Supplementary Material to "Bayesian Fixed-domain Asymptotics for Covariance Parameters in a Gaussian Process Model"

Cheng Li *[1]

[1]Department of Statistics and Data Science, National University of Singapore

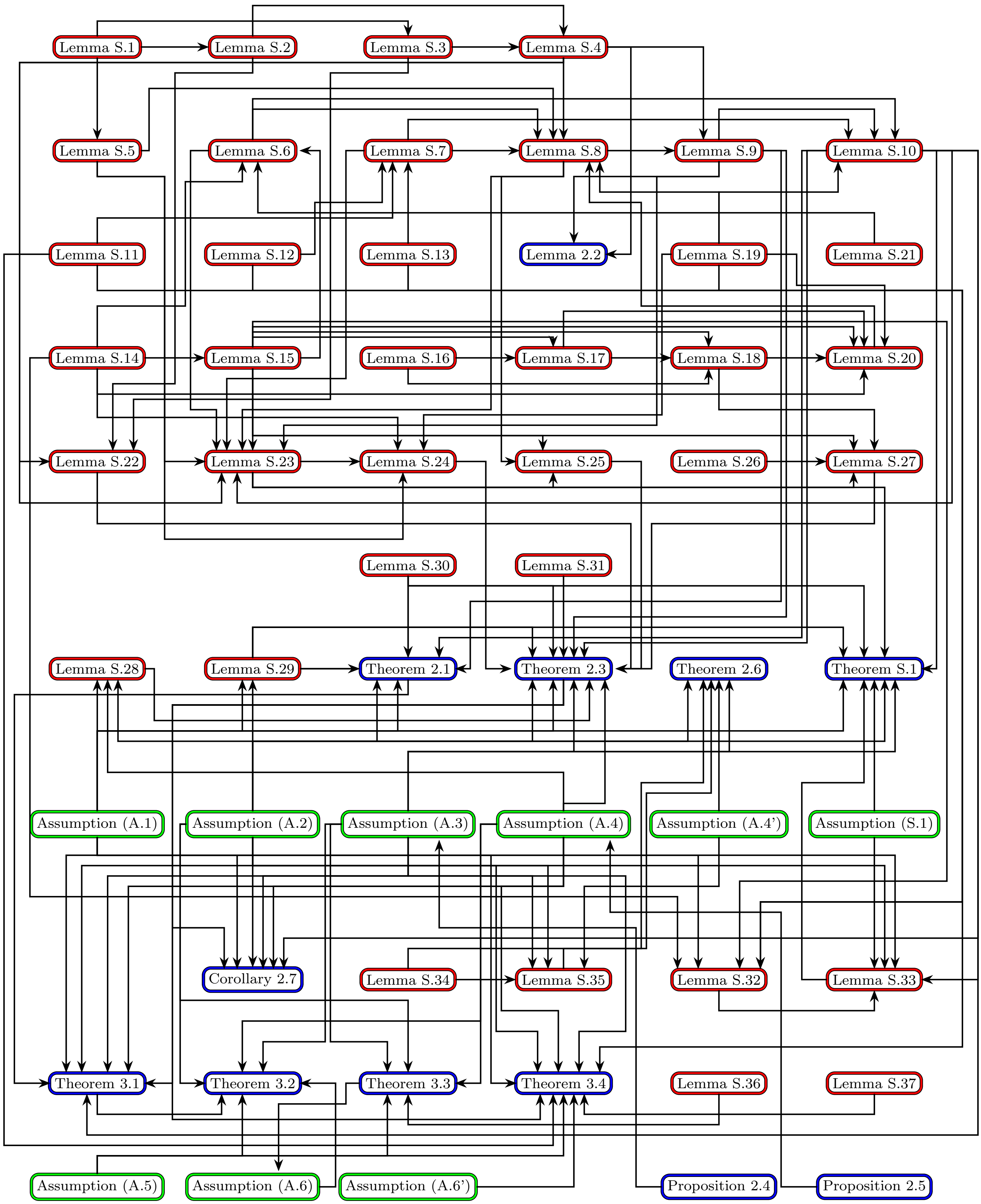
This supplementary material includes more simulation results and all technical proofs of the theorems, lemmas, propositions, and corollaries in the main text. The contents are organized as follows.

Section S1 provides the proof of the monotonicity and uniform convergence of REML in Lemma 2.2 of the main text, as well as auxiliary results on RKHS theory and spectral analysis of Matérn covariance functions. Section S2 includes technical lemmas for the profile likelihood function. Section S3 presents the proof of Theorem 2.1 and Theorem 2.3 of the main text, as well as the theory for $d \geq 5$. Section S4 presents the proof of Propositions 2.4 and 2.5 of the main text. Section S5 presents the proof of Theorem 2.6 and Corollary 2.7. Section S6 presents the proof of all theorems in Section 3 of the main text, including Theorems 3.1, 3.2, 3.3, and 3.4. Section S7 includes the additional simulation results for the model with regression terms for $\nu = 1/2, 1/4, 3/2$ in both $d = 1$ and $d = 2$ cases. To keep consistency, every lemma in this supplementary material is immediately followed by its proof.

