Submitted to the Annals of Statistics

SUPPLEMENTARY MATERIALS FOR "NETWORK VECTOR AUTOREGRESSION"

By Xuening Zhu, Rui Pan, Guodong Li, Yuewen Liu and Hansheng Wang

Peking University, Central University of Finance and Economics, University of Hong Kong and Xi'an Jiaotong University

This is a supplementary material that contains the verification of (2.6) and (2.7), proofs of Theorem 1, Theorem 4, Theorem 5, two useful lemmas, and Proposition 2. Lastly, the numerical verification of conditions (C2)–(C3) are also included.

1. Verification of (2.6) and (2.7). First, by Taylor's expansion, $(I - \beta_1 W - \beta_2 I)^{-1} = (1-\beta_2)^{-1}\{I + (1-\beta_2)^{-1}\beta_1 W + (1-\beta_2)^{-2}\beta_1^2 W^2 + \cdots\}$. As a result the stationary mean in (2.4) can be written as

$$\begin{aligned}\mu &= \frac{1}{1-\beta_2}\big\{I + \frac{\beta_1}{1-\beta_2}W + \frac{\beta_1^2}{(1-\beta_2)^2}W^2 + \cdots\big\}\mathcal{B}_0\\ &\approx \frac{1}{1-\beta_2}\big(I + \frac{\beta_1}{1-\beta_2}W\big)\mathcal{B}_0,\end{aligned}$$

where the last equation is approximated by the first order Taylor's expansion and $\max_i|\lambda_i(W)| \le 1$. Next, one can verify that $\mathbb{Y}_t = \mathcal{B}_0 + G\mathbb{Y}_{t-1} + \mathcal{E}_t = \mu + \mathcal{E}_t + G\mathcal{E}_{t-1} + G^2\mathcal{E}_{t-2} + \cdots + G^k\mathcal{E}_{t-k} + \cdots = \mu + \sum_{k=0}^\infty G^k\mathcal{E}_{t-k}$. Recall $\mu = \mathcal{B}_0 + G\mathcal{B}_0 + \cdots = (I-G)^{-1}\mathcal{B}_0$. Then, we have $\Gamma(0) = \mathrm{cov}(\sum_{k=0}^\infty G^k\mathcal{E}_{t-k}) =$

$$\begin{aligned}\sum_{k=0}^\infty \sigma^2 G^k(G^k)^\top &= \sum_{k=0}^\infty \sigma^2(I-\beta_1 W - \beta_2 I)^k(I - \beta_1 W^\top - \beta_2 I)^k\\ &\approx \frac{1}{1-\beta_2^2}I + \frac{\beta_1\beta_2}{(1-\beta_2^2)^2}(W + W^\top),\end{aligned}$$

where the last equation is approximated by the first order Taylor's expansion and $\max_i|\lambda_i(W)| \le 1$. This completes the proof.

2. Proof of Theorem 1. Denote $\lambda_i(M)$ as the ith eigenvalue of any arbitrary matrix $M \in \mathbb{R}^{N\times N}$. We first verify that the solution given by (2.3) satisfies strict stationarity. To this end, note that $\max_i|\lambda_i(W)| \le 1$ [1], and

$$\rho = \max_{1\le i\le N}|\lambda_i(G)| \le |\beta_1|\max_{1\le i\le N}|\lambda_i(W)| + |\beta_2| < 1. \tag{2.1}$$

It holds that $\lim_{m\to\infty}\sum_{j=0}^{m} G^j\mathcal{E}_{t-j}$ exists, and then $\{\mathbb{Y}_t\}$ defined in (2.3) is a strictly stationary process. It is straightforward to verify that $\{\mathbb{Y}_t\}$ satisfies the NAR model (2.2). Next, we verify the uniqueness of the strictly stationarity solution (2.3). Assume that $\{\widetilde{\mathbb{Y}}_t\}$ is another strictly stationary solution to the NAR model with $E\|\widetilde{\mathbb{Y}}_t\| < \infty$. Then $\widetilde{\mathbb{Y}}_t = \sum_{j=1}^{m-1} G^j(\mathcal{B}_0 + \mathcal{E}_{t-j}) + G^m\widetilde{\mathbb{Y}}_{t-m}$ for any positive integer m. Hence by (2.1), $E\|\mathbb{Y}_t - \widetilde{\mathbb{Y}}_t\| = E\|\sum_{j=m}^{\infty} G^j(\mathcal{B}_0 + \mathcal{E}_{t-j}) - G^m\widetilde{\mathbb{Y}}_{t-m}\| \le C\rho^m$, where C is a constant independent of t and m. Note that m is chosen arbitrarily. Hence, 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 theorem proof.

3. Proof of Theorem 4. Note that $E|\mathcal{B}_0^* + \mathcal{E}_{t-j}^*|_e \preccurlyeq (|\beta_0| + E|Z_i^\top\gamma| + E|\varepsilon_{it}|)\mathbf{1}_{Np}$. Let $\rho = \sum_{m=1}^{p}\rho_m$, where $\rho_m = |\alpha_m| + |\beta_m|$. We then can verify that $|G^*|_e\mathbf{1}_{Np} = (\rho\mathbf{1}_N^\top, \mathbf{1}_{N(p-1)}^\top)^\top$, and $|G^{*2}|_e\mathbf{1}_{Np} = \{(\rho\rho_1 + \sum_{m=2}^{p}\rho_m)\mathbf{1}_N^\top, \rho\mathbf{1}_N^\top, \mathbf{1}_{N(p-2)}^\top\}^\top \preccurlyeq (\rho\mathbf{1}_{2N}^\top, \mathbf{1}_{N(p-2)}^\top)^\top$. Similarly, $|G^{*n}|_e\mathbf{1}_{Np} \preccurlyeq \rho\mathbf{1}_{Np}$ for $n = 3, ..., p$. As a result, the rest follows the proofs of Theorems 1 and 2 by noting that $\rho < 1$ and $(I - G^*)^{-1}\mathcal{B}_0^* = (I - \widetilde{G})^{-1}\mathcal{B}_0$.

