Supplementary Materials of Bayesian Inference of Spatially Varying Correlations via the Thresholded Correlation Gaussian Process

Moyan Li[1], Lexin Li[2] and Jian Kang[1]

[1]*University of Michigan, Ann Arbor and* [2]*University of California, Berkeley*

In this supplement, we first present the proofs of all the theoretical results in the paper, along with a number of useful lemmas. We next derive the full conditional distributions of the model parameters, and present some additional numerical results.

S1. Proofs

S1.1 Proof of Proposition 1

Given $\tau_1^2(v)$ and $\tau_2^2(v)$, if $\pi(Y_{+,i}(v), Y_{-,i}(v) \mid \theta) = \pi\left(Y_{+,i}(v), Y_{-,i}(v) \mid \theta'\right)$, for any $i = 1, \ldots, n$, $v \in \mathcal{B}_m$, and since $\{Y_{+,i}(v), Y_{-,i}(v)\}$ follows a bivariate normal distribution, we have that $\mu_{+,i}(v) = \mu'_{+,i}(v)$, and $\mu_{-,i}(v) = \mu'_{-,i}(v)$, i.e., $s\{\rho(v)\}E_{+,i}(v) = s\{\rho'(v)\}E'_{+,i}(v)$, and $s\{-\rho(v)\}E_{-,i}(v) = s\{-\rho'(v)\}E'_{-,i}(v)$, for any $i = 1, \ldots, n$, $v \in \mathcal{B}_m$.

Furthermore, we have that,

$$
\begin{aligned}
0 &= \sum_{i=1}^{n} \left[s\{\rho(v)\} E_{+,i}(v) - s\{\rho'(v)\} E'_{+,i}(v) \right]^2 \\
&= \sum_{i=1}^{n} \left[s\{\rho(v)\}^2 E_{+,i}(v)^2 - 2s\{\rho(v)\} s\{\rho'(v)\} E_{+,i}(v) E'_{+,i}(v) + s\{\rho'(v)\}^2 E'_{+,i}(v)^2 \right] \\
&= \left[s\{\rho(v)\} - s\{\rho'(v)\} \right]^2 \sum_{i=1}^{n} E^2_{+,i}(v) + s\{\rho'(v)\} s\{\rho(v)\} \sum_{i=1}^{n} \{E_{+,i}(v) - E'_{+,i}(v)\}^2 \\
&\quad + s\{\rho'(v)\} \left[s\{\rho(v)\} - s\{\rho'(v)\} \right] \sum_{i=1}^{n} \left\{ E_{+,i}(v)^2 - E'_{+,i}(v)^2 \right\}
\end{aligned}
$$

By Definition (4), we have $\sum_{i=1}^{n} E_{+,i}(v)^2 = \sum_{i=1}^{n} E'_{+,i}(v)^2$.

When $v \in \mathcal{V}(\rho) \cup \mathcal{V}(\rho')$, we have $s\{\rho(v)\} \geq 0$, $s\{\rho'(v)\} \geq 0$, and at least one of $s\{\rho(v)\}$ and $s\{\rho'(v)\}$ is not equal to 0. Therefore, $s\{\rho(v)\} = s\{\rho'(v)\}$, and $E_{+,i}(v) = E'_{+,i}(v)$, for any $i = 1, \ldots, n,\ v \in \mathcal{B}_m$. On the other hand, if $v \notin \mathcal{V}(\rho) \cup \mathcal{V}(\rho')$, then $s\{\rho(v)\} = s\{\rho'(v)\} = 0$. Similarly, we have $E_{-,i}(v) = E'_{-,i}(v) = 0$, for any $i = 1, \ldots, n$, $v \in \mathcal{B}_m$.

Since $s(\cdot)$ is a monotonic function, we have $\rho(v) = \rho'(v)$ for all $v \in \mathcal{B}_m$. This completes the proof of Proposition 1. $\square$

S1.2 Proof of Theorem 1

By Lemma S1, we have $\rho(v) = T_\omega\{\xi(v)\} = H[R_\omega\{\xi(v)\}]$, where $H(t) = t^2/(t^2 + 1)$ when $\xi(v) > \omega$, $H(t) = -t^2/(t^2 + 1)$ when $\xi(v) < -\omega$, and $H(t) = 0$ otherwise, and $R_\omega(x) = G_\omega(x) - G_\omega(-x)$ is the hard thresholded function. Therefore, we have that,

$$pr\left(\|\rho-\rho_0\|_\infty<\varepsilon\right)=pr\left(\|H[R_\omega\{\xi(v)\}]-H[R_\omega\{\xi_0(v)\}]\|<\epsilon\right)$$
$$\geq pr\left(\|R_\omega\{\xi(v)\}-R_\omega\{\xi_0(v)\}\|<\epsilon\right),$$

by the Lipschitz continuity of $H(\cdot)$. Given the assumptions for $\rho_0(v)$, we have that $\xi(v)$ is bounded away from 0 for $v\notin\mathcal{R}_0$. Henceforth,

$$\begin{aligned} &pr(\|R_\omega(\xi(v))-R_\omega(\xi_0(v))\|<\epsilon)\\ &\geq pr\left(\sup_{v\notin\mathcal{R}_0}|\xi(v)-\xi_0(v)|<\epsilon,\inf_{v\notin\mathcal{R}_0}|\xi(v)|>\omega,\sup_{v\in\mathcal{R}_0}|\xi(v)|\leq\omega\right). \end{aligned} \tag{S1}$$

Without loss of generality, we only consider $0<\epsilon<\omega-\omega_0$, where $\omega_0=\inf_{v\notin\mathcal{R}_0}|\rho(v)|$. Note that for all $v\notin\mathcal{R}_0$, $|\xi(v)-\xi_0(v)|<\epsilon$ and $|\xi_0(v)|\geq\omega_0$, which implies that $|\xi(v)|\geq\omega_0-\epsilon>\omega$. Then (S1) is equivalent to

$$pr(\|\rho(v)-\rho_0(v)\|<\epsilon)\geq pr\left(\sup_{v\notin\mathcal{R}_0}|\xi(v)-\xi_0(v)|<\epsilon,\sup_{v\in\mathcal{R}_0}|\xi(v)|\leq\omega\right).$$

Let $\psi_l(v)$ and λ_l be the normalized eigenfunctions and eigenvalues of the kernel function $\kappa(\cdot,\cdot)$. The KL expansions of $\xi(v)$ and $\xi_0(v)$ are $\xi(v)=\sum_{l=1}^{\infty}c_l\psi_l(v)$, $\xi_0(v)=\sum_{l=1}^{\infty}c_{l0}\psi_l(v)$.

For $v\notin\mathcal{R}_0$, we have that,

$$\sup_{v\notin\mathcal{R}_0}|\xi(v)-\xi_0(v)|\leq\sup_{v\notin\mathcal{R}_0}|\xi_L(v)-\xi_L^0(v)|+\sup_{v\notin\mathcal{R}_0}|\xi(v)-\xi_L(v)|+\sup_{v\notin\mathcal{R}_0}|\xi_L^0(v)-\xi_0(v)|.$$

Since the RKHS of $\kappa(\cdot,\cdot)$ is the space of the continuous functions on $\mathcal{R}$, $\xi(v)$ is uniformly continuous on $\mathcal{B}\backslash\mathcal{R}_0$ with probability 1. Then by Theorem 3.1.2 of Adler and Taylor (2009), $\lim_{L\to\infty}\sup_{v\notin\mathcal{R}_0}|\xi(v)-\xi_L(v)|=0$ with probability 1. By the uniform convergence of the series $\sum_{l=1}^{L}c_{l0}\psi_l(v)$ to $\xi_0(v)$ on $\mathcal{B}\backslash\mathcal{R}_0$, as $L\to\infty$, we have $\lim_{L\to\infty}\sup_{v\notin\mathcal{R}_0}|\xi_0(v)-\xi_L^0(v)|=0$. Then we can find a finite integer L', such that, for all $L>L'$, $\sup_{v\notin\mathcal{R}_0}|\xi(v)-\xi_L(v)|<\epsilon/3$ with probability 1, and $\sup_{v\notin\mathcal{R}_0}|\xi_0(v)-\xi_L^0(v)|<\epsilon/3$. Since $\psi_l(v), l=1,\ldots,L$, are all continuous functions in $\mathcal{R}$, we have $\max_{1\le l\le L}\|\psi_l(v)\|_\infty<M_{\psi,L}$, for some constant $M_{\psi,L}$. When $|c_l-c_{l0}|<\epsilon/(3LM_{\psi,L})$ for all $l=1,\ldots,L$, we have $\sup_{v\notin\mathcal{R}_0}|\xi_L(v)-\xi_L^0(v)|\le\epsilon/3$. Therefore, $|c_l-c_{l0}|<\epsilon/(3LM_{\psi,L})$, $l=1,\ldots,L$, guarantees that $\sup_{v\notin\mathcal{R}_0}|\xi(v)-\xi_0(v)|\le\epsilon$ with probability one.

For $v\in\mathcal{R}_0$, we have that,

$$\sup_{v\in\mathcal{R}_0}|\xi(v)|\le\sup_{v\in\mathcal{R}_0}|\xi(v)-\xi_L(v)|+\sup_{v\in\mathcal{R}_0}|\xi_L(v)|.$$

Similarly, we can find L and $M_{\psi,L}$, such that $|c_l|\le\omega/(2LM_{\psi,L})$, $l=1,\ldots,L$, guarantees that $\sup_{v\in\mathcal{R}_0}|\xi(v)|\le\omega$ with probability 1.

Then we have that,

$$pr\left(\|\rho-\rho_0\|_\infty<\varepsilon\right)\ge pr\Bigg(\{|c_l-c_{l0}|<\frac{\epsilon}{3LM_{\psi,L}}:L=1,2,\ldots,L\;when\;v\notin\mathcal{R}_0\}\\ \cup\{|c_l|\le\frac{\omega}{2LM_{\psi,L}}:L=1,2,\ldots,L\;when\;v\in\mathcal{R}_0\}\Bigg).$$

This completes the proof of Theorem 1. □

S1.3 Proof of Theorem 2

Based on Theorem 1, Lemma S3 shows the positivity of prior neighborhoods. We then construct sieves for $\theta(v)$ as follows:

$$\begin{aligned}\Theta_n = \Bigg\{ & \rho \in \Theta_\rho, E_+, E_- \in \Theta_E : \\ & \|\rho\|_\infty \leq H\left(m^{1/(2d)}\right), \sup_{v \in \mathcal{R}_1 \cup \mathcal{R}_{-1}} |D^\tau \rho(v)| \leq m^{1/(2d)},\ 1 \leq \|\tau\|_1 \leq \alpha \\ & \|E_{+,i}\|_\infty \leq m^{1/(2d)}, \sup_{v \in \mathcal{R}_1 \cup \mathcal{R}_{-1}} |D^\tau E_{+,i}(v)| \leq m^{1/(2d)}, \\ & \|E_{-,i}\|_\infty \leq m^{1/(2d)}, \sup_{v \in \mathcal{R}_1 \cup \mathcal{R}_{-1}} |D^\tau E_{-,i}(v)| \leq m^{1/(2d)},\ \text{for } i = 1, \ldots, n \Bigg\},\end{aligned} \tag{S2}$$

where α and m are defined in Assumption 3.