We first define some universal notation that will used throughout the proofs. Let $\mathbb{R}^+ = (0, +\infty)$ and $\mathbb{Z}^+$ be the set of all positive integers. For any $x = (x_1, \ldots, x_d)^\top \in \mathbb{R}^d$, we let $\|x\| = \sqrt{\sum_{i=1}^d x_i^2}$, $\|x\|_1 = \sum_{i=1}^d |x_i|$, and $\|x\|_\infty = \max(x_1, \ldots, x_d)$. For two positive sequences a_n and b_n, we use $a_n \prec b_n$ and $b_n \succ a_n$ to denote the relation $\lim_{n\to\infty} a_n/b_n = 0$, $a_n \preceq b_n$ and $b_n \succeq a_n$ to denote the relation $\limsup_{n\to\infty} a_n/b_n < +\infty$, and $a_n \asymp b_n$ to denote the relation $a_n \preceq b_n$ and $a_n \succeq b_n$. For any integers k, m, we let I_k be the $k \times k$ identity matrix, 0_k and 1_k be the k-dimensional column vectors of all zeros and all ones, $0_{k\times m}$ be the $k \times m$ zero matrix. For any generic matrix A, cA denotes the matrix of A with all entries multiplied by the number c, and $|A|$ denotes the determinant of A. For a square matrix A, $\mathrm{tr}(A)$ denotes the trace of A. If A is symmetric positive semidefinite, then $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ denote the smallest and largest eigenvalues of A, and $A^{1/2}$ denotes a symmetric positive semidefinite square root of A. For two symmetric positive semidefinite matrix A and B, we use $A \leq B$ and $B \geq A$ to denote the relation that $B - A$ is symmetric positive semidefinite, and use $A < B$ and $B > A$ to denote the relation that $B - A$ is symmetric positive definite. For any matrix A, $\|A\|_{\mathrm{op}} = \sqrt{\lambda_{\max}(A^\top A)}$ denotes the operator norm of A. Let $\mathcal{N}(\mu, \Sigma)$ be the normal distribution with mean μ and covariance matrix Σ. Sometimes to highlight the random variable $Z \sim \mathcal{N}(\mu, \Sigma)$, we also write the normal measure as $\mathcal{N}(\mathrm{d}z|\mu, \Sigma)$. $\Pr(\cdot)$ denotes the probability under true probability measure $P_{(\beta_0, \sigma_0^2, \alpha_0)}$. The convergence in distribution is denoted by $\xrightarrow{\mathcal{D}}$. The acronym i.i.d. stands for "independent and identically distributed".

The following diagram shows the dependence relations among all assumptions, lemmas, theorems, propositions and corollaries. An arrow from A to B in the diagram represents the relation that A is used in the proof of B. The only exception is that many lemmas depend on Assumption (A.1) and we do not plot all the dependence on (A.1) in the diagram.

*stalic@nus.edu.sg


Lemma S.1
Lemma S.2
Lemma S.3
Lemma S.4
Lemma S.5
Lemma S.6
Lemma S.7
Lemma S.8
Lemma S.9
Lemma S.10
Lemma S.11
Lemma S.12
Lemma S.13
Lemma 2.2
Lemma S.19
Lemma S.21
Lemma S.14
Lemma S.15
Lemma S.16
Lemma S.17
Lemma S.18
Lemma S.20
Lemma S.22
Lemma S.23
Lemma S.24
Lemma S.25
Lemma S.26
Lemma S.27
Lemma S.30
Lemma S.31
Lemma S.28
Lemma S.29
Theorem 2.1
Theorem 2.3
Theorem 2.6
Theorem S.1
Assumption (A.1)
Assumption (A.2)
Assumption (A.3)
Assumption (A.4)
Assumption (A.4')
Assumption (S.1)
Corollary 2.7
Lemma S.34
Lemma S.35
Lemma S.32
Lemma S.33
Theorem 3.1
Theorem 3.2
Theorem 3.3
Theorem 3.4
Lemma S.36
Lemma S.37
Assumption (A.5)
Assumption (A.6)
Assumption (A.6')
Proposition 2.4
Proposition 2.5

S1 Proof of Monotonicity and Uniform Convergence in Lemma 2.2

This section is organized as follows.

Subsection S1.1 contains Lemmas S.1, S.2, S.3, and S.4 for showing the monotonicity of REML $\widetilde{\theta}_\alpha$ in Part (i) of Lemma 2.2 in the main text. The main proof is given in the strengthened Lemma S.4.

Subsection S1.2 contains Lemmas S.5, S.6, S.7, S.8, S.9, and S.10, for showing the uniform convergence of REML $\widetilde{\theta}_\alpha$ in Part (ii) of Lemma 2.2 in the main text. We start with a decomposition of the REML $\widetilde{\theta}_\alpha$ in Lemma S.5, and then provide detailed concentration inequalities for each terms in Lemmas S.6, S.7, and S.8. The uniform convergence is proved in Lemma S.9. Lemma S.10 includes the proof of asymptotic normality of the REML $\widetilde{\theta}_\alpha$ in Theorem 2.1, as well as a concentration error bound for $\widetilde{\theta}_{\alpha_0}$, which will be used as a crucial result in the proof of Theorem 2.1 in Section S3.

Subsection S1.3 introduces the RKHS theory with the technical Lemmas S.11, S.12, and S.13. They are used for proving Lemma S.7 and also later for proving Theorem 3.4.

Subsection S1.4 includes the spectral analysis of Matérn covariance function, with the technical Lemmas S.14, S.15, S.16, S.17, S.18, and S.20. Lemma S.20 is used for proving the concentration inequality in Lemma S.9. We also cite the two-sided chi-square concentration inequality from [Laurent and Massart, 2000] in Lemma S.19 and the Hanson-Wright inequality from [Hsu et al., 2012] in Lemma S.21.

We assume Assumptions (A.1) throughout this section. We recall that the universal kriging model (1) in the main text implies that the underlying true model is $Y_n = M_n\beta_0 + X_n$ with $X_n \sim \mathcal{N}(0_n, \sigma_0^2 R_{\alpha_0})$, where R_α is the $n \times n$ Matérn correlation matrix on $\mathcal{S}_n = \{s_1, \ldots, s_n\}$ indexed by α with the (i,j)th entry $R_{\alpha,ij} = K_{\alpha,\nu}(s_i - s_j)$, for $i,j \in \{1,\ldots,n\}$. The REML $\widetilde{\theta}_\alpha$ is defined as

$$\widetilde{\theta}_\alpha = \frac{\alpha^{2\nu} Y_n^\top \left[R_\alpha^{-1} - R_\alpha^{-1} M_n \left(M_n^\top R_\alpha^{-1} M_n + \Omega_\beta\right)^{-1} M_n^\top R_\alpha^{-1}\right] Y_n}{n-p}. \tag{S.1}$$

We emphasize that all the proofs below apply to any symmetric positive semidefinite matrix Ω_β, including the special case $\Omega_\beta = 0_{p\times p}$ corresponding to the noninformative improper prior $\pi(\beta) \propto 1$.