4. Proof of Theorem 5. From Theorem 4, we have $\mathbb{Y}_t = \mathfrak{I}\mathbb{Y}_t^* = (I - \widetilde{G})^{-1}\mathcal{B}_0 + \sum_{j=0}^{\infty}\mathfrak{I}G^{*j}\mathcal{E}_{t-j}^* = (I - \widetilde{G})^{-1}\mathcal{B}_0 + \sum_{j=0}^{\infty}\mathfrak{I}G^{*j}\mathfrak{I}^\top\mathcal{E}_{t-j}$, and hence

$$\mathbb{Y}_t = c_\beta^*\mathbf{1}_p + (I - \widetilde{G})^{-1}\mathbb{Z}\gamma + \sum_{j=0}^{\infty}\mathfrak{I}G^{*j}\mathfrak{I}^\top\mathcal{E}_{t-j}, \tag{4.1}$$

which is in the same form as the decomposition of (A.3) for the NAR model. We then establish similar results as in Lemma 2, therefore Theorem 5 can be proved subsequently as Theorem 3. Note that

$$|G^*|_e = \begin{pmatrix} |\alpha_1|W + |\beta_1|I & \cdots & |\alpha_p|W + |\beta_p|I \\ & I_{N(p-1)} & O_{N(p-1),N} \end{pmatrix}.$$

Let $\overline{\mathfrak{I}} = \mathbf{1}_p^\top \otimes I_N$, and then $\overline{\mathfrak{I}}|G^*|_e^n\overline{\mathfrak{I}}^\top$ has a polynomial form of

$$\overline{\mathfrak{I}}|G^*|_e^n\overline{\mathfrak{I}}^\top = \sum_{m=1}^{n} a_m^{(n)}W^n, \tag{4.2}$$

where $W^0 = I$ and $a_m^{(n)}$s $(1 \le k \le p)$ are nonnegative coefficients.

We then derive an upper bound for $\sum_{m=1}^{n} a_m^{(n)}$, thus the upper bound for $|\mathfrak{I}G^{*n}\mathfrak{I}^\top|_e$ and $|\mathfrak{I}G^{*n}\mathfrak{I}^\top\mathfrak{I}(G^{*\top})^n\mathfrak{I}^\top|_e$ can be established. To this end, define a $p \times p$ matrix function

$$\mathfrak{G} = \begin{pmatrix} |\alpha_1| + |\beta_1| & \cdots & |\alpha_p| + |\beta_p| \\ & I_{p-1} & 0 \end{pmatrix},$$

and it can be verified that $\mathbf{1}_p^\top \mathfrak{G}^n \mathbf{1}_p = \sum_{m=1}^n a_m^{(n)}$, where the coefficients $a_m^{(n)}$s are the same as in (4.2). Moreover, we have $\sum_{m=1}^n a_m^{(n)} = \mathbf{1}_p^\top \mathfrak{G}^n \mathbf{1}_p \leq C\rho^{n/p}$, where C is a constant. Note that $\mathfrak{I} \preccurlyeq \overline{\mathfrak{I}}$. Together with (5.7), we can establish the upper bound for $|\mathfrak{I}G^{*n}\mathfrak{I}^\top \mathfrak{I}(G^{*\top})^n \mathfrak{I}^\top|_e$, which is similar to (B.1) in the proof of Lemma 2.

To extend (b) and (c) in Lemma 2, note that for integers k_1, k_2, and $n \geq 0$, define

$$g^*_{j,k_1,k_2}(G^*, W) = |W^{k_1}\{\mathfrak{I}G^{*j}\mathfrak{I}^\top \mathfrak{I}(G^{*\top})^j \mathfrak{I}^\top\}^{k_2}(W^\top)^{k_1}|_e \in \mathbb{R}^{N\times N},$$

and

$$f^*_{k_1,k_2}(W, \widetilde{Q}) = |W^{k_1}\widetilde{Q}^{k_2}(W^\top)^{k_1}|_e \in \mathbb{R}^{N\times N}.$$

It is noteworthy that $g^*_{j,k_1,k_2}(G^*, W)$ and $f^*_{k_1,k_2}(W, \widetilde{Q})$ are $g_{j,k_1,k_2}(G, W)$ and $f_{k_1,k_2}(W, Q)$ by replacing G^j and Q by $\mathfrak{I}G^{*j}\mathfrak{I}^\top$ and $\widetilde{Q}$, respectively. Similarly we can establish the same results as (b) and (c) in Lemma 2, and the rest follows the proofs of Theorem 3.

5. Two 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* $n^{-2}tr(A A^\top) \to 0$ *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) $(nT)^{-1}\sum_{t=1}^T \widetilde{\mathbb{Y}}_t^\top A \widetilde{\mathbb{Y}}_t^\top \to_p \sigma^2 \lim_{n\to\infty} n^{-1}tr\{A\Gamma(0)\}$ *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* $n^{-1}\sum_{j=0}^\infty [tr\{AG^j UU^\top (G^\top)^j A^\top\}]^{1/2} \to 0$ *as* $n \to \infty$.

Lemma 2. *Assume the stationary condition $|\beta_1| + |\beta_2| < 1$. 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 $|M|_e$ as $|M|_e = (|m_{ij}|) \in \mathbb{R}^{n\times p}$ for any arbitrary matrix $M = (m_{ij}) \in \mathbb{R}^{n\times p}$. Then there exists $K > 0$, such that*
(a) for any integer $n > 0$, we have

$$|G^n(G^\top)^n|_e \preccurlyeq n^K(|\beta_1| + |\beta_2|)^{2n} MM^\top, \tag{5.1}$$

where $M = C\mathbf{1}\pi^\top + \sum_{j=0}^K W^j$, $C > 1$ is a constant, and π is defined in (C2.1).
(b) For integers k_1, k_2, and $j \ge 0$, define $g_{j,k_1,k_2}(G, W) = |W^{k_1}\{G^j(G^\top)^j\}^{k_2}(W^\top)^{k_1}|_e \in \mathbb{R}^{N\times N}$. For integers $0 \le k_1, k_2, m_1, m_2 \le 1$, as $N \to \infty$ we have

$$N^{-1}\sum_{j=0}^{\infty}\Big\{\mathbf{1}^\top g_{j,k_1,k_2}(G, W)\mathbf{1}\Big\}^{1/2} \to 0, \tag{5.2}$$

$$N^{-1}\sum_{i,j=0}^{\infty}\Big[tr\Big\{g_{i,k_1,k_2}(G, W)g_{j,m_1,m_2}(G, W)\Big\}\Big]^{1/2} \to 0. \tag{5.3}$$

