# Supplementary Material for "Mixed Domain Asymptotics for Geostatistical Processes"

Tingjin Chu

*School of Mathematics and Statistics, University of Melbourne*

## S1 Appendix A: Proof of Theorem 1–2

Here, we denote $\ell'(\boldsymbol{\theta}) = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}}$, $\ell'(\boldsymbol{\theta}_0) = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$, $\ell''(\boldsymbol{\theta}, \boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}}$, and $\ell''(\boldsymbol{\theta}_0, \boldsymbol{\theta}_0) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$ for ease of presentation. For any square $\boldsymbol{A}$ and $\boldsymbol{B}$, the two inequalities $\|\boldsymbol{AB}\|_F \leq \|\boldsymbol{A}\|_2 \|\boldsymbol{B}\|_F$ and $|\mathrm{tr}(\boldsymbol{AB})| \leq \|\boldsymbol{A}\|_F \|\boldsymbol{B}\|_F$ hold.

*Proof of Theorem 1.* To prove consistency, it suffices to show that, for any given constant $\epsilon > 0$, there is a constant $C$, such that

$$P\left\{ \sup_{|u_k| = C t_{kk,n}^{-1/2}} \ell(\boldsymbol{\theta}_0 + \boldsymbol{u}) < \ell(\boldsymbol{\theta}_0) \right\} \geq 1 - \epsilon \tag{S1.1}$$

holds for a sufficiently large $N_n$, where $\boldsymbol{u} = (u_1, \ldots, u_q)$. Through Taylor's expansion,

$$\ell(\boldsymbol{\theta}_0 + \boldsymbol{u}) - \ell(\boldsymbol{\theta}_0) = \ell'(\boldsymbol{\theta}_0)^\top \boldsymbol{u} - \frac{1}{2} \boldsymbol{u}^\top \ell''(\widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\theta}}) \boldsymbol{u}, \tag{S1.2}$$

where $\widetilde{\boldsymbol{\theta}}$ is between $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_0 + \boldsymbol{u}$.

First, the mean and variance of $\ell'(\theta_{k0})$ are

$$\begin{aligned} E\ell'(\theta_{k0}) &= 0, \\ \mathrm{Var}\{\ell'(\theta_{k0})\} &= t_{kk,n}. \end{aligned}$$

Therefore, $\ell'(\boldsymbol{\theta}_0) = \mathcal{O}_p(t_{kk,n}^{1/2})$, and the first term of (S1.2) is $\ell'(\boldsymbol{\theta}_0)^\top \boldsymbol{u} = \mathcal{O}_p(1)$. Next, we quantify the second term of (S1.2). First, by (A2) and (A4), we have the following equations

$$\max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_0\|_2 \leq \max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \sum_{k=1}^{q} \|\boldsymbol{\Sigma}_k\|_2 \; |\theta_k - \theta_{k0}| = o(1), \tag{S1.3a}$$

$$\max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0\|_2 \leq \max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}\|_2 \; \|\boldsymbol{\Sigma}^{-1}\|_2 + 1 = \mathcal{O}(1), \tag{S1.3b}$$

$$\max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_k - \boldsymbol{\Sigma}_{k0}\|_F \leq 2 \max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \sum_{k'=1}^{q} \|\boldsymbol{\Sigma}_{kk'}\|_F \; |\theta_{k'} - \theta_{k'0}| = o(t_{kk,n}^{1/2}), \tag{S1.3c}$$

where $\mathcal{V}(\boldsymbol{\theta}_0) = \{\boldsymbol{\theta} : \|\theta_k - \theta_{k0}\|_2 \leq C t_{kk,n}^{-1/2}\}$. Furthermore, for $\boldsymbol{\theta} \in \mathcal{V}(\boldsymbol{\theta}_0)$, by matrix calculus,

$$\begin{aligned}
\|\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_0)\|_F^2 &\leq \sum_{k=1}^{q} \max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \left|\mathrm{tr}\{2\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_0)\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_k\}\right| \; |\theta_k - \theta_{k0}| \\
&\leq \sum_{k=1}^{q} 2 \max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_0)\|_F \; \max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_k\|_F \; |\theta_k - \theta_{k0}|, \\
\max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_k\|_F &\leq \max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k0}\|_F + \max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\Sigma}_k - \boldsymbol{\Sigma}_{k0})\|_F = \mathcal{O}(t_{kk,n}^{1/2}).
\end{aligned}$$

Combing the above two inequalities, we obtain

$$\max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_0)\|_F = \mathcal{O}(1). \tag{S1.4}$$

Together with (S1.3a)–(S1.3c), it yields

$$\begin{aligned}
\max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_k\|_F &\leq \max_{\boldsymbol{\theta}\in\mathcal{V}(\boldsymbol{\theta}_0)} \left\{\|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0\|_2 \; (\|\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\Sigma}_k - \boldsymbol{\Sigma}_{k0})\|_F + \|\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k0}\|_F)\right\} \\
&= \mathcal{O}(t_{kk,n}^{1/2}).
\end{aligned} \tag{S1.5}$$

The $(k, k')$th element of $\ell''(\boldsymbol{\theta}, \boldsymbol{\theta})$ is $\ell''(\theta_k, \theta_{k'}) = \frac{1}{2}\{tr(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{kk'} + \boldsymbol{\Sigma}^k\boldsymbol{\Sigma}_{k'}) + \boldsymbol{y}^\top\boldsymbol{\Sigma}^{kk'}\boldsymbol{y}\}$, and $\mathrm{var}\{\ell''(\theta_k, \theta_{k'})\} = \mathrm{tr}\{(\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}^{kk'})^2\}$. Let $\boldsymbol{A}_1 = \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_k\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{k'}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0$ and $\boldsymbol{A}_2 = \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{kk'}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0$, we have $\mathrm{tr}\{(\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}^{kk'})^2\} \leq 8\|\boldsymbol{A}_1\|_F^2 + 2\|\boldsymbol{A}_2\|_F^2$. For any $\boldsymbol{\theta} \in \mathcal{B}(\boldsymbol{\theta}_0)$, by

(A2), (A4), (S1.3b) and (S1.5),

$$\|\boldsymbol{A}_1\|_F^2 \leq \|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0\|_2^2\ \|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_k\|_2\|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{k'}\|_2\ t_{kk,n}^{1/2}t_{k'k',n}^{1/2} \leq o(t_{kk,n}t_{k'k',n}),$$
$$\|\boldsymbol{A}_2\|_F \leq \|\boldsymbol{\Sigma}^{-1}\|_2\ \|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0\|_2\ \|\boldsymbol{\Sigma}_{kk'}\|_F \leq o\left(t_{kk,n}^{1/2}t_{k'k',n}^{1/2}\right).$$

Therefore, $(t_{kk,n}t_{k'k',n})^{-1/2}[\ell''(\theta_k,\theta_{k'}) - E\{\ell''(\theta_k,\theta_{k'})\}] \xrightarrow{p} 0$.

Moreover, the $(k,k')$th element of $E\{\ell''(\boldsymbol{\theta},\boldsymbol{\theta})\}$ is $E\{\ell''(\theta_k,\theta_{k'})\} = \frac{1}{2}tr(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{kk'} + \boldsymbol{\Sigma}^k\boldsymbol{\Sigma}_{k'} + \boldsymbol{\Sigma}^{kk'}\boldsymbol{\Sigma}_0)$, and the $(k,k')$th element of $2\left[E\{\ell''(\theta_k,\theta_{k'})\} - E\{\ell''(\theta_{k0},\theta_{k'0})\}\right]$ is

$$\begin{aligned}
&= tr\{\boldsymbol{\Sigma}^k\boldsymbol{\Sigma}_{k'} + \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{kk'} + \boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}_k\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{k'} + \boldsymbol{\Sigma}_{k'}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_k - \boldsymbol{\Sigma}_{kk'})\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0\} + tr(\boldsymbol{\Sigma}_0^{k'}\boldsymbol{\Sigma}_{k0}) \\
&= tr\{\boldsymbol{\Sigma}^k\boldsymbol{\Sigma}_{k'}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}-\boldsymbol{\Sigma}_0)\} + tr\{\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{kk'}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}-\boldsymbol{\Sigma}_0)\} - tr(\boldsymbol{\Sigma}^{k'}\boldsymbol{\Sigma}_k\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_0^{k'}\boldsymbol{\Sigma}_{k0}) \\
&= (I_1) + (I_2) + (I_3),
\end{aligned}$$

