Grouped Network Vector Autoregression

Xuening Zhu[1] and Rui Pan[2]

[1]*Fudan University,* [2]*Central University of Finance and Economics*

Supplementary Material

We present here the detailed technical proofs of Lemma 1–Lemma 4 in Appendix A. Next, the proofs of Theorem 1 to Theorem 3 are given in Appendix B, C, and D respectively.

Appendix A. Four Useful Lemmas

Lemma 1. *Let $X = (X_1, \cdots, X_n)^\top \in \mathbb{R}^n$, where X_is are independent and identically distributed random variables with mean zero, variance σ_X^2 and finite fourth order moment. Let $\widetilde{\mathbb{Y}}_t = \sum_{j=0}^{\infty} G^j U \mathcal{E}_{t-j}$, where $G \in \mathbb{R}^{n\times n}$, $U \in \mathbb{R}^{n\times N}$, and $\{\mathcal{E}_t\}$ satisfy Condition (C1) and are independent of $\{X_i\}$. Then for a matrix $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ and a vector $B = (b_1, \cdots, b_n)^\top \in \mathbb{R}^n$, it holds that*

(a) $n^{-1}B^\top X \to_p 0$ *if* $n^{-2}B^\top B \to 0$ *as* $n \to \infty$.

(b) $n^{-1}X^\top AX \to_p \sigma_X^2 \lim_{n\to\infty} n^{-1} tr(A)$ *if the limit exists, and* $n^{-2}tr(AA^\top) \to 0$ *as* $n \to \infty$.

(c) $(nT)^{-1}\sum_{t=1}^{T} B^\top \widetilde{\mathbb{Y}}_t \to_p 0$ *if* $n^{-1}\sum_{j=0}^{\infty}(B^\top G^j UU^\top (G^\top)^j B)^{1/2} \to 0$ *as* $n \to \infty$.

(d) $(nT)^{-1}\sum_{t=1}^{T} \widetilde{\mathbb{Y}}_t^\top A \widetilde{\mathbb{Y}}_t^\top \to_p \lim_{n\to\infty} n^{-1} tr\{A\Gamma(0)\}$ *if the limit exists, and* $n^{-1}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}[tr\{U^\top (G^\top)^i AG^j UU^\top (G^\top)^j A^\top G^i U\}]^{1/2} \to 0$ *as* $n \to \infty$.

(e) $(nT)^{-1}\sum_{t=1}^{T} X^\top A \widetilde{\mathbb{Y}}_t^\top \to_p 0$ *if* $n^{-1}\sum_{j=0}^{\infty}[tr\{AG^jUU^\top(G^\top)^jA^\top\}]^{1/2} \to 0$ *as* $n \to \infty$.

Proof: The detailed proof can be found in Lemma 1 of Zhu et al. (2017).

Lemma 2. *Assume* $\min_k N_k = O(N^\delta)$ *and the stationary condition* $c_\beta < 1$*, where* $c_\beta = \max_k(|\beta_{1k}| + |\beta_{2k}|)$*. Further assume Conditions (C1)-(C3) hold. For matrices* $M_1 = (m_{ij}^{(1)}) \in \mathbb{R}^{n\times p}$ *and* $M_2 = (m_{ij}^{(2)}) \in \mathbb{R}^{n\times p}$*, define* $M_1 \preccurlyeq M_2$ *as* $m_{ij}^{(1)} \le m_{ij}^{(2)}$ *for* $1 \le i \le n$ *and* $1 \le j \le p$*. In addition, define* $|M|_e = (|m_{ij}|) \in \mathbb{R}^{n\times p}$ *for any arbitrary matrix* $M = (m_{ij}) \in \mathbb{R}^{n\times p}$*. Then there exists* $J > 0$*, such that*

(a) for any integer $n > 0$*, we have*

$$|\mathcal{G}^n(\mathcal{G}^\top)^n|_e \preccurlyeq n^J c_\beta^{2n} MM^\top, \tag{A.1}$$

$$|\mathcal{G}^n\Sigma_Y|_e \preccurlyeq \alpha n^J c_\beta^{n} MM^\top, \tag{A.2}$$

where $M = C\mathbf{1}\pi^\top + \sum_{j=0}^{J} W^j$*,* $C > 1$ *is a constant,* π *is defined in (C3.1), and* α *is a finite constant.*

(b) For positive integers $k_1 \le 1$*,* $k_2 \le 1$*, and* $j \ge 0$*, define* $g_{j,k_1,k_2}(\mathcal{G}, W^{(k)}) = |(W^{(k)})^{k_1}\{\mathcal{G}^j(\mathcal{G}^\top)^j\}^{k_2}(W^{(k)\top})^{k_1}|_e \in \mathbb{R}^{N\times N}$*. In addition, define* $(W^{(k)})^0 = \mathcal{I}_k = (I_{N_k}, \mathbf{0}) \in \mathbb{R}^{N_k\times N}$*. For integers* $0 \le k_1, k_2, m_1, m_2 \le 1$*, as* $N \to \infty$ *we have*

$$N^{-1}\sum_{j=0}^{\infty}\Big\{\mu^\top g_{j,k_1,k_2}(\mathcal{G}, W^{(k)})\mu\Big\}^{1/2} \to 0, \tag{A.3}$$

$$N^{-1}\sum_{i,j=0}^{\infty}\Big[tr\Big\{g_{i,k_1,k_2}(\mathcal{G}, W^{(k)})g_{j,m_1,m_2}(\mathcal{G}, W^{(k)})\Big\}\Big]^{1/2} \to 0, \tag{A.4}$$

where $|\mu|_e \preccurlyeq c_\mu\mathbf{1}$ *and* c_μ *is a finite constant.*

(c) For integers $0 \le k_1, k_2 \le 1$*, define* $f_{k_1,k_2}(W^{(k)}, Q) = |(W^{(k)})^{k_1}Q^{k_2}(W^{(k)\top})^{k_1}|_e \in \mathbb{R}^{N\times N}$*,*

where Q is given in (C3). Then for integers $0 \le k_1, k_2, m_1, m_2 \le 1$, as $N \to \infty$ we have

$$N^{-2}\mu^\top f_{k_1,k_2}(W^{(k)}, Q)\mu \to 0, \tag{A.5}$$

$$N^{-2}tr\Big\{f_{k_1,k_2}(W^{(k)}, Q)f_{m_1,m_2}(W^{(k)}, Q)\Big\} \to 0, \tag{A.6}$$

$$N^{-1}\sum_{j=0}^{\infty}\Big[tr\Big\{f_{k_1,k_2}(W^{(k)}, Q)g_{j,m_1,m_2}(\mathcal{G}, W^{(k)})\Big\}\Big]^{1/2} \to 0, \tag{A.7}$$

where $|\mu|_e \preccurlyeq c_\mu \mathbf{1}$ and c_μ is a finite constant.

Proof: The proof is similar in spirit to Zhu et al. (2017). Therefore, we give the guideline of the proof and skip some similar details. Without loss of generality, we let $c_\beta = |\beta_{11}| + |\beta_{21}|$ (i.e., $k = 1$). Consequently, we have $|\mathcal{G}|_e \preccurlyeq |\beta_{11}|W + |\beta_{21}|I$. Let $G = |\beta_{11}|W + |\beta_{21}|I$. Follow similar technique in part (a) in Lemma 2 of Zhu et al. (2017), it can be verified

$$|\mathcal{G}^n|_e \preccurlyeq n^J c_\beta^n M, \tag{A.8}$$

where $M = C\mathbf{1}\pi^\top + \sum_{j=0}^J W^j$ is defined in (a) of Lemma 2. Subsequently, the result (A.1) can be readily obtained. Next, recall that $\Sigma_Y = (I-\mathcal{G})^{-1}\Sigma_{\mathbb{Z}}(I-\mathcal{G}^\top)^{-1} + \sum_{j=0}^\infty \mathcal{G}^j\Sigma_e(\mathcal{G}^\top)^j = (\sum_{j=0}^\infty \mathcal{G}^j)\Sigma_{\mathbb{Z}}(\sum_{j=0}^\infty (\mathcal{G}^\top)^j) + \sum_{j=0}^\infty \mathcal{G}^j\Sigma_e(\mathcal{G}^\top)^j$. Let $\sigma_z^2 = \max_k\{\gamma_k^\top\Sigma_z\gamma_k\}$ and $\sigma_e^2 = \max_k\{\sigma_k^2\}$. Then we have $|\mathcal{G}^n\Sigma_Y|_e \preccurlyeq \sigma_z^2(\sum_{j=0}^\infty |\mathcal{G}^{n+j}|_e)(\sum_{j=0}^\infty |(\mathcal{G}^\top)^j|_e) + \sigma_e^2\sum_{j=0}^\infty |\mathcal{G}^{n+j}|_e|(\mathcal{G}^\top)^j|_e$. Subsequently, (A.2) can by obtained by applying (A.8). Next, we give the proof of (b) in the following. The conclusion (c) can be proved by similar techniques, which is omitted here to save space.

Let $k_1 = k_2 = 1$. Then we have $g_{j,1,1}(\mathcal{G}, W^{(k)}) = |W^{(k)}\mathcal{G}^j\mathcal{G}^jW^{(k)\top}|$. Recall that $W^{(k)} = (w_{ij} : i \in \mathcal{M}_k, 1 \le j \le N) \in \mathbb{R}^{N_k\times N}$. Since we have $|\mu|_e \preccurlyeq c_\mu\mathbf{1}$, then it suffices to show $\sum_{j=0}^\infty N_k^{-1}\big\{\mathbf{1}^\top g_{j,1,1}(\mathcal{G}, W^{(k)})\mathbf{1}\big\}^{1/2} \to 0$. We first prove (A.3). By (A.8) we have $|W^{(k)}\mathcal{G}^j|_e \preccurlyeq$

$j^K(|\beta_1|+|\beta_2|)^j W^{(k)} M$. As a result, we have

$$|W^{(k)}\mathcal{G}^j(\mathcal{G}^\top)^j W^{(k)\top}|_e \preccurlyeq j^{2K}(|\beta_1|+|\beta_2|)^{2j}\mathcal{M}, \tag{A.9}$$

where $\mathcal{M}$ is defined as $\mathcal{M} = W^{(k)}MM^\top W^{(k)\top}$. As a result, we have $\sum_{j=0}^\infty N_k^{-1}\big\{\mathbf{1}^\top W^{(k)}\mathcal{G}^j(\mathcal{G}^\top)^j W^{(k)\top}\mathbf{1}\big\}^{1/2} \le N_k^{-1}\alpha_1(\mathbf{1}^\top\mathcal{M}\mathbf{1})^{1/2}$, where $\alpha_1 = \sum_{j=0}^\infty j^K c_\beta^j < \infty$. Then it leads to show $N_k^{-2}\mathbf{1}^\top\mathcal{M}\mathbf{1} \to 0$. It can be shown $\mathbf{1}^\top\mathcal{M}\mathbf{1} = N_k^2 C\sum_j \pi_j^2 + \sum_{j=1}^K \mathbf{1}^\top W^{(k)}W^j(W^\top)^j W^{(k)\top}\mathbf{1} + 2N_k C\sum_j \pi^\top(W^\top)^j W^{(k)\top}\mathbf{1} + \sum_{i\neq j}\mathbf{1}^\top W^{(k)}W^i(W^\top)^j W^{(k)\top}\mathbf{1}$. For the last two terms of $\mathbf{1}^\top\mathcal{M}\mathbf{1}$, by Cauchy inequality, we have

$$N_k\sum_j \pi^\top(W^\top)^j W^{(k)\top}\mathbf{1} \quad\le\quad N_k\Big(\sum_j \pi_j^2\Big)^{1/2}\Big\{\mathbf{1}^\top W^{(k)}W^j(W^\top)^j W^{(k)\top}\mathbf{1}\Big\}^{1/2},$$

and $\sum_{i\neq j}\mathbf{1}^\top W^{(k)}W^i(W^\top)^j W^{(k)\top}\mathbf{1} \le \sum_{i\neq j}\big\{\mathbf{1}^\top W^{(k)}W^i(W^\top)^i W^{(k)\top}\mathbf{1}\big\}^{1/2}\big\{\mathbf{1}^\top W^{(k)}W^j(W^\top)^j W^{(k)\top}\mathbf{1}\big\}^{1/2}$. As a result, it leads to show

$$\sum_{j=1}^N \pi_j^2 \to 0 \quad\text{and}\quad N_k^{-2}\mathbf{1}^\top W^{(k)}W^j(W^\top)^j W^{(k)\top}\mathbf{1} \to 0 \tag{A.10}$$

for $1 \le j \le K+1$. As the first convergence in (A.10) is implied by (C2.1), we next prove $N_k^{-2}\mathbf{1}^\top W^{(k)}W^j(W^\top)^j W^{(k)\top}\mathbf{1} \to 0$ $(1 \le j \le K)$. Recall that $W^* = W + W^\top$. Therefore, we have $N_k^{-2}\mathbf{1}^\top W^{(k)}W^j(W^\top)^j W^{(k)\top}\mathbf{1} \le N^{-2}\mathbf{1}^\top W^{(k)}W^{*2j}W^{(k)\top}\mathbf{1}$. Then it suffices to show $N^{-2}\mathbf{1}^\top W^{(k)}W^{*2j}W^{(k)\top}\mathbf{1} \to 0$. By eigenvalue-eigenvector decomposition of W^* we have $W^* = \sum_k \lambda_k(W^*)u_k u_k^\top$, where $\lambda_k(W^*)$ and $u_k \in \mathbb{R}^N$ are the kth eigenvalue and eigenvector of W^* respectively. As a result, we have $N_k^{-2}\mathbf{1}^\top W^{(k)}W^{*2j}W^{(k)\top}\mathbf{1} \le N_k^{-2}\lambda_{\max}(W^*)^{2j}(\mathbf{1}^\top W^{(k)}W^{(k)\top}\mathbf{1})$ $(1 \le j \le K)$. Further we have $\mathbf{1}^\top W^{(k)}W^{(k)\top}\mathbf{1} \le N_k\lambda_{\max}(W^{(k)}W^{(k)\top})$. Note that $W^{(k)}W^{(k)\top}$

is a sub-matrix of $WW^\top$ with row and column index in $\mathcal{M}_k$. Therefore, by Cauchy's interlacing Theorem, we have $\lambda_{\max}(W^{(k)}W^{(k)\top}) \le \lambda_{\max}(WW^\top) = O(N^{\delta'})$ for $\delta' < \delta$. Since we have $\min_k N_k = N^\delta$ for $\delta > 0$, then we have $N_k^{-1}\lambda_{\max}(W^{(k)}W^{(k)\top}) \to 0$ as $N \to \infty$. As a consequence, the second term in (A.10) holds. Similarly, it can be proved that (A.10) holds for all $0 \le k_1, k_2 \le 1$. As a result, we have (A.3) holds.

We next prove (A.4) with $k_1 = k_2 = m_1 = m_2 = 1$, and $g_{i,1,1}(\mathcal{G}, W^{(k)})g_{j,1,1}(\mathcal{G}, W^{(k)}) = |W^{(k)}\mathcal{G}^i(\mathcal{G}^\top)^iW^{(k)\top}W^{(k)}\mathcal{G}^j(\mathcal{G}^\top)^jW^{(k)\top}|_e$. Then it can be similarly proved for other cases (i.e., $0 \le k_1, k_2, m_1, m_2 \le 1$). Note that by (A.9), we have

$$\Big[\mathrm{tr}\Big\{W^{(k)}\mathcal{G}^i(\mathcal{G}^\top)^iW^{(k)\top}W^{(k)}\mathcal{G}^j(\mathcal{G}^\top)^jW^{(k)\top}\Big\}\Big]^{1/2} \le i^Kj^K(|\beta_1| + |\beta_2|)^{i+j}\mathrm{tr}\big\{\mathcal{M}^2\big\}^{1/2}.$$

It then can be derived that $N_k^{-1}\sum_{i,j=0}^\infty[\mathrm{tr}\{W^{(k)}\mathcal{G}^i(\mathcal{G}^\top)^iW^{(k)\top}W^{(k)}\mathcal{G}^j(\mathcal{G}^\top)^jW^{(k)\top}\}]^{1/2} \le \alpha^2N_k^{-1}$ $\mathrm{tr}\big\{\mathcal{M}^2\big\}^{1/2}$. In order to obtain (A.4), it suffices to show that

$$N_k^{-2}\mathrm{tr}\{\mathcal{M}^2\} \to 0. \tag{A.11}$$

Equivalently, by Cauchy inequality, it suffices to prove $(\sum \pi_j^2)^2 \to 0$, and $N_k^{-2}\mathrm{tr}\{W^{(k)}W^jW^{j\top}W^{(k)\top}$ $W^{(k)}W^jW^{j\top}W^{(k)\top}\} \to 0$ holds for $1 \le j \le K$. It can be easily verified the first term holds by (C2.1). For the second one, we have $N_k^{-2}\mathrm{tr}\{W^{(k)}W^jW^{j\top}W^{(k)\top}W^{(k)}W^jW^{j\top}W^{(k)\top}\} \le$ $N_k^{-2}\mathrm{tr}\{W^{(k)}(W^*)^{4j}W^{(k)\top}\} \le N_k^{-2}\lambda_{\max}(W^*)^{4j}\mathrm{tr}(W^{(k)}W^{(k)\top}) \le N_k^{-2}N_k\lambda_{\max}(W^*)^{4K}\lambda_{\max}(WW^\top)$. Similarly, due to that $\lambda_{\max}(W^*) = O(\log N)$ and $\lambda_{\max}(WW^\top) = O(N^{\delta'})$ in (C2.2), we have $N_k^{-1}\lambda_{\max}(W^*)^{4K}\lambda_{\max}(WW^\top) \to 0$ as $N \to \infty$. Consequently, we have (A.11) and then (A.4) holds. This completes the proof of (b).

Lemma 3. *Let* $\{X_{it} : 1 \le t \le T\}$ *and* $\{Y_{it} : 1 \le t \le T\}$ *be random sub-Gaussian time* *series with mean 0, var*$(X_{it}) = \sigma_{i,xx}$*, var*$(Y_{it}) = \sigma_{i,yy}$*, and cov*$(X_{it}, Y_{it}) = \sigma_{i,xy}$*. Let* $\sigma_{xi,t_1t_2} =$ *cov*(X_{it_1}, X_{it_2}) *and* $\Sigma_{xi} = (\sigma_{xi,t_1t_2} : 1 \leq t_1, t_2 \leq T) \in \mathbb{R}^{T\times T}$*. Similarly, define* σ_{yi,t_1t_2} *and* $\Sigma_{yi} \in \mathbb{R}^{T\times T}$*. Then we have*

$$P\Big(\Big|T^{-1}\sum_{t=1}^{T} X_{it}Y_{it} - \sigma_{i,xy}\Big| > \nu\Big) \leq c_1\Big\{\exp(-c_2\sigma_{xi}^{-2}T^2\nu^2) + \exp(-c_2\sigma_{yi}^{-2}T^2\nu^2)\Big\} \qquad \text{(A.12)}$$

for $|\nu| \leq \delta$*, where* $\sigma_{xi}^2 = tr(\Sigma_{xi}^2)$*,* $\sigma_{yi}^2 = tr(\Sigma_{yi}^2)$*,* c_1*,* c_2*, and* δ *are finite constants.*

Proof: Let $X_i = (X_{i1}, \cdots, X_{iT})^\top \in \mathbb{R}^T$ and $Y_i = (Y_{i1}, \cdots, Y_{iT})^\top \in \mathbb{R}^T$. In addition, let $Z_i = Z_i + Y_i$. Therefore, we have $Z_i^\top Z_i = 2^{-1}(Z_i^\top Z_i - X_i^\top X_i - Y_i^\top Y_i)$. It can be derived that

$$P\{|T^{-1}(X_i^\top Y_i) - \sigma_{i,xy}| \geq \nu\} \leq P\{|T^{-1}(Z_i^\top Z_i) - (\sigma_{i,xx} + \sigma_{i,yy} + 2\sigma_{i,xy})| \geq \nu_1\}$$
$$+ P\{|T^{-1}(X_i^\top X_i) - \sigma_{i,xx}| \geq \nu_1\} + P\{|T^{-1}(Y_i^\top Y_i) - \sigma_{i,yy}| \geq \nu_1\}, \qquad \text{(A.13)}$$

where $\nu_1 = 2\nu/3$. Next, we derive the upper bound for the right side of (A.13). Note that $\mathbb{X}_i^\top \mathbb{X}_i$, $Y_i^\top Y_i$, and $Z_i^\top Z_i$ all take quadratic form. Therefore the proofs are similar. For the sake of simplicity, we take $Y_i^\top Y_i$ for an example and derive the upper bound for $P\{|n^{-1}(Y_i^\top Y_i) - \sigma_{i,yy}| \geq \nu_1\}$. Similar results can be obtained for the other two terms.

First we have $Y_i^\top Y_i = Y_i^\top \Sigma_{yi}^{-1/2}\Sigma_{yi}\Sigma_{yi}^{-1/2}Y_i = \widetilde{Y}_i^\top \Sigma_{yi}\widetilde{Y}_i$, where $\widetilde{Y}_i = \Sigma_{yi}^{-1/2}Y_i$ follows sub-Gaussian distribution. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_T$ be the eigenvalues of Σ_{yi}. Since Σ_{yi} is a non-negative definite matrix, The eigenvalue decomposition can be applied to obtain $\Sigma_{yi} = U^\top \Lambda U$, where $U = (U_1, \cdots, U_T)^\top \in \mathbb{R}^{T\times T}$ is an orthogonal matrix and $\Lambda = \mathrm{diag}\{\lambda_1, \cdots, \lambda_T\}$. As a consequence, we have $Y_t^\top Y_t = \sum_t \lambda_t \zeta_t^2$, where $\zeta_t = U_t^\top \widetilde{Y}_t$ and ζ_ts are independent and identically distributed as standard sub-Gaussian. It can be verified $\zeta_t^2 - 1$ satisfies sub-exponential distribution and $T^{-1}(\sum_t \lambda_t) = \sigma_{i,yy}$. In addition, the sub-exponential distribution satisfies condition

(P) on page 45 of Saulis and Statuleviveccius (2012). There exists constants c_1, c_2, and δ such that $P\{|T^{-1}(Y_i^\top Y_i) - \sigma_{i,yy}| \geq \nu_1\} = P\{\sum_t \lambda_t(\zeta_t^2 - 1)| \geq T\nu_1\} \leq c_1 \exp\{-c_2(\sum_t \lambda_t^2)^{-1} T^2\nu^2\} = c_1 \exp\{-c_2\sigma_{yi}^{-1} T^2\nu^2\}$ for $|\nu| < \delta$ by the Theorem 3.3 of Saulis and Statuleviveccius (2012). Consequently, (A.12) can be obtained by appropriately chosen c_1, c_2, and δ.

Lemma 4. *Assume Y_{it} follows the GNAR model (2.4) and $|c_\beta| < 1$. Then there exists finite constants c_1, c_2, and δ, for $\nu < \delta$ we have*

$$P\Big\{\Big|T^{-1}\sum_{t=1}^{T} Y_{it}^2 - \mu_i^2 - e_i^\top \Sigma_Y e_i\Big| > \nu\Big\} \leq \delta_T, \tag{A.14}$$

$$P\Big\{\Big|T^{-1}\sum_{t=1}^{T} Y_{it}(w_i^\top \mathbb{Y}_t) - \mu_{Yi}(w_i^\top \mu_Y) - w_i^\top \Sigma_Y e_i\Big| > \nu\Big\} \leq \delta_T \tag{A.15}$$

$$P\Big\{\Big|T^{-1}\sum_{t=1}^{T} Y_{i(t-1)}\varepsilon_{it}\Big| > \nu\Big\} \leq \delta_T, \quad P\Big\{\Big|T^{-1}\sum_{t=1}^{T} (w_i^\top \mathbb{Y}_{t-1})\varepsilon_{it}\Big| > \nu\Big\} \leq \delta_T, \tag{A.16}$$

$$P\Big\{\Big|T^{-1}\sum_{t=1}^{T} Y_{i(t-1)} - \mu_i\Big| > \nu\Big\} \leq \delta_T, \quad P\Big\{\Big|T^{-1}\sum_{t=1}^{T} w_i^\top \mathbb{Y}_t - w_i^\top \mu_Y\Big| > \nu\Big\} \leq \delta_T, \tag{A.17}$$

where $\delta_T = c_1 \exp(-c_2 T\nu^2)$, $e_i \in \mathbb{R}^N$ is an N-dimensional vector with all elements being 0 but the ith element being 1, and $\mu_i = e_i^\top \mu_Y$.

Proof: For the similarity of proof procedure, we only prove (A.14) in the following. Without loss of generality, let $\mu_Y = \mathbf{0}$. Recall that the group information is denoted as $\boldsymbol{Z} = \{z_{ik} : 1 \leq i \leq N, 1 \leq k \leq K\}$. Define $P^*(\cdot) = P(\cdot|\boldsymbol{Z})$, $E^*(\cdot) = E(\cdot|\boldsymbol{Z})$, and $\text{cov}^*(\cdot) = \text{cov}(\cdot|\boldsymbol{Z})$. Write $\mathcal{Y}_i = (Y_{i1}, \cdots, Y_{iT})^\top \in \mathbb{R}^T$. Given $\boldsymbol{Z}$, $\mathcal{Y}_i$ is a sub-Gaussian random vector with $\text{cov}(\mathcal{Y}_i) = \Sigma_i = (\sigma_{i,t_1t_2}) \in \mathbb{R}^{T\times T}$, where $\sigma_{i,t_1t_2} = e_i^\top \mathcal{G}^{t_1-t_2}\Sigma_Y e_i$ for $t_1 \geq t_2$, $\sigma_{i,t_1t_2} = e_i^\top \Sigma_Y (\mathcal{G}^\top)^{t_2-t_1} e_i$, and $\mathcal{G}$ is pre-defined in (2.6) as $\mathcal{G} = \mathcal{B}_1 W + \mathcal{B}_2$. It can be derived $\text{var}^*(\mathcal{Y}_i^\top \mathcal{Y}_i) \leq c\text{tr}(\Sigma_i^2)$, where c is a positive constant and $\text{tr}(\Sigma_i^2) = T(e_i^\top \Sigma_Y e_i)^2 + 2\sum_{t=1}^{T-1}(T-t)(e_i^\top \mathcal{G}^t \Sigma_Y e_i)^2$. It can be derived $|\Sigma_Y|_e \preccurlyeq \alpha M M^\top$ and $|\mathcal{G}^t \Sigma_Y|_e \preccurlyeq \alpha_1 t^J c_\beta^t M M^\top$ by (A.2) of Lemma 2, where c_β, J and M are defined in Lemma 2, α and α_1 are finite constants. In addition, it can be verified $\sum_{t=1}^{T-1}(T-t)t^{2J}c_\beta^{2t} \le \alpha_2 T$, where α_2 is a finite constant. Therefore we have $\mathrm{tr}(\Sigma_i^2) \le T(\alpha + 2\alpha_1\alpha_2)\{(e_i^\top MM^\top e_i)^2\}$. Since we have $e_i^\top MM^\top e_i \le (J+1)e_i^\top M\mathbf{1} \le (J+1)^2 = O(1)$, it can be concluded that $\mathrm{tr}(\Sigma_i^2) \le T\alpha_3$, where $\alpha_3 = (\alpha+2\alpha_1\alpha_2)(J+1)^2$. By Lemma 3, the (A.14) can be obtained.

Appendix B. Proof of Theorem 1

Let $\lambda_i(M)$ be the ith eigenvalue of $M \in \mathbb{R}^{N\times N}$. We first verify that the solution (2.7) is strictly stationary. By Banerjee et al. (2014), we have $\max_i |\lambda_i(W)| \le 1$. Hence we have

$$\max_{1\le i\le N} |\lambda_i(\mathcal{G})| \le \big(\max_{1\le k\le K} |\beta_{1k}|\big)\big(\max_{1\le i\le N} |\lambda_i(W)|\big) + \max_{1\le k\le K} |\beta_{2k}| < 1. \tag{A.18}$$

Consequently, we have $\lim_{m\to\infty}\sum_{j=0}^m \mathcal{G}^j\mathcal{E}_{t-j}$ exists and $\{\mathbb{Y}_t\}$ given by (2.7) is a strictly stationary process. In addition, one could directly verify that $\{\mathbb{Y}_t\}$ satisfies the GNAR model (2.4).

Next, we verify that the strictly stationary solution (2.7) is unique. Assume $\{\widetilde{\mathbb{Y}}_t\}$ is another strictly stationary solution to the GNAR model (2.4) with $E\|\widetilde{\mathbb{Y}}_t\| < \infty$. Then we have $\widetilde{\mathbb{Y}}_t = \sum_{j=1}^{m-1}\mathcal{G}^j(\mathcal{B}_0 + \mathcal{E}_{t-j}) + \mathcal{G}^m\widetilde{\mathbb{Y}}_{t-m}$ for any positive integer m. Let $\rho = \max_k(|\beta_{1k}| + |\beta_{2k}|)$. Then one could verify $E\|\mathbb{Y}_t - \widetilde{\mathbb{Y}}_t\| = E\|\sum_{j=m}^\infty \mathcal{G}^j(\mathcal{B}_0 + \mathcal{E}_{t-j}) - \mathcal{G}^m\widetilde{\mathbb{Y}}_{t-m}\| \le C\rho^m$, where C is a finite constant unrelated to t and m. Note that m can be chosen arbitrarily. As a result, we have that $E\|\mathbb{Y}_t - \widetilde{\mathbb{Y}}_t\| = 0$, i.e. $\mathbb{Y}_t = \widetilde{\mathbb{Y}}_t$ with probability one. This completes the proof.

Appendix C. Proof of Theorem 2

According to (3.9), $\widehat{\theta}_k$ can be explicitly written as $\widehat{\theta}_k = \theta_k + \widehat{\Sigma}_k^{-1}\widehat{\zeta}_k$, where $\widehat{\Sigma}_k = (N_kT)^{-1}\sum_{t=1}^T \mathbb{X}_{t-1}^{(k)\top}\mathbb{X}_{t-1}^{(k)}$ and $\widehat{\zeta}_k = (N_kT)^{-1}\sum_{t=1}^T \mathbb{X}_{t-1}^{(k)\top}\mathcal{E}_t^{(k)}$. Without loss of generality, we assume $\sigma_k^2 = 1$ for

$k = 1, \cdots, K$. Let $\Sigma_k = \lim_{N\to\infty} E(\widehat{\Sigma}_k)$. As a result, it suffices to show that

$$\widehat{\Sigma}_k \to_p \Sigma_k, \tag{A.19}$$

$$\sqrt{N_k T}\widehat{\zeta}_k = O_p(1), \tag{A.20}$$

as $\min\{N, T\} \to \infty$. Subsequently, we prove (A.19) in Step 1 and (A.20) in Step 2.

Step 1. Proof of (A.19). Define $Q = (I - \mathcal{G})^{-1}\Sigma_{\mathbb{V}}(I - \mathcal{G}^\top)^{-1}$. In this step, we intend to show that $\widehat{\Sigma}_k =$

$$\frac{1}{N_k T}\sum_{t=1}^{T} \mathbb{X}_{t-1}^{(k)\top}\mathbb{X}_{t-1}^{(k)} = \begin{pmatrix} 1 & \mathbb{S}_{12} & \mathbb{S}_{13} & \mathbb{S}_{14} \\ & \mathbb{S}_{22} & \mathbb{S}_{23} & \mathbb{S}_{24} \\ & & \mathbb{S}_{33} & \mathbb{S}_{34} \\ & & & \mathbb{S}_{44} \end{pmatrix} \to_p \begin{pmatrix} 1 & c_{1\beta} & c_{2\beta} & \mathbf{0}^\top \\ & \Sigma_1 & \Sigma_2 & \kappa_8\gamma^\top\Sigma_z \\ & & \Sigma_3 & \kappa_3\gamma^\top\Sigma_z \\ & & & \Sigma_z \end{pmatrix} = \Sigma_k,$$

where

$$\mathbb{S}_{12} = \frac{1}{N_k T}\sum_{t=1}^{T}\sum_{i\in\mathcal{M}_k} w_i^\top \mathbb{Y}_{t-1}, \quad \mathbb{S}_{13} = \frac{1}{N_k T}\sum_{t=1}^{T}\sum_{i\in\mathcal{M}_k} Y_{i(t-1)}, \quad \mathbb{S}_{14} = \frac{1}{N_k}\sum_{i\in\mathcal{M}_k} V_i^\top,$$

$$\mathbb{S}_{22} = \frac{1}{N_k T}\sum_{t=1}^{T}\sum_{i\in\mathcal{M}_k} (w_i^\top \mathbb{Y}_{t-1})^2, \quad \mathbb{S}_{23} = \frac{1}{N_k T}\sum_{t=1}^{T}\sum_{i\in\mathcal{M}_k} w_i^\top \mathbb{Y}_{t-1} Y_{i(t-1)},$$

$$\mathbb{S}_{24} = \frac{1}{N_k T}\sum_{t=1}^{T}\sum_{i\in\mathcal{M}_k} w_i^\top \mathbb{Y}_{t-1} V_i^\top, \quad \mathbb{S}_{33} = \frac{1}{N_k T}\sum_{t=1}^{T}\sum_{i\in\mathcal{M}_k} Y_{i(t-1)}^2,$$

$\mathbb{S}_{34} = (N_k T)^{-1}\sum_{t=1}^{T}\sum_{i\in\mathcal{M}_k} Y_{i(t-1)} V_i^\top$, $\mathbb{S}_{44} = N_k^{-1}\sum_{i\in\mathcal{M}_k} V_i V_i^\top$. By (2.7), we have

$$\mathbb{Y}_t = (I - \mathcal{G})^{-1} b_0 + (I - \mathcal{G})^{-1} b_v + \widetilde{\mathbb{Y}}_t, \tag{A.21}$$

where $b_0 = \sum_k D_k B_{0k}$, $b_v = \sum_k D_k \mathbb{V}\gamma_k$, and $\widetilde{\mathbb{Y}}_t = \sum_{j=0}^{\infty} \mathcal{G}^j \mathcal{E}_{t-j}$. By the law of large numbers, one could directly obtain that $\mathbb{S}_{44} \to_p \Sigma_v$ and $\mathbb{S}_{14} \to_p \mathbf{0}^\top$. Subsequently, we only show the convergence of $\mathbb{S}_{12}$ and $\mathbb{S}_{23}$ in $\widehat{\Sigma}_k$ as follows.

Convergence of $\mathbb{S}_{12}$. It can be derived that

$$\mathbb{S}_{12} = \frac{1}{N_k T} \sum_{t=1}^{T} \mathbf{1}^\top W^{(k)} \mathbb{Y}_{t-1} = \frac{\mathbf{1}^\top W^{(k)} \mu_Y}{N_k} + \mathbb{S}_{12a} + \mathbb{S}_{12b},$$

where $\mathbb{S}_{12a} = N_k^{-1} \mathbf{1}^\top W^{(k)} (I - \mathcal{G})^{-1} b_v$ and $\mathbb{S}_{12b} = (N_k T)^{-1} \sum_{t=1}^{T} \mathbf{1}^\top W^{(k)} \widetilde{\mathbb{Y}}_{t-1}$. Then by (A.5) and (A.3) in Lemma 2, we have $N_k^{-2} \mathbf{1}^\top W^{(k)} Q W^{(k)\top} \mathbf{1} \to 0$ and $N_k^{-1} \sum_{j=0}^{\infty} \big\{ \mathbf{1}^\top W^{(k)} \mathcal{G}^j (\mathcal{G}^\top)^j W^{(k)\top} \mathbf{1} \big\}^{1/2} \to 0$, as $N \to \infty$. As a result, it is implied by Lemma 1 (a) and (c) that $\mathbb{S}_{12a} \to_p 0$ and $\mathbb{S}_{12b} \to_p 0$.

Convergence of $\mathbb{S}_{23}$. Note that

$$\begin{aligned}\mathbb{S}_{23} &= \frac{1}{N_k T} \sum_{t=1}^{T} \sum_{i \in \mathcal{M}_k} w_i^\top \mathbb{Y}_{t-1} Y_{i(t-1)} = \frac{1}{N_k T} \sum_{t=1}^{T} \mathbb{Y}_{t-1}^{(k)\top} W^{(k)} \mathbb{Y}_{t-1} \\ &= \frac{\mu_Y^{(k)\top} W^{(k)} \mu_Y}{N_k} + \mathbb{S}_{23a} + \mathbb{S}_{23b} + \mathbb{S}_{23c} + \mathbb{S}_{23d} + \mathbb{S}_{23e},\end{aligned}$$

where $\mathbb{S}_{23a} = N_k^{-1} \widetilde{b}_v^\top \mathcal{I}_k^\top W^{(k)} \widetilde{b}_v$, $\mathbb{S}_{23b} = N_k^{-1} T^{-1} \sum_{t=1}^{T} \widetilde{\mathbb{Y}}_{t-1}^{(k)\top} W^{(k)} \widetilde{\mathbb{Y}}_{t-1}$ and $\mathbb{S}_{23c} = N_k^{-1} T^{-1} \sum_{t=1}^{T} (\widetilde{b}_v^\top \mathcal{I}_k^\top W^{(k)} \widetilde{\mathbb{Y}}_{t-1} + \widetilde{\mathbb{Y}}_{t-1}^\top \mathcal{I}_k^\top W^{(k)} \widetilde{b}_v)$, $\mathbb{S}_{23d} = N_k^{-1} (\widetilde{b}_v^\top \mathcal{I}_k^\top \widetilde{\mu}_Y + \mu_Y^\top \mathcal{I}_k^\top \widetilde{b}_v)$, $\mathbb{S}_{23e} = N_k^{-1} T^{-1} \sum_{t=1}^{T} (\mathbb{Y}_{t-1}^{(k)\top} \widetilde{\mu}_Y + \mu_Y^\top \mathcal{I}_k^\top W^{(k)} \mathbb{Y}_{t-1})$, where $\widetilde{\mu}_Y = W^{(k)} \mu_Y$ and $\widetilde{b}_v = (I - \mathcal{G})^{-1} b_v$.

We next look at the terms one by one. First we have $N_k^{-2} \mathrm{tr}(\mathcal{I}_k Q \mathcal{I}_k^\top W^{(k)} Q W^{(k)\top}) \to 0$ by (A.6) in Lemma 2 (c). Therefore, by (b) in Lemma 1, we have $\mathbb{S}_{23a} \to_p s_{23a}$, where $s_{23a} = \lim_{N_k \to \infty} E(\mathbb{S}_{23a})$. Next, for $\mathbb{S}_{23b}$ we have $N_k^{-1} \sum_{i,j=0}^{\infty} \mathrm{tr}\{\mathcal{I}_k \mathcal{G}^i (\mathcal{G}^\top)^i \mathcal{I}_k^\top W^{(k)} \mathcal{G}^j (\mathcal{G}^\top)^j W^{(k)\top}\} \to 0$ by (A.4) in Lemma 2 (b). Therefore, by (d) in Lemma 1, we have $\mathbb{S}_{23b} \to_p s_{23b}$, where $s_{23b} = \lim_{N_k \to \infty} E(\mathbb{S}_{23b})$. Next, let $\mathbb{S}_{23c} = \mathbb{S}_{23c}^{(1)} + \mathbb{S}_{23c}^{(2)}$, where $\mathbb{S}_{23c}^{(1)} = N_k^{-1} T^{-1} \sum_{t=1}^{T} \widetilde{b}_z^\top \mathcal{I}_k^\top W^{(k)} \widetilde{\mathbb{Y}}_{t-1}$ and

$\mathbb{S}_{23c}^{(2)} = N_k^{-1}T^{-1}\sum_{t=1}^T \widetilde{\mathbb{Y}}_{t-1}^\top \mathcal{I}_k^\top W^{(k)}\widetilde{b}_v$. Note that we have $N_k^{-1}\sum_{j=0}^\infty \mathrm{tr}\{W^{(k)}\mathcal{G}^j(\mathcal{G}^\top)^j W^{(k)\top}\mathcal{I}_k Q$ $\mathcal{I}_k^\top\} \to 0$ and $N_k^{-1}\sum_{j=0}^\infty \mathrm{tr}\{\mathcal{I}_k\mathcal{G}^j(\mathcal{G}^\top)^j\mathcal{I}_k^\top W^{(k)}QW^{(k)\top}\} \to 0$ by (A.7) in Lemma 2 (c). Therefore, $\mathbb{S}_{23c} \to_p s_{23c}$ by (e) in Lemma 1, where $s_{23c} = \lim_{N_k\to\infty} E(\mathbb{S}_{23c})$. Next, by similar proof to the convergence of $\mathbb{S}_{13}$, we have that $\mathbb{S}_{23d} \to_p 0$ and $\mathbb{S}_{23e} \to_p 0$. As a consequence, we have $\mathbb{S}_{23} \to_p \Sigma_2$.

Step 2. Proof of (A.20). It can be verified that $\sqrt{N_k T}E(\widehat{\zeta}_k) = 0$. In addition, we have $\mathrm{var}\{\sqrt{N_k T}\widehat{\zeta}_k\} = E(\widehat{\Sigma}_k) \to \Sigma_k$ as $N_k \to \infty$. Consequently, we have $\sqrt{N_k T}\widehat{\zeta}_k = O_p(1)$.

Appendix D. Proof of Theorem 3

Let $\widehat{\Sigma}_x^{(i)} = T^{-1}\sum_{t=1}^T \boldsymbol{X}_{i(t-1)}\boldsymbol{X}_{i(t-1)}^\top = (\widehat{\sigma}_{x,ij}) \in \mathbb{R}^{3\times 3}$, and $\widehat{\Sigma}_{xe}^{(i)} = T^{-1}(\sum_{t=1}^T \boldsymbol{X}_{i(t-1)}\delta_i$ $\varepsilon_{it})$. We then have

$$\widehat{b}_i - b_i = (\widehat{\Sigma}_x^{(i)})^{-1}\Sigma_{xe}^{(i)}.$$

Let $\widehat{\Sigma}_x^{(i)} = (\widehat{\sigma}_{x,j_1j_2} : 1 \le l_1, l_2 \le 3) \in \mathbb{R}^{3\times 3}$, where the index i of $\widehat{\sigma}_{x,l_1l_2}$ is omitted. Specifically, $\widehat{\sigma}_{x,11} = 1$, $\widehat{\sigma}_{x,12} = T^{-1}\sum_t w_i^\top \mathbb{Y}_{t-1}$, $\widehat{\sigma}_{x,13} = T^{-1}\sum_t e_i^\top \mathbb{Y}_{t-1}$, $\widehat{\sigma}_{x,22} = T^{-1}\sum_t Y_{i(t-1)}^2$, $\widehat{\sigma}_{x,23} = T^{-1}\sum_t Y_{i(t-1)}(w_i^\top \mathbb{Y}_{t-1})$, $\widehat{\sigma}_{x,33} = T^{-1}\sum_t (w_i^\top \mathbb{Y}_{t-1})^2$. Mathematically, it can be computed $(\widehat{\Sigma}_x^{(i)})^{-1} = |\widehat{\Sigma}_x^{(i)}|^{-1}\widehat{\Sigma}_x^{*(i)}$, where $|\widehat{\Sigma}_x^{(i)}|$ is the determinant of $\widehat{\Sigma}_x^{(i)}$, and $\widehat{\Sigma}_x^{*(i)}$ is the adjugate matrix of $\widehat{\Sigma}_x^{(i)}$, and $\Sigma_x^{*(i)} = (\widehat{\sigma}_{x,l_1l_2}^*)$, where $\widehat{\sigma}_{x,11}^* = \widehat{\sigma}_{x,22}\widehat{\sigma}_{x,33} - \widehat{\sigma}_{x,23}^2$, $\widehat{\sigma}_{x,12}^* = \widehat{\sigma}_{x,13}\widehat{\sigma}_{x,32} - \widehat{\sigma}_{x,12}\widehat{\sigma}_{x,33}$ $\widehat{\sigma}_{x,13}^* = \widehat{\sigma}_{x,21}\widehat{\sigma}_{x,32} - \widehat{\sigma}_{x,22}\widehat{\sigma}_{x,31}$, $\widehat{\sigma}_{x,22}^* = \widehat{\sigma}_{x,11}\widehat{\sigma}_{x,33} - \widehat{\sigma}_{x,13}^2$, $\widehat{\sigma}_{x,23}^* = \widehat{\sigma}_{x,13}\widehat{\sigma}_{x,32} - \widehat{\sigma}_{x,12}\widehat{\sigma}_{x,33}$, and $\widehat{\sigma}_{x,33}^* = \widehat{\sigma}_{x,11}\widehat{\sigma}_{x,22} - \widehat{\sigma}_{x,12}^2$. It can be derived $|\widehat{\Sigma}_x^{(i)}| = \widehat{\sigma}_{x,11}(\widehat{\sigma}_{x,22}\widehat{\sigma}_{x,33} - \widehat{\sigma}_{x,23}^2) - \widehat{\sigma}_{x,12}(\widehat{\sigma}_{x,12}\widehat{\sigma}_{33} - \widehat{\sigma}_{13}\widehat{\sigma}_{23}) + \widehat{\sigma}_{13}(\widehat{\sigma}_{12}\widehat{\sigma}_{23} - \widehat{\sigma}_{22}\widehat{\sigma}_{13})$. By the maximum inequality, we have

$$P(\sup_i \|\widehat{b}_i - b_i\| > \nu) \le \sum_{i=1}^N P(\|\widehat{b}_i - b_i\| > \nu). \tag{A.22}$$

In addition, we have

$$P(\|\widehat{b}_i - b_i\| > \nu) \leq P\big(\big||\widehat{\Sigma}_x^{(i)}| - \sigma_x^{(i)}\big| \geq \delta_i\big) + P\big(\big|\widehat{\Sigma}_x^{*(i)}\widehat{\Sigma}_{xe}^{(i)}\big| \geq \delta_i \nu\big), \tag{A.23}$$

where $\sigma_x^{(i)} = \sigma_{x,11}(\sigma_{x,22}\sigma_{x,33} - \sigma_{x,23}^2) - \sigma_{x,12}(\sigma_{x,12}\sigma_{33} - \sigma_{13}\sigma_{23}) + \sigma_{13}(\sigma_{12}\sigma_{23} - \sigma_{22}\sigma_{13}) = (e_i^\top \Sigma_Y e_i)(w_i^\top \Sigma_Y w_i) - (e_i^\top \Sigma_Y w_i)^2$, $\delta_i = \sigma_x^{(i)}/2$. By lemma 4, for each component of $|\widehat{\Sigma}_x^{(i)}|$ we have $P(|\widehat{\sigma}_{x,l_1 l_2} - \sigma_{x,l_1 l_2}| > \nu_0) \leq c_1 \exp(-c_2 T \nu_0^2)$, where $\sigma_{x,l_1 l_2} = E(\widehat{\sigma}_{x,l_1 l_2})$ and ν_0 is a finite positive constant. Moreover, by the conditions of Theorem 3, we have $\sigma_x^{(i)} \geq \tau$ with probability tending to 1. Consequently, it is not difficult to obtain the result $P(||\widehat{\Sigma}_x^{(i)}| - \sigma_x^{(i)}| \geq \delta_i) \leq c_1^* \exp(-c_2^* T\tau^2)$, where c_1^*, c_2^* are finite constants. Subsequently, we have $P(|\widehat{\Sigma}_x^{*(i)}\widehat{\Sigma}_{xe}^{(i)}| \geq \delta_i \nu) \leq P(|\widehat{\Sigma}_x^{*(i)}\widehat{\Sigma}_{xe}^{(i)}| \geq \tau\nu/2)$. By similar technique, one could verify that each element of $\widehat{\Sigma}_x^{*(i)}$ and $\widehat{\Sigma}_{xe}^{(i)}$ converge with probability and the tail probability can be controlled, where the basic results are given in Lemma 4. Consequently, there exists constants c_3^* and c_4^* such that $P(|\widehat{\Sigma}_x^{*(i)}\widehat{\Sigma}_{xe}^{(i)}| \geq \tau\nu/2) \leq c_3^* \exp(-c_4^* T\tau^2\nu^2)$. Consequently, we have $P(\|\widehat{b}_i - b_i\| > \nu) \leq c_1^* \exp(-c_2^* T\tau^2) + c_3^* \exp(-c_4^* T\tau^2\nu^2)$ by (A.23). By the condition $N = o(\exp(T))$, the right side of (A.22) goes to 0 as $N \to \infty$. This completes the proof.

References

Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2014), *Hierarchical modeling and analysis for spatial data*, Crc Press.

Saulis, L. and Statuleviveccius, V. (2012), *Limit theorems for large deviations*, vol. 73, Springer Science & Business Media.

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