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

$$N^{-2}\mathbf{1}^\top f_{k_1,k_2}(W, Q)\mathbf{1} \to 0, \tag{5.4}$$

$$N^{-2}tr\Big\{f_{k_1,k_2}(W, Q)f_{m_1,m_2}(W, Q)\Big\} \to 0, \tag{5.5}$$

$$N^{-1}\sum_{j=0}^{\infty}\Big[tr\Big\{f_{k_1,k_2}(W, Q)g_{j,m_1,m_2}(G, W)\Big\}\Big]^{1/2} \to 0, \tag{5.6}$$

5.1. *Proof of Lemma 1.* We prove the five conclusions in Lemma 1 one by one.
Proof of (a). Since $E(n^{-1}B^\top X)^2 = n^{-2}\sigma_X^2 B^\top B$, conclusion (a) holds.
Proof of (b). note that $X^\top AX = \sum_{i=1}^n \sum_{j=1}^n a_{ij}X_iX_j$. It can be verified

that $E(X^\top AX) = \sigma_X^2 \sum_{j=1}^n a_{jj} = \sigma_X^2 \mathrm{tr}(A)$, and

$$
\begin{aligned}
E(X^\top AX)^2 &= \sum_{i=1}^n \sum_{j=1}^n \sum_{l=1}^n \sum_{m=1}^n a_{ij}a_{lm}E(X_iX_jX_lX_m) \\
&= \sigma_X^4 \sum_{i\neq j} a_{ii}a_{jj} + \sigma_X^4 \sum_{i\neq j} a_{ij}^2 + \sigma_X^4 \sum_{i\neq j} a_{ij}a_{ji} + \sum_{i=1}^n a_{ii}^2 EX_i^4 \\
&\leq \sigma_X^4 \Big(\sum_{i,j} a_{ii}a_{jj}\Big) + 2\sigma_X^4 \Big(\sum_{i,j} a_{ij}^2\Big) + \sum_i a_{ii}^2 \Big\{EX_i^4 - 3\sigma_X^4\Big\}
\end{aligned}
$$

where the last inequality is due to $\sum_{i\neq j} a_{ij}a_{ji} \leq 0.5\sum_{i\neq j}(a_{ij}^2 + a_{ji}^2) = \sum_{i\neq j} a_{ij}^2$. Since we have $\sigma_X^4 \sum_{i,j} a_{ii}a_{jj} = \{E(X^\top AX)\}^2$, $\sum_{i,j} a_{ij}^2 = \mathrm{tr}(AA^\top)$, and $\sum_{i=1}^n a_{ii}^2 \leq \mathrm{tr}(AA^\top)$, then it can be derived that $E(X^\top AX)^2 \leq \{E(X^\top AX)\}^2 + C_1\mathrm{tr}(AA^\top)$, where $C_1 = 2\sigma_X^4 + |EX_i^4 - 3\sigma_X^4|$. Then $n^{-2}\mathrm{var}(X^\top AX) \leq C_1 n^{-2}\mathrm{tr}(AA^\top) \to 0$, and (b) holds.

Proof of (c). Denote $\mathcal{S}_a = (nT)^{-1}\sum_{t=1}^T B^\top \widetilde{\mathbb{Y}}_t^\top = (nT)^{-1}\sum_{j=0}^\infty \sum_{t=1}^T B^\top G^j U\mathcal{E}_{t-j}$, then it holds that $E(\sum_{t=1}^T B^\top G^j U \mathcal{E}_{t-j})^2 = TB^\top G^j UU^\top (G^\top)^j B$. By applying Minkowski inequality, as $n \to \infty$ we have

$$
\|\mathcal{S}_a\|_2 \leq \frac{1}{nT}\sum_{j=0}^\infty \Big\|\sum_{t=1}^T B^\top G^j U \mathcal{E}_{t-j}\Big\|_2 = \frac{\sigma^2}{n\sqrt{T}}\sum_{j=0}^\infty \big\{B^\top G^j UU^\top (G^\top)^j B\big\}^{1/2} \to 0,
$$

which implies conclusion (c), where $\|X\|_2 = (EX^2)^{1/2}$ is the L_2 norm.

Proof of (d). Let $\mathcal{S}_b = (nT)^{-1}\sum_{t=1}^T \widetilde{\mathbb{Y}}_t^\top A \widetilde{\mathbb{Y}}_t = (nT)^{-1}\sum_{i,j=0}^\infty \sum_{t=1}^T \xi_t(i,j)$, where $\xi_t(i,j) = \mathcal{E}_{t-i}^\top U^\top (G^\top)^i A G^j U \mathcal{E}_{t-j}$. Similar to the proof of conclusion (b), we can show that $E\{\xi_t(j,j)\} = \sigma^2 \mathrm{tr}\{U^\top (G^\top)^j A G^j U\}$ and $\mathrm{var}\{\xi_t(j,j)\} \leq C_2 \mathrm{tr}\{U^\top (G^\top)^j A G^j UU^\top (G^\top)^j A^\top G^j U\}$, where $C_2 = 2\sigma^4 + |E(\varepsilon_{it}^4) - 3\sigma^4|$. When $i \neq j$, it can be verified that $E\xi_t(i,j) = 0$ and $E\{\xi_t(i,j)^2\} = \sigma^4 \mathrm{tr}\{U^\top (G^\top)^i A G^j UU^\top (G^\top)^j A^\top G^i U\} \leq C_2 \mathrm{tr}\{U^\top (G^\top)^i A G^j UU^\top (G^\top)^j A^\top G^i U\}$. As a result, we have $E\{\sum_{t=1}^T \xi_t(i,j) - E\xi_t(i,j)\}^2 \leq TC_2 \mathrm{tr}\{U^\top (G^\top)^i A G^j UU^\top (G^\top)^j A^\top G^i U\}$. Then, by Minkowski inequality

$$
\begin{aligned}
\|\mathcal{S}_b - E\mathcal{S}_b\|_2 &\leq \frac{1}{nT}\sum_{i=0}^\infty \sum_{j=0}^\infty \Big\|\sum_{t=1}^T \xi_t(i,j) - E\xi_t(i,j)\Big\|_2 \\
&\leq \frac{\sqrt{C_2}}{n\sqrt{T}}\sum_{i=0}^\infty \sum_{j=0}^\infty \Big[\mathrm{tr}\big\{U^\top (G^\top)^i A G^j UU^\top (G^\top)^j A^\top G^i U\big\}\Big]^{1/2} \to 0
\end{aligned}
$$

as $n \to \infty$. In addition, we have $E\mathcal{S}_b = n^{-1}\sigma^2 \sum_{j=0}^\infty \mathrm{tr}\{U^\top (G^\top)^j A G^j U\} = n^{-1}\mathrm{tr}\{A\Gamma(0)\}$. Thus conclusion (d) holds.

Proof of (e). It is straight forward that $\mathcal{S}_c = (nT)^{-1}\sum_{t=1}^{T} X^\top A\widetilde{\mathbb{Y}}_t = (nT)^{-1}\sum_{j=0}^{\infty}\sum_{t=1}^{T} X^\top AG^jU\mathcal{E}_{t-j}$. note that $E(\sum_{t=1}^{T} X^\top AG^jU\mathcal{E}_{t-j})^2 = T\sigma^2$ $E\{X^\top AG^jUU^\top(G^\top)^jA^\top X\} = T\sigma^2\sigma_X^2\mathrm{tr}\{AG^jUU^\top(G^\top)^jA^\top\}$. Similar to the proof of conclusion (c), by Minkowski inequality, $\|\mathcal{S}_c\|_2 \le (n\sqrt{T})^{-1}\sigma\sigma_X$ $\sum_{j=0}^{\infty}[\mathrm{tr}\{AG^jUU^\top(G^\top)^jA^\top\}]^{1/2} \to 0$ as $n\to\infty$. As a result, conclusion (e) holds.

5.2. *Proof of Lemma 2.* Firstly, by [2], for the irreducible and aperiodic Markov chains in (C2.1) with transition probability matrix W, we have $\lim_{n\to\infty} W^n = \mathbf{1}\pi^\top$, where π is the stationary distribution vector defined in (C2.1). As a result, it can be concluded that there exists an integer $K>0$, for $n>K$ we have

$$W^n \preccurlyeq C\mathbf{1}\pi^\top, \tag{5.7}$$

where $C>1$ is a constant. We then prove (a)-(c) in Lemma 2 one by one as follows.

Proof of (a). Firstly, for any integer $n>0$, we have $G^n = (\beta_1W+\beta_2I)^n = \sum_{j=0}^{n} C_n^j\beta_1^j\beta_2^{n-j}W^j$, where $C_n^j = n!/\{j!(n-j)!\}$. Since W is an element-wise non-negative matrix, $|G^n|_e \preccurlyeq \sum_{j=0}^{n} C_n^j|\beta_1|^j|\beta_2|^{n-j}W^j$. Then for $n>K$ we have $|G^n|_e \preccurlyeq$

$$\begin{aligned}
\sum_{j=0}^{n} C_n^j|\beta_1|^j|\beta_2|^{n-j}W^j &= \sum_{j=K+1}^{n} C_n^j|\beta_1|^j|\beta_2|^{n-j}W^j + \sum_{j=0}^{K} C_n^j|\beta_1|^j|\beta_2|^{n-j}W^j\\
&\preccurlyeq \sum_{j=K+1}^{n} C_n^j|\beta_1|^j|\beta_2|^{n-j}C\mathbf{1}\pi^\top + \sum_{j=0}^{K} C_n^j|\beta_1|^j|\beta_2|^{n-j}W^j\\
&\preccurlyeq \Big(\sum_{j=0}^{n} C_n^j|\beta_1|^j|\beta_2|^{n-j}\Big)C\mathbf{1}\pi^\top + \sum_{j=0}^{K} C_n^j|\beta_1|^j|\beta_2|^{n-j}W^j\\
&= (|\beta_1|+|\beta_2|)^nC\mathbf{1}\pi^\top + \sum_{j=0}^{K} C_n^j|\beta_1|^j|\beta_2|^{n-j}W^j, \qquad (5.8)
\end{aligned}$$

where the second inequality is due to $W^n \preccurlyeq C\mathbf{1}\pi^\top$. Further note that $|\beta_1|^j|\beta_2|^{n-j} < (|\beta_1|+|\beta_2|)^n$ $(0\le j\le n)$, and $C_n^K \le n^K$. As a result, for $n>K$ we have,

$$|G^n|_e \preccurlyeq n^K(|\beta_1|+|\beta_2|)^nM, \tag{5.9}$$

where $M = C\mathbf{1}\pi^\top + \sum_{j=0}^{K} W^j$ is defined in (a) of Lemma 2. It is easy to verify that (5.9) also holds for $n = 1,\cdots,K-1$. Then we have $|G^n(G^\top)^n|_e \preccurlyeq$

$n^{2K}(|\beta_1|+|\beta_2|)^{2n}MM^\top$ for any positive integer n. As a result, (5.1) can be proved.

PROOF OF (b). First of all, let $k_1 = k_2 = 1$ and then $g_{j,1,1}(G,W) = |WG^j(G^\top)^jW^\top|_e$. We first prove (5.2). By (5.9) we have $|WG^j|_e \preccurlyeq j^K(|\beta_1|+|\beta_2|)^jWM$ and $WM = C\mathbf{1}\pi^\top + \sum_{j=1}^{K+1}W^j$, where the equality is due to $W\mathbf{1}\pi^\top = \mathbf{1}\pi^\top$. As a result, we have

$$|WG^j(G^\top)^jW^\top|_e \preccurlyeq j^{2K}(|\beta_1|+|\beta_2|)^{2j}\mathcal{M}, \tag{5.10}$$

where $\mathcal{M}$ is defined as $\mathcal{M} = WMM^\top W^\top$. As a result, we have $\sum_{j=0}^\infty N^{-1}\big\{\mathbf{1}^\top WG^j(G^\top)^jW^\top\mathbf{1}\big\}^{1/2} \le N^{-1}\alpha(\mathbf{1}^\top\mathcal{M}\mathbf{1})^{1/2}$, where $\alpha = \sum_{j=0}^\infty j^K(|\beta_1|+|\beta_2|)^j < \infty$. Then it leads to show $N^{-2}\mathbf{1}^\top\mathcal{M}\mathbf{1} \to 0$. It can be verified $\mathbf{1}^\top\mathcal{M}\mathbf{1} = N^2C\sum_j\pi_j^2+\sum_{j=1}^{K+1}\mathbf{1}^\top W^j(W^\top)^j\mathbf{1}+2NC\sum_j\pi^\top(W^\top)^j\mathbf{1}+\sum_{i\neq j}\mathbf{1}^\top W^i(W^\top)^j\mathbf{1}$. For the last two terms of $\mathbf{1}^\top\mathcal{M}\mathbf{1}$, by Cauchy inequality, we have

$$\begin{aligned} N\sum_j\pi^\top(W^\top)^j\mathbf{1} &\le N\Big(\sum_j\pi_j^2\Big)^{1/2}\Big\{\mathbf{1}^\top W^j(W^\top)^j\mathbf{1}\Big\}^{1/2}, \\ \sum_{i\neq j}\mathbf{1}^\top W^i(W^\top)^j\mathbf{1} &\le \sum_{i\neq j}\Big\{\mathbf{1}^\top W^i(W^\top)^i\mathbf{1}\Big\}^{1/2}\Big\{\mathbf{1}^\top W^j(W^\top)^j\mathbf{1}\Big\}^{1/2}. \end{aligned}$$

As a result, it leads to show

$$\sum_{j=1}^N\pi_j^2 \to 0 \quad \text{and} \quad N^{-2}\mathbf{1}^\top W^j(W^\top)^j\mathbf{1} \to 0 \tag{5.11}$$

for $1 \le j \le K+1$. As the first convergence in (5.11) is implied by (C2.1), we next prove $N^{-2}\mathbf{1}^\top W^j(W^\top)^j\mathbf{1} \to 0$ $(1 \le j \le K+1)$. Recall that $W^* = W + W^\top$. Thus it can be easily verified $W \preccurlyeq W^*$ and $W^\top \preccurlyeq W^*$. As a result, we have $N^{-2}\mathbf{1}^\top W^j(W^\top)^j\mathbf{1} \le N^{-2}\mathbf{1}^\top W^{*2j}\mathbf{1}$. Then it suffices to show $N^{-2}\mathbf{1}^\top W^{*2j}\mathbf{1} \to 0$. By eigendecomposition of W^* we have $W^* = \sum_k\lambda_k(W^*)u_ku_k^\top$, 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 $N^{-2}\mathbf{1}^\top W^{*2j}\mathbf{1} = N^{-2}\sum_k\lambda_k(W^*)^{2j}(\mathbf{1}^\top u_k)^2 \le N^{-2}\lambda_{\max}(W^*)^{2j}\sum_k(\mathbf{1}^\top u_k)^2 = N^{-1}\lambda_{\max}(W^*)^{2j}$ $(1 \le j \le K+1)$, where the last equality is due to $\sum_k(\mathbf{1}^\top u_k)^2 = N$. Since we have $\lambda_{\max}(W^*) = O(\log N)$, it then leads to the conclusion that $N^{-1}\lambda_{\max}(W^*)^{2j} \to 0$ for $1 \le j \le K+1$. As a consequence, the second term in (5.11) holds. Similarly, it can be proved that (5.11) holds for all $0 \le k_1, k_2 \le 1$. As a result, we have (5.2) holds.

We next prove (5.3) with $k_1 = k_2 = m_1 = m_2 = 1$, and $g_{i,1,1}(G,W)g_{j,1,1}(G,W) = |WG^i(G^\top)^iW^\top WG^j(G^\top)^jW^\top|_e$. 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 (5.10), we have

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

It then can be derived that $N^{-1}\sum_{i,j=0}^{\infty}[\mathrm{tr}\{WG^i(G^\top)^iW^\top WG^j(G^\top)^jW^\top\}]^{1/2}$ $\le \alpha^2 N^{-1}\mathrm{tr}\big\{\mathcal{M}^2\big\}^{1/2}$. In order to obtain (5.3), it suffices to show that

$$N^{-2}\mathrm{tr}\{\mathcal{M}^2\} \to 0. \tag{5.12}$$

Equivalently, by Cauchy inequality, it suffices to prove $(\sum \pi_j^2)^2 \to 0$, and $N^{-2}\mathrm{tr}\{W^j{W^j}^\top W^j{W^j}^\top\} \to 0$ holds for $1 \le j \le K+1$. It can be easily verified the first term holds by (C2.1). For the second one, we have $N^{-2}\mathrm{tr}\{W^j{W^j}^\top W^j{W^j}^\top\} \le N^{-2}\mathrm{tr}\{(W^*)^{4j}\} = N^{-2}\sum_k \lambda_k(W^*)^{4j} \le N^{-1}$ $\lambda_{\max}(W^*)^{4j} \le N^{-1}\lambda_{\max}(W^*)^{4(K+1)}$. Similarly, due to that $\lambda_{\max}(W^*) = O(\log N)$ in (C2.2), we have $N^{-1}\lambda_{\max}(W^*)^{4(K+1)} \to 0$ as $N \to \infty$. Consequently, we have (5.12) and then (5.3) holds. This completes the proof of (b).