where $(I_1) = tr\{\boldsymbol{\Sigma}^k\boldsymbol{\Sigma}_{k'}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}-\boldsymbol{\Sigma}_0)\}$, $(I_2) = tr\{\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{kk'}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}-\boldsymbol{\Sigma}_0)\}$, $(I_3) = -tr(\boldsymbol{\Sigma}^{k'}\boldsymbol{\Sigma}_k\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_0^{k'}\boldsymbol{\Sigma}_{k0})$, and $\boldsymbol{\Sigma}_0^{k'} = \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k'0}\boldsymbol{\Sigma}_0^{-1}$. By (A2), (A4), (S1.3a), (S1.4), and (S1.5), we have

$$\begin{aligned}
\left|(I_1)\right| &\leq \|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_k\|_F\ \|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{k'}\|_F\|\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}-\boldsymbol{\Sigma}_0)\|_2 = o\left(t_{kk,n}^{1/2}t_{k'k',n}^{1/2}\right), \\
\left|(I_2)\right| &\leq \|\boldsymbol{\Sigma}^{-1}\|_2\ \|\boldsymbol{\Sigma}_{kk'}\|_F\ \|\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}-\boldsymbol{\Sigma}_0)\|_F = o\left(t_{kk,n}^{1/2}t_{k'k',n}^{1/2}\right).
\end{aligned}$$

For $(I_3)$, it can be written as

$$\begin{aligned}
(I_3) &= tr(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{k'}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_k\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k'0}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k0}) \\
&= tr\{\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{k'}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_k(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}_0^{-1})\boldsymbol{\Sigma}_0\} + tr\{\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{k'}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}_k - \boldsymbol{\Sigma}_{k0})\} \\
&+ tr\{\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{k'}(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}_0^{-1})\boldsymbol{\Sigma}_{k0}\} + tr\{\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}_{k'} - \boldsymbol{\Sigma}_{k'0})\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k0}\} \\
&+ tr\{(\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}_0^{-1})\boldsymbol{\Sigma}_{k'0}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k0}\} \equiv (I_{3,1}) + (I_{3,2}) + (I_{3,3}) + (I_{3,4}) + (I_{3,5}).
\end{aligned}$$

Using (A2), (A4), and (S1.3a)–(S1.5), the following results are obtained

$$
\begin{aligned}
\left|(I_{3,1})\right| &\leq \|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{k'}\|_F \ \|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_k\|_F \ \|(\boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}_0^{-1})\boldsymbol{\Sigma}_0\|_2 \\
&\leq \|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{k'}\|_F \ \|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_k\|_F \ \|\boldsymbol{\Sigma}^{-1}\|_2 \ \|\boldsymbol{\Sigma}-\boldsymbol{\Sigma}_0\|_2 = o\left(t_{kk,n}^{1/2}t_{k'k',n}^{1/2}\right), \\
\left|(I_{3,2})\right| &\leq \|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{k'}\|_F \ \|\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}_k-\boldsymbol{\Sigma}_{k0})\|_F = o\left(t_{kk,n}^{1/2}t_{k'k',n}^{1/2}\right), \\
\left|(I_{3,3})\right| &\leq \|\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_{k'}\|_F \ \|(\boldsymbol{\Sigma}-\boldsymbol{\Sigma}_0)\boldsymbol{\Sigma}^{-1}\|_2 \ \|\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k0}\|_F = o\left(t_{kk,n}^{1/2}t_{k'k',n}^{1/2}\right), \\
\left|(I_{3,4})\right| &\leq \|\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Sigma}_{k'}-\boldsymbol{\Sigma}_{k'0})\|_F \ \|\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k0}\|_F = o\left(t_{kk,n}^{1/2}t_{k'k',n}^{1/2}\right), \\
\left|(I_{3,5})\right| &\leq \|(\boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}_0^{-1})\boldsymbol{\Sigma}_{k'0}\|_F \ \|\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k0}\|_F = o\left(t_{kk,n}^{1/2}t_{k'k',n}^{1/2}\right).
\end{aligned}
$$

Therefore, $(I_3) = o\left(t_{kk,n}^{1/2}t_{k'k',n}^{1/2}\right)$, and $E\{\ell''(\theta_k,\theta_{k'})\} - E\{\ell''(\theta_{k0},\theta_{k'0})\} = o\left(t_{kk,n}^{1/2}t_{k'k',n}^{1/2}\right)$.

Let $\boldsymbol{D} = diag(t_{11,n}^{1/2},\ldots,t_{qq,n}^{1/2})$, and by definition, $E\{\ell''(\boldsymbol{\theta}_0,\boldsymbol{\theta}_0)\} = \boldsymbol{D}\boldsymbol{\Omega}_n\boldsymbol{D}$. By (A3), we have $(\boldsymbol{D}\boldsymbol{D})^{-1}\{\ell''(\boldsymbol{\theta},\boldsymbol{\theta}) - E\ell''(\boldsymbol{\theta},\boldsymbol{\theta})\} \xrightarrow{p} 0$. Because $\widetilde{\boldsymbol{\theta}} \in \mathcal{B}(\boldsymbol{\theta}_0)$, the second term of (S1.2) is $\frac{1}{2}\boldsymbol{u}^\top\ell''(\widetilde{\boldsymbol{\theta}},\widetilde{\boldsymbol{\theta}})\boldsymbol{u} = \mathcal{O}(1)$, which dominates the first term for large enough $C$. Thus, $(S1.1)$ holds, and the consistency of $\hat{\boldsymbol{\theta}}_n$ is proved.

□

*Proof of Theorem 2.* For $\hat{\boldsymbol{\theta}}_n$, it satisfies $\frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_n} = 0$, which implies

$$
\frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} + \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}\Big|_{\boldsymbol{\theta}=\bar{\boldsymbol{\theta}}_n}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_0) = 0, \tag{S1.6}
$$

where $\bar{\boldsymbol{\theta}}_n$ is between $\boldsymbol{\theta}_0$ and $\hat{\boldsymbol{\theta}}_n$. For the second term of (S1.6), recall that

$$
\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}\Big|_{\boldsymbol{\theta}=\bar{\boldsymbol{\theta}}_n} = E\{\ell''(\boldsymbol{\theta}_0,\boldsymbol{\theta}_0)\}(1+o_p(1)) = \mathcal{J}_{\boldsymbol{\theta}_0}\{1+o_p(1)\}. \tag{S1.7}
$$

Therefore, we obtain

$$
\mathcal{J}_{\boldsymbol{\theta}_0}(\hat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}_0) \xrightarrow{p} -\ell'(\boldsymbol{\theta}_0).
$$

Next, we use the Cramer-Wold theorem to prove $\ell'(\boldsymbol{\theta}_0) \xrightarrow{p} \mathcal{N}(\mathbf{0}, \mathcal{J}_{\boldsymbol{\theta}_0})$. That is, we will prove that for any $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_q)^\top \in \mathbb{R}^q$,

$$-\boldsymbol{\alpha}^\top \ell'(\boldsymbol{\theta}_0) \xrightarrow{D} \mathcal{N}(0, \boldsymbol{\alpha}^\top \mathcal{J}_{\boldsymbol{\theta}_0} \boldsymbol{\alpha}).$$

Let $-\boldsymbol{\alpha}^\top \ell'(\boldsymbol{\theta}_0) = \sum_{k=1}^q \boldsymbol{e}^\top \boldsymbol{A}(\boldsymbol{\alpha})\boldsymbol{e} - \operatorname{tr}\{\boldsymbol{A}(\boldsymbol{\alpha})\}$, where $\boldsymbol{B}_{k0} = \boldsymbol{\Sigma}_0^{1/2}\boldsymbol{\Sigma}_0^k\boldsymbol{\Sigma}_0^{1/2}$ and $\boldsymbol{A}(\boldsymbol{\alpha}) = \frac{1}{2}\sum_{k=1}^q \alpha_k \boldsymbol{B}_{k0}$.