S1.1 Proof of Monotonicity in Part (i) of Lemma 2.2

Lemma S.1. *Suppose that $A_1, A_2 \in \mathbb{R}^{n\times n}$ are two symmetric positive definite matrices and $A_2 - A_1$ is also positive (semi)definite. Then $A_1^{-1} - A_2^{-1}$ is symmetric positive (semi)definite.*

Proof of Lemma S.1. The lemma follows from Theorem 7.7.3 and Corollary 7.7.4 in [Horn and Johnson, 1985]. □

Lemma S.2. *Suppose that $A_1, A_2 \in \mathbb{R}^{n\times n}$ are two symmetric positive definite matrices and $A_2 - A_1$ is also positive definite. Then for any $p \times p$ symmetric positive semidefinite matrix Ω and any full-rank $n \times p$ matrix G, the matrix*

$$\Delta A = \left[A_2 - A_2 G(G^\top A_2 G + \Omega)^{-1} G^\top A_2\right] - \left[A_1 - A_1 G(G^\top A_1 G + \Omega)^{-1} G^\top A_1\right]. \tag{S.2}$$

is symmetric positive semidefinite.

Proof of Lemma S.2. For any $t > 0$, we let $\Omega_t = \Omega + tI_p$. Then Ω_t is symmetric positive definite and hence invertible.

By the Sherman-Morrison-Woodbury formula, we have that for $i = 1, 2$,

$$A_i - A_i G(G^\top A_i G + \Omega_t)^{-1} G^\top A_i = \left(A_i^{-1} + G\Omega_t^{-1} G^\top\right)^{-1}. \tag{S.3}$$

Since $A_2 - A_1$ is symmetric positive definite, by Lemma S.1, we have that $A_1^{-1} - A_2^{-1}$ is symmetric positive definite. But $A_1^{-1} - A_2^{-1} = \left(A_1^{-1} + G\Omega_t^{-1}G^\top\right) - \left(A_2^{-1} + G\Omega_t^{-1}G^\top\right)$ and $A_i^{-1} + G\Omega_t^{-1}G^\top$ for both $i = 1, 2$ are also symmetric positive definite. Therefore, we apply Lemma S.1 again to $A_i^{-1} + G\Omega_t^{-1}G^\top$ for $i = 1, 2$ to conclude that

$$\left(A_2^{-1} + G\Omega_t^{-1}G^\top\right)^{-1} - \left(A_1^{-1} + G\Omega_t^{-1}G^\top\right)^{-1}$$

is a symmetric positive definite matrix. This together with (S.3) implies that

$$\begin{aligned}
&\left(A_2^{-1} + G\Omega_t^{-1}G^\top\right)^{-1} - \left(A_1^{-1} + G\Omega_t^{-1}G^\top\right)^{-1} \\
=& \left[A_2 - A_2 G(G^\top A_2 G + \Omega_t)^{-1} G^\top A_2\right] - \left[A_1 - A_1 G(G^\top A_1 G + \Omega_t)^{-1} G^\top A_1\right] \\
=& \left[A_2 - A_2 G(G^\top A_2 G + \Omega + tI_p)^{-1} G^\top A_2\right] - \left[A_1 - A_1 G(G^\top A_1 G + \Omega + tI_p)^{-1} G^\top A_1\right] \quad \text{(S.4)}
\end{aligned}$$

is symmetric positive definite. The eigenvalues of the last matrix in (S.4) are continuous functions of t. We take $t \to 0+$ and conclude that all eigenvalues of the matrix

$$\left[A_2 - A_2 G(G^\top A_2 G + \Omega)^{-1} G^\top A_2\right] - \left[A_1 - A_1 G(G^\top A_1 G + \Omega)^{-1} G^\top A_1\right]$$

are nonnegative. Therefore, this matrix is symmetric positive semidefinite. □

Lemma S.3. *For all $d \in \mathbb{Z}^+$, $\nu \in \mathbb{R}^+$, for any $0 < \alpha_1 < \alpha_2 < \infty$, the two matrices $\alpha_2^{2\nu} R_{\alpha_2}^{-1} - \alpha_1^{2\nu} R_{\alpha_1}^{-1}$ and $\alpha_2^d R_{\alpha_2} - \alpha_1^d R_{\alpha_1}$ are always positive definite as long as the n points $\{s_1, \ldots, s_n\}$ are distinct in the domain $\mathcal{S} = [0, T]^d$.*

Proof of Lemma S.3. We first define the matrix $\Omega^\dagger = \alpha_1^{-2\nu} R_{\alpha_1} - \alpha_2^{-2\nu} R_{\alpha_2}$. Then the entries of $\Omega^\dagger$ can be expressed in terms of a function $\widetilde{K}_{\Omega^\dagger} : \mathbb{R}^d \to \mathbb{R}$, with

$$\Omega^\dagger_{ij} = \widetilde{K}_{\Omega^\dagger}(s_i - s_j) = \alpha_1^{-2\nu} K_{\alpha_1,\nu}(s_i - s_j) - \alpha_2^{-2\nu} K_{\alpha_2,\nu}(s_i - s_j),$$

for $i, j \in \{1, \ldots, n\}$. The matrix $\Omega^\dagger$ is positive definite if $\widetilde{K}_{\Omega^\dagger}$ is a positive definite function.

From (S.56) in Section S1.4, for the isotropic Matérn covariance function $\sigma^2 K_{\alpha,\nu}$ defined in (2) of the main text, its spectral density is

$$f_{\sigma,\alpha}(\omega) = \frac{\Gamma(\nu + d/2)}{\Gamma(\nu)} \cdot \frac{\sigma^2 \alpha^{2\nu}}{\pi^{d/2} \left(\alpha^2 + \|\omega\|^2\right)^{\nu + d/2}},$$

for any $\omega \in \mathbb{R}^d$. Therefore, we can compute the spectral density of $\widetilde{K}_{\Omega^\dagger}$:

$$\begin{aligned}
f_{\Omega^\dagger}(\omega) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \mathrm{e}^{-\imath\omega^\top x} \widetilde{K}_{\Omega^\dagger}(x) \mathrm{d}x \\
&= \frac{1}{(2\pi)^d} \left\{ \alpha_1^{-2\nu} \int_{\mathbb{R}^d} \mathrm{e}^{-\imath\omega^\top x} K_{\alpha_1,\nu}(x) \mathrm{d}x - \alpha_2^{-2\nu} \int_{\mathbb{R}^d} \mathrm{e}^{-\imath\omega^\top x} K_{\alpha_2,\nu}(x) \mathrm{d}x \right\}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\Gamma(\nu+d/2)}{\pi^{d/2}\Gamma(\nu)}\left\{\alpha_1^{-2\nu}\cdot\frac{\alpha_1^{2\nu}}{\left(\alpha_1^2+\|\omega\|^2\right)^{\nu+d/2}}-\alpha_2^{-2\nu}\cdot\frac{\alpha_2^{2\nu}}{\left(\alpha_2^2+\|\omega\|^2\right)^{\nu+d/2}}\right\}\\
&= \frac{\Gamma(\nu+d/2)}{\pi^{d/2}\Gamma(\nu)}\left\{\frac{1}{\left(\alpha_1^2+\|\omega\|^2\right)^{\nu+d/2}}-\frac{1}{\left(\alpha_2^2+\|\omega\|^2\right)^{\nu+d/2}}\right\}\\
&> 0, \text{ for all } \omega\in\mathbb{R}^d, \qquad\qquad \text{(S.5)}
\end{aligned}
$$

where the last step follows because $0<\alpha_1<\alpha_2$. This has shown that $\widetilde{K}_{\Omega^\dagger}$ is indeed a positive definite function. Therefore, $\Omega^\dagger=\alpha_1^{-2\nu}R_{\alpha_1}-\alpha_2^{-2\nu}R_{\alpha_2}$ is a positive definite matrix. Since $\{s_1,\ldots,s_n\}$ are distinct, both R_{α_1} and R_{α_2} are positive definite matrices. By Lemma S.1, $\alpha_2^{2\nu}R_{\alpha_2}^{-1}-\alpha_1^{2\nu}R_{\alpha_1}^{-1}$ is a positive definite matrix.

Next, we define the matrix $\Omega^\ddagger=\alpha_2^dR_{\alpha_2}-\alpha_1^dR_{\alpha_1}$. Then the entries of $\Omega^\ddagger$ can be expressed in terms of a function $\widetilde{K}_{\Omega^\ddagger}:\mathbb{R}^d\to\mathbb{R}$, with

$$
\Omega^\ddagger_{ij}=\widetilde{K}_{\Omega^\ddagger}(x_i-x_j)=\alpha_2^dK_{\alpha_2,\nu}(x_i-x_j)-\alpha_1^dK_{\alpha_1,\nu}(x_i-x_j),
$$

for $i,j\in\{1,\ldots,n\}$. The matrix $\Omega^\ddagger$ is positive definite if $\widetilde{K}_{\Omega^\ddagger}$ is a positive definite function. We compute the spectral density of $\widetilde{K}_{\Omega^\ddagger}$:

$$
\begin{aligned}
f_{\Omega^\ddagger}(\omega) &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\mathrm{e}^{-\imath\omega^\top x}\widetilde{K}_{\Omega^\ddagger}(x)\mathrm{d}x\\
&= \frac{1}{(2\pi)^d}\left\{\alpha_2^d\int_{\mathbb{R}^d}\mathrm{e}^{-\imath\omega^\top x}K_{\alpha_2,\nu}(x)\mathrm{d}x-\alpha_1^d\int_{\mathbb{R}^d}\mathrm{e}^{-\imath\omega^\top x}K_{\alpha_1,\nu}(x)\mathrm{d}x\right\}\\
&= \frac{\Gamma(\nu+d/2)}{\pi^{d/2}\Gamma(\nu)}\left\{\alpha_2^d\cdot\frac{\alpha_2^{2\nu}}{\left(\alpha_2^2+\|\omega\|^2\right)^{\nu+d/2}}-\alpha_1^d\cdot\frac{\alpha_1^{2\nu}}{\left(\alpha_1^2+\|\omega\|^2\right)^{\nu+d/2}}\right\}\\
&= \frac{\Gamma(\nu+d/2)}{\pi^{d/2}\Gamma(\nu)}\left\{\frac{1}{\left(1+\alpha_2^{-2}\|\omega\|^2\right)^{\nu+d/2}}-\frac{1}{\left(1+\alpha_1^{-2}\|\omega\|^2\right)^{\nu+d/2}}\right\}\\
&> 0, \text{ for all } \omega\in\mathbb{R}^d, \qquad\qquad \text{(S.6)}
\end{aligned}
$$

where the last step follows because $0<\alpha_1<\alpha_2$. This has shown that $\widetilde{K}_{\Omega^\ddagger}$ is indeed a positive definite function. Therefore, $\Omega^\ddagger=\alpha_2^dR_{\alpha_2}-\alpha_1^dR_{\alpha_1}$ is a positive definite matrix. □

We restate and strengthen the monotonicity in Part (i) of Lemma 2.2 in the main text as the following lemma.

Lemma S.4 (Monotonicity of $\widetilde{\theta}_\alpha$ in Lemma 2.2 in the Main Text)**.** *Both $\widetilde{\theta}_\alpha$ defined in* (S.1) *and $\widetilde{\theta}_\alpha^{(1)}$ defined in* (S.10) *are non-decreasing functions in α for all $\alpha\in\mathbb{R}^+$, all $d\in\mathbb{Z}^+$, all $\nu\in\mathbb{R}^+$, for any symmetric positive semidefinite matrix Ω_β.*

Proof of Lemma S.4. We first show that $\widetilde{\theta}_\alpha$ is a non-decreasing function in α. We notice that M_n is full-rank by Assumption (A.1) and Ω_β is positive semidefinite. Consider two generic values $0<\alpha_1<\alpha_2$. By Lemma S.3, we have that $\alpha_2^{2\nu}R_{\alpha_2}^{-1}-\alpha_1^{2\nu}R_{\alpha_1}^{-1}$ is positive definite.

Therefore, in Lemma S.2, we can set $A_1=\alpha_1^{2\nu}R_{\alpha_1}^{-1}$, $A_2=\alpha_2^{2\nu}R_{\alpha_2}^{-1}$, $G=M_n$, $\Omega=\alpha_1^{2\nu}\Omega_\beta$, then the conclusion of Lemma S.2 implies that the matrix ΔA should be positive semidefinite, which implies that

$$
\begin{aligned}
0_{n\times n}\overset{(i)}{\preceq}\Delta A &= \left[\alpha_2^{2\nu}R_{\alpha_2}^{-1}-\alpha_2^{2\nu}R_{\alpha_2}^{-1}M_n\left(\alpha_2^{2\nu}M_n^\top R_{\alpha_2}^{-1}M_n+\alpha_1^{2\nu}\Omega_\beta\right)^{-1}M_n^\top\left(\alpha_2^{2\nu}R_{\alpha_2}^{-1}\right)\right]\\
&\quad-\left[\alpha_1^{2\nu}R_{\alpha_1}^{-1}-\alpha_1^{2\nu}R_{\alpha_1}^{-1}M_n\left(\alpha_1^{2\nu}M_n^\top R_{\alpha_1}^{-1}M_n+\alpha_1^{2\nu}\Omega_\beta\right)^{-1}M_n^\top\left(\alpha_1^{2\nu}R_{\alpha_1}^{-1}\right)\right]
\end{aligned}
$$

$$\begin{aligned}
&\overset{(ii)}{\leq} \Big[\alpha_2^{2\nu}R_{\alpha_2}^{-1} - \alpha_2^{2\nu}R_{\alpha_2}^{-1}M_n\big(\alpha_2^{2\nu}M_n^\top R_{\alpha_2}^{-1}M_n + \alpha_2^{2\nu}\Omega_\beta\big)^{-1}M_n^\top\big(\alpha_2^{2\nu}R_{\alpha_2}^{-1}\big)\Big] \\
&\quad - \Big[\alpha_1^{2\nu}R_{\alpha_1}^{-1} - \alpha_1^{2\nu}R_{\alpha_1}^{-1}M_n\big(\alpha_1^{2\nu}M_n^\top R_{\alpha_1}^{-1}M_n + \alpha_1^{2\nu}\Omega_\beta\big)^{-1}M_n^\top\big(\alpha_1^{2\nu}R_{\alpha_1}^{-1}\big)\Big] \\
&= \alpha_2^{2\nu}\Big[R_{\alpha_2}^{-1} - R_{\alpha_2}^{-1}M_n\big(M_n^\top R_{\alpha_2}^{-1}M_n + \Omega_\beta\big)^{-1}M_n^\top R_{\alpha_2}^{-1}\Big] \\
&\quad - \alpha_1^{2\nu}\Big[R_{\alpha_1}^{-1} - R_{\alpha_1}^{-1}M_n\big(M_n^\top R_{\alpha_1}^{-1}M_n + \Omega_\beta\big)^{-1}M_n^\top R_{\alpha_1}^{-1}\Big], \qquad \text{(S.7)}
\end{aligned}$$

where the $\leq$ relation in the inequalities (i) and (ii) of (S.7) means that if $A \leq B$ for two positive semidefinite matrices A, B, then $B - A$ is positive semidefinite; (i) follows from Lemma S.2, and (ii) follows from replacing $\alpha_1^{2\nu}\Omega_\beta$ inside the first inverse by $\alpha_2^{2\nu}\Omega_\beta$. This implies that the right-hand side of (S.7) is positive semidefinite. Therefore, together with the form of $\widetilde{\theta}_\alpha$ in (S.1), we have proved that if $0 < \alpha_1 < \alpha_2$, then

$$\begin{aligned}
0 &\leq \alpha_2^{2\nu}Y_n^\top\Big[R_{\alpha_2}^{-1} - R_{\alpha_2}^{-1}M_n\big(M_n^\top R_{\alpha_2}^{-1}M_n + \Omega_\beta\big)^{-1}M_n^\top R_{\alpha_2}^{-1}\Big]Y_n/(n-p) \\
&\quad - \alpha_1^{2\nu}Y_n^\top\Big[R_{\alpha_1}^{-1} - R_{\alpha_1}^{-1}M_n\big(M_n^\top R_{\alpha_1}^{-1}M_n + \Omega_\beta\big)^{-1}M_n^\top R_{\alpha_1}^{-1}\Big]Y_n/(n-p) \\
&= \widetilde{\theta}_{\alpha_2} - \widetilde{\theta}_{\alpha_1}, \qquad \text{(S.8)}
\end{aligned}$$

so $\widetilde{\theta}_{\alpha_1} \leq \widetilde{\theta}_{\alpha_2}$, i.e., $\widetilde{\theta}_\alpha$ is a non-decreasing function in α.

For $\widetilde{\theta}_\alpha^{(1)} = \alpha^{2\nu}X_n^\top R_\alpha^{-1}X_n/(n-p)$ from (S.10), since $\alpha_2^{2\nu}R_{\alpha_2}^{-1} - \alpha_1^{2\nu}R_{\alpha_1}^{-1}$ is positive definite by Lemma S.3, we have that for any $X_n \in \mathbb{R}^n$, $\widetilde{\theta}_{\alpha_2}^{(1)} \geq \widetilde{\theta}_{\alpha_1}^{(1)}$, i.e., $\widetilde{\theta}_\alpha^{(1)}$ is a non-decreasing function in α. □

S1.2 Proof of Uniform Convergence in Part (ii) of Lemma 2.2

We prove Part (ii) of Lemma 2.2 in this subsection. We first restate the important quantities of $\underline{\kappa}, \overline{\kappa}, \underline{\alpha}_n, \overline{\alpha}_n$ as in (12) of the main text. We also define the constant $\tau \in (0, 1/2)$:

$$\begin{aligned}
\underline{\kappa} &= \frac{1}{2}\min\left\{\frac{0.9}{(2d+0.94)(8\nu+3d-0.9)},\ \frac{1}{4(3\nu+d)},\ 0.01\right\}, \quad \underline{\alpha}_n = n^{-\underline{\kappa}}, \\
\overline{\kappa} &= \frac{1}{2}\min\left\{\frac{0.9}{(2d+0.94)(8\nu+5d+0.9)},\ \frac{1}{2(2\nu+d)},\ 0.01\right\}, \quad \overline{\alpha}_n = n^{-\overline{\kappa}}, \\
\tau &= \frac{1}{2}\min\left\{\frac{0.9}{4d+1.88} - (4\nu+5d+0.45)\overline{\kappa},\ \frac{15}{98} - 5.95\overline{\kappa},\ \frac{1}{2} - (2\nu+d)\overline{\kappa},\ \frac{1}{2} - 5\overline{\kappa},\right. \\
&\qquad \left.\frac{0.9}{4d+1.88} - (4\nu+1.5d-0.45)\underline{\kappa},\ \frac{15}{98} - 4.05\underline{\kappa},\ \frac{1}{2} - 2(3\nu+d)\underline{\kappa},\ \frac{1}{2} - 5\underline{\kappa}\right\}. \qquad \text{(S.9)}
\end{aligned}$$

Lemma S.5. *For all $d \in \mathbb{Z}^+, \nu \in \mathbb{R}^+, \alpha \in \mathbb{R}^+$, the REML $\widetilde{\theta}_\alpha$ in* (S.1) *can be decomposed into three terms:*

$$\begin{aligned}
\widetilde{\theta}_\alpha &= \widetilde{\theta}_\alpha^{(1)} - \widetilde{\theta}_\alpha^{(2)} + \widetilde{\theta}_\alpha^{(3)}, \\
\widetilde{\theta}_\alpha^{(1)} &= \frac{\alpha^{2\nu}X_n^\top R_\alpha^{-1}X_n}{n-p}, \\
\widetilde{\theta}_\alpha^{(2)} &= \frac{\alpha^{2\nu}X_n^\top R_\alpha^{-1}M_n\big(M_n^\top R_\alpha^{-1}M_n\big)^{-1}M_n^\top R_\alpha^{-1}X_n}{n-p},
\end{aligned}$$

$$\widetilde{\theta}_\alpha^{(3)} = \frac{\alpha^{2\nu} Y_n^\top R_\alpha^{-1} M_n \left[\left(M_n^\top R_\alpha^{-1} M_n\right)^{-1} - \left(M_n^\top R_\alpha^{-1} M_n + \Omega_\beta\right)^{-1} \right] M_n^\top R_\alpha^{-1} Y_n}{n-p}. \tag{S.10}$$

Furthermore,

$$0 \le \widetilde{\theta}_\alpha^{(2)} \le \widetilde{\theta}_\alpha^{(1)}, \qquad \widetilde{\theta}_\alpha^{(3)} \ge 0.$$

Proof of Lemma S.5. The universal kriging model (1) implies that $Y_n = M_n\beta_0 + X_n$ with $X_n \sim \mathcal{N}(0_n, \sigma_0^2 R_{\alpha_0})$. Therefore, the REML $\widetilde{\theta}_\alpha$ defined in (S.1) can be rewritten as

$$\begin{aligned}
\widetilde{\theta}_\alpha &= \frac{\alpha^{2\nu} Y_n^\top \left[R_\alpha^{-1} - R_\alpha^{-1} M_n \left(M_n^\top R_\alpha^{-1} M_n + \Omega_\beta\right)^{-1} M_n^\top R_\alpha^{-1} \right] Y_n}{n-p} \\
&= \frac{\alpha^{2\nu} (M_n\beta_0 + X_n)^\top \left[R_\alpha^{-1} - R_\alpha^{-1} M_n \left(M_n^\top R_\alpha^{-1} M_n\right)^{-1} M_n^\top R_\alpha^{-1} \right] (M_n\beta_0 + X_n)}{n-p} \\
&\quad + \frac{\alpha^{2\nu} Y_n^\top R_\alpha^{-1} M_n \left[\left(M_n^\top R_\alpha^{-1} M_n\right)^{-1} - \left(M_n^\top R_\alpha^{-1} M_n + \Omega_\beta\right)^{-1} \right] M_n^\top R_\alpha^{-1} Y_n}{n-p} \\
&\overset{(i)}{=} \frac{\alpha^{2\nu} X_n^\top R_\alpha^{-1} X_n}{n-p} - \frac{\alpha^{2\nu} X_n^\top R_\alpha^{-1} M_n \left(M_n^\top R_\alpha^{-1} M_n\right)^{-1} M_n^\top R_\alpha^{-1} X_n}{n-p} \\
&\quad + \frac{\alpha^{2\nu} Y_n^\top R_\alpha^{-1} M_n \left[\left(M_n^\top R_\alpha^{-1} M_n\right)^{-1} - \left(M_n^\top R_\alpha^{-1} M_n + \Omega_\beta\right)^{-1} \right] M_n^\top R_\alpha^{-1} Y_n}{n-p} \\
&= \widetilde{\theta}_\alpha^{(1)} - \widetilde{\theta}_\alpha^{(2)} + \widetilde{\theta}_\alpha^{(3)},
\end{aligned} \tag{S.11}$$

where in (i), we use the relation $\left[R_\alpha^{-1} - R_\alpha^{-1} M_n \left(M_n^\top R_\alpha^{-1} M_n\right)^{-1} M_n^\top R_\alpha^{-1} \right] M_n = 0_{n\times p}$.

Since for any $\alpha > 0$,

$$\begin{aligned}
&R_\alpha^{-1} - R_\alpha^{-1} M_n \left(M_n^\top R_\alpha^{-1} M_n\right)^{-1} M_n^\top R_\alpha^{-1} \\
&= R_\alpha^{-1/2} \left[I_n - R_\alpha^{-1/2} M_n \left(M_n^\top R_\alpha^{-1} M_n\right)^{-1} M_n^\top R_\alpha^{-1/2} \right] R_\alpha^{-1/2},
\end{aligned}$$

where $I_n - R_\alpha^{-1/2} M_n \left(M_n^\top R_\alpha^{-1} M_n\right)^{-1} M_n^\top R_\alpha^{-1/2}$ is an idempotent matrix, it follows that $0 \le \widetilde{\theta}_\alpha^{(2)} \le \widetilde{\theta}_\alpha^{(1)}$.

Since Ω_β is symmetric positive semidefinite, by Lemma S.1, $\left(M_n^\top R_\alpha^{-1} M_n\right)^{-1} - \left(M_n^\top R_\alpha^{-1} M_n + \Omega_\beta\right)^{-1}$ is positive semidefinite. Therefore, $\widetilde{\theta}_\alpha^{(3)} \ge 0$ for any $\alpha > 0$. □

Lemma S.6. *For $\widetilde{\theta}_\alpha^{(2)}$ defined in* (S.10), *for $d \in \mathbb{Z}^+$ and $\nu \in \mathbb{R}^+$, there exists a large integer N_1' that only depends on $\nu, d, T, \theta_0, \alpha_0$, such that for all $n > N_1'$ and $\underline{\alpha}_n, \overline{\alpha}_n, \tau$ defined in* (S.9),

$$\Pr\left(\sqrt{n}\widetilde{\theta}_{\alpha_0}^{(2)} > \theta_0 n^{-\tau}/16\right) \le \exp(-16\log^2 n), \tag{S.12}$$

$$\Pr\left(\sqrt{n}\widetilde{\theta}_{\underline{\alpha}_n}^{(2)} > \theta_0 n^{-\tau}/16\right) \le \exp(-16\log^2 n), \tag{S.13}$$

$$\Pr\left(\sqrt{n}\widetilde{\theta}_{\overline{\alpha}_n}^{(2)} > \theta_0 n^{-\tau}/16\right) \le \exp(-16\log^2 n). \tag{S.14}$$

Proof of Lemma S.6. We first prove (S.14) below. Then the proofs of (S.12) and (S.13) follow similarly.

For $\sqrt{n}\widetilde{\theta}_{\overline{\alpha}_n}^{(2)}$, we notice that by Lemma S.15, $\lambda_{i,n}(\overline{\alpha}_n) \ge (\alpha_0/\overline{\alpha}_n)^{2\nu+d}$ for all $i = 1, \ldots, n$, so $\lambda_{\max}\left(\Lambda_{\overline{\alpha}_n}^{-1}\right) \le (\overline{\alpha}_n/\alpha_0)^{2\nu+d}$.

Using Lemma S.14, we have $\overline{\alpha}_n^{2\nu} R_{\overline{\alpha}_n}^{-1} = \theta_0 U_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1} U_{\overline{\alpha}_n}^\top$. For any $\alpha > 0$, we define $Z_n(\alpha) = (Z_{1,n}(\alpha), \ldots, Z_{n,n}(\alpha))^\top = U_\alpha^\top X_n$. Since $X_n \sim \mathcal{N}(0_n, \sigma_0^2 R_{\alpha_0})$, by Lemma S.14, we have $Z_n(\alpha) \sim \mathcal{N}(0_n, I_n)$ for any $\alpha > 0$. We can then write $X_n = U_{\overline{\alpha}_n}^{-\top} Z_n(\overline{\alpha}_n)$. Since Ω_β is positive semidefinite, we can upper bound $\widetilde{\theta}_{\overline{\alpha}_n}^{(2)}$ by

$$
\begin{aligned}
\widetilde{\theta}_{\overline{\alpha}_n}^{(2)} &= (n-p)^{-1} \overline{\alpha}_n^{2\nu} X_n^\top R_{\overline{\alpha}_n}^{-1} M_n \big(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n + \Omega_\beta\big)^{-1} M_n^\top R_{\overline{\alpha}_n}^{-1} X_n \\
&\leq (n-p)^{-1} \overline{\alpha}_n^{2\nu} X_n^\top R_{\overline{\alpha}_n}^{-1} M_n \big(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big)^{-1} M_n^\top R_{\overline{\alpha}_n}^{-1} X_n \\
&= (n-p)^{-1} \theta_0 X_n^\top \left(\overline{\alpha}_n^{2\nu} R_{\overline{\alpha}_n}^{-1}\right) M_n \big[M_n^\top \left(\overline{\alpha}_n^{2\nu} R_{\overline{\alpha}_n}^{-1}\right) M_n\big]^{-1} M_n^\top \left(\overline{\alpha}_n^{2\nu} R_{\overline{\alpha}_n}^{-1}\right) X_n \\
&\leq \frac{2\theta_0}{n} Z_n(\overline{\alpha}_n)^\top \Lambda_{\overline{\alpha}_n}^{-1} U_{\overline{\alpha}_n}^\top M_n \big(M_n^\top U_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1} U_{\overline{\alpha}_n}^\top M_n\big)^{-1} M_n^\top U_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1} Z_n(\overline{\alpha}_n) \\
&= \frac{2\theta_0}{n} Z_n(\overline{\alpha}_n)^\top \Lambda_{\overline{\alpha}_n}^{-1/2} H_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1/2} Z_n(\overline{\alpha}_n), \qquad \text{(S.15)}
\end{aligned}
$$

where $H_{\overline{\alpha}_n} = \Lambda_{\overline{\alpha}_n}^{-1/2} U_{\overline{\alpha}_n}^\top M_n \big(M_n^\top U_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1} U_{\overline{\alpha}_n}^\top M_n\big)^{-1} M_n^\top U_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1/2}$ is an $n \times n$ idempotent matrix of rank p (i.e., $H_{\overline{\alpha}_n}^2 = H_{\overline{\alpha}_n}$), since $\mathrm{rank}(M_n) = p \ll n$ as $n \to \infty$. Hence $\mathrm{tr}(H_{\overline{\alpha}_n}) = p$.

We are going to apply the Hanson-Wright inequality in Lemma S.21 to (S.15), with $Z = Z_n(\overline{\alpha}_n)$, $z = 16 \log^2 n$, and $\Sigma = \Lambda_{\overline{\alpha}_n}^{-1/2} H_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1/2}$. For this purpose, we need to find upper bounds for $\mathrm{tr}(\Sigma)$, $\mathrm{tr}(\Sigma^2)$, and $\|\Sigma\|_{\mathrm{op}}$ in Lemma S.21. We first notice that for two generic $n \times n$ symmetric positive semidefinite matrices A and B,

$$
\mathrm{tr}(BA) = \mathrm{tr}(AB) = \mathrm{tr}(B^{1/2} A B^{1/2}) \leq \mathrm{tr}\{B^{1/2} (\lambda_{\max}(A) I) B^{1/2}\} \leq \lambda_{\max}(A)\, \mathrm{tr}(B).
$$

Therefore, using $\lambda_{\max}(\Lambda_{\overline{\alpha}_n}^{-1}) \leq (\overline{\alpha}_n/\alpha_0)^{2\nu+d}$, we apply the inequality above repeatedly to obtain that

$$
\begin{aligned}
\mathrm{tr}\left(\Lambda_{\overline{\alpha}_n}^{-1/2} H_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1/2}\right) &= \mathrm{tr}\left(\Lambda_{\overline{\alpha}_n}^{-1} H_{\overline{\alpha}_n}\right) \leq \lambda_{\max}\left(\Lambda_{\overline{\alpha}_n}^{-1}\right) \mathrm{tr}\left(H_{\overline{\alpha}_n}\right) \leq p(\overline{\alpha}_n/\alpha_0)^{2\nu+d}, \\
\mathrm{tr}\left[\left(\Lambda_{\overline{\alpha}_n}^{-1/2} H_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1/2}\right)^2\right] &= \mathrm{tr}\left(\Lambda_{\overline{\alpha}_n}^{-1} H_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1} H_{\overline{\alpha}_n}\right) \leq \lambda_{\max}\left(\Lambda_{\overline{\alpha}_n}^{-1}\right) \cdot \mathrm{tr}\left(H_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1} H_{\overline{\alpha}_n}\right) \\
&\leq \lambda_{\max}\left(\Lambda_{\overline{\alpha}_n}^{-1}\right) \cdot \mathrm{tr}\left(\Lambda_{\overline{\alpha}_n}^{-1} H_{\overline{\alpha}_n}^2\right) \leq \lambda_{\max}\left(\Lambda_{\overline{\alpha}_n}^{-1}\right)^2 \cdot \mathrm{tr}\left(H_{\overline{\alpha}_n}^2\right) \\
&= \lambda_{\max}\left(\Lambda_{\overline{\alpha}_n}^{-1}\right)^2 \cdot \mathrm{tr}\left(H_{\overline{\alpha}_n}\right) = p(\overline{\alpha}_n/\alpha_0)^{2(2\nu+d)}, \\
\left\|\Lambda_{\overline{\alpha}_n}^{-1/2} H_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1/2}\right\|_{\mathrm{op}} &\leq \left[\lambda_{\max}\left\{\left(\Lambda_{\overline{\alpha}_n}^{-1/2} H_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1/2}\right)^2\right\}\right]^{1/2} \\
&\leq \left[\mathrm{tr}\left\{\left(\Lambda_{\overline{\alpha}_n}^{-1/2} H_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1/2}\right)^2\right\}\right]^{1/2} \leq \sqrt{p}(\overline{\alpha}_n/\alpha_0)^{2\nu+d}. \qquad \text{(S.16)}
\end{aligned}
$$

Therefore, for $z = 16 \log^2 n$ and $\Sigma = \Lambda_{\overline{\alpha}_n}^{-1/2} H_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1/2}$, given the choice of τ in (S.9), $\tau < 1/2 - (2\nu + d)\overline{\kappa}$, so we have that for all sufficiently large n,

$$
\begin{aligned}
&\mathrm{tr}(\Sigma) + 2\sqrt{\mathrm{tr}(\Sigma^2) z} + 2\|\Sigma\|_{\mathrm{op}} z \\
&\leq p(\overline{\alpha}_n/\alpha_0)^{2\nu+d} + 8\sqrt{p}(\overline{\alpha}_n/\alpha_0)^{2\nu+d} \log n + 32\sqