bar{d}_0}$$

for all $s \in [0, \infty)$, where $C_1 > 0$ is a finite constant. Thus,

$$S_{ji,n}^+ \leq C_1 \zeta^{d_{ji}/\bar{d}_0}$$

for all $j, i \in \{1, \dots, n\}$.

Under Assumption 3(ii), by Eq.(S.33) in the Supplementary Material of Wu et al. (2024), we have

$$\sup_{n,j} \sum_{i \in \{1, \dots, n\}: d_{ji} \geq s} S_{ji,n}^+ \leq C_2 s^{-(\alpha-d)} (\log(2s))^{\alpha-d}$$

for all $s \in [1, \infty)$, where $C_2 > 0$ is a finite constant. Thus,

$$S_{ji,n}^+ \leq C_2 d_{ji}^{-(\alpha-d)} (\log(2d_{ji}))^{\alpha-d}$$

for all $j, i \in \{1, \dots, n\}$. □

Proof of Proposition 3.3. Since $\zeta = \sup_n \|W_n\|_\infty L |\lambda| = L |\lambda|$ and $|W_n| = W_n$, by Neumann's expansion,

$$S_n^+ = (S_{ji,n}^+)_n \times n = L (I_n - L |\lambda W_n|)^{-1} = L \sum_{\ell=0}^{\infty} \zeta^\ell W_n^\ell.$$

We now consider any $i_0 \neq j \in D_n$. Let $k = d_{i_0 j} \geq 1$, then

$$S_{i_0 j, n}^+ = L \sum_{\ell=1}^{\infty} \zeta^\ell \sum_{i_1, \dots, i_{\ell-1}=1}^n w_{i_0 i_1, n} w_{i_1 i_2, n} \cdots w_{i_{\ell-1} j, n}$$

$$= L \sum_{\ell=k}^{\infty} \zeta^{\ell} \sum_{i_1, \dots, i_{\ell-1}=1}^n w_{i_0 i_1, n} w_{i_1 i_2, n} \cdots w_{i_{\ell-1} j, n} \leq L \sum_{\ell=k}^{\infty} \zeta^{\ell} = \frac{L \zeta^k}{1 - \zeta} = \frac{L}{1 - \zeta} \zeta^{d_{i_0 j}},$$

where the second equality follows from $\sum_{i_1, \dots, i_{\ell-1}=1}^n w_{i_0 i_1, n} w_{i_1 i_2, n} \cdots w_{i_{\ell-1} j, n} = 0$ for $\ell < k = d_{i_0 j}$ and the inequality follows from $\sum_{i_1, \dots, i_{\ell-1}=1}^n w_{i_0 i_1, n} w_{i_1 i_2, n} \cdots w_{i_{\ell-1} j, n} \leq \|W_n^\ell\|_\infty \leq \|W_n\|_\infty^\ell = 1$ as W_n is row-normalized. For any $i \in D_n$, we have

$$S_{ii, n}^+ \leq \|S_n^+\|_\infty \leq L \sum_{\ell=0}^{\infty} \zeta^{\ell} \|W_n^\ell\|_\infty \leq L \sum_{\ell=0}^{\infty} \zeta^{\ell} = \frac{L}{1 - \zeta} = \frac{L}{1 - \zeta} \zeta^{d_{ii}}.$$

Then the proof is completed. \square

F. Proofs for Lemmas 4.3-4.4

Proof of Lemma 4.3. Let $\kappa = \frac{2}{3}\alpha$. Then $\alpha - d > \kappa > 2d$ because of $\alpha > 3d$. By Proposition 3.2, we always have $S_{ji, n}^+ \leq C d_{ij}^{-\kappa}$ for some finite constant $C > 0$. Then for all $i \in \{1, \dots, n\}$ and $n \geq 1$, we have