The term $\boldsymbol{e}^\top \boldsymbol{A}(\boldsymbol{\alpha})\boldsymbol{e}$ is known as the generalized quadratic form and various conditions has been imposed to ensure that the generalized quadratic form converges to a normal distribution (De Jong, 1987; Kelejian and Prucha, 2001, 2010; Shao and Zhang, 2019). One key condition often imposed is that diagonal elements of $\boldsymbol{A}(\boldsymbol{\alpha})$ are zero, which does not hold here. However, for Gaussian processes, we have

$$\boldsymbol{e}^\top \boldsymbol{A}(\boldsymbol{\alpha})\boldsymbol{e} - \operatorname{tr}\{\boldsymbol{A}(\boldsymbol{\alpha})\} = \sum_{i=1}^{N_n} \lambda_i (Z_i - 1),$$

where $\lambda_i$ is the $i$th largest eigenvalue of $\boldsymbol{A}(\boldsymbol{\alpha})$, and $Z_i, \dots, Z_{N_n}$ is a sequence of independent centered chi-square distributed random variables with 1 degree of freedom. Following De Jong (1987), a necessary and sufficient condition for the asymptotic normality is,

$$\frac{\lambda_1^2}{\sum_{i=1}^{N_n} \lambda_i^2} \to 0, \tag{S1.8}$$

as $N_n \to \infty$.

By Cauchy-Schwarz inequality, $\lambda_1^2 \le (\sum_{k=1}^q |\alpha_k| \|\boldsymbol{B}_{k0}\|_2)^2 \le \|\boldsymbol{\alpha}\|_2^2 (\sum_{k=1}^q \|\boldsymbol{B}_{k0}\|_2^2)$.

Furthermore,

$$\begin{aligned}\sum_{i=1}^{N_n}\lambda_i^2 &= \operatorname{tr}\{\boldsymbol{A}(\boldsymbol{\alpha})^\top\boldsymbol{A}(\boldsymbol{\alpha})\} = \boldsymbol{\alpha}^\top \begin{pmatrix} t_{11,n} & \dots & t_{1q,n} \\ \vdots & \ddots & \vdots \\ tq1,n & \dots & t_{qq,n} \end{pmatrix} \boldsymbol{\alpha} \\ &\geq \boldsymbol{\alpha}^\top\boldsymbol{\alpha}\ \lambda_{\min}(\boldsymbol{\Omega}_n) \min_{1\leq k\leq q} t_{kk,n},\end{aligned}$$

where $\lambda_{\min}(\boldsymbol{\Omega}_n)$ is the smallest eigenvalue of $\boldsymbol{\Omega}_n$. By (A3) and (A5), equation (S1.8) is verified, the asymptotic normality of $\hat{\boldsymbol{\theta}}$ is shown.

Last, we show $\mathcal{J}_{\boldsymbol{\theta}_0}^{-1}\mathcal{J}_{\hat{\boldsymbol{\theta}}} \xrightarrow{p} I_q$. The $(k,k')$th element of $\hat{t}_{kk',n} - t_{kk',n}$ is

$$\begin{aligned}(II) &= tr(\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Sigma}}_k\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Sigma}}_{k'}) - tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k'0}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k0}) \\ &= tr\{\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Sigma}}_k\hat{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\Sigma}}_{k'} - \boldsymbol{\Sigma}_{k'0})\} + tr\{\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Sigma}}_k(\hat{\boldsymbol{\Sigma}}^{-1} - \boldsymbol{\Sigma}_0^{-1})\boldsymbol{\Sigma}_{k'0}\} \\ &+ tr\{\hat{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\Sigma}}_{k'} - \boldsymbol{\Sigma}_{k'0})\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k'0}\} + tr\{(\hat{\boldsymbol{\Sigma}}^{-1} - \boldsymbol{\Sigma}_0^{-1})\boldsymbol{\Sigma}_{k0}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{k'0}\} \\ &\equiv (II_1) + (II_2) + (II_3) + (II_4).\end{aligned}$$

Following the similar arguments as $(I_{3,2})$–$(I_{3,5})$ in the proof of Theorem 1, it can be shown that $(II) = o_p\left(t_{kk,n}^{1/2}t_{k'k',n}^{1/2}\right)$. Recall that $\mathcal{J}_{\boldsymbol{\theta}_0} = \boldsymbol{D}\boldsymbol{\Omega}_n\boldsymbol{D}$, we obtain

$$\|\mathcal{J}_{\boldsymbol{\theta}_0}^{-1}\mathcal{J}_{\hat{\boldsymbol{\theta}}} \overset{p}{-} I_q\|_2 = \|\boldsymbol{\Omega}_n\|_2\|\boldsymbol{D}^{-1}(\mathcal{J}_{\hat{\boldsymbol{\theta}}} - \mathcal{J}_{\boldsymbol{\theta}_0})\boldsymbol{D}^{-1}\|_2 = o_p(1).$$

□

## S2 Appendix B: Proof of Theorem 3–Theorem 5

*Proof of Theorem 3.* First, we show that there exists a constant $c_2 > 0$, such that for any $\boldsymbol{\theta} \in \mathcal{B}(\boldsymbol{\theta}_0)$,

$$\max_{k,k'=1,\dots,q} \left\{ \|\boldsymbol{\Sigma}\|_2, \|\boldsymbol{\Sigma}_k\|_2, \|\boldsymbol{\Sigma}_{kk'}\|_2 \right\} \leq c_2 N_n^{1-\alpha l} \tag{S2.1}$$

holds for sufficiently large $N_n$.

Let $d_{ii'} = \|\boldsymbol{s}_i - \boldsymbol{s}_{i'}\|_2$, and $B_m = \{i' : m\delta < d_{ii'} \leq (m+1)\delta\}$, where $\delta$ is independent of $n$. For large enough $N_n$, by Assumption (S1), the sampling density of any subset of $\mathcal{R}_n$ is bounded by $\rho N_n^{1-\alpha l}$, where $\rho > 0$ is a constant. Thus, the number of elements in $B_m$ is at the rate of $N_n^{1-\alpha l} m^{l-1} \delta^l$. Consequently, we have

$$\begin{aligned} \max_{1 \leq i \leq N_n} \sum_{i'=1}^{N_n} \gamma(d_{ii'}, \boldsymbol{\theta}) &= \max_{1 \leq i \leq N_n} \sum_{m=0}^{\infty} \sum_{i' \in B_m} \gamma(d_{ii'}, \boldsymbol{\theta}) \\ &\leq \max_{1 \leq i \leq N_n} \sum_{m=0}^{\infty} \left\{ \mathcal{O}(N_n^{1-\alpha l}) m^{l-1} \delta^l \max_{m\delta < d \leq (m+1)\delta} \gamma(d, \boldsymbol{\theta}) \right\}. \end{aligned} \tag{S2.2}$$

Let $\delta \to 0$, by the definition of Riemann integral, we have

$$\sum_{m=0}^{\infty} m^{l-1} \delta^l \max_{m\delta < d \leq (m+1)\delta} \gamma(d, \boldsymbol{\theta}) \to \int_0^{\infty} u^{l-1} \gamma(u; \boldsymbol{\theta}) du.$$

By (C1), we have $\max_{\boldsymbol{\theta} \in \mathcal{B}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}\|_\infty = \mathcal{O}(N_n^{1-\alpha l})$, and thus, $\max_{\boldsymbol{\theta} \in \mathcal{B}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}\|_2 = \mathcal{O}(N_n^{1-\alpha l})$. Similarly, we can show that $\max_{\boldsymbol{\theta} \in \mathcal{B}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_k\|_2 = \mathcal{O}(N_n^{1-\alpha l})$, $\max_{\boldsymbol{\theta} \in \mathcal{B}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_{kk'}\|_2 = \mathcal{O}(N_n^{1-\alpha l})$ and $\max_{\boldsymbol{\theta} \in \mathcal{B}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_{kk'}\|_F^2 = \mathcal{O}(N_n^{1+(1-\alpha l)})$.

Next, we show that as $N_n \to \infty$, the rate of $t_{kk,n}$ is larger or equal to $N_n^{1-(1-\alpha l)}$. In Section 3, we have shown $t_{22,n}$ is at the rate of $N_n$. Here, we show that for $k \neq 2$, the rate of $t_{kk,n}$ is larger or equal to $N_n^{1-(1-\alpha l)}$, as $N_n \to \infty$. Following the inequality $t_{kk,n} = \mathrm{tr}(\boldsymbol{\Sigma}_0^{-1} \boldsymbol{\Sigma}_{k0} \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\Sigma}_{k0}) \geq \|\boldsymbol{\Sigma}_0\|_2^{-2} \ \|\boldsymbol{\Sigma}_{k0}\|_F^2$ and (S2.1), it suffices to show that

$\|\mathbf{\Sigma}_{k0}\|_F^2$ has the rate of $N^{1+(1-\alpha l)}$.