We can then find an upper bound for the tail probability, and construct the uniform consistent tests in Lemmas S4, S5, S6 and S8. These lemmas verify the three key conditions in Theorem A1 of Choudhuri et al. (2004), which leads to the posterior consistency. That is, by Lemmas S4, S5, S6 and S8, as $n \to \infty$, $m \to \infty$, we have that,

$$\begin{aligned}& \mathbb{E}_{\theta_0}\left(\Psi_n\right) \to 0, \\ & \sup_{\theta \in \mathcal{U}_\epsilon^C \cap \Theta_n} \mathbb{E}_\theta\left(1 - \Psi_n\right) \leq C_0 \exp\left(-C_1 n\right), \\ & pr\left(\Theta_n^C\right) \leq K \exp\left(-b m^{1/d}\right) \leq K \exp\left(-C_3 n\right).\end{aligned}$$

where $\mathcal{U}_\epsilon = \{\theta \in \Theta : \|\theta - \theta_0\|_1 < \epsilon\}$ for any $\epsilon > 0$, and Ψ_n is the test statistic defined in (S7). This completes the proof of Theorem 2. □

S1.4 Proof of Theorem 3

Let $\mathcal{R}_0 = \{v : \rho_0(v) = 0\}$, $\mathcal{R}_1 = \{v : \rho_0(v) > 0\}$, and $\mathcal{R}_{-1} = \{v : \rho_0(v) < 0\}$. For any $\mathcal{A} \subset \mathcal{B}$ and any integer $k \geq 1$, define

$$\mathcal{F}_k(\mathcal{A}) = \left\{\rho \in \Theta_\rho : \int_{\mathcal{A}} |\rho(v) - \rho_0(v)| \, \mathrm{d}v < \frac{1}{k}\right\}.$$

Then $\mathcal{F}_{k+1}(\mathcal{A}) \subseteq \mathcal{F}_k(\mathcal{A})$ for all k, and $\mathcal{F}_k(\mathcal{B}) \subseteq \mathcal{F}_k(\mathcal{A})$. Consider

$$\mathcal{F}_k\left(\mathcal{R}_0\right) = \left\{\rho \in \Theta_\rho : \int_{\mathcal{R}_0} |\rho(v)| \mathrm{d}v < \frac{1}{k}\right\}.$$

Define $\mathcal{U}_\epsilon^\rho = \{\rho \in \Theta_\rho : \|\rho - \rho_0\|_1 < \epsilon\}$. By Theorem 2 and the fact that $\mathcal{U}_{1/k}^\rho = \mathcal{F}_k(\mathcal{B})$, we have

$$pr\left\{\mathcal{F}_k\left(\mathcal{R}_0\right) \mid Y_+, Y_-\right\} \geq pr\left(\mathcal{U}_{1/k}^\rho \mid Y_+, Y_-\right) \to 1, \text{ as } n \to \infty.$$

In addition,

$$\{\rho(v) = 0, \text{ for all } v \in \mathcal{R}_0\} = \left\{\int_{\mathcal{R}_0} |\rho(v)| \mathrm{d}v = 0\right\} = \bigcap_{k=1}^{\infty} \mathcal{F}_k\left(\mathcal{R}_0\right).$$

By the monotonic continuity of the probability measure, we have,

$$pr\big\{\rho(v) = 0, \text{ for all } v \in \mathcal{R}_0 \mid Y_+, Y_-\big\} = \lim_{k\to\infty} pr\big\{\mathcal{F}_k\left(\mathcal{R}_0\right) \mid Y_+, Y_-\big\} = 1, \text{ as } n \to \infty.$$

For any $v_0 \in \mathcal{R}_1$ and any integer $k \geq 1$, there exists $\delta_0 > 0$, such that $|\rho(v_1) - \rho(v_0)| < 1/2k$, for any $v_1 \in \mathcal{B}(v_0, \delta_0) = \{v : \|v_1 - v_0\|_1 < \delta_0\}$. As $\mathcal{R}_1$ is an open set, there exists $\delta_1 > 0$, such that $\mathcal{B}(v_0, \delta_1) \subseteq \mathcal{R}_1$. Let $\delta = \min\{\delta_1, \delta_0\} > 0$, we have that,

$$\begin{aligned}
&\left\{\rho(v_0) > -\frac{1}{k},\ \text{for all } v_0 \in \mathcal{R}_1\right\} \\
&\supseteq \left\{\rho(v_0) > \rho(v_1) - \frac{1}{2k} \text{ and } \rho(v_1) > -\frac{1}{2k},\ \text{for some } v_1 \in \mathcal{B}(v_0, \delta),\ \text{for all } v_0 \in \mathcal{R}_1\right\} \\
&\supseteq \left\{\int_{\mathcal{B}(v_0,\delta)} \rho(v)dv > -\frac{1}{2k},\ \text{for all } v_0 \in \mathcal{R}_1\right\} \\
&\supseteq \left\{\int_{\mathcal{B}(v_0,\delta)} \rho(v)dv > \int_{\mathcal{B}(v_0,\delta)} \rho_0(v)dv - \frac{1}{2k},\ \text{for all } v_0 \in \mathcal{R}_1\right\} \\
&\supseteq \mathcal{F}_{2k}\left[\mathcal{B}(v_0, \delta)\right] \supseteq \mathcal{U}^{\rho}_{1/2k}.
\end{aligned}$$

Therefore,

$$pr\left\{\rho(v_0) > -1/k,\ \text{for all } v_0 \in \mathcal{R}_1 \mid Y_+, Y_-\right\} \geq pr\left(\mathcal{U}^{\rho}_{1/2k} \mid Y_+, Y_-\right) \to 1,$$

as $n \to \infty$. By the monotonic continuity of the probability measure, we have that,

$$pr\left\{\rho(v) > 0,\ \text{for all } v \in \mathcal{R}_1 \mid Y_+, Y_-\right\} = \lim_{k\to\infty} pr\left\{\rho(v_0) > -\tfrac{1}{k},\ \text{for all } v_0 \in \mathcal{R}_1 \mid Y_+, Y_-\right\} \to 1,$$

as $n \to \infty$. Similarly, we can obtain that $pr\left\{\rho(v) < 0, \text{for all } v \in \mathcal{R}_{-1} \mid Y_+, Y_-\right\} \to 1, n \to \infty$. This completes the proof of Theorem 3. □

S1.5 Proof of Proposition 2

We prove this proposition by sorting all the thresholding values, and derive the unnormalized density on each interval, respectively. We then obtain the full conditional density function of θ by normalizing the function on each interval as the density function.

We sort $(L_1, \ldots, L_P, U_1, \ldots, U_K)$ in ascending order, which leads to $P+K+1$ intervals, and denoted them as $I_1, I_2, \ldots, I_{P+K+1}$. For each interval I_i, $i = 1, \ldots, P + K + 1$, the full conditional distribution of θ is proportional to $\exp(-D_i\theta^2 - E_i\theta - F_i)$. We initialize $D_i = E_i = F_i = 0$, then loop through $p = 1, \ldots, P$ and $k = 1, \ldots, K$ to update D_i, E_i and F_i. More specifically, if $I_i \subset [L_p, +\infty)$, we update $D_i = D_i + a_{1p}$, $E_i = E_i + a_{2p}$, and $F_i = F_i + a_{3p}$. If $I_i \subset (-\infty, U_k]$, we update $D_i = D_i + b_{1k}$, $E_i = E_i + b_{2k}$, and $F_i = F_i + b_{3k}$. We consider three specific cases.

- If at least one of $\{a_{1p}, \ldots, a_{1P}, b_{1k}, \ldots, b_{1K}\}$ is not equal to 0, then $D_i \neq 0$, for any $i = 1, \ldots, P + K + 1$. Therefore, when $\theta \in I_i$, the full conditional distribution of θ is $\mathrm{N}\{-E_i/(2D_i), -1/(2D_i)\}$. Incorporating the normalizing constant M_i for each interval, which is independent of θ, the full conditional distribution of θ is the mixture of truncated normal distributions, $\sum_{i=1}^{P+K+1} M_i \cdot \mathrm{TruncatedNormal}_{I_i}\{-E_i/(2D_i), -1/(2D_i)\}$.

- If $a_{1p} = b_{1k} = 0$, for any $p = 1, \ldots, P$ and $k = 1, \ldots, K$, and at least one of $\{a_{2p}, \ldots, a_{2P}, b_{2k}, \ldots, b_{2K}\}$ is not equal to 0, then $D_i = 0$ and $E_i \neq 0$, for any $i = 1, \ldots, P + K + 1$. Therefore, when $\theta \in I_i$, the full conditional distribution of θ is the exponential distribution $\text{Exp}(E_i)$. Incorporating the normalizing constant M_i, the full conditional distribution of θ is $\sum_{i=1}^{P+K+1} M_i \cdot \text{Exponential}_{I_i}(E_i)$.

- If $a_{1p} = b_{1k} = a_{2p} = b_{2k} = 0$, for any $p = 1, \ldots, P$ and $k = 1, \ldots, K$, and at least one of $\{a_{3p}, \ldots, a_{3P}, b_{3k}, \ldots, b_{3K}\}$ is not equal to 0, then $D_i = E_i = 0$, and at least one of $F_i \neq 0$, for any $i = 1, \ldots, P + K + 1$. Therefore, when $\theta \in I_i$, the full conditional distribution of θ is proportional to the uniform distribution on $I_i = [u_{1i}, u_{2i}]$. Incorporating the normalizing constant M_i, the full conditional distribution of θ is $\sum_{i=1}^{P+K+1} M_i \cdot \text{U}(u_{1i}, u_{2i})$.

This completes the proof of Proposition 2. □

S2. Additional Lemmas

Lemma S1 *Rewrite* $\rho(v) = T_\omega\{\xi(v); \tau_1^2(v), \tau_2^2(v)\}$ *in Equation* (2.6). *Then* $T_\omega(\cdot)$ *is a piecewise Lipschitz continuous function for any* ω.

Proof: From Equation (2.6), it is straightforward to verify that $\rho(v)$ can be written as

$$\begin{aligned}
\rho(v) &= \text{Corr}\{Y_{1,i}(v), Y_{2,i}(v)\} \\
&= \frac{G_\omega^2\{\xi(v)\} - G_\omega^2\{-\xi(v)\}}{\left[G_\omega^2\{\xi(v)\} + G_\omega^2\{-\xi(v)\} + \tau_1^2(v)\right]^{1/2}\left[G_\omega^2\{\xi(v)\} + G_\omega^2\{-\xi(v)\} + \tau_2^2(v)\right]^{1/2}} \\
&= \frac{\text{sgn}\{\xi(v)\} R_\omega^2\{\xi(v)\}}{\left[R_\omega^2\{\xi(v)\} + \tau_1^2(v)\right]^{1/2}\left[R_\omega^2\{\xi(v)\} + \tau_2^2(v)\right]^{1/2}},
\end{aligned}$$

where $R_\omega(x) = G_\omega(x) - G_\omega(-x)$. Without loss of generality, suppose $\tau_1^2(v)$ and $\tau_2^2(v)$ are both equal to one. Then $T_\omega(x) = H\{R_\omega(x)\}$, where $H(t) = t^2/(t^2+1)$ when $\xi(v) > \omega$, $H(t) = -t^2/(t^2+1)$ when $\xi(v) < -\omega$, and $H(t) = 0$ otherwise. Since $H(t)$ is continuous and $|H'(t)| \leq 1/(2\omega)$, $H(t)$ is Lipschitz continuous. As $R_\omega(x)$ is the hard thresholding function, which is piecewise Lipschitz continuous function, $T_\omega(x) = H\{R_\omega(x)\}$ is also a piecewise Lipschitz continuous function. This completes the proof of Lemma S1. □

Lemma S2 *Given $\rho(v) = T_\omega\{\xi(v); \tau_1^2(v), \tau_2^2(v)\}$ in (6), there exist a piecewise Lipschitz continuous function $s(\cdot)$, such that $G_\omega\{\xi(v)\} = s\{\rho(v); \tau_1^2(v), \tau_2^2(v)\}$.*

Proof: It is straightforward to show that $G_\omega\{\xi(v)\} = s\left\{\rho(v); \tau_1^2(v), \tau_2^2(v)\right\}$, and $G_\omega\{-\xi(v)\} = s\left\{-\rho(v); \tau_1^2(v), \tau_2^2(v)\right\}$, where $s(x; t_1, t_2)$ is as given in (7). Therefore, $s(\cdot)$ is a piecewise Lipschitz continuous function. This completes the proof of Lemma S2. □

Lemma S3 *Let $\Pi_{n,i}(\cdot; \theta)$ denote the density function of $Z_{n,i} = (Y_{+,i}, Y_{-,i})$. Define $\Lambda_{n,i}(\cdot; \theta_0, \theta)$ $= \log \pi_{n,i}(\cdot; \theta) - \log \pi_{n,i}(\cdot; \theta_0)$,*
$K_{n,i}(\theta_0, \theta) = \mathbb{E}_{\theta_0}\left\{\Lambda_{n,i}\left(Z_{n,i}; \theta_0, \theta\right)\right\}$, and $V_{n,i}\left(\theta_0, \theta\right) = \mathrm{var}_{\theta_0}\left\{\Lambda_{n,i}(Z_{n,i}; \theta_0, \theta)\right\}$. There exists a set O with $\Pi(O) > 0$, such that, for any $\epsilon > 0$,

$$\liminf_{n\to\infty} \Pi\left[\left\{\theta \in O, n^{-1}\sum_{i=1}^{n} K_{n,i}\left(\theta_0, \theta\right) < \epsilon\right\}\right] > 0 \text{ and } n^{-2}\sum_{i=1}^{n} V_{n,i}\left(\theta_0, \theta\right) \to 0 \text{ for } \theta \in O.$$