$$\begin{aligned} & \sum_{j=1}^n \sum_{k=1}^n \min \{S_{ki, n}^+, S_{kj, n}^+\} = \sum_{k=1}^n \sum_{j=1}^n \min \{S_{ki, n}^+, S_{kj, n}^+\} \\ &= \sum_{s=1}^{\infty} \sum_{k: s \leq d_{ik} < s+1} \sum_{t=1}^{\infty} \sum_{j: t \leq d_{kj} < t+1} \min \{S_{ki, n}^+, S_{kj, n}^+\} \\ &\leq C \sum_{s=1}^{\infty} \sum_{k: s \leq d_{ik} < s+1} \sum_{t=1}^{\infty} \sum_{j: t \leq d_{kj} < t+1} \min \{d_{ik}^{-\kappa}, d_{kj}^{-\kappa}\} \\ &\leq C^2 \sum_{s=1}^{\infty} \sum_{k: s \leq d_{ki} < s+1} \sum_{t=1}^{\infty} t^{d-1} \min \{d_{ik}^{-\kappa}, t^{-\kappa}\} \\ &= C^2 \sum_{t=1}^{\infty} t^{d-1} \sum_{s=1}^{\infty} \sum_{k: s \leq d_{ki} < s+1} \min \{d_{ik}^{-\kappa}, t^{-\kappa}\} \leq C^3 \sum_{t=1}^{\infty} t^{d-1} \sum_{s=1}^{\infty} s^{d-1} \min \{s^{-\kappa}, t^{-\kappa}\} \\ &= C^3 \sum_{t=1}^{\infty} t^{d-1} \sum_{s=1}^{\infty} s^{d-1} \{s^{-\kappa} 1(s > t) + t^{-\kappa} 1(s \leq t)\} \\ &= C^3 \sum_{t=1}^{\infty} t^{d-1} \left\{ \sum_{s=t+1}^{\infty} s^{d-1-\kappa} + \sum_{s=1}^t s^{d-1} t^{-\kappa} \right\} = C^3 \sum_{t=1}^{\infty} t^{d-1} \{O(t^{d-\kappa}) + O(t^{d-\kappa})\} < \infty, \end{aligned}$$

where the last equality follows from Lemma S.12 in the Supplementary Material in [Wu et al. \(2024\)](#) and we have used the fact that $\sup_{j \in \mathbb{R}^d} |\{i \in \mathbb{R}^d : m \leq d_{ji} < m+1\}| \leq Cm^{d-1}$ (Lemma A.1 in [Jenish and Prucha, 2009](#)) in the last two inequalities. Thus Eq.(4.9) holds. The second conclusion is straightforward and we omit its proof. \square

Proof of Lemma 4.4. Eq.(4.10) follows from

$$\begin{aligned}
& \sup_{i,n} \sum_{k=1}^n \sum_{j=1}^n \min \{S_{ki,n}^+, S_{kj,n}^+\} = \sup_{i,n} \sum_{t=0}^{\infty} \sum_{k \in \mathcal{N}_n^{\partial}(i;t)}^n \sum_{\ell=0}^{\infty} \min \{S_{ki,n}^+, S_{kj,n}^+\} \\
& \leq \frac{L}{1-\zeta} \sup_{i,n} \sum_{t=0}^{\infty} \sum_{k \in \mathcal{N}_n^{\partial}(i;t)}^n \sum_{\ell=0}^{\infty} |\mathcal{N}_n^{\partial}(k;\ell)| \min \{S_{ki,n}^+, \zeta^\ell\} \\
& \leq \frac{LC}{1-\zeta} \sup_{i,n} \sum_{t=0}^{\infty} \sum_{k \in \mathcal{N}_n^{\partial}(i;t)}^n \sum_{\ell=0}^{\infty} \eta^\ell \min \{S_{ki,n}^+, \zeta^\ell\} \\
& \leq \frac{L^2 C}{(1-\zeta)^2} \sup_{i,n} \sum_{\ell=0}^{\infty} \eta^\ell \sum_{t=0}^{\infty} |\mathcal{N}_n^{\partial}(i;t)| \min \{\zeta^t, \zeta^\ell\} \leq \frac{L^2 C^2}{(1-\zeta)^2} \sum_{\ell=0}^{\infty} \eta^\ell \sum_{t=0}^{\infty} \eta^t \min \{\zeta^t, \zeta^\ell\} \\
& = \frac{L^2 C^2}{(1-\zeta)^2} \sum_{\ell=0}^{\infty} \eta^\ell \left\{ \sum_{t=0}^{\ell} \eta^t \zeta^\ell + \sum_{t=\ell+1}^{\infty} \eta^t \zeta^t \right\} = \frac{L^2 C^2}{(1-\zeta)^2} \sum_{\ell=0}^{\infty} \eta^\ell \left\{ \zeta^\ell \frac{\eta^{\ell+1}-1}{\eta-1} + \frac{(\eta\zeta)^{\ell+1}}{1-\eta\zeta} \right\} < \infty,
\end{aligned}$$

where the first and third inequalities follow from Proposition 3.3, the second and the fourth inequalities follow from $\sup_{i,n} |\mathcal{N}_n^{\partial}(i;\ell)| \leq C\eta^\ell$, and the last step follows from $0 < \eta^2\zeta < 1$. The second conclusion is straightforward and we omit its proof. \square

G. Proofs for Section 5

Proof of Proposition 5.1. By Eq.(5.1) and $B_{j,n}(y, y^\bullet) \leq C$, we have

$$\begin{aligned}
& \delta_{p,n}^{(Z)}(j, i, \mathcal{C}_n) = \|H_{j,n}(Y_{j,n}) - H_{j,n}(Y_{j,n,i})\|_{L^p, \mathcal{C}_n} \leq \|B_{j,n}(Y_{j,n}, Y_{j,n,i})\|_{L^p, \mathcal{C}_n} |Y_{j,n} - Y_{j,n,i}| \\
& \leq C \|Y_{j,n} - Y_{j,n,i}\|_{L^p, \mathcal{C}_n} = C \delta_{p,n}^{(Y)}(j, i, \mathcal{C}_n)
\end{aligned}$$

a.s. \square

Proof of Proposition 5.2. By Eq.(5.1) and $B_{j,n}(y, y^\bullet) \leq C_1(|y|^a + |y^\bullet|^a + 1)$, we have

$$\delta_{p,n}^{(Z)}(j, i, \mathcal{C}_n) = \|H_{j,n}(Y_{j,n}) - H_{j,n}(Y_{j,n,i})\|_{L^p, \mathcal{C}_n}$$

$$\begin{aligned} &\leq C_1 \|(|Y_{j,n}|^a + |Y_{j,n,i}|^a + 1) \cdot |Y_{j,n} - Y_{j,n,i}| \|_{L^p, \mathcal{C}_n} \\ &\leq C_1 \| |Y_{j,n}|^a + |Y_{j,n,i}|^a + 1 \|_{L^r, \mathcal{C}_n} \|Y_{j,n} - Y_{j,n,i}\|_{L^q, \mathcal{C}_n} \leq C_1 (2 \|Y\|_{L^{ar}, \mathcal{C}_n}^a + 1) \delta_{q,n}^{(Y)}(j, i, \mathcal{C}_n), \end{aligned}$$

where the second inequality follows from conditional generalized Hölder's inequality (because $p^{-1} = q^{-1} + r^{-1}$), and the third inequality follows from conditional Minkowski's inequality. \square

Proof of Proposition 5.3. Denote $Z_{j,n,i} \equiv H_{j,n}(Y_{j,n,i})$, $B \equiv |Y_{j,n}|^a + |Y_{j,n,i}|^a + 1$, $\rho \equiv |Y_{j,n} - Y_{j,n,i}|$, $r = q/(a+1) > p$. By conditional Lyapunov's inequality and conditional Minkowski's inequality, we have

$$\|B\|_{L^{p/(p-1)}, \mathcal{C}_n} \leq \|B\|_{L^{q/a}, \mathcal{C}_n} \leq \|Y_{j,n}^a\|_{L^{q/a}, \mathcal{C}_n} + \||Y_{j,n}^*|^a\|_{L^{q/a}, \mathcal{C}_n} + 1 \leq 2 \|Y\|_{L^q, \mathcal{C}_n}^a + 1 < \infty$$

a.s., and

$$\|\rho\|_{L^q, \mathcal{C}_n} \leq 2 \|Y_{j,n}\|_{L^q, \mathcal{C}_n} < 2 \|Y\|_{L^q, \mathcal{C}_n} < \infty \text{ a.s.}$$

Using the above two results, by conditional generalized Hölder's inequality ($\frac{1}{r} = \frac{a+1}{q} = \frac{a}{q} + \frac{1}{q}$), we have

$$\|B\rho\|_{L^r, \mathcal{C}_n} \leq \|B\|_{L^{q/a}, \mathcal{C}_n} \|\rho\|_{L^q, \mathcal{C}_n} \leq \left(2 \|Y\|_{L^q, \mathcal{C}_n}^a + 1\right) 2 \|Y\|_{L^q, \mathcal{C}_n} < \infty \text{ a.s.}$$

Thus, by Lemma S.7 in the Supplementary Material of Wu et al. (2024), for all i, j , and $n \geq 1$, it holds that

$$\begin{aligned} \delta_{p,n}^{(Z)}(j, i, \mathcal{C}_n) &= \|Z_{j,n} - Z_{j,n,i}\|_{L^p, \mathcal{C}_n} \leq C_1 \|B\rho\|_{L^p, \mathcal{C}_n} \\ &\leq 2C_1 \left(\|B\|_{L^{p/(p-1)}, \mathcal{C}_n}^{r-p} \|\rho\|_{L^p, \mathcal{C}_n}^{r-p} \|B\rho\|_{L^r, \mathcal{C}_n}^{(p-1)r} \right)^{1/(pr-p)} \\ &\leq C_2(\mathcal{C}_n) \left(\|Y_{j,n} - Y_{j,n,i}\|_{L^p, \mathcal{C}_n} \right)^{(q-ap-p)/(pq-ap-p)} \\ &= C_2(\mathcal{C}_n) [\delta_{p,n}^{(Y)}(j, i, \mathcal{C}_n)]^{(q-ap-p)/(pq-ap-p)}, \end{aligned}$$

where $C_2(\mathcal{C}_n) < \infty$ a.s. \square

Proof of Proposition 5.4. For any $\epsilon > 0$, let $B = \{|Y_{j,n}| < \epsilon, |Y_{j,n,i}| < \epsilon\}$. Since

$$|\mathbf{1}(x_1 > 0) - \mathbf{1}(x_1 > 0)| \leq \frac{|x_1 - x_2|}{\epsilon} \mathbf{1}(|x_1| \geq \epsilon \text{ or } |x_2| \geq \epsilon) + \mathbf{1}(|x_1| < \epsilon \text{ and } |x_2| < \epsilon),$$

⁹we have

$$\begin{aligned}
\delta_{p,n}^{(Z)}(j, i, \mathcal{C}_n) &= \|\mathbf{1}(Y_{j,n} > 0) - \mathbf{1}(Y_{j,n,i} > 0)\|_{L^p, \mathcal{C}_n} \\
&\leq \left\{ \frac{1}{\epsilon^p} \int_{B^c} |Y_{j,n} - Y_{j,n,i}|^p d\mathbb{P} + \mathbb{P}(B|\mathcal{C}_n) \right\}^{1/p} \\
&\leq \frac{\|Y_{j,n} - Y_{j,n,i}\|_{L^p, \mathcal{C}_n}}{\epsilon} + \mathbb{P}(|Y_{j,n}| < \epsilon|\mathcal{C}_n)^{1/p} \leq \frac{\delta_{p,n}^{(Y)}(j, i, \mathcal{C}_n)}{\epsilon} + C_2 \epsilon^{1/p}
\end{aligned}$$

for some constant $C_2 > 0$, where the second inequality follows from the fact that $(a^p + b^p)^{1/p} \leq a+b$ for arbitrary non-negative numbers a, b and $p \geq 1$, and the last one follows from $\sup_{n \geq 1, 1 \leq j \leq n} \sup_y f_{j,n}(y | \mathcal{C}_n) < C_1$. By letting $\epsilon = [\delta_{p,n}^{(Y)}(j, i, \mathcal{C}_n)]^{p/(p+1)}$, we have $\delta_{p,n}^{(Z)}(j, i, \mathcal{C}_n) \leq (1 + C_2) [\delta_{p,n}^{(Y)}(j, i, \mathcal{C}_n)]^{1/(p+1)}$. \square

Proof of Proposition 5.6. We have

$$\begin{aligned}
\delta_{p,n}^{(YZ)}(j, i, \mathcal{C}_n) &= \|Y_{j,n}Z_{j,n} - Y_{j,n,i}Z_{j,n,i}\|_{L^p, \mathcal{C}_n} \\
&= \|Y_{j,n}Z_{j,n} - Y_{j,n,i}Z_{j,n} + Y_{j,n,i}Z_{j,n} - Y_{j,n,i}Z_{j,n,i}\|_{L^p, \mathcal{C}_n} \\
&\leq \|Y_{j,n}Z_{j,n} - Y_{j,n,i}Z_{j,n}\|_{L^p, \mathcal{C}_n} + \|Y_{j,n,i}Z_{j,n} - Y_{j,n,i}Z_{j,n,i}\|_{L^p, \mathcal{C}_n} \\
&= \|(Y_{j,n} - Y_{j,n,i})Z_{j,n}\|_{L^p, \mathcal{C}_n} + \|Y_{j,n,i}(Z_{j,n} - Z_{j,n,i})\|_{L^p, \mathcal{C}_n} \\
&\leq \|Z_{j,n}\|_{L^{r_1}, \mathcal{C}_n} \|Y_{j,n} - Y_{j,n,i}\|_{L^{q_1}, \mathcal{C}_n} + \|Y_{j,n}\|_{L^{r_2}, \mathcal{C}_n} \|Z_{j,n} - Z_{j,n,i}\|_{L^{q_2}, \mathcal{C}_n} \\
&\leq \|Z\|_{L^{r_1}, \mathcal{C}_n} \delta_{q_1, n}^{(Y)}(j, i, \mathcal{C}_n) + \|Y\|_{L^{r_2}, \mathcal{C}_n} \delta_{q_2, n}^{(Z)}(j, i, \mathcal{C}_n),
\end{aligned}$$

where the first inequality follows from conditional Minkowski inequality, and the last inequality follows from conditional Hölder's inequality. \square

Proof of Proposition 5.7. The proof is similar to that of Proposition 5.3. Let $B = |Z_{j,n}|$, $\rho = |Y_{j,n} - Y_{j,n,i}|$, $r = \frac{q}{2} > p$, by Lyapunov's inequality and Minkowski's inequality, we have

$$\|B\|_{L^{p/(p-1)}, \mathcal{C}_n} \leq \|B\|_{L^q, \mathcal{C}_n} = \|Z_{j,n}\|_{L^q, \mathcal{C}_n} \leq \|Z\|_{L^q, \mathcal{C}_n} < \infty \text{ a.s.},$$

and

$$\|\rho\|_{L^q, \mathcal{C}_n} = \|Y_{j,n} - Y_{j,n,i}\|_{L^q, \mathcal{C}_n} \leq 2 \|Y_{j,n}\|_{L^q, \mathcal{C}_n} \leq 2 \|Y\|_{L^q, \mathcal{C}_n} < \infty \text{ a.s.}$$

⁹See the proof of Proposition 2 in Xu and Lee (2015).