By (S1), for sufficiently large $N_n$, the number of elements in $B_m$ is larger than $\rho_2 N_n^{1-\alpha l} m^{l-1}\delta^l$, as $N_n \to \infty$, where $\rho_2 > 0$ is a constant. For each $i$, we have

$$\begin{aligned}\sum_{i'=1}^{N_n}\{\gamma_k(d_{ii'},\boldsymbol{\theta}_0)\}^2 &= \sum_{m=0}^{\infty}\sum_{i'\in B_m}\{\gamma_k(d_{ii'},\boldsymbol{\theta}_0)\}^2 \\ &\geq \sum_{m=0}^{\infty}\rho_2 N_n^{1-\alpha l} m^{l-1}\delta^l \min_{m\delta<d\leq(m+1)\delta}\{\gamma_k(d,\boldsymbol{\theta}_0)\}^2. \end{aligned} \tag{S2.3}$$

Let $\delta \to 0$, by the definition of Riemann integral, we have

$$\sum_{m=0}^{\infty} m^{l-1}\delta^l \min_{m\delta<d\delta\leq(m+1)\delta}\{\gamma_k(d_{ii'},\boldsymbol{\theta}_0)\}^2 \to \int_0^\infty u^{l-1}\{\gamma_k(u;\boldsymbol{\theta}_0)\}^2 du > 0.$$

By (C3), we have

$$\|\mathbf{\Sigma}_{k0}\|_F^2 = \sum_{i=1}^{N_n}\sum_{i'=1}^{N_n}\{\gamma_k(d_{ii'},\boldsymbol{\theta}_0)\}^2 \geq \rho_2 N_n^{1+(1-\alpha l)}\int_0^\infty u^{l-1}\{\gamma_k(u;\boldsymbol{\theta}_0)\}^2 du,$$

which is at the rate of $N^{1+(1-\alpha l)}$.

Based on the above results, we discuss the sufficient conditions for (A4) and (A5). For (A4), since $\max_{\boldsymbol{\theta}\in\mathcal{B}(\boldsymbol{\theta}_0)}\|\mathbf{\Sigma}_k\|_2$ is at the rate of $N^{1-\alpha l}$ and $t_{kk,n}$ is at the rate larger or equal to $N_n^{1-(1-\alpha l)}$, a sufficient condition for $\max_{\boldsymbol{\theta}\in\mathcal{B}(\boldsymbol{\theta}_0)}\|\mathbf{\Sigma}_k\|_2 = o(t_{kk,n}^{1/2})$ is $(1-\alpha l) < 1/3$. Similarly, the sufficient condition for $\max_{\boldsymbol{\theta}\in\mathcal{B}(\boldsymbol{\theta}_0)}\|\mathbf{\Sigma}_{kk'}\|_F = o(t_{kk,n}^{1/2}t_{k'k',n}^{1/2})$ is $(1-\alpha l) < 1/3$, and the sufficient condition for (A5) is $(1-\alpha l) < 1/3$.

□

**Lemma 1.** *Covariance functions (2.2)–(2.4) satisfy conditions (C1) and (C3).*

*Proof of Lemma 1.* For the Matérn class in (2.2), straightforward calculation shows that

$$\frac{\partial \gamma(d;\boldsymbol{\theta})}{\partial \theta_1} = -\theta_3 \frac{2}{\Gamma(\kappa)} \left(\frac{\theta_1 d}{2}\right)^{\kappa} K_{\kappa-1}(\theta_1 d) d,$$

$$\frac{\partial \gamma(d;\boldsymbol{\theta})}{\partial \theta_2} = \begin{cases} 1, & \text{if} \quad d = 0, \\ 0, & \text{if} \quad d > 0, \end{cases}$$

$$\frac{\partial \gamma(d;\boldsymbol{\theta})}{\partial \theta_3} = \frac{2}{\Gamma(\kappa)} \left(\frac{\theta_1 d}{2}\right)^{\kappa} K_{\kappa}(\theta_1 d),$$

$$\frac{\partial^2 \gamma(d;\boldsymbol{\theta})}{\partial \theta_k \partial \theta_{k'}} = \begin{cases} \theta_3 \frac{2}{\Gamma(\kappa)} \left(\frac{\theta_1 d}{2}\right)^{\kappa} K_{\kappa-2}(\theta_1 d) d^2, & \text{if} \quad k = k' = 1, \\ -\frac{2}{\Gamma(\kappa)} \left(\frac{\theta_1 d}{2}\right)^{\kappa} K_{\kappa-1}(\theta_1 d) d, & \text{if} \quad (k,k') = (1,3) \text{ or } (k,k') = (3,1), \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see that $\int_0^\infty u^{l-1}|\gamma_2(u;\boldsymbol{\theta})|du = 0$ and $\int_0^\infty u^{l-1}|\gamma_{2k'}(u;\boldsymbol{\theta})|du = 0$. Moreover, since $K_\nu(d) \propto e^{-d} d^{-1/2}\{1+\mathcal{O}(1/d)\}$ when $d \to \infty$, $\int_0^\infty u^{l-1}|\gamma(u;\boldsymbol{\theta})|du$, $\int_0^\infty u^{l-1}|\gamma_k(u;\boldsymbol{\theta})|du$ and $\int_0^\infty u^{l-1}|\gamma_{kk'}(u;\boldsymbol{\theta})|du$ are finite and continuous with respect to $\boldsymbol{\theta}$, for $k, k' \neq 2$. Therefore, (C1) is satisfied. Moreover, for $k \neq 2$, $\int_0^\infty u^{l-1}\{\gamma_k(u;\boldsymbol{\theta}_0)\}^2 du > 0$, and (C3) is satisfied.

For the Gaussian covariance function in (2.3), we have

$$\frac{\partial \gamma(d;\boldsymbol{\theta})}{\partial \theta_1} = 2\theta_3 \frac{d^2}{\theta_1^3} \exp\left\{-\left(\frac{d}{\theta_1}\right)^2\right\}$$

$$\frac{\partial \gamma(d;\boldsymbol{\theta})}{\partial \theta_2} = \begin{cases} 1, & \text{if} \quad d = 0, \\ 0, & \text{if} \quad d > 0, \end{cases}$$

$$\frac{\partial \gamma(d;\boldsymbol{\theta})}{\partial \theta_3} = \exp\left\{-\left(\frac{d}{\theta_1}\right)^2\right\}.$$

For the powered exponential covariance function in (2.4), we have

$$\frac{\partial\gamma(d;\boldsymbol{\theta})}{\partial\theta_1} = \theta_3 \exp\left\{-\left(\frac{d}{\theta_1}\right)^{\theta_4}\right\}(\theta_4 d^{\theta_4}\theta_1^{-\theta_4-1})$$

$$\frac{\partial\gamma(d;\boldsymbol{\theta})}{\partial\theta_2} = \begin{cases} 1, & \text{if} \quad d = 0, \\ 0, & \text{if} \quad d > 0, \end{cases}$$

$$\frac{\partial\gamma(d;\boldsymbol{\theta})}{\partial\theta_3} = \exp\left\{-\left(\frac{d}{\theta_1}\right)^{\theta_4}\right\},$$

$$\frac{\partial\gamma(d;\boldsymbol{\theta})}{\partial\theta_4} = \begin{cases} \theta_3 \exp\left\{-\left(\frac{d}{\theta_1}\right)^{\theta_4}\right\}(-d^{\theta_4}\theta_1^{-\theta_4}\log(d/\theta_1)), & \text{if} \quad d = 0, \\ 0, & \text{if} \quad d > 0. \end{cases}$$

Moreover, for both the Gaussian and the powered exponential covariance functions, it is easy to see that the second derivatives of $\gamma(d;\boldsymbol{\theta})$ will be either zero or decay at exponentially rate. Thus, by the similar argument as covariance function (2.2), covariance functions (2.3)–(2.4) also satisfy (C1) and (C3).

□

*Proof of Theorem 4.* First, we show that under the fixed sampling design (S1), if covariance functions satisfy (C2), then there exists a constant $c_3 > 0$, such that for any $\boldsymbol{\theta} \in \mathcal{B}(\boldsymbol{\theta}_0)$,

$$\max_{k,k'=1,\ldots,q}\left\{\|\boldsymbol{\Sigma}\|_2, \|\boldsymbol{\Sigma}_k\|_2, \|\boldsymbol{\Sigma}_{kk'}\|_2\right\} \leq c_3 N_n^{1-\alpha\zeta} \tag{S2.4}$$

holds for sufficiently large $N_n$.

Recall that $d_{ii'} = \|\boldsymbol{s}_i - \boldsymbol{s}_{i'}\|_2$ and $B_m = \{i' : m\delta < d_{ii'} \leq (m+1)\delta\}$. Similar to the proof of Theorem 3, we have

$$\begin{aligned}\max_{1\le i\le N_n}\sum_{i'=1}^{N_n}\gamma(d_{ii'},\boldsymbol{\theta}) &= \max_{1\le i\le N_n}\sum_{m=0}^{\lceil\lambda_n/2\delta\rceil}\sum_{i'\in B_m}\gamma(d_{ii'},\boldsymbol{\theta}) \\ &= \max_{1\le i\le N_n}\sum_{m=0}^{\lceil\lambda_n/2\delta\rceil}\left\{\mathcal{O}(N_n^{1-\alpha l})m^{l-1}\delta^l\max_{m\delta<d\le(m+1)\delta}\gamma(d,\boldsymbol{\theta})\right\}, \qquad \text{(S2.5)}\end{aligned}$$

where $\lceil\cdot\rceil$ is the ceiling function.

Let $\delta\to 0$, by the definition of Riemann integral, we have

$$\sum_{m=0}^{\lceil\lambda_n/2\delta\rceil} m^{l-1}\delta^l\max_{m\delta<d\le(m+1)\delta}\gamma(d,\boldsymbol{\theta})\to\int_0^{\lceil\lambda_n/2\rceil}u^{l-1}\gamma(u;\boldsymbol{\theta})du=\mathcal{O}(N_n^{\alpha(l-\zeta)}).$$

The last equation holds since the $\lambda_n$ is at the rate of $N_n^{\alpha}$. By (C2), we have $\max_{\boldsymbol{\theta}\in\mathcal{B}(\boldsymbol{\theta}_0)}\|\boldsymbol{\Sigma}\|_\infty = \mathcal{O}(N_n^{1-\alpha\zeta})$, and thus, $\max_{\boldsymbol{\theta}\in\mathcal{B}(\boldsymbol{\theta}_0)}\|\boldsymbol{\Sigma}\|_2=\mathcal{O}(N_n^{1-\alpha\zeta})$. Similarly, we can show that $\max_{\boldsymbol{\theta}\in\mathcal{B}(\boldsymbol{\theta}_0)}\|\boldsymbol{\Sigma}_k\|_2 = \mathcal{O}(N_n^{1-\alpha\zeta})$ and $\max_{\boldsymbol{\theta}\in\mathcal{B}(\boldsymbol{\theta}_0)}\|\boldsymbol{\Sigma}_{kk'}\|_2=\mathcal{O}(N_n^{1-\alpha\zeta})$.

For $\max_{\boldsymbol{\theta}\in\mathcal{B}(\boldsymbol{\theta}_0)}\|\boldsymbol{\Sigma}_{kk'}\|_F^2$, we have

$$\begin{aligned}\max_{1\le i\le N_n}\sum_{i'=1}^{N_n}\gamma(d_{ii'},\boldsymbol{\theta}) &= \max_{1\le i\le N_n}\sum_{m=0}^{\infty}\sum_{i'\in B_m}\gamma^2(d_{ii'},\boldsymbol{\theta}) \\ &= \max_{1\le i\le N_n}\sum_{m=0}^{\infty}\left\{\mathcal{O}(N_n^{1-\alpha l})m^{l-1}\delta^l\max_{m\delta<d\le(m+1)\delta}\gamma^2(d,\boldsymbol{\theta})\right\}, \qquad \text{(S2.6)}\end{aligned}$$

Therefore, by (C2), we have $\max_{\boldsymbol{\theta}\in\mathcal{B}(\boldsymbol{\theta}_0)}\|\boldsymbol{\Sigma}_{kk'}\|_F^2=\mathcal{O}(N_n^{1+(1-\alpha l)})$.

Under (C3), we can show that as $N_n\to\infty$, the rate of $t_{kk,n}$ is larger or equal to $N_n^{\alpha(2\zeta-l)}$. It is suffice to show that for $k\neq 2$, the rate of $t_{kk,n}$ is larger or equal to $N_n^{1-(1-\alpha l)}$, as $N_n\to\infty$. The proof is the same as that of Theorem 3, and we omit the proof here. Based on the above results, it can be calculate that $(4\zeta-l)\alpha>2$ is a sufficient condition for (A4) and (A5). □

**Lemma 2.** *If $l/4\le\kappa\le l/2$, Covariance function (2.5) satisfies conditions (C2)–(C3).*

*Proof of Lemma 2.* The first-order and second-order derivatives of Cauchy covariance function are,

$$\gamma(d,\boldsymbol{\theta}) = \begin{cases} \theta_3 \left\{1+\left(\frac{d}{\theta_1}\right)^2\right\}^{-\kappa}, & \text{if} \quad d>0, \\ \theta_3+\theta_2, & \text{if} \quad d=0, \end{cases} \tag{S2.7}$$

straightforward calculation shows that

$$\frac{\partial\gamma(d;\boldsymbol{\theta})}{\partial\theta_1} = \frac{2\kappa\theta_3 d^2}{\theta_1^3}\left\{1+\left(\frac{d}{\theta_1}\right)^2\right\}^{-\kappa-1},$$

$$\frac{\partial\gamma(d;\boldsymbol{\theta})}{\partial\theta_2} = \begin{cases} 1, & \text{if} \quad d=0, \\ 0, & \text{if} \quad d>0, \end{cases}$$

$$\frac{\partial\gamma(d;\boldsymbol{\theta})}{\partial\theta_3} = \left\{1+\left(\frac{d}{\theta_1}\right)^2\right\}^{-\kappa},$$

$$\frac{\partial^2\gamma(d;\boldsymbol{\theta})}{\partial\theta_k\partial\theta_{k'}} = \begin{cases} \frac{2\kappa\theta_3 d^2\left\{(2\kappa-1)d^2\theta_1^{-2}-3\right\}}{\theta_1^4}\left\{1+\left(\frac{d}{\theta_1}\right)^2\right\}^{-\kappa-2}, & \text{if} \quad k=k'=1, \\ \frac{2\kappa d^2}{\theta_1^3}\left\{1+\left(\frac{d}{\theta_1}\right)^2\right\}^{-\kappa-1}, & \text{if} \quad (k,k')=(1,3) \text{ or } (k,k')=(3,1), \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see that $\int_0^\infty u^{l-1}|\gamma_2(u;\boldsymbol{\theta})|du = 0$ and $\int_0^\infty u^{l-1}|\gamma_{2k'}(u;\boldsymbol{\theta})|du = 0$. Moreover, for other terms, it decays at the rate of $d^{-2\kappa}$. Setting $\zeta = 2\kappa$, (C1) is satisfied. Moreover, for $k \neq 2$, $\int_0^\infty u^{l-1}\{\gamma_k(u;\boldsymbol{\theta}_0)\}^2 du > 0$, and (C3) is satisfied.

□

# S3 Appendix C: The relationship with increasing domain asymptotics

Here, we discuss the relationship between Theorems 1–2 and increasing domain asymptotics in Mardia and Marshall (1984). Under the increasing domain framework, it is well-known that Mardia and Marshall (1984) established asymptotic normality of $\hat{\boldsymbol{\theta}}_n$. Similar to Theorem 2, (A1)–(A3) are also assumed in Mardia and Marshall (1984). However, instead of (A4)–(A5), the following conditions (B1)–(B2) are assumed by Mardia and Marshall (1984) for the increasing domain framework.

(B1) There exists a constant $c_2 > 0$, such that for any $\boldsymbol{\theta} \in \mathcal{B}(\boldsymbol{\theta}_0)$,

$$\max_{k,k'=1,\ldots,q} \left\{ \|\boldsymbol{\Sigma}\|_2, \|\boldsymbol{\Sigma}_k\|_2, \|\boldsymbol{\Sigma}_{kk'}\|_2 \right\} \leq c_2$$

holds for sufficiently large $N_n$.

(B2) As $N_n \to \infty$, we have $\max_{\boldsymbol{\theta}\in\mathcal{B}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_k\|_F^{-2} = \mathcal{O}(N^{-1/2-\tau})$, for some $\tau > 0$ and $k = 1, \ldots, q$.

The main result of $\hat{\boldsymbol{\theta}}_n$ in Mardia and Marshall (1984) can be restated in the following Theorem A.

**Theorem A.** *Under (A1)–(A3) and (B1)–(B2), $\hat{\boldsymbol{\theta}}_n$ has asymptotically normal distribution, that is,*

$$\mathcal{J}_{\boldsymbol{\theta}_0}^{1/2}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{D} \mathcal{N}\left(0, I_q\right).$$

In the following Corollary A, it is shown that under the same conditions as Theorem A, Theorems 1–2 hold for the increasing domain framework. Thus, Theorems 1–2 are consistent with previous studies in the increasing domain framework by Mardia and Marshall (1984), and hold for both the mixed domain asymptotic framework and the increasing domain asymptotic framework.

**Corollary A.** *Under (A1)–(A3) and (B1)–(B2), Theorems 1–2 hold.*

*Proof of Corollary A..* By (A2), (B1) and inequality $t_{kk,n} \geq \|\boldsymbol{\Sigma}_0^{-2}\|_2 \, \|\boldsymbol{\Sigma}_{k0}\|_F^2$, the rate of $t_{kk,n}$ is at least $N_n^{1/2+\tau}$. Since $\max_{\boldsymbol{\theta} \in \mathcal{B}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_k\|_2 = \mathcal{O}(1)$ and $\max_{\boldsymbol{\theta} \in \mathcal{B}(\boldsymbol{\theta}_0)} \|\boldsymbol{\Sigma}_{kk'}\|_F \leq \max_{\boldsymbol{\theta} \in \mathcal{B}(\boldsymbol{\theta}_0)} N_n^{1/2} \|\boldsymbol{\Sigma}_{kk'}\|_2 = \mathcal{O}(N_n^{1/2})$, Assumptions (A4)–(A5) hold, and Corollary A is proved. □

Assumption (B1) is closely related to asymptotic framework, and it holds for Type-I covariance function under the increasing domain framework, as shown in (S2.1) of Appendix B. However, Assumption (B1) does not hold for Type-II covariance functions, due to stronger spatial dependence, as shown in (S2.4) of Appendix B. Therefore, Theorem A holds for Type-I covariance function under increasing domain asymptotics, while Theorems 1–2 here hold for both types of covariance functions under increasing or mixed domain asymptotics.

## S4 Appendix D: Additional Simulation Results

Here, we present the additional simulation results. Similarly to Zhang and Zimmerman (2005), we plot the $0.05 + 0.1(i-1)$ quantiles for each parameter. For the exponential covariance function, the plot is also drawn for $\phi = \theta_3/\theta_1$.

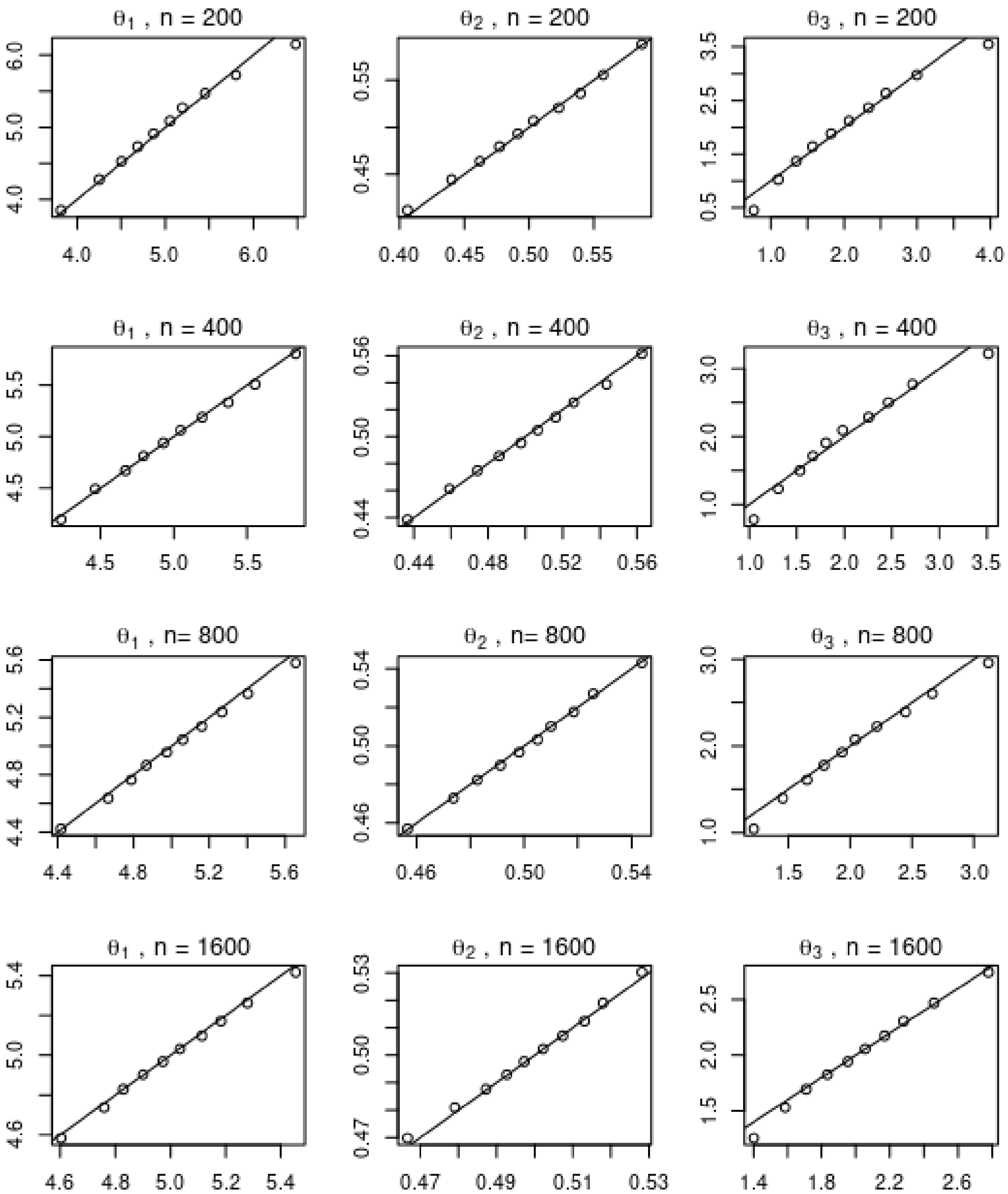


Figure A: Quantiles of $\hat{\boldsymbol{\theta}}$ for the Gaussian covariance function with $\alpha = 0.4$

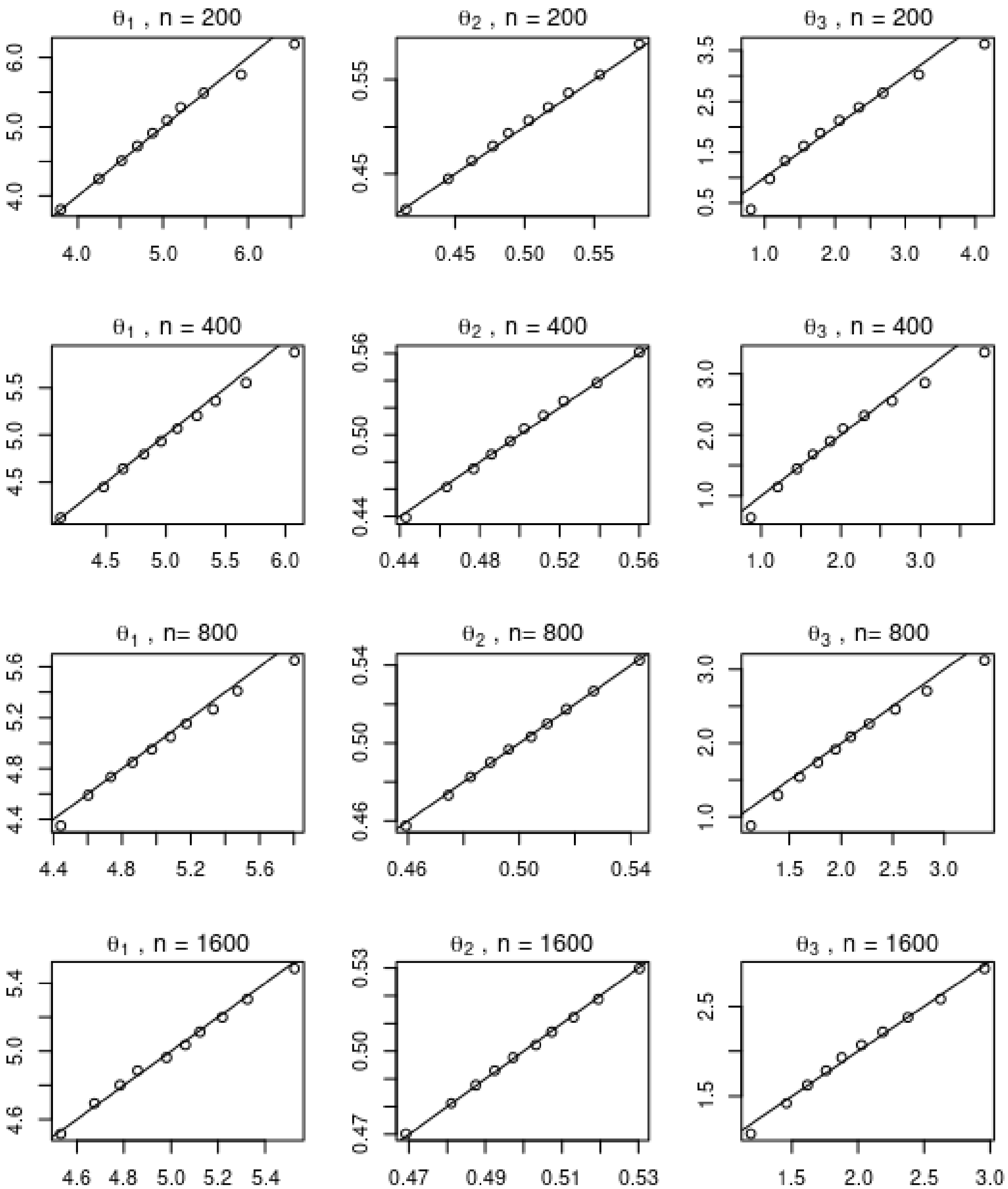


Figure B: Quantiles of $\hat{\boldsymbol{\theta}}$ for the Gaussian covariance function with $\alpha = 0.3$

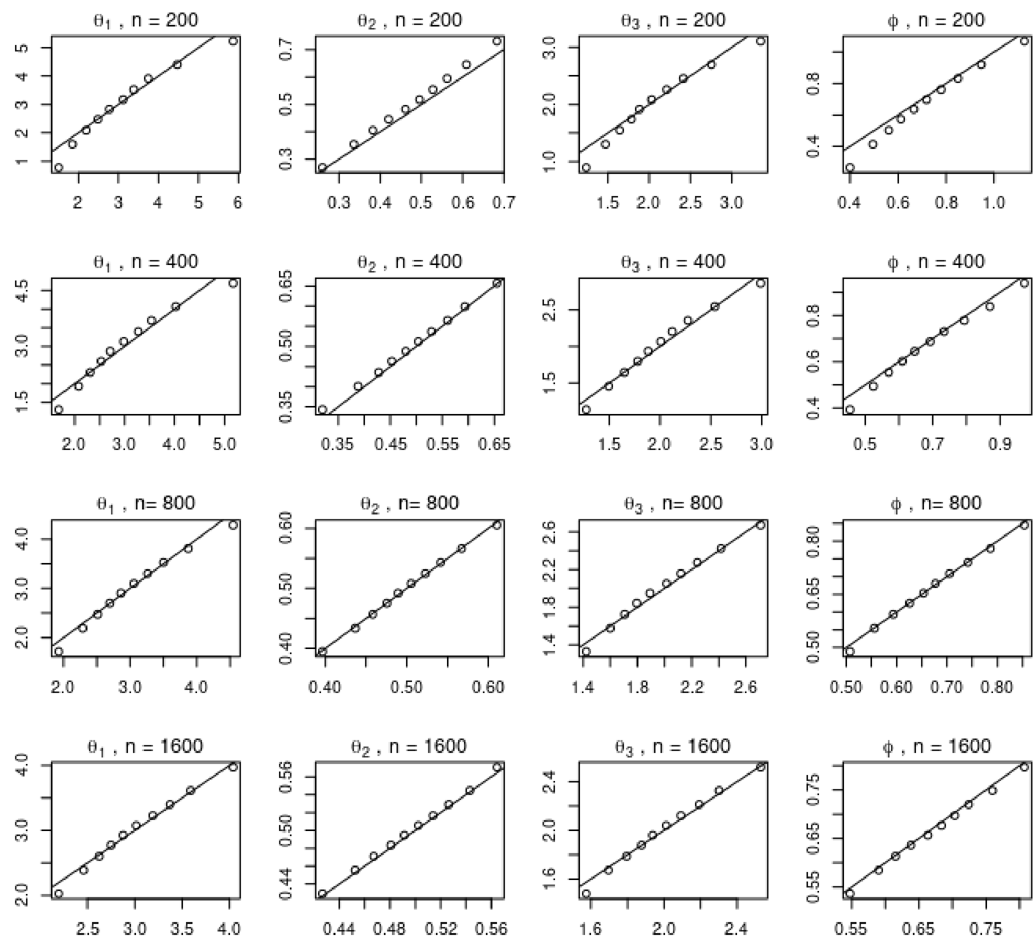


Figure C: Quantiles of $\hat{\boldsymbol{\theta}}$ and $\hat{\phi}$ for the exponential covariance function with $\alpha = 0.4$

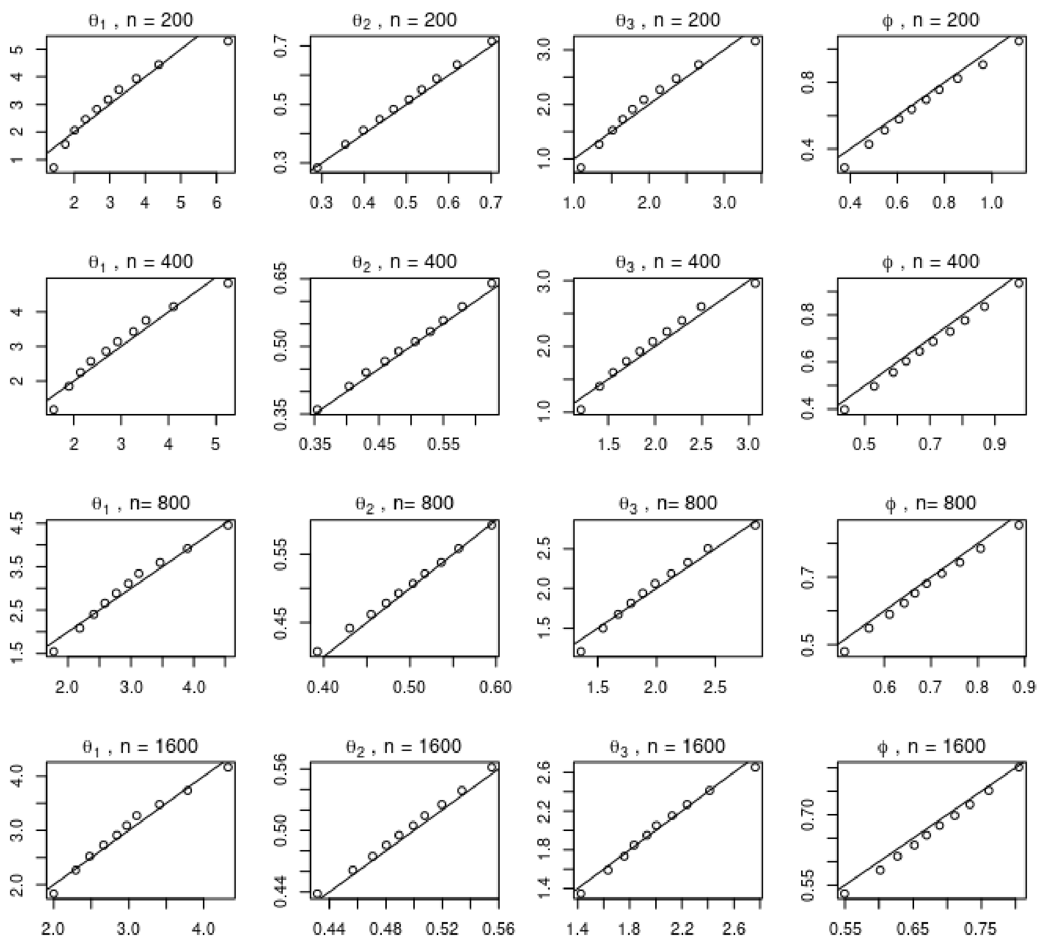


Figure D: Quantiles of $\hat{\boldsymbol{\theta}}$ and $\hat{\phi}$ for the exponential covariance function with $\alpha = 0.3$