Proof: The density function is of the form,

$$\Pi_{n,i}(Z_{n,i}; \theta) = \sum_{v\in\mathcal{B}_m} \frac{1}{2\pi u^2(v)\{1-r^2(v)\}^{1/2}} \cdot \exp\left[-\frac{W_i(v)}{2\{1-r^2(v)\}u^2(v)}\right],$$

where $W_i(v) = \{Y_{+,i}(v) - \mu_{+,i}(v)\}^2 + \{Y_{-,i}(v) - \mu_{-,i}(v)\}^2 + 2r(v)\{Y_{+,i}(v)\mu_{-,i}(v) + Y_{-,i}(v)\mu_{+,i}(v)\}$, $r(v) = \{\tau_1^2(v) - \tau_2^2(v)\}/\{\tau_1^2(v) + \tau_2^2(v)\}$, and $u^2(v) = \{\tau_1^2(v) + \tau_2^2(v)\}/4$. Therefore, we have,

$$\begin{aligned}
\Lambda_{n,i}\left(Z_{n,i};\theta_0,\theta\right) &= \log\Pi(Z_{n,i};\theta) - \log\Pi(Z_{n,i};\theta_0) \\
&= \sum_{v\in\mathcal{B}_m}\left[-\frac{1}{2\{1-r^2(v)\}u^2(v)}\right]\Big[\mu_{+,i}^2(v) - \mu_{+,i,0}^2(v) + \mu_{-,i}^2(v) - \mu_{+,i,0}^2(v) \\
&\quad + 2Y_{+,i}(v)\{\mu_{+,i,0}(v) - \mu_{+,i}(v)\} + 2Y_{-,i}(v)\{\mu_{-,i,0}(v) - \mu_{-,i}(v)\}(v) \\
&\quad + 2rY_{+,i}(v)\{\mu_{-,i}(v) - \mu_{-,i,0}(v)\} + 2rY_{-,i}(v)\{\mu_{+,i}(v) - \mu_{+,i,0}(v)\}\Big],
\end{aligned}$$

$$\begin{aligned}
K_{n,i}\left(\theta_0,\theta\right) &= \mathbb{E}_{\theta_0}\left\{\Lambda_{n,i}\left(Z_{n,i};\theta_0,\theta\right)\right\} \\
&= \sum_{v\in\mathcal{B}_m}\Big(-\frac{1}{2\{1-r^2(v)\}u^2(v)}\Big[\{\mu_{+,i}(v) - \mu_{+,i,0}(v)\}^2 + \{\mu_{-,i}(v) - \mu_{-,i,0}(v)\}^2 \\
&\quad + 2r(v)\mu_{+,i,0}(v)\mu_{-,i}(v) + 2r(v)\mu_{-,i,0}(v)\mu_{+,i}(v) \\
&\quad - 2r(v)\mu_{+,i,0}(v)\mu_{-,i,0}(v) - 2r(v)\mu_{-,i,0}(v)\mu_{+,i,0}(v)\Big]\Big).
\end{aligned}$$

Given any $\zeta > 0$, let $O(\zeta) = \{\theta : \|\theta - \theta_0\|_\infty < \zeta\}$, with

$$\|\theta-\theta_0\|_\infty = \max_{v\in\mathcal{V}(\rho)\cup\mathcal{V}(\rho_0)}\left\{\|\rho-\rho_0\|_\infty, \max_{1\le i\le n}\|E_{-,i} - E_{-,i,0}\|_\infty, \max_{1\le i\le n}\|E_{+,i} - E_{+,i,0}\|_\infty\right\},$$

and $\mathcal{V}(\rho) = \{v : \rho(v) \neq 0\}$, $\mathcal{V}(\rho_0) = \{v : \rho_0(v) \neq 0\}$, then, for any $v \in O(\zeta)$,

$$\begin{aligned}
&|\mu_{i,+}(v) - \mu_{i,+,0}(v)| \le |s\{\rho(v)\}E_{+,i}(v) - s\{\rho_0(v)\}E_{i,+,0}(v)| \\
&\le |E_{+,i}(v)\left(s\{\rho(v)\} - s\{\rho_0(v)\}\right)| + |s\{\rho_0(v)\}\left(E_{+,i}(v) - E_{i,+,0}(v)\right)| \le K_1\zeta,
\end{aligned}$$

where the last inequality is due to the compactness and convexity of $\mathcal{B}_m$, and

$$K_1 = \max_{v \in \mathcal{V}(\rho) \cup \mathcal{V}(\rho_0)} \left\{ E_{+,i}(v), s\{\rho_0(v)\} \right\}.$$

Similarly, we have $|\mu_{i,-}(v) - \mu_{i,-,0}(v)| \leq K_2 \zeta$, for any v, where

$$K_2 = \max_{v \in \mathcal{V}(\rho) \cup \mathcal{V}(\rho_0)} \left\{ E_{-,i}(v), s\{-\rho_0(v)\} \right\}.$$

Therefore, we have that,

$$\begin{aligned}
\left| \sum_{i=1}^{n} K_{n,i}(\theta, \theta_0) \right| \leq & \sum_{v \in \mathcal{V}(\rho) \cup \mathcal{V}(\rho_0)} \frac{1}{2\left\{1 - r^2(v)\right\} u^2(v)} \left(\sum_{i=1}^{n} |\mu_{i,+}(v) - \mu_{i,+,0}(v)|^2 \right. \\
& + \sum_{i=1}^{n} |\mu_{i,-}(v) - \mu_{i,-,0}(v)|^2 \\
& \left. + 2r(v) M \sum_{i=1}^{n} |\mu_{i,-}(v) - \mu_{i,-,0}(v)| + 2r(v) M \sum_{i=1}^{n} |\mu_{i,+}(v) - \mu_{i,+,0}(v)| \right) \\
\leq & \sum_{v \in \mathcal{V}(\rho) \cup \mathcal{V}(\rho_0)} \frac{1}{2\{1 - r^2(v)\} u^2(v)} \left(n K_1^2 \zeta^2 + n K_2^2 \zeta^2 + 2|r(v)| M n (K_1 + K_2) \zeta \right) \\
\leq & A n \zeta^2 + B n \zeta,
\end{aligned}$$

where

$$\begin{aligned}
M &= \max_{v \in \mathcal{V}(\rho) \cup \mathcal{V}_0(\rho_0), \forall i} \{\mu_{+,i,0}(v), \mu_{-,i,0}(v)\}, \\
A &= (K_1^2 + K_2^2) \sum_{v \in \mathcal{V}(\rho) \cup \mathcal{V}(\rho_0)} \frac{1}{2\{1 - r^2(v)\} u^2(v)}, \\
B &= M(K_1 + K_2) \sum_{v \in \mathcal{V}(\rho) \cup \mathcal{V}(\rho_0)} \frac{|r(v)|}{2\{1 - r^2(v)\} u^2(v)}.
\end{aligned}$$

Henceforth, for any $\epsilon > 0$, we obtain that,

$$\liminf_{n\to\infty} \Pi\left[\left\{\theta \in O, n^{-1}\sum_{i=1}^{n} K_{n,i}\left(\theta_0,\theta\right) < \epsilon\right\}\right] > 0.$$

Similarly, we have that,

$$\begin{aligned} V_{n,i}\left(\theta_0,\theta\right) = & \sum_{v\in\mathcal{V}(\rho)\cup\mathcal{V}(\rho_0)} \frac{1}{\{1-r^2(v)\}u^2(v)} \left[\{\mu_{+,i}(v)-\mu_{+,i,0}(v)\}^2 + \{\mu_{-,i}(v)-\mu_{-,i,0}(v)\}^2\right. \\ & \left. +\{r^3(v)-3r(v)\}\{\mu_{+,i}(v)-\mu_{+,i,0}(v)\}\{\mu_{-,i}(v)-\mu_{-,i,0}(v)\}\right], \\ |V_{n,i}\left(\theta_0,\theta\right)| \leq & \sum_{v\in\mathcal{V}(\rho)\cup\mathcal{V}(\rho_0)} \frac{1}{\{1-r^2(v)\}u^2(v)} \left(K_1^2\zeta^2 + K_2^2\zeta^2 + |r^3(v)-3r(v)|K_1K_2\zeta^2\right) \leq C\zeta^2, \end{aligned}$$

where

$$C = (K_1^2+K_2^2)\sum_{v\in\mathcal{V}(\rho)\cup\mathcal{V}(\rho_0)} \frac{1}{\{1-r^2(v)\}u^2(v)} + K_1K_2 \sum_{v\in\mathcal{V}(\rho)\cup\mathcal{V}(\rho_0)} \frac{|r^3(v)-3r(v)|}{\{1-r^2(v)\}u^2(v)}.$$

Henceforth, we obtain that,

$$\left|\sum_{i=1}^{n} V_{n,i}\left(\theta_0,\theta\right)\right| \leq nC\zeta^2 \text{ and } \frac{1}{n^2}\sum_{i=1}^{n} V_{i,n}\left(\theta_0,\theta\right) \to 0, \text{as } n\to\infty.$$

This completes the proof of Lemma S3. □

Given the sieves we construct in Equation (S2), we next derive an upper bound for the tail probability, and construct the uniform consistent tests in Lemmas S4, S5, S6 and S8.

Lemma S4 *Suppose* $\rho \sim \mathrm{TCGP}(\omega_0,\kappa)$ *with* $\omega_0 > 0$*, the kernel function* κ *satisfies Assumption 2, and* $E_{+,i}, E_{-,i} \sim \mathcal{GP}(0, I)$*, for* $i = 1,\ldots,n$*. Then there exist constants* K *and* *b, such that pr* $\left(\Theta_n^C\right) \leq K\exp(-C_3 n)$.

Proof: Following the same notation as that in the proof of Lemma S1, we have $\rho(v) = T_\omega\{\xi(v)\} = H[R_\omega\{\xi(v)\}]$. Let $\mathcal{R}_1 = \{v : \rho(v) > 0\}$, and $\mathcal{R}_{-1} = \{v : \rho(v) < 0\}$. We have $R_\omega\{\xi(v)\} = \xi(v) > \omega$ when $v \in \mathcal{R}_1$, and $R_\omega\{\xi(v)\} = \xi(v) < -\omega$ when $v \in \mathcal{R}_{-1}$. Then

$$pr\left(\Theta_n^C\right) \leq pr\left\{\sup_{v\in\mathcal{R}_1\cup\mathcal{R}_{-1}} |H(\xi(v))| > H\left(m^{1/2d}\right)\right\} \tag{S3}$$

$$+ \sum_{\tau:1\leq\|\tau\|_1\leq\alpha} pr\left\{\sup_{v\in\mathcal{R}_1\cup\mathcal{R}_{-1}} |D^\tau H(\xi(v))| > m^{1/2d}\right\}$$
$$+ \sum_{i=1}^n pr\left\{\sup_{v\in\mathcal{R}_1\cup\mathcal{R}_{-1}} |E_{+,i}| > m^{1/2d}\right\} + \sum_{i=1}^n pr\left\{\sup_{v\in\mathcal{R}_1\cup\mathcal{R}_{-1}} |E_{-,i}| > m^{1/2d}\right\}$$
$$+ \sum_{i=1}^n \sum_{\tau:1\leq\|\tau\|_1\leq\alpha} pr\left\{\sup_{v\in\mathcal{R}_1\cup\mathcal{R}_{-1}} |D^\tau E_{+,i}| > m^{1/2d}\right\}$$
$$+ \sum_{i=1}^n \sum_{\tau:1\leq\|\tau\|_1\leq\alpha} pr\left\{\sup_{v\in\mathcal{R}_1\cup\mathcal{R}_{-1}} |D^\tau E_{-,i}| > m^{1/2d}\right\}. \tag{S4}$$

Since $H(t)$ is a monotonic function,

$$pr\left\{\sup_{v\in\mathcal{R}_1\cup\mathcal{R}_{-1}} |H(\xi(v))| > H(m^{1/2d})\right\} \leq pr\left\{\sup_{v\in\mathcal{R}_1\cup\mathcal{R}_{-1}} |\xi(v)| > m^{1/2d}\right\}$$
$$\leq K_1\exp\left(-b_1 m^{1/d}\right) + K_{-1}\exp\left(-b_{-1} m^{1/d}\right),$$

where the existence of K_1, K_{-1}, b_1, b_{-1} in the second inequality is ensured by Theorem 5 of Ghosal and Roy (2006).

We next consider the second term in (S3). Since $|H'(t)| \leq 1$ and $|H''(x)| \leq 2$, we have,

$$\sum_{\tau:1\leq\|\tau\|_1\leq\alpha} pr\left\{\sup_{v\in\mathcal{R}_1\cup\mathcal{R}_{-1}} |D^\tau H(\xi(v)-\omega)| > m^{1/2d}\right\}$$
$$\leq pr\left\{\sup_{v\in\mathcal{R}_1\cup\mathcal{R}_{-1}} |D^\tau \xi(v)| > m^{1/2d}\right\} + pr\left\{\sup_{v\in\mathcal{R}_1\cup\mathcal{R}_{-1}} |2\cdot D^\tau \xi(v)| > m^{1/2d}\right\}$$
$$\leq \sum_{\tau:0<\|\tau\|_1\leq\alpha} K_\tau \exp\left(-b_\tau m^{1/d}\right).$$

Denote the sum of the last four terms in (S3) as S_E. By Theorem 5 of Ghosal and Roy (2006) again, there exist K_{E_+}, b_{E_+},K_{E_-}, b_{E_-}, K_{E_τ} and b_{E_τ}, such that

$$S_E \leq K_{E_+}\exp(-b_{E_+}m^{1/d}) + K_{E_-}\exp(-b_{E_-}m^{1/d}) + \sum_{\tau:0<\|\tau\|_1\leq\alpha} K_{E_\tau}\exp\left(-b_{E_\tau}m^{1/d}\right).$$

Taking $K = K_{-1} + K_1 + K_{E_+} + K_{E_-} + \sum_{\tau:0<\|\tau\|\leq\alpha} K_\tau + \sum_{\tau:0<\|\tau\|\leq\alpha} K_{E_\tau}$, and $b = \min\left\{b_{-1}, b_1, b_{E_+}, b_{E_-}, \min_{1\leq|\tau|\leq\alpha} b_\tau, \min_{1\leq|\tau|\leq\alpha} b_{E_\tau}\right\}$, we have,

$$pr\left(\Theta_n^C\right) \leq K\exp\left(-bm^{1/d}\right) \leq K\exp\left(-C_3 n\right).$$

This completes the proof of Lemma S4. □

Lemma S5 *Suppose Assumption 1 holds. The hypothesis testing problem,*

$$H_0 : \rho(v) = \rho_0(v), \quad E_{\pm,i}(v) = E_{\pm,i,0}(v), \quad i = 1,\ldots,n,\ v \in \mathcal{V}(\rho_1)\cup\mathcal{V}(\rho_0),$$
$$H_1 : \rho(v) = \rho_1(v), \quad E_{\pm,i}(v) = E_{\pm,i,1}(v),$$

is equivalent to the hypothesis testing problem,

$$H_0^*: \mu_{\pm,i}(v) = \mu_{\pm,i,0}(v), \quad i = 1, \ldots, n,\ v \in \mathcal{V}(\rho_1) \cup \mathcal{V}(\rho_0),$$

$$H_1^*: \mu_{\pm,i}(v) = \mu_{\pm,i,1}(v),$$

where $\mathcal{V}(\rho_1) = \{v : \rho_1(v) \neq 0\}$ and $\mathcal{V}(\rho_0) = \{v : \rho_0(v) \neq 0\}$.