So, by the generalized Hölder's inequality,

$$\|B\rho\|_{L^r, \mathcal{C}_n} \leq \|B\|_{L^q, \mathcal{C}_n} \|\rho\|_{L^q, \mathcal{C}_n} \leq 2 \|Y\|_{L^q, \mathcal{C}_n} \|Z\|_{L^q, \mathcal{C}_n} < \infty \text{ a.s.}$$

Hence, by Lemma S.7 in the Supplementary Material of Wu et al. (2024),

$$\begin{aligned} \||Y_{j,n} - Y_{j,n,i}| \cdot |Z_{j,n}|\|_{L^p, \mathcal{C}_n} &\leq 2 \left(\|B\|_{L^{p/(p-1)}, \mathcal{C}_n}^{r-p} \|\rho\|_{L^p, \mathcal{C}_n}^{r-p} \|B\rho\|_{L^r, \mathcal{C}_n}^{(p-1)r} \right)^{1/(pr-p)} \\ &\leq C_1(\mathcal{C}_n) \left(\|Y_{j,n} - Y_{j,n,i}\|_{L^p, \mathcal{C}_n} \right)^{(q-2p)/(pq-2p)} \end{aligned}$$

where $C_1(\mathcal{C}_n) < \infty$ a.s. Similarly, we have

$$\||Y_{j,n,i}| \cdot |Z_{j,n} - Z_{j,n,i}|\|_{L^p, \mathcal{C}_n} \leq C_2(\mathcal{C}_n) \left(\|Z_{j,n} - Z_{j,n,i}\|_{L^p, \mathcal{C}_n} \right)^{(q-2p)/(pq-2p)}$$

where $C_2(\mathcal{C}_n) < \infty$ a.s. By the above two inequalities, we have

$$\begin{aligned} \delta_{p,n}^{(YZ)}(j, i, \mathcal{C}_n) &= \|Y_{j,n}Z_{j,n} - Y_{j,n,i}Z_{j,n,i}\|_{L^p, \mathcal{C}_n} \\ &= \|Y_{j,n}Z_{j,n} - Y_{j,n,i}Z_{j,n} + Y_{j,n,i}Z_{j,n} - Y_{j,n,i}Z_{j,n,i}\|_{L^p, \mathcal{C}_n} \\ &\leq \|Y_{j,n}Z_{j,n} - Y_{j,n,i}Z_{j,n}\|_{L^p, \mathcal{C}_n} + \|Y_{j,n,i}Z_{j,n} - Y_{j,n,i}Z_{j,n,i}\|_{L^p, \mathcal{C}_n} \\ &= \||Y_{j,n} - Y_{j,n,i}| \cdot |Z_{j,n}|\|_{L^p, \mathcal{C}_n} + \||Y_{j,n,i}| \cdot |Z_{j,n} - Z_{j,n,i}|\|_{L^p, \mathcal{C}_n} \\ &\leq C_1(\mathcal{C}_n) \left(\|Y_{j,n} - Y_{j,n,i}\|_{L^p, \mathcal{C}_n} \right)^{(q-2p)/(pq-2p)} + C_2(\mathcal{C}_n) \left(\|Z_{j,n} - Z_{j,n,i}\|_{L^p, \mathcal{C}_n} \right)^{(q-2p)/(pq-2p)} \\ &= C_1(\mathcal{C}_n) [\delta_{p,n}^{(Y)}(j, i, \mathcal{C}_n)]^{(q-2p)/(pq-2p)} + C_2(\mathcal{C}_n) [\delta_{p,n}^{(Z)}(j, i, \mathcal{C}_n)]^{(q-2p)/(pq-2p)} \end{aligned}$$

a.s. □