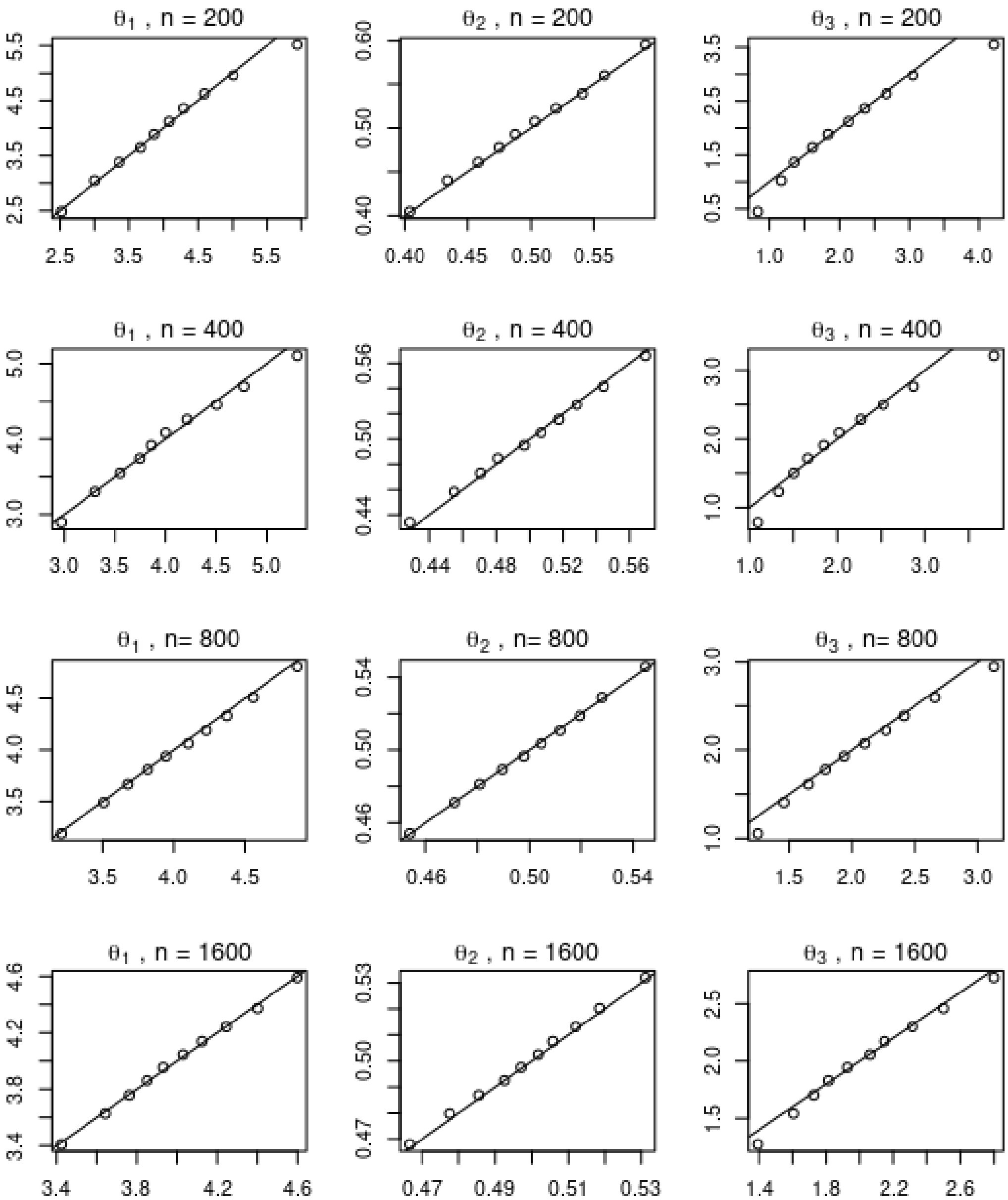


Figure E: Quantiles of $\hat{\boldsymbol{\theta}}$ for the Cauchy covariance function with $\alpha = 0.4$

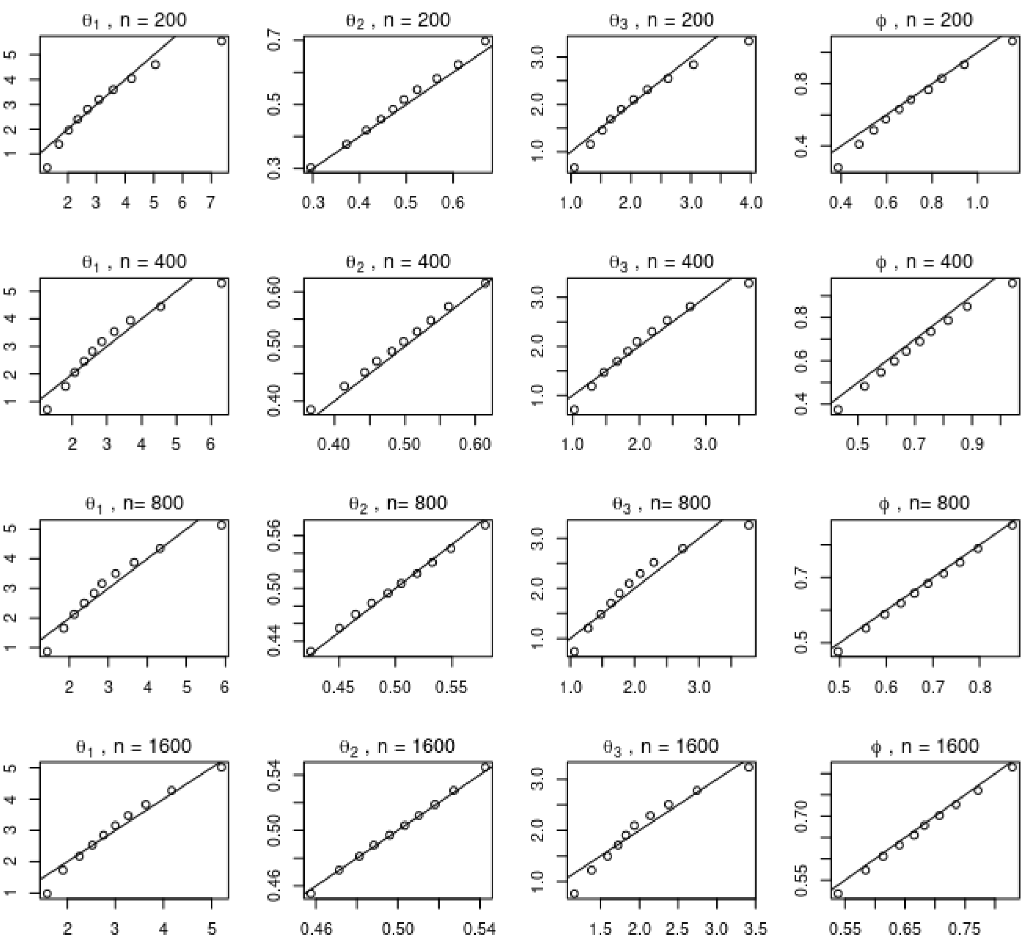


Figure F: Quantiles of $\hat{\boldsymbol{\theta}}$ and $\hat{\phi}$ for the exponential covariance function with $\alpha = 0$

## Bibliography

De Jong, P. (1987). A central limit theorem for generalized quadratic forms. Probability Theory and Related Fields 75, 261–277.

Kelejian, H. H. and I. R. Prucha (2001). On the asymptotic distribution of the Moran I test statistic with applications. Journal of Econometrics 104, 219–257.

Kelejian, H. H. and I. R. Prucha (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. Journal of Econometrics 157, 53–67.

Mardia, K. V. and R. J. Marshall (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. Biometrika 71, 135–146.

Shao, Q.-M. and Z.-S. Zhang (2019). Berry–Esseen bounds of normal and nonnormal approximation for unbounded exchangeable pairs. The Annals of Probability 47, 61–108.

Zhang, H. and D. L. Zimmerman (2005). Towards reconciling two asymptotic frameworks in spatial statistics. Biometrika 92, 921–936.