Proof of (c). Similarly, let $k_1 = k_2 = 1$, and $f_{1,1}(W,Q) = |WQW^\top|_e = |W(I-G)^{-1}(I-G^\top)^{-1}W^\top|_e$. We first establish the upper bound for $|WQW^\top|_e$. Note that $(I-G)^{-1} = \sum_{j=0}^{\infty} G^j$. As a result, for any integer $k > 0$ we have $|W(I-G)^{-1}|_e =$

$$|W(\sum_{j=0}^{\infty} G^j)|_e \preccurlyeq \sum_{j=0}^{\infty}|WG^j|_e \preccurlyeq \sum_{j=0}^{\infty} j^K(|\beta_1|+|\beta_2|)^j WM = \alpha WM,$$

where the last inequality is due to (5.9). Then we have $|W(I-G)^{-1}|_e \preccurlyeq \alpha WM$. As a result, it can be derived that

$$\begin{aligned}|WQW^\top|_e &\preccurlyeq |W(I-G)^{-1}|_e|W(I-G)^{-1}|_e^\top \\ &\preccurlyeq \alpha^2 WMM^\top W^\top = \alpha^2\mathcal{M}.\end{aligned} \tag{5.13}$$

Consequently, (5.4) and (5.5) can be proved due to (5.11) and (5.12).

At last, we only prove (5.6) with $k_1 = k_2 = m_1 = m_2 = 1$. By (5.10) and (5.13), it can be obtained that

$$\begin{aligned}\frac{1}{N}\sum_j \Big[\mathrm{tr}\Big\{WG^j(G^\top)^jW^\top WQW^\top\Big\}\Big]^{1/2} &\le \tfrac{\alpha}{N}\textstyle\sum_j j^K(|\beta_1|+|\beta_2|)^j\mathrm{tr}\big\{\mathcal{M}^2\big\}^{1/2} \\ &= \alpha^2\Big[N^{-2}\mathrm{tr}\big\{\mathcal{M}^2\big\}\Big]^{1/2} \to 0,\end{aligned}$$

where the last convergence is due to (5.12). As a result, one can obtain (5.6). This completes the proof of Lemma 2.

6. Proof of Proposition 2. Let $\zeta_i = N^{-1/2}T^{-1}\eta^\top \sum_t X_{i(t-1)}^\top \varepsilon_{it}$, $\mathbb{S}_N = \sum_{i=1}^N \zeta_i$, and $\mathcal{F}_i = \sigma\{\varepsilon_{jt}, 1 \le j \le i, 0 \le t \le T\}$. Therefore we have $E(\zeta_{i+1}|\mathcal{F}_i) = 0$. Consequently, $\{\mathbb{S}_i, \mathcal{F}_i, 1 \le i \le N\}$ is a martingale sequence. By a method similar to that of Step 2 in Appendix B.2, we can show that $\mathbb{S}_N = \sqrt{N}\hat{\Sigma}_{xe} \to_d N(0, \sigma^2 T^{-1}\Sigma)$ as $N \to \infty$, which, together with (A.1), completes the proof.

7. Verification of Conditions (C2)-(C3). We devote this section to verify the technical conditions (C2) and (C3) for the simulation studies and the real data analysis.

7.1. *Connectivity Analysis (I).* Condition (C2.1) assumes that the network W is fully connected within a finite number of steps. To check this, we conduct a simulation analysis as follows. Fix $N = 1000$. We then generate the network structure as the three simulation examples in Section 4.1. Specifically, for stochastic block model we set block number $K = 4$, and for power-law distribution model we set $\alpha = 3$. Other parameters remains the same as in Section 4.1. Write $W^n = (w_{ij}^{(n)})$, where $w_{ij}^{(n)}$ represents the probability that the i, j could connect with each other in the n step. We then calculate the network density as $\mathrm{ND}_n = \sum_{i,j} I(w_{ij}^{(n)} > 0)/N^2$. For a reliable evaluation, we replicate the experiment 1000 times and the ND_n value are averaged for each $1 \le n \le 10$. We then report these averaged values in Figure 1. We find that, even though the original network density (i.e., ND_1) is very low, however, it increases rapidly. After (for example) 6 steps, the whole network becomes pretty much fully connected with ND_n values extremely close to 100%. This corroborates with the six degrees of separation theory well [3]. Similar analysis is also conducted for the real data analysis and the findings are extremely similar.

7.2. *Connectivity Analysis (II).* Condition (C2.1) also assumes that $\sum_j \pi_j^2$ should converge to 0 as $N \to \infty$. We then simulate network structures as in Appendix C.1. However, the difference is that we allow N to increase from $N = 200$ to $N = 1000$. For each simulated network with size N, we compute its $\Pi(N) = \sum_j \pi_j^2$ value. Then, the experiment is randomly replicated for 1000 times and the resulting $\Pi(N)$ values are boxplotted for each N. The detailed results are given in Figure 2. We do find that a clear pattern for $\Pi(N) \to 0$ as $N \to \infty$. For the real data example, the network size N is fixed. However, we can still compute its $\Pi(N)$ value, which is given by 0.002 and seems extremely small.

7.3. *Uniformity.* Condition (C3) requires that $\lambda_{\max}(W^*) < C \log N$. We then replicate the three simulation examples as in Appendix C.2. However, for each simulated network structure, we compute $\tau_{\max}(N) = \lambda_{\max}(W^*)/\log(N)$ value. Then, the 1000 randomly replicated $\tau_{\max}(N)$ values are boxplotted for each N in Figure 3. We find that they are all well bounded by a constant $C = 2$. For the real dataset, the $\tau_{\max}(N)$ values is given by 0.41. It is also bounded by $C = 2$.

7.4. *Law of Large Numbers.* We then consider how to verify the law of large number type conditions in (C3). Once again, we simulate the data according to the three simulation examples in Section 4.2 with identical parameter setup. In order to verify this condition, one natural way is to compute those κ-values according to the analytical formula given in (C3). However, we find this task computationally extremely challenging. This is because it involves $\Gamma(0)$. The computation of $\Gamma(0)$ depends on $G \otimes G$, which is a $N^2 \times N^2$ matrix; see (2.5). This makes the computation extremely expensive. One natural solution to this problem is to replace $\Gamma(0)$ by its sample estimate, that is $\hat{\Gamma}(0) = T^{-1}\sum_t(\mathbb{Y}_t - \bar{\mathbb{Y}})(\mathbb{Y}_t - \bar{\mathbb{Y}})^\top$ and $\bar{\mathbb{Y}} = \sum_t \mathbb{Y}_t/T$. By treating $\hat{\Gamma}(0)$ as if it were $\Gamma(0)$, then the quantities in condition (C3) (e.g., $N^{-1}\mathrm{tr}\{\Gamma(0)\}$ and so on) can be computed. We randomly replicate the experiment for 1000 times, and then boxplot these quantities in Figure 4, 5, and 6. A clear convergence pattern can be detected.

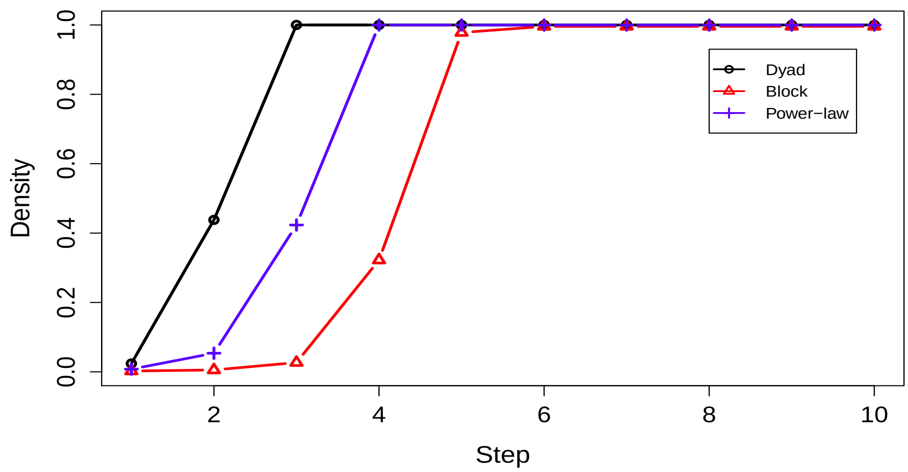

FIG 1. *Connectivity Analysis: black (○) for dyad independence model, red (△) for stochastic block model, and blue (+) for power-law distribution model. After 6 steps, the whole network becomes fully connected with ND values close to 100%.*

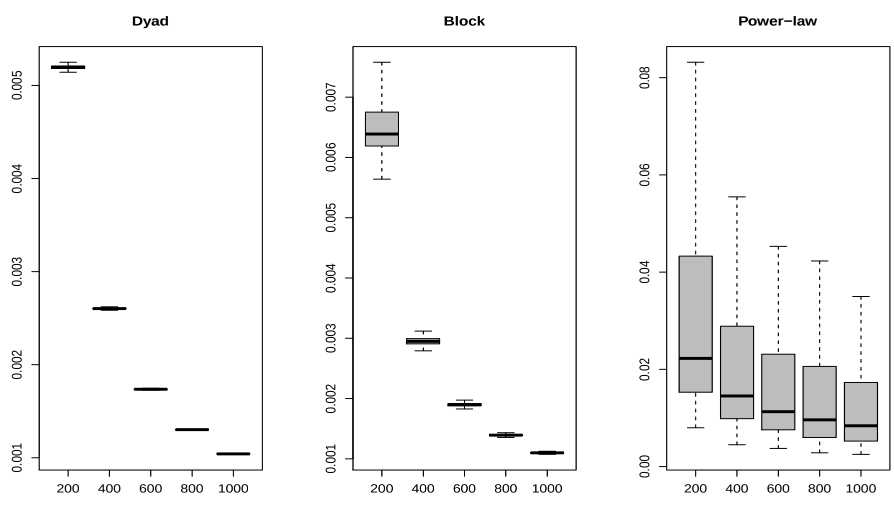

FIG 2. $\Pi(N)$ *versus* N. *The left panel for dyad independence model; The middle panel for stochastic block model; The right panel for power-law distribution model. There is a clear pattern that* $\Pi(N) \to 0$ *as* $N \to \infty$.

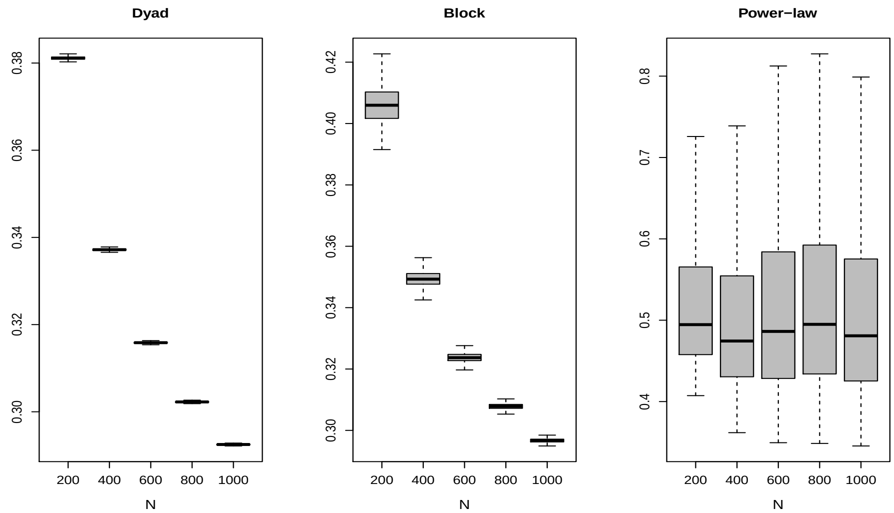

Fig 3. *$\tau_{\max}(N)$ versus N. The left panel for dyad independence model; The middle panel for stochastic block model; The right panel for power-law distribution model. They are all well bounded by a constant $C = 2$.*