Proof: For any $k \in \{0, 1\}$, it is straightforward to see that if H_k holds, then H_k^* also holds. We show that, if H_k^* holds, then H_k also holds. For any $v \in \mathcal{B}_m$,

$$\begin{aligned}
0 &= \sum_{i=1}^{n} [s\{\rho(v)\}E_{+,i}(v) - s\{\rho_k(v)\}E_{+,i,k}(v)]^2 \\
&= \sum_{i=1}^{n} \left[s\{\rho(v)\}^2 E_{+,i}(v)^2 - 2s\{\rho(v)\}s\{\rho_k(v)\}E_{+,i,1}(v)E_{+,i,k}(v) + s\{\rho_k(v)\}^2 E_{+,i,0}(v)^2\right] \\
&= [s\{\rho(v)\} - s\{\rho_0(v)\}]^2 \sum_{i=1}^{n} E_{+,i,k}^2(v) + s\{\rho_k(v)\}s\{\rho(v)\} \sum_{i=1}^{n} \{E_{+,i}(v) - E_{+,i,k}(v)\}^2 \\
&\quad + s\{\rho_0(v)\}[s\{\rho(v)\} - s\{\rho_0(v)\}] \sum_{i=1}^{n} \left\{E_{+,i}(v)^2 - E_{+,i,k}(v)^2\right\},
\end{aligned}$$

By Definition 4, we have $\sum_{i=1}^n E_{+,i}(v)^2 = \sum_{i=1}^n E_{+,i,0}(v)^2 = \sum_{i=1}^n E_{+,i,1}(v)^2$. When $v \in \mathcal{V}(\rho_1) \cup \mathcal{V}(\rho_0)$, $s\{\rho_0(v)\} \geq 0$, $s\{\rho_1(v)\} \geq 0$, and at least one of $s\{\rho_0(v)\}$ and $s\{\rho_1(v)\}$ is not equal to 0,

$$s\{\rho(v)\} - s\{\rho_k(v)\} = 0, \quad E_{+,i}(v) - E_{+,i,k}(v) = 0, \quad i = 1, \ldots, n.$$

Similarly, we have that $E_{-,i}(v) - E_{-,i,k}(v) = 0$ for any $v \in \mathcal{V}(\rho_1) \cup \mathcal{V}(\rho_0)$, $i = 1, \ldots, n$. Since $s(\cdot)$ is a monotonic function, $\rho(v) = \rho_k(v)$ for any $v \in \mathcal{B}_m$, which ccompletes the proof of Lemma S5. $\square$

Lemma S6 *For the hypothesis testing problem,*

$$H_0 : \mu_{\pm,i}(v_j) = \mu_{\pm,i,0}(v_j), \quad i = 1, \ldots, n,\ v_j \in \mathcal{V}(\rho_1) \cup \mathcal{V}(\rho_0),\ j = 1, \ldots, m,$$

$$H_1 : \mu_{\pm,i}(v_j) = \mu_{\pm,i,1}(v_j),$$

construct the testing statistic, $\Psi_n = \Psi_{+n} + \Psi_{-n} - \Psi_{+n}\Psi_{-n}$*, where*

$$\Psi_{\pm n} = \max_{i=1,\ldots,n} \left\{ I\left(\sum_{j=1}^{m} \delta_{\pm,i}(v_j)(Y_{\pm,i}(v_j) - \mu_{\pm,i,0}(v_j)) > 2\left(\frac{m}{C_0}\right)^{\frac{\nu}{d}+\frac{1}{2d}} \right) \right\},$$

$\delta_{\pm,i}(v_j) = 2I\{\mu_{\pm,i,1}(v_j) \geq \mu_{\pm,i,0}(v_j)\} - 1$*,* $\nu_0/2 < \nu < 1/2$*, and* ν_0, d*,* C_0 *are as defined in Assumption 3. Write* $\mu = \{\mu_{i,\pm}(v_j)\}$*, and* $\mu_k = \{\mu_{i,\pm,k}(v_j)\}$ *for* $k = 0, 1$*. Then, for any* $\epsilon_0 > 0$*, there exist constants* C_0*,* C_1 *and* $i_* \in \{1, \ldots, n\}$*, such that, for any* μ_1 *and* μ_0 *satisfying that* $\sum_{j=1}^{m} |\mu_{+,i_*,1}(v_j) - \mu_{+,i_*,0}(v_j)| > m\epsilon_0$*, or* $\sum_{j=1}^{m} |\mu_{-,i_*,1}(v_j) - \mu_{-,i_*,0}(v_j)| > m\epsilon_0$*, and* μ *satisfying that* $\|\mu - \mu_1\|_\infty < \epsilon_0/4$*, we have* $\mathbb{E}_{\mu_0}(\Psi_n) < C_0 \exp(-2n^{2\nu})$ *and* $\mathbb{E}_{\mu}(\Psi_n) < C_0 \exp(-C_1 n)$*.*

Proof: To bound the type I error, we have $\mathbb{E}_{\mu_0}(\Psi_n) \leq \mathbb{E}_{\mu_0}(\Psi_{+n}) + \mathbb{E}_{\mu_0}(\Psi_{-n})$. By Assumption 3, we have $(m/C_0)^{\nu/d} \geq n^\nu$. By the definition of Ψ_{+n}, we have that,

$$\begin{aligned}
\mathbb{E}_{\mu_0}(\Psi_{+n}) &\leq pr\left(\sum_{j=1}^{m} \delta_{+,i_*}(v_j)\{Y_{+,i_*}(v_j) - \mu_{+,i_*,0}(v_j)\} > 2\left(\frac{m}{C_0}\right)^{\frac{\nu}{d}+\frac{1}{2d}} \right) \\
&= pr\left(\sqrt{\frac{C_0}{m^d}} \sum_{j=1}^{m} \delta_{+,i_*}(v_j)\{Y_{+,i_*}(v_j) - \mu_{+,i_*,0}(v_j)\} > 2\left(\frac{m}{C_0}\right)^{\frac{\nu}{d}} \right) \\
&= 1 - \Phi\left(2\left(\frac{m}{C_0}\right)^{\frac{\nu}{d}} \right) \leq 1 - \Phi\left(2n^\nu\right) \leq \frac{\phi(2n^\nu)}{2n^\nu} = \frac{1}{2\sqrt{2\pi}} \frac{\exp(-2n^{2\nu})}{n^\nu}.
\end{aligned}$$

Similarly, we have that $\mathbb{E}_{\mu_0}(\Psi_{-n}) \leq \dfrac{1}{2\sqrt{2\pi}} \dfrac{\exp(-2n^{2\nu})}{n^{\nu}}$. Therefore,

$$\mathbb{E}_{\mu_0}(\Psi_n) \leq \frac{1}{\sqrt{2\pi}} \frac{\exp(-2n^{2\nu})}{n^{\nu}}.$$

To bound the type II error, we have that,

$$\mathbb{E}_{\mu}\left[1 - \Psi_n\right] \leq \min\left\{\mathbb{E}_{\mu}\left(1 - \Psi_{+n}\right), \mathbb{E}_{\mu}\left(1 - \Psi_{-n}\right)\right\}.$$

As such, we only need to show that at least one of the type II error probabilities for Ψ_{+n} and Ψ_{-n} is exponentially small. Suppose $\sum_{j=1}^{m} |\mu_{+,i_*,0}(v_j) - \mu_{+,i_*,1}(v_j)| > m\epsilon_0$. Since $\sum_{j=1}^{m} |\mu_{+,i_*}(v_j) - \mu_{+,i_*,1}(v_j)| < m\epsilon_0/4$, we have,

$$\begin{aligned}
&\mathbb{E}_{\mu}(1 - \Psi_{+n}) \\
&\leq pr\left(\sum_{j=1}^{m} \delta_{+,i_*}(v_j)\{Y_{+,i_*}(v_j) - \mu_{i_*,+,0}(v_j)\} > 2\left(\frac{m}{C_0}\right)^{\frac{\nu}{d}+\frac{1}{2d}}\right) \\
&= pr\left(\sqrt{\frac{C_0}{m^d}}\sum_{j=1}^{m} \delta_{+,i_*}(v_j)\{Y_{+,i}(v_j) - \mu_{+,i,0}(v_j)\} \leq 2\left(\frac{m}{C_0}\right)^{\frac{\nu}{d}}\right) \\
&= pr\left(\sqrt{\frac{C_0}{m^d}}\sum_{j=1}^{m} \delta_{+,i_*}(v_j)\{Y_{+,i}(v_j) - \mu_{+,i}(v_j)\} + \sqrt{\frac{C_0}{m^d}}\sum_{j=1}^{m} \delta_{+,i_*}(v_j)\{\mu_{+,i}(v_j) - \mu_{+,i,1}(v_j)\}\right. \\
&\qquad \left. + \sqrt{\frac{C_0}{m^d}}\sum_{j=1}^{m} \delta_{+,i_*}(v_j)\{\mu_{+,i,1}(v_j) - \mu_{+,i,0}(v_j)\} < 2(m/C_0)^{\nu/d}\right) \\
&\leq pr\left(\sqrt{\frac{C_0}{m^d}}\sum_{j=1}^{m} \delta_{+,i_*}(v_j)\{Y_{+,i}(v_j) - \mu_{+,i}(v_j)\} \leq \frac{C_0\epsilon_0 m^{1/2d}}{4} - C_0\epsilon_0 m^{1/2d} + 2(m/C_0)^{\nu/d}\right).
\end{aligned}$$

Since $\nu < 1/2$, there exists $N > N_0$, such that, for all $n \geq N$, $(m/C_0)^{\nu/d} < C_0 m^{1/2d}\epsilon_0/4$.

By Assumption 3, this further implies that,

$$\mathbb{E}_\mu(1-\Psi_{+n}) \le pr\left(\sqrt{\frac{C_0}{m^d}}\sum_{j=1}^m \delta_{+,i_*}(v_j)\{Y_{+,i_*}(v_j)-\mu_{+,i_*}(v_j)\} \le -\frac{C_0\epsilon_0 m^{1/2d}}{4}\right)$$
$$\le \Phi\left(-\frac{C_0\epsilon_0 m^{1/2d}}{4}\right) \le \Phi\left(-\frac{\epsilon_0 n^{1/2}}{4}\right) \le \frac{4}{\epsilon_0(2\pi n)^{1/2}}\exp\left(-\frac{n\epsilon_0^2}{32}\right).$$

Taking $C_0 = \max\left\{2^{-1}(2\pi)^{-1/2}, 4\epsilon_0^{-1}(2\pi)^{-1/2}\right\}$ and $C_1 = \epsilon_0^2/32$ completes the proof of Lemma S6. □

Lemma S7 *Suppose Assumption 1, 2 and 3 hold. For any $\epsilon > 0$, there exist N, i and $\epsilon_0 > 0$, such that, for all $n \ge N$ and all $\theta \in \Theta_n$ that $\|\theta-\theta_0\|_1 > \varepsilon$, we have $\sum_{j=1}^m |\mu_{\pm,i}(v_j) - \mu_{\pm,i,0}(v_j)| > \epsilon_0 m$.*

Proof: We first note that,

$$\|\theta-\theta_0\|_1 = \sum_{v\in\mathcal{V}(\rho)\cup\mathcal{V}(\rho_0)} |\rho(v)-\rho_0(v)| + \max_{i=1,\dots,n}\sum_{v\in\mathcal{V}(\rho)\cup\mathcal{V}(\rho_0)} |E_{+,i}(v)-E_{+,i,0}(v)|$$
$$+ \max_{i=1,\dots,n}\sum_{v\in\mathcal{V}(\rho)\cup\mathcal{V}(\rho_0)} |E_{-,i}(v)-E_{-,i,0}(v)| \tag{S5}$$

Since $\|\theta-\theta_0\|_1 > \epsilon$, at lease one of the three terms in (S5) is greater than $\epsilon/3$. Without loss of generality, suppose $\max_{i=1,\dots,n}\left\{\sum_{v\in\mathcal{V}(\rho)\cup\mathcal{V}(\rho_0)} |E_{+,i}(v)-E_{+,i,0}(v)|\right\} > \epsilon/3$. Then there exist i, such that $\sum_{v\in\mathcal{V}(\rho)\cup\mathcal{V}(\rho_0)} |E_{+,i}(v)-E_{+,i,0}(v)| > \epsilon/3$. Therefore,

$$\sum_{j=1}^m |\mu_{\pm,i}(v_j)-\mu_{\pm,i,0}(v_j)| = \sum_{j=1}^m |s\{\rho(v_j)\}E_{+,i}(v) - s\{\rho_0(v_j)\}E_{+,i,0}(v)|$$
$$= \sum_{j=1}^m |s\{\rho(v_j)\}\{E_{+,i}(v)-E_{+,i,0}(v)\} + E_{+,i,0}(v)[s\{\rho(v_j)\}-s\{\rho_0(v_j)\}]| \tag{S6}$$
$$> \sum_{j=1}^m |s\{\rho(v_j)\}|\,|E_{+,i}(v_j)-E_{+,i,0}(v_j)| - \sum_{j=1}^m |E_{+,i,0}(v_j)|\,|s(\rho(v_j))-s(\rho_0(v_j))|$$

By Definition 3, there exists $C_\rho > 0$, such that $|s\{\rho(v_j)\}| > C_\rho$ when $v_j \in \mathcal{V}(\rho) \cup \mathcal{V}(\rho_0)$. By the compactness of $\mathcal{V}(\rho) \cup \mathcal{V}(\rho_0)$, there exists C, such that $\max_{j=1,\ldots,m} |E_{+,i,0}(v_j)| \, |s(\rho(v_j)) - s(\rho_0(v_j))| < C$. Therefore,

$$\sum_{j=1}^{m} |\mu_{+,i}(v_j) - \mu_{+,i,0}(v_j)| > C_\rho m\epsilon/3 - mC$$

Taking $\epsilon_0 = C_\rho \epsilon/3 - C$ completes the proof of Lemma S7. □

Lemma S8 *For any $\epsilon^\star > 0$ and $\nu_0 < \nu < \frac{1}{2}$, there exist N, C_0, C_1 and C_2, such that, for all $n > N$ and $\theta \in \Theta_n$, if $\|\theta - \theta_0\|_1 > \epsilon^\star$, a test function Ψ_n can be constructed satisfying that $\mathbb{E}_{\theta_0}(\Psi_n) \leq C_0 \exp(-C_2 n^{2\nu})$ and $\mathbb{E}_\theta(1 - \Psi_n) \leq C_0 \exp(-C_1 n)$, where ν_0 is as defined in Assumption 3.*

Proof: Let N_t be the t covering number of Θ_n in the supremum norm. Let $\theta^1, \ldots, \theta^{N_t} \in \Theta_n$ satisfy that, for each $\theta \in \Theta_n$, there exist at least one l such that $\left\|\theta - \theta^l\right\|_\infty < t$. For any $\theta \in \Theta_n$, define

$$\Psi_n = \max_{1 \leq l \leq N_t} \Psi_n\left(\theta_0, \theta^l\right), \tag{S7}$$

where $\Psi_n\left(\theta_0, \theta^l\right)$ is the test statistic constructed in Lemma S6 for the hypothesis testing problem $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta^l$. If $\|\theta - \theta_0\|_1 > \epsilon^\star$, then for θ^l satisfying that $\left\|\theta - \theta^l\right\|_1 < t \leq \epsilon^\star/2$, we have $\left\|\theta^l - \theta_0\right\|_1 > \epsilon^\star/2$. By Lemma S7, there exist N_0^*, i and $\epsilon > 0$, such that $\sum_{j=1}^m \left|\mu_{+,i}^l(v_j) - \mu_{+,i,0}(v_j)\right| > \epsilon m$. By Lemma S6, we can choose ϵ_0, such that

$$\mathbb{E}_{\theta_0}\left\{\Psi_n\left(\theta_0, \theta^l\right)\right\} \leq C_0 \exp\left(-2n^{2\nu}\right), \quad \text{and} \quad \mathbb{E}_\theta\left\{1 - \Psi_n\left(\theta_0, \theta^l\right)\right\} \leq C_0 \exp\left(-C_1 n\right).$$

Furthermore, we have,

$$\begin{aligned}\mathbb{E}_{\theta_0}(\Psi_n) &\le \sum_{l=1}^{N_t} \Psi_n\left(\theta_0, \theta^l\right) \le C_0 N_t \exp\left(-2n^{2\nu}\right) = C_0 \exp\left(\log N_t - 2n^{2\nu}\right) \\ &\le C_0 \exp\left\{C n^{1/(2\alpha)} t^{-d/\alpha} - 2n^{2\nu}\right\} \le C_0 \exp\left(C n^{\nu_0} t^{-d/\alpha} - 2n^{2\nu}\right) \\ &= C_0 \exp\left\{-\left(2 - C n^{\nu_0 - 2\nu} t^{-d/\alpha}\right) n^{2\nu}\right\}.\end{aligned}$$

When $Ct^{-d/\alpha} < 2$, $\mathbb{E}_{\theta_0}(\Psi_n) \le C_0 \exp\left\{-\left(2 - Ct^{-d/\alpha}\right) n^{2\nu}\right\}$. When $Ct^{-d/\alpha} \ge 2$, since $\nu_0 - 2\nu < 0$, there exists $N_1^\star$, such that, for all $n > N_1^*$, $Cn^{\nu_0 - 2\nu} t^{-d/\alpha} < 1$. Then $\mathbb{E}_{\theta_0}(\Psi_n) \le C_0 \exp\{-n^{2\nu}\}$. In addition,

$$\mathbb{E}_{\theta}(1 - \Psi_n) = \mathbb{E}_{\theta}\left[\min_{1 \le l \le N_t} \left\{1 - \Psi_n\left(\theta_0, \theta^l\right)\right\}\right] \le \mathbb{E}_{\theta}\left[\left\{1 - \Psi_n\left(\theta_0, \theta^l\right)\right\}\right] \le C_0 \exp(-C_1 n)$$

Taking $C_2 = \left(2 - Ct^{-d/\alpha}\right) I\left(Ct^{-d/\alpha} < 2\right) + I\left(Ct^{-d/\alpha} \ge 2\right) > 0$, and $N = \max\{N_1^*, N_0^*\}$ completes the proof of Lemma S8. □

S3. Derivations of Posterior Computation

S3.1 Full conditional distribution

We first summarize in Algorithm S1 the general procedure of deriving the full conditional distribution of θ using Proposition 2. The main steps are to first rewrite the density of θ in the form of (15), where $\{L_p\}_{p=1}^P$, $\{U_k\}_{k=1}^K$, $\{f_p(\theta)\}_{p=1}^P$, $\{h_k(\theta)\}_{k=1}^K$ are the input to Algorithm S1. We then sort $(L_1, \dots, L_P, U_1, \dots, U_K)$ in ascending order, which leads to

Algorithm S1. Full conditional distribution of θ

Input: $\{L_p\}_{p=1}^{P}$, $\{U_k\}_{k=1}^{K}$, $\{f_p(\theta)\}_{p=1}^{P}$, $\{h_k(\theta)\}_{k=1}^{K}$.
Output: the full conditional distribution of θ.
Sort $(L_1, \ldots, L_P, U_1, \ldots, U_K)$ in ascending order, which leads to $P+K+1$ intervals, denoted as $I_1, I_2, \ldots, I_{P+K+1}$.
for *interval* I_i, $i = 1, \ldots, P+K+1$ **do**
    Initialize $D_i = E_i = F_i = 0$
    for $p = 1, \ldots, P$, $k = 1, \ldots, K$ **do**
        if $I_i \subset [L_p, +\infty)$ **then**
            $D_i = D_i + a_{1p}$, $E_i = E_i + a_{2p}$, $F_i = F_i + a_{3p}$.
        if $I_i \subset (-\infty, U_k]$ **then**
            $D_i = D_i + b_{1k}$, $E_i = E_i + b_{2k}$, $F_i = F_i + b_{3k}$.
    end
    Write $H_i(\theta) = D_i\theta^2 + E_i\theta + F_i$.
end
if *there exists* i*, such that* $D_i \neq 0$ **then**
    the full conditional distribution of θ is a mixture of truncated normal distributions.
if $D_i = 0$ *for all* i*, and there exists* i*, such that* $E_i \neq 0$ **then**
    the full conditional distribution of θ is a mixture of truncated exponential distributions.
if $D_i = E_i = 0$ *for all* i*, and there exists* i*, such that* $F_i \neq 0$ **then**
    the full conditional distribution of θ is a mixture of uniform distributions.

$P + K + 1$ intervals. We next loop through all the intervals, and update the coefficient of $H_i(\theta)$. Finally, after obtaining the unnormalized conditional density function of θ on each interval, we derive the full conditional density of θ by incorporating the corresponding normalizing constants.

S3.2 Full conditional distribution of c_l

Without loss of generality, we only consider c_1 in the following discussion. By model (9) and the Karhunen-Loève expansion, we have

$\mu_{\pm,i}(v) = G_\omega\left\{\pm\xi(v)\right\} E_{\pm,i}(v)$, $\xi(v) = \sum_{l=1}^{L} c_l\psi_l(v)$,

and $E_{\pm,i}(v) = \sum_{l=1}^{L} e_{i,l,\pm}\psi_l(v)$. Given Y_+, Y_-, $\tilde{\Theta}_{\backslash c_1}$, the full conditional density of c_1 is,

$$\pi(c_1 \mid Y_+, Y_-, \tilde{\Theta}_{\backslash c_1}) \propto \exp\left(-\sum_{v\in\mathcal{B}_m} \frac{\sum_{i=1}^{n} W_i(v)}{K(v)}\right) \cdot \exp\left(-\frac{c_1^2}{2\lambda_l}\right), \tag{S8}$$

where $W_i(v) = \{Y_{+,i}(v) - \mu_{+,i}(v)\}^2 + \{Y_{-,i}(v) - \mu_{-,i}(v)\}^2 +$

$2r(v)\{Y_{+,i}(v)\mu_{-,i}(v) + Y_{-,i}(v)\mu_{+,i}(v)\}$, and $K(v) = 2\{1 - r^2(v)\}u^2(v)$,

with $r(v) = \{\tau_1^2(v) - \tau_2^2(v)\}/\{\tau_1^2(v) + \tau_2^2(v)\}$ and $u^2(v) = \{\tau_1^2(v) + \tau_2^2(v)\}/4$. Write $T_\pm(v) = \{\pm\lambda_1 - \sum_{l=2}^{L} c_l\psi_l(v)\}/\{\psi_1(v)\}$. According to the sign of $\psi_1(v)$, we have two different representations of $\sum_{i=1}^{n} W_i(v)$.

When $\psi_1(v) > 0$,

$$\sum_{i=1}^{n} W_i(v) = \{A_+(v)c_1^2 + B_+(v)c_1 + C_+(v)\}I\{c_1 > T_+(v)\}$$
$$+ \{A_-(v)c_1^2 + B_-(v)c_1 + C_-(v)\}I\{c_1 < T_-(v)\}.$$

When $\psi_1(v) < 0$,

$$\sum_{i=1}^{n} W_i(v) = \{A_+(v)c_1^2 + B_+(v)c_1 + C_+(v)\}I\{c_1 < T_+(v)\}$$
$$+ \{A_-(v)c_1^2 + B_-(v)c_1 + C_-(v)\}I\{c_1 > T_-(v)\}.$$

Algorithm S2. Full conditional distribution of c_l

Input: $P = K = m$, where m is the number of spatial locations,
$L_p = \begin{cases} T_+(v_j) & \text{if } \psi_l(v_j) > 0 \\ T_-(v_j) & \text{if } \psi_l(v_j) < 0 \end{cases}, U_k = \begin{cases} T_-(v_j) & \text{if } \psi_l(v_j) > 0 \\ T_+(v_j) & \text{if } \psi_l(v_j) < 0 \end{cases},$
$f_p(\theta) = \begin{cases} g_+(c_l; v_j) & \text{if } \psi_l(v_j) > 0 \\ g_-(c_l; v_j) & \text{if } \psi_l(v_j) < 0 \end{cases}, h_k(\theta) = \begin{cases} g_-(c_l; v_j) & \text{if } \psi_l(v_j) > 0 \\ g_+(c_l; v_j) & \text{if } \psi_l(v_j) < 0 \end{cases}.$
Output: the full conditional distribution of c_l.
Follow the procedure in Algorithm S1.

where $A_\pm(v), B_\pm(v), C_\pm(v)$ are all functions of $\tilde{\Theta}_{\backslash c_1}$, and are of the form,

$$
\begin{aligned}
A_\pm(v) &= \left\{ \sum_{i=1}^{n} E_{\pm,,i}(v)^2 \right\} \cdot \psi_1^2(v), \\
B_\pm(v) &= 2\psi_1(v) \left[\left\{ \sum_{l=2}^{L} c_l \psi_1(v) \right\} \left\{ \sum_{i=1}^{n} E_{\pm,i}(v)^2 \right\} \right. \\
&\quad \left. \mp \sum_{i=1}^{n} \{ Y_{\pm,i}(v) \cdot E_{\pm,i}(v) \} \mp r(v) \sum_{i=1}^{n} \{ Y_{\mp,i}(v) \cdot E_{\pm,i}(v) \} \right], \\
C_\pm(v) &= \left\{ \sum_{l=2}^{L} c_l \psi_1(v) \right\}^2 \left\{ \sum_{i=1}^{n} E_{\pm,i}(v)^2 \mp \frac{2 \cdot \sum_{i=1}^{n} Y_{\pm,i}(v) E_{\pm,i}(v)}{\sum_{l=2}^{L} c_l \psi_1(v)} \pm \frac{2r(v) \sum_{i=1}^{n} Y_{\mp,i}(v) E_{\pm,i}(v)}{\sum_{l=2}^{L} c_l \psi_1(v)} \right\}.
\end{aligned}
$$

Therefore, given Y_+, Y_-, $\tilde{\Theta}_{\backslash c_1}$ and the eigenfunctions $\{\psi_1(v_j)\}_{j=1}^m$ evaluated on $\mathcal{B}_m$,

$$
\begin{aligned}
\pi(c_1 \mid Y_+, Y_-, \tilde{\Theta}_{\backslash c_1}) \propto \exp \Bigg(&\sum_{\substack{j=1 \\ \psi_1(v_j) > 0}}^{m} [g_+(c_1; v_j) I\{c_1 > T_+(v_j)\} + g_-(c_1; v_j) I\{c_1 < T_-(v_j)\}] \\
&+ \sum_{\substack{j=1 \\ \psi_1(v_j) < 0}}^{m} [g_+(c_1; v_j) I\{c_1 < T_+(v_j)\} + g_-(c_1; v_j) I\{c_1 > T_-(v_j)\}] \Bigg),
\end{aligned}
$$

where

$$
g_\pm(c_1; v_j) = \left\{ -\frac{A_\pm(v_j)}{K(v_j)} - \frac{1}{2\lambda_1^2} \right\} c_1^2 + \frac{B_\pm(v_j)}{K(v_j)} c_1 + \frac{C_\pm(v_j)}{K(v_j)}.
$$

By Proposition 1, the full conditional distribution of c_1 is a mixture of truncated normal distributions. We summarize the procedure of obtaining this distribution in Algorithm S2.

S3.3 Full conditional distribution of ω

Recall that the prior of ω is the uniform distribution on $[a_\omega, b_\omega]$. Then we have,

$$\pi(\omega \mid Y_+, Y_-, \tilde{\Theta}_{\setminus\omega}) \propto \exp\left\{-\sum_{v\in\mathcal{B}_m} \frac{\sum_{i=1}^n W_i(v)}{K(v)}\right\} \cdot \frac{1}{b_\omega - a_\omega} I(a_\omega \leq \omega \leq b_\omega), \tag{S9}$$

where $W_i(v)$ is defined as in (S8). Then,

$$\sum_{i=1}^n W_i(v) = Q_+(v) I\{\omega < \xi(v)\} + Q_-(v) I\{\omega < -\xi(v)\},$$

where

$$\begin{aligned} Q_\pm(v) =& \xi(v)^2 \left\{\sum_{i=1}^n E_{\pm,i}(v)^2\right\} \mp 2\xi(v) \left\{\sum_{i=1}^n Y_{\pm,i}(v) E_{\pm,i}(v)\right\} \\ &\pm 2r(v)\xi(v) \left\{\sum_{i=1}^n Y_{\mp,i}(v) E_{\pm,i}(v)\right\}. \end{aligned}$$

Algorithm S3. Full conditional distribution of ω

Input: $P = 0$, $K = 2m$,

$U_k = \begin{cases} \xi(v_j), & \text{if } a_\omega < \xi(v_j) < b_\omega \\ -\xi(v_j) & \text{if } a_\omega < -\xi(v_j) < b_\omega \end{cases}$, $h_k(\theta) = \begin{cases} C_+(v_j), & \text{if } U_k = \xi(v_j) \\ C_-(v_j) & \text{if } U_k = -\xi(v_j) \end{cases}$.

Output: the full conditional distribution of ω Follow the procedure in Algorithm S1

Therefore, given $Y_+, Y_-, \tilde{\Theta}_{\setminus\omega}$ and the eigenfunctions $\psi_l(v_j), j = 1, \ldots, m, l = 1, \ldots, L$, evaluated on $\mathcal{B}_m$, we have,

$$\pi(\omega \mid Y_+, Y_-, \tilde{\Theta}_{\setminus\omega})$$
$$\propto \exp\left[\sum_{\substack{j=1\\ a_\omega<\xi(v_j)<b_\omega}}^{m} C_+(v_j)I\{\omega<\xi(v_j)\} + \sum_{\substack{j=1\\ a_\omega<-\xi(v_j)<b_\omega}}^{m} C_-(v_j)I\{\omega<-\xi(v_j)\}\right],$$

where $C_\pm(v_j) = -\dfrac{Q_\pm(v_j)}{K(v_j)} - \log(b_\omega - a_\omega)$, and we only consider those $\xi(v_j)$ and $-\xi(v_j)$ that are between a_ω and b_ω.

By Proposition 1, the full conditional distribution of ω is a mixture of uniform distributions. We summarize the procedure of obtaining this distribution in Algorithm S3.

S3.4 Full conditional distribution of $e_{i,l\pm}$

Since $e_{i,l,+}$ only exist in $\mu_{+,i}(v)$, we can rewrite $\mu_{+,i}(v)$ as $\mu_{+,i}(v) = a_{+,i}(v)+b_{+,i}(v)$, where $a_{+,i}(v) = G_\omega\left\{\sum_{l=1}^{L} c_l\psi_l(v)\right\} e_{i,l,+}\psi_l(v) = C_{l,+}(v)\cdot e_{i,l,+}$, and $b_{+,i}(v) = G_\omega\left\{\sum_{l=1}^{L} c_l\psi_l(v)\right\}$ $\sum_{l'\neq l} e_{i,l',+}\psi_{l'}(v)$. Note that $b_{+,i}(v)$ does not depend on $e_{i,l,+}$. Henceforth, we have that,

$$\{Y_{+,i}(v) - \mu_{+,i}(v)\}^2 =$$
$$Y_{+,i}^2(v) + a_{+,i}^2(v) + b_{+,i}^2(v) + 2a_{+,i}(v)b_{+,i}(v) - 2Y_{+,i}(v)a_{+,i}(v) - 2Y_{+,i}(v)b_{+,i}(v),$$
$$\{Y_{+,i}(v) - \mu_{+,i}(v)\}\{Y_{-,i}(v) - \mu_{-,i}(v)\} =$$
$$Y_{+,i}(v)\{Y_{-,i}(v) - \mu_{-,i}(v)\} - a_{+,i}(v)\{Y_{-,i}(v) - \mu_{-,i}(v)\} - b_{+,i}(v)\{Y_{-,i}(v) - \mu_{-,i}(v)\}.$$

Ignoring the terms $\{Y_{-,i}(v) - \mu_{-,i}(v)\}^2$ that do not contain $e_{i,l,+}$, we have,

$$\begin{aligned}
&\pi(e_{i,l+} \mid Y_+, Y_-, \tilde{\Theta}_{\setminus e_{i,l+}}) \\
&\propto \prod_{v \in \mathcal{B}_m} \exp\left(-\frac{a_{+,i}^2(v) + 2a_{+,i}(v)[b_{+,i}(v) - Y_{+,i}(v) - r(v)\{Y_{-,i}(v) - \mu_{-,i}(v)\}]}{2\{1 - r^2(v)\}u^2(v)} \right) \\
&\quad \cdot \exp\left(-\frac{e_{i,l,+}^2}{2\lambda_l} \right) \\
&\propto \exp\left[-\frac{1}{2} \frac{\{e_{i,l,+} - M_{i,l,+}\}^2}{V_{i,l,+}^2} \right].
\end{aligned}$$

where the mean and the variance are

$$\begin{aligned}
M_{i,l,\pm} &= \sum_{v \in \mathcal{B}_m} \left[\{\lambda_l m_{i,l,\pm}(v)\} / \{\lambda_l + \sigma_{i,l,\pm}^2(v)\} \right], \\
V_{i,l,\pm}^2 &= \sum_{v \in \mathcal{B}_m} \left[\lambda_l \sigma_{i,l,\pm}^2(v) / \{\lambda_l + \sigma_{i,l,\pm}^2(v)\} \right],
\end{aligned}$$

with $m_{i,l,\pm}(v) = -\left[\{Y_{\pm,i}(v) - b_{\pm,i}(v)\} - r(v) \cdot \{Y_{\pm,i}(v) - \mu_{\pm,i}(v)\}\right] / C_{l,\pm}(v)$, and $\sigma_{i,l,\pm}^2(v)$ $= \{1 - r^2(v)\}u^2(v)/C_{l,\pm}^2(v)$. Therefore, $e_{i,l\pm}$ follows a normal distribution, i.e.,

$$e_{i,l\pm} \mid Y_+, Y_-, \tilde{\Theta}_{\setminus e_{i,l\pm}} \sim \mathrm{N}(M_{i,l,\pm}, V_{i,l,\pm}^2).$$

S3.5 Full conditional distribution of $\tau_1^2(v)$ and $\tau_2^2(v)$

For a given $v_0 \in \mathcal{B}_m$, we have,

$$
\begin{aligned}
&\pi\left\{\tau_1^2(v_0) \mid Y_+, Y_-, \tilde{\Theta}_{\setminus \tau_1^2(v_0)}\right\} \\
&\propto \prod_{i=1}^{n} \frac{1}{\sqrt{\tau_1^2}} \cdot \exp\left[-\frac{1}{2}\left(\frac{1}{\tau_1^2}+\frac{1}{\tau_2^2}\right)\left\{\tilde{Y}_{+,i}(v_0)^2+\tilde{Y}_{-,i}(v_0)^2-2\frac{\tau_1^2-\tau_2^2}{\tau_1^2+\tau_2^2}\tilde{Y}_{+,i}(v_0)\tilde{Y}_{-,i}(v_0)\right\}\right] \\
&\cdot \Gamma^{-1}_{\tau_1^2}(a_\tau, b_\tau) \\
&\propto \left\{\frac{1}{\tau_1^2(v_0)}\right\}^{\frac{n}{2}} \exp\left[-\frac{1}{2\tau_1^2(v_0)}\sum_{i=1}^{n}\left\{Y_{+,i}(v_0)-\mu_{+,i}(v_0)+Y_{-,i}(v_0)-\mu_{-,i}(v_0)\right\}^2\right]
\end{aligned}
$$

where $\tilde{Y}_{\pm,i}(v_0) = Y_{\pm,i}(v_0) - \mu_{\pm,i}(v_0)$. Therefore, we have,

$$
\tau_1^2(v_0) \mid Y_+, Y_-, \tilde{\Theta}_{\setminus \tau_1^2(v_0)} \sim \mathrm{IG}\left(a_\tau + \frac{n}{2}, \frac{\sum_{i=1}^{n}\{\tilde{Y}_{+,i}(v_0)+\tilde{Y}_{-,i}(v_0)\}^2}{2} + nb_\tau\right).
$$

Similarly, we have,

$$
\tau_2^2(v_0) \mid Y_+, Y_-, \tilde{\Theta}_{\setminus \tau_2^2(v_0)} \sim \mathrm{IG}\left(a_\tau + \frac{n}{2}, \frac{\sum_{i=1}^{n}\left\{\tilde{Y}_{+,i}(v_0)-\tilde{Y}_{-,i}(v_0)\right\}^2}{2} + nb_\tau\right).
$$

S3.6 Derivation of hybrid mini-batch MCMC

We derive the acceptance ratio in the hybrid mini-batch MCMC. Let $Y = \{Y_{1i}(v), Y_{2i}(v), i = 1, \ldots, n, v \in \mathcal{B}_m\}$, $Y_{m_s} = \{Y_{1i}(v), Y_{2i}(v), i = 1, \ldots, n, v \in \mathcal{B}_{m_s},\}$, and $\tilde{\Theta} = \{\theta, \tilde{\Theta}_{\setminus\theta}\}$, where $m_s < m$, and henceforth $\mathcal{B}_{m_s} \subset \mathcal{B}_m$. In the Gibbs sampler, we use the full conditional distribution $P(\theta|Y, \tilde{\Theta}_{\setminus\theta})$ as the proposal function, with the acceptance ratio equal to 1. In the hybrid mini-batch MCMC, we use $P(\theta|Y_{m_s}, \tilde{\Theta}_{\setminus\theta})$ as the proposal function, and the acceptance ratio becomes,

$$
\begin{aligned}
\phi(\theta', \theta) &= \min\left\{1, \frac{P(Y|\theta', \tilde{\Theta}_{\backslash\theta})}{P(Y|\theta, \tilde{\Theta}_{\backslash\theta})} \frac{P(\theta|Y_{m_s}, \tilde{\Theta}_{\backslash\theta})}{P(\theta'|Y_{m_s}, \tilde{\Theta}_{\backslash\theta})}\right\} \\
&= \min\left\{1, \frac{\prod_{v\in\mathcal{B}_m} P(Y(v)|\theta', \tilde{\Theta}_{\backslash\theta})}{\prod_{v\in\mathcal{B}_m} P(Y(v)|\theta, \tilde{\Theta}_{\backslash\theta})} \cdot \frac{\prod_{v\in\mathcal{B}_{m_s}} P(Y(v)|\theta, \tilde{\Theta}_{\backslash\theta})p(\theta)}{\prod_{v\in\mathcal{B}_{m_s}} P(Y(v)|\theta', \tilde{\Theta}_{\backslash\theta})p(\theta')}\right\} \\
&= \min\left\{1, \frac{\prod_{v\notin\mathcal{B}_{m_s}} P(Y(v)|\theta', \tilde{\Theta}_{\backslash\theta})}{\prod_{v\notin\mathcal{B}_{m_s}} P(Y(v)|\theta, \tilde{\Theta}_{\backslash\theta})}\right\}.
\end{aligned}
$$

S3.7 Posterior computation algorithms

We summarize the Gibbs sampling for the TCGP in Algorithm S4, and the hybrid mini-batch MCMC procedure in Algorithm S5

Algorithm S4. Gibbs sampling for TCGP

Input: the observed imaging data $Y = \{\{Y_{1,i}(v), Y_{2,i}(v)\}_{i=1}^n, v \in \mathcal{B}_m\}$,
the kernel function $\kappa(\cdot,\cdot)$,
the Karhunen-Loève truncation number L,
the prior hyperparameters $a_\tau, b_\tau, a_\omega, b_\omega$.

Output: the posterior samples of
$\tilde{\Theta} = \{\{c_l\}_{l=1}^L, \{e_{i,l,\pm}\}_{l=1,i=1}^{L,n}, \{\tau_1^2(v), \tau_2^2(v)\}_{v\in\mathcal{B}_m}, \omega\}$.

Initialize $\tilde{\Theta}$: sample $\tilde{\Theta}$ from the prior distribution.

for $t = 1, \cdots, T$ **do**
  parallel sample $\tau_k^2(v)$ from the inverse Gamma distribution, $v \in \mathcal{B}_m, k = 1, 2$.
end
for $l = 1, \ldots, L$ **do**
  sample c_l from the mixture of truncated normal distributions.
  sample ω from the mixture of uniform distributions.
  sample $e_{i,l,\pm}$ from the normal distribution, $i = 1, \ldots, n$.
end

Algorithm S5. Hybrid mini-batch MCMC for TCGP.

Input: the observed imaging data $Y = \{\{Y_{1,i}(v), Y_{2,i}(v)\}_{i=1}^{n}, v \in \mathcal{B}_m\}$,

the kernel function $\kappa(\cdot,\cdot)$,

the Karhunen-Loève truncation number L,

the prior hyperparameters $a_\tau, b_\tau, a_\omega, b_\omega$.

Output: the posterior samples of

$\tilde{\Theta} = \{\{c_l\}_{l=1}^{L}, \{e_{i,l,\pm}\}_{l=1,i=1}^{L,n}, \{\tau_1^2(v), \tau_2^2(v)\}_{v\in\mathcal{B}_m}, \omega\}$.

Initialize $\tilde{\Theta}$: sample $\tilde{\Theta}$ from the prior distribution.

for $t = 1, \cdots, T$ **do**

parallel sample $\tau_k^2(v)$ from the inverse Gamma distribution, for all $v \in \mathcal{B}_m, k = 1, 2$.

random sample m_s locations from $\mathcal{B}_m$ and form $\mathcal{B}_{m_s}$ and Y_{m_s}.

end

for $l = 1, \cdots, L$ **do**

if $t \mod T_0 = 0$ **then**

sample c_l from the mixture of truncated normal distributions based on Y.

sample ω from the mixture of uniform distributions based on Y.

else

sample $c_l^{(t)}$ from the mixture of truncated normal distributions based on Y_{m_s}.

accept $c_l^{(t)}$ with probability $\phi(c_l^{(t)}, c_l^{(t-1)})$.

sample $\omega^{(t)}$ from the mixture of uniform distributions based on Y_{m_s}.

accept $\omega^{(t)}$ with probability $\phi(\omega^{(t)}, \omega^{(t-1)})$.

parallel sample $e_{i,l,\pm}$ from the normal distribution, $i = 1, \ldots, n$.

end

S4. Additional numerical results

S4.1 2D image simulation

We simulate the data from model (2.1), with the sample size $n = 50$, and the image resolution $m = 64 \times 64$. We simulate the mean $\mu_{k,i}$ from (2.2) and (2.3), $k = 1, 2$, with $\kappa(v, v') = \exp -0.1(v^2 + v'^2) - 10(v - v')^2$, $\sigma_+^2(v) = \zeta_+ \sum_{j=1}^{3} I(\|v - u_{+,j}\|_1 < 0.1)$, where $u_{+,1} = (0.3, 0.7)$, $u_{+,2} = (0.7, 0.7)$, $u_{+,3} = (0.3, 0.3)$, and $\sigma_-^2(v) = \zeta_-\{I(\|v - u_{-,1}\|_1 < 0.1) + I(\|v - u_{-,2}\|_2 < 0.1)\}$, where $u_{-,1} = (0.5, 0.5)$, $u_{-,2} = (0.7, 0.3)$. Here (ζ_+, ζ_-) controls the signal strength, and we consider two settings, with $(\zeta_+, \zeta_-) = (0.15, 0.25)$ for a weak signal, and $(\zeta_+, \zeta_-) = (0.75, 0.85)$ for a strong signal. We simulate the noise $\varepsilon_{k,i}$ from the normal distribution with mean zero and variance $\tau_k^2(v)$, and simulate $\log(\tau_k^2(v))$ from a Gaussian process with mean zero and correlation kernel $\kappa(v, v')$, $k = 1, 2$.

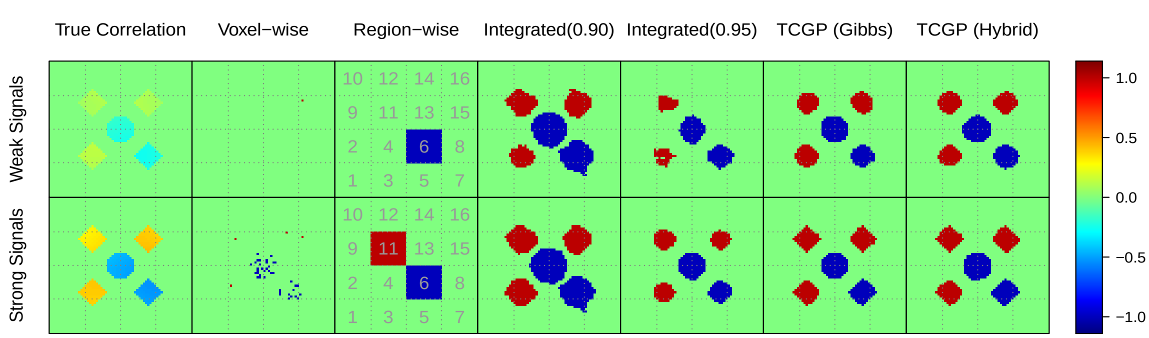

Figure S1: Results of 2D image simulations. The first row is for a weak signal and the second row a strong signal. The panels from left to right show the true correlation map, the significantly positively (red) and negatively (blue) correlated regions selected by different methods. TCGP represents the proposed Thresholded Correlation Gaussian Process.

Table S1: Results of 2D image simulations. Reported are the average sensitivity, specificity, and FDR, with standard error in the parenthesis, based on 100 data replications. Six methods are compared: the voxel-wise analysis, the region-wise analysis, the integrated method of Li et al. (2019) with two thresholding values, 0.95 and 0.90, and the proposed Bayesian method Thresholded Correlation Gaussian Process (TCGP) with the Gibbs sampler and the hybrid mini-batch MCMC.

Signal	Method	Positive Correlation			Negative Correlation		
		Sensitivity	Specificity	FDR	Sensitivity	Specificity	FDR
Weak	Voxel-wise	0.000 (0.000)	1.000 (0.000)	0.020 (0.010)	0.000 (0.001)	1.000 (0.001)	0.010 (0.001)
	Region-wise	0.238 (0.001)	0.953 (0.002)	0.447 (0.002)	0.473 (0.002)	0.956 (0.003)	0.629 (0.004)
	Integrated (0.95)	0.612 (0.001)	0.994 (0.000)	0.134 (0.010)	0.844 (0.003)	0.993 (0.000)	0.131 (0.003)
	Integrated (0.90)	0.821 (0.001)	0.971 (0.000)	0.341 (0.010)	0.963 (0.003)	0.966 (0.000)	0.398 (0.006)
	TCGP (Gibbs)	0.855 (0.003)	0.996 (0.001)	0.057 (0.008)	0.997 (0.002)	0.993 (0.001)	0.108 (0.005)
	TCGP (Hybrid)	0.851 (0.006)	0.993 (0.001)	0.092 (0.010)	0.993 (0.002)	0.992 (0.001)	0.126 (0.005)
Strong	Voxel-wise	0.062 (0.002)	1.000 (0.000)	0.000 (0.014)	0.091 (0.002)	1.000 (0.000)	0.000 (0.006)
	Region-wise	0.741 (0.002)	0.852 (0.003)	0.747 (0.004)	0.479 (0.002)	0.950 (0.002)	0.645 (0.003)
	Integrated (0.95)	0.773 (0.001)	0.998 (0.000)	0.036 (0.002)	0.933 (0.002)	0.996 (0.000)	0.067 (0.001)
	Integrated (0.90)	0.996 (0.020)	0.959 (0.000)	0.378 (0.017)	0.999 (0.020)	0.953 (0.000)	0.468 (0.001)
	TCGP (Gibbs)	0.976 (0.002)	0.999 (0.000)	0.015 (0.004)	1.000 (0.001)	0.999 (0.000)	0.018 (0.001)
	TCGP (Hybrid)	0.960 (0.003)	0.997 (0.001)	0.049 (0.005)	0.990 (0.001)	0.999 (0.000)	0.023 (0.002)

Table S1 reports the results averaged over 100 data replications, and Figure S1 visualizes the result for one data replication. We see that our proposed method clearly outperforms the alternative solutions. We observe essentially the same patterns as in the 3D example. In addition, the proposed Bayesian method is also capable of statistical inference, in that we can simulate the entire posterior distribution, compute the posterior inclusion probability, and quantify the uncertainty for the spatially varying correlation. Figure S2 shows the probability map of the identified positively and negatively correlated regions, which are close to the truth.

We then vary the sample size $n = \{30, 50, 100\}$ while fixing the image resolution $m = 64 \times 64$, or vary $m = \{32 \times 32, 64 \times 64, 100 \times 100\}$ while fixing $n = 50$. Table S2 reports the

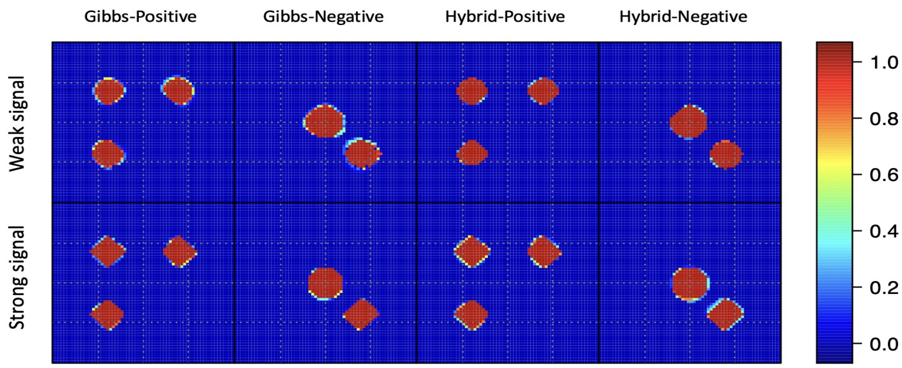

Figure S2: Results of 2D image simulations. The posterior inclusion probability map of the positive and negative spatially-varying correlations using the Gibbs sampler and the hybrid mini-batch MCMC.

results averaged over 100 data replications. We see that our proposed method performs the best across different values of n and m. Meanwhile, it maintains a competitive performance even when n is relatively small or when m is relatively large.

S4.2 Additional 3D simulations

We conduct additional simulations for the 3D image example. We fix the sample size $n = 904$ follow the Human Connectome Project Data and vary the signal to noise ratio with $\zeta_k = 5$ for weak signal and $\zeta_k = 0.5$ for strong signal. Table S3 reports the results averaged over 100 data replications. We see that the proposed method outperforms other methods with different signal to noise ratio.

S4.2 Additional 3D simulations

Table S2: The 2D simulation example with the varying sample size n and the varying image resolution m. Reported are the average sensitivity, specificity, and FDR, with standard error in the parenthesis, based on 100 data replications. Six methods are compared: the voxel-wise analysis, the region-wise analysis, the integrated method of Li et al. (2019) with two thresholding values, 0.95 and 0.90, and the proposed Bayesian method Thresholded Correlation Gaussian Process (TCGP) with the Gibbs sampler and the hybrid mini-batch MCMC.

	Method	Positive Correlation			Negative Correlation		
		Sensitivity	Specificity	FDR	Sensitivity	Specificity	FDR
$n = 30$	Voxel-wise	0.080(0.002)	1.000(0.000)	0.004(0.003)	0.102(0.002)	1.000(0.000)	0.001(0.002)
	Region-wise	0.148(0.005)	0.971(0.002)	0.326(0.003)	0.473(0.006)	0.957(0.003)	0.624(0.003)
	Integrated(0.95)	0.518(0.005)	0.992(0.003)	0.199(0.008)	0.781(0.003)	0.993(0.004)	0.146(0.009)
	Integrated(0.90)	0.855(0.007)	0.960(0.005)	0.378(0.010)	0.871(0.004)	0.937(0.004)	0.392(0.011)
	TCGP (Gibbs)	0.910(0.004)	0.991(0.003)	0.109(0.005)	0.990(0.002)	0.993(0.001)	0.065(0.007)
	TCGP (Hybrid)	0.890(0.005)	0.990(0.003)	0.111(0.008)	0.983(0.004)	0.990(0.002)	0.110(0.008)
$n = 50$	Voxel-wise	0.098(0.002)	1.000(0.000)	0.002(0.001)	0.150(0.002)	1.000(0.000)	0.003(0.001)
	Region-wise	0.438(0.004)	0.953(0.005)	0.547(0.010)	0.573(0.003)	0.956(0.001)	0.629(0.010)
	Integrated(0.95)	0.659(0.003)	0.995(0.002)	0.130(0.008)	0.899(0.005)	0.997(0.001)	0.110(0.009)
	Integrated(0.90)	0.959(0.009)	0.970(0.005)	0.308(0.009)	0.969(0.003)	0.969(0.003)	0.355(0.010)
	TCGP (Gibbs)	0.941(0.004)	0.995(0.002)	0.081(0.005)	0.996(0.002)	0.992(0.001)	0.063(0.005)
	TCGP (Hybrid)	0.931(0.005)	0.993(0.003)	0.092(0.005)	0.993(0.002)	0.992(0.002)	0.086(0.006)
$n = 100$	Voxel-wise	0.102(0.004)	1.000(0.001)	0.002(0.003)	0.198(0.001)	1.000(0.000)	0.003(0.001)
	Region-wise	0.617(0.010)	0.881(0.003)	0.744(0.004)	0.476(0.005)	0.955(0.002)	0.631(0.010)
	Integrated(0.95)	0.714(0.005)	0.998(0.003)	0.099(0.005)	0.898(0.004)	0.997(0.002)	0.099(0.008)
	Integrated(0.90)	0.980(0.010)	0.969(0.010)	0.300(0.010)	0.975(0.003)	0.971(0.003)	0.298(0.011)
	TCGP (Gibbs)	0.953(0.002)	0.997(0.002)	0.041(0.002)	0.999(0.001)	0.997(0.001)	0.033(0.001)
	TCGP (Hybrid)	0.945(0.003)	0.997(0.002)	0.069(0.003)	0.993(0.003)	0.996(0.001)	0.085(0.002)
$m = 32 \times 32$	Voxel-wise	0.017(0.001)	1.000(0.000)	0.005(0.001)	0.040(0.002)	1.000(0.000)	0.004(0.002)
	Region-wise	0.297(0.005)	0.945(0.005)	0.531(0.010)	0.472(0.003)	0.957(0.002)	0.617(0.010)
	Integrated(0.95)	0.620(0.005)	0.989(0.004)	0.138(0.005)	0.852(0.004)	0.989(0.001)	0.198(0.009)
	Integrated(0.90)	0.933(0.010)	0.971(0.006)	0.287(0.008)	0.944(0.004)	0.957(0.005)	0.300(0.011)
	TCGP (Gibbs)	0.931(0.003)	0.993(0.002)	0.083(0.003)	0.991(0.004)	0.992(0.003)	0.065(0.004)
	TCGP (Hybrid)	0.922(0.005)	0.992(0.002)	0.082(0.005)	0.991(0.005)	0.991(0.002)	0.089(0.005)
$m = 64 \times 64$	Voxel-wise	0.098(0.002)	1.000(0.000)	0.002(0.001)	0.150(0.002)	1.000(0.000)	0.003(0.001)
	Region-wise	0.438(0.004)	0.953(0.005)	0.547(0.010)	0.573(0.003)	0.956(0.001)	0.629(0.010)
	Integrated(0.95)	0.659(0.003)	0.995(0.002)	0.130(0.008)	0.899(0.005)	0.997(0.001)	0.110(0.009)
	Integrated(0.90)	0.959(0.009)	0.970(0.005)	0.308(0.009)	0.969(0.003)	0.969(0.003)	0.355(0.010)
	TCGP (Gibbs)	0.941(0.004)	0.995(0.002)	0.081(0.005)	0.996(0.002)	0.992(0.001)	0.063(0.004)
	TCGP (Hybrid)	0.931(0.005)	0.993(0.003)	0.092(0.005)	0.993(0.002)	0.992(0.002)	0.086(0.006)
$m = 100 \times 100$	Voxel-wise	0.005(0.001)	1.000(0.000)	0.004(0.002)	0.011(0.001)	1.000(0.000)	0.000(0.001)
	Region-wise	0.627(0.002)	0.861(0.003)	0.763(0.005)	0.462(0.002)	0.948(0.002)	0.663(0.008)
	Integrated(0.95)	0.843(0.002)	0.998(0.003)	0.039(0.005)	0.952(0.004)	0.997(0.002)	0.052(0.007)
	Integrated(0.90)	0.960(0.003)	0.965(0.005)	0.300(0.012)	0.977(0.002)	0.965(0.004)	0.298(0.011)
	TCGP (Gibbs)	0.971(0.001)	0.999(0.000)	0.029(0.002)	0.997(0.001)	0.998(0.002)	0.031(0.001)
	TCGP (Hybrid)	0.964(0.001)	0.999(0.000)	0.033(0.001)	0.995(0.001)	0.997(0.002)	0.033(0.002)

Table S3: Simulation results of the 3D image example with the varying signal to noise ratio. Reported are the average sensitivity, specificity, and FDR, with standard error in the parenthesis, based on 100 data replications. Six methods are compared: the voxel-wise analysis, the region-wise analysis, the integrated method of Li et al. (2019) with two thresholding values, 0.95 and 0.90, and the proposed Bayesian method with the Gibbs sampler and the hybrid mini-batch MCMC.

Signal	Method	Positive Correlation			Negative Correlation		
		Sensitivity	Specificity	FDR	Sensitivity	Specificity	FDR
Weak	Voxel-wise	0.082 (0.003)	0.999 (0.005)	0.001 (0.006)	0.101 (0.005)	0.998 (0.000)	0.002 (0.003)
	Region-wise	0.366 (0.001)	0.865 (0.002)	0.573 (0.011)	0.472 (0.002)	0.892 (0.003)	0.453 (0.004)
	Integrated(0.95)	0.487 (0.002)	0.981 (0.001)	0.160 (0.010)	0.582 (0.001)	0.952 (0.005)	0.101 (0.003)
	Integrated(0.90)	0.873 (0.008)	0.934 (0.001)	0.230 (0.009)	0.831 (0.004)	0.946 (0.005)	0.270 (0.004)
	TCGP (Gibbs)	0.890 (0.005)	0.987 (0.001)	0.070 (0.007)	0.890 (0.002)	0.975 (0.003)	0.075 (0.001)
	TCGP (Hybrid)	0.884 (0.002)	0.978 (0.001)	0.078 (0.006)	0.871 (0.004)	0.965 (0.003)	0.089 (0.002)
Strong	Voxel-wise	0.220 (0.005)	0.999 (0.002)	0.001 (0.001)	0.237 (0.004)	0.999 (0.000)	0.002 (0.001)
	Region-wise	0.641 (0.003)	0.765 (0.001)	0.587 (0.010)	0.627 (0.006)	0.824 (0.005)	0.532 (0.003)
	Integrated(0.95)	0.550 (0.005)	0.992 (0.000)	0.066 (0.005)	0.882 (0.005)	0.970 (0.000)	0.101 (0.002)
	Integrated(0.90)	0.934 (0.010)	0.974 (0.003)	0.244 (0.007)	0.933 (0.010)	0.955 (0.001)	0.233 (0.003)
	TCGP (Gibbs)	0.951 (0.002)	0.997 (0.001)	0.052 (0.003)	0.951 (0.002)	0.991 (0.001)	0.041 (0.002)
	TCGP (Hybrid)	0.949 (0.003)	0.995 (0.001)	0.058 (0.004)	0.950 (0.003)	0.989 (0.001)	0.050 (0.001)

S4.3 Sensitivity analysis

In our hybrid mini-batch MCMC, we sample a subset of m_s voxels and use the full dataset after every T_0 iterations of using the mini-batch data. We next carry out a sensitivity analysis to study the effect of m_s and T_0. Table S4 reports the results averaged over 100 data replications. We see that the results are relatively stable for different values of m_s and T_0.

S4.4 Prior specification for the HCP data analysis

In our HCP data analysis, we set the prior for ω as $\mathrm{U}(a_\omega, b_\omega)$, and we choose a_ω and b_ω as the 75% quantile and 100% quantile of $\{|\xi(v)|\}_{v\in\mathcal{B}}$, respectively. The choice of a_ω is based on the belief that at most 25% voxels have non-zero correlations. Here we vary

Table S4: The sensitivity analysis of the batch size m_s and the number of iterations T_0 for the hybrid mini-batch MCMC. Reported are the average sensitivity, specificity, and False Discorvery Rate (FDR), with standard error in the parenthesis, based on 100 data replications.

m_s	T_0	Positive Correlation			Negative Correlation		
		Sensitivity	Specificity	FDR	Sensitivity	Specificity	FDR
$m/32$	20	0.950(0.003)	1.000(0.001)	0.015(0.003)	0.991(0.002)	0.989(0.003)	0.050(0.005)
$m/16$	20	0.953(0.003)	0.996(0.001)	0.061(0.002)	0.991(0.003)	0.997(0.001)	0.049(0.005)
$m/4$	20	0.955(0.002)	0.997(0.001)	0.058(0.002)	0.990(0.001)	0.997(0.001)	0.047(0.003)
$m/16$	50	0.948(0.003)	0.998(0.001)	0.045(0.002)	0.990(0.002)	0.990(0.003)	0.062(0.003)
$m/16$	20	0.953(0.003)	0.996(0.001)	0.061(0.002)	0.991(0.003)	0.997(0.001)	0.049(0.005)
$m/16$	10	0.953(0.001)	0.995(0.001)	0.059(0.003)	0.993(0.002)	0.998(0.001)	0.041(0.004)

$a_\omega = \{0.73, 0.75, 0.77\}$, and investigate the corresponding performance of our proposed method. Table S5 reports the results, which we see that are relatively stable across different choices of a_ω.

Table S5: Prior specification for the Human Connectome Project data under different choices of a_ω. Reported are the activation regions containing more than 100 voxels that are declared having a nonzero correlation.

Lingual-R				
a_ω	cluster size	Activation center	overlap rate	mean correlation
0.73	151	(-10.0, -74.5, -4.0)	0.931	0.35
0.75	144	(-10.4, -75.3, -4.5)	1.000	0.35
0.77	140	(-10.6, -75.8, -5.4)	0.905	0.38
Angular-R				
a_ω	cluster size	cluster center	overlap rate	mean correlation
0.73	215	(-45.9, -60.1, 45.5)	0.910	0.41
0.75	209	(-46.9, -60.2, 44.7)	1.000	0.43
0.77	200	(-46.0, -59.9, 43.9)	0.911	0.43
Temporal-Mid-L				
a_ω	cluster size	cluster center	overlap rate	mean correlation
0.73	110	(62.1, -24.9, 1.3)	0.940	0.42
0.75	104	(63.1, -25.7, 1.4)	1.000	0.41
0.77	99	(62.7, -25.5, 1.3)	0.921	0.43
Precentral-L				
a_ω	cluster size	cluster center	overlap rate	mean correlation
0.73	130	(29.1, -23.0, 64.5)	0.930	-0.41
0.75	115	(28.6, -23.1, 65.4)	1.000	-0.44
0.77	107	(28.8, -23.1, 65.8)	0.931	-0.42
Occipital-Inf-R				
a_ω	cluster size	cluster center	overlap rate	mean correlation
0.73	130	(-38.1, -81.0, -3.9)	0.910	-0.45
0.75	122	(-38.8, -81.7, -3.2)	1.000	-0.44
0.77	107	(-38.5, -80.0, -4.0)	0.901	-0.43

References

Adler, R. and J. Taylor (2009). *Random Fields and Geometry*. Springer Monographs in Mathematics. Springer New York.

Choudhuri, N., S. Ghosal, and A. Roy (2004). Bayesian estimation of the spectral density of a time series. *Journal of the American Statistical Association 99*(468), 1050–1059.

Ghosal, S. and A. Roy (2006). Posterior consistency of gaussian process prior for nonparametric binary regression. *The Annals of Statistics 34*(5), 2413–2429.

Li, L., J. Kang, S. N. Lockhart, J. Adams, and W. J. Jagust (2019). Spatially adaptive varying correlation analysis for multimodal neuroimaging data. *IEEE transactions on medical imaging 38*(1), 113–123.