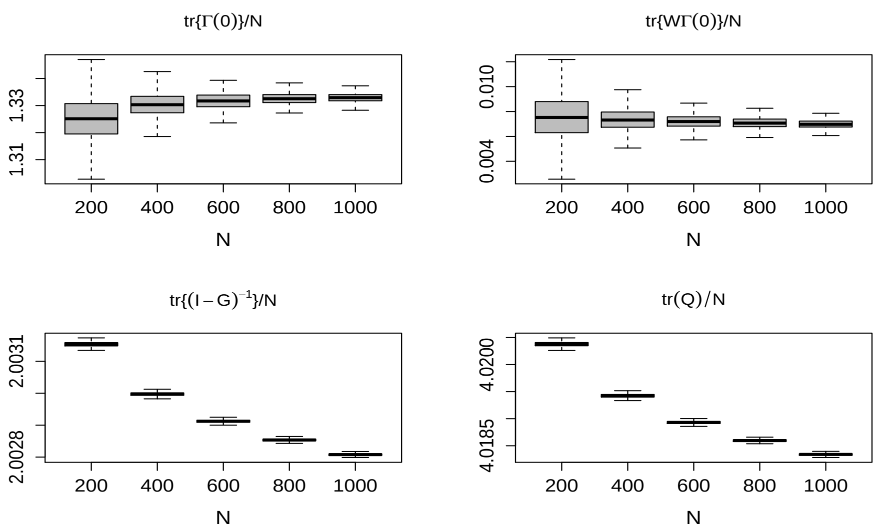

Fig 4. *Quantities of (C3) for dyad independence model. The top left panel for $N^{-1}tr\{\Gamma(0)\}$; The top right panel for $N^{-1}tr\{W\Gamma(0)\}$; The bottom left panel for $N^{-1}tr\{(I-G)^{-1}\}$; The bottom right panel for $N^{-1}tr(Q)$. A clear convergence pattern can be detected as $N \to \infty$.*

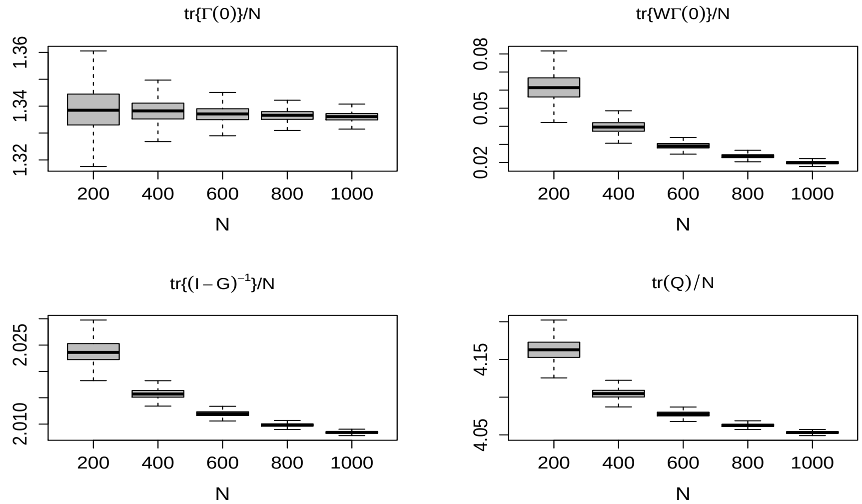

Fig 5. *Quantities of (C3) for stochastic block model. The top left panel for* $N^{-1}tr\{\Gamma(0)\}$*; The top right panel for* $N^{-1}tr\{W\Gamma(0)\}$*; The bottom left panel for* $N^{-1}tr\{(I-G)^{-1}\}$*; The bottom right panel for* $N^{-1}tr(Q)$*. A clear convergence pattern can be detected as* $N \to \infty$*.*

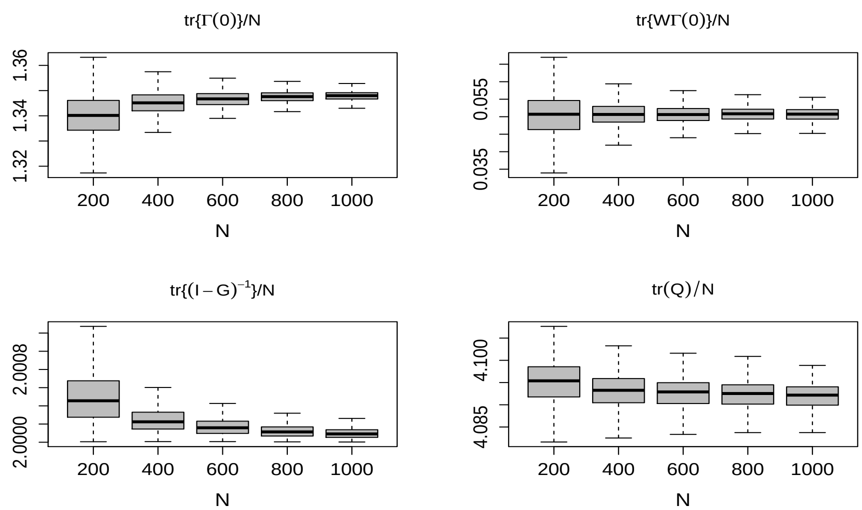

Fig 6. *Quantities of (C3) for power-law distribution model. The top left panel for* $N^{-1}tr\{\Gamma(0)\}$*; The top right panel for* $N^{-1}tr\{W\Gamma(0)\}$*; The bottom left panel for* $N^{-1}tr\{(I-G)^{-1}\}$*; The bottom right panel for* $N^{-1}tr(Q)$*. A clear convergence pattern can be detected as* $N \to \infty$*.*

References.

[1] Banerjee, S., Carlin, B. P. and Gelfand, A. E. (2014). *Hierarchical Modeling and Analysis for Spatial Data*, London, CRC Press.

[2] Meyn, S. P. and Tweedie, R. L. (2012). *Markov Chains and Stochastic Stability*, New York, Springer Science and Business Media.

[3] Newman, M., Barabasi, A. L. and Watts, D. J. (2006). *The Structure and Dynamics of Networks*, New Jersey, Princeton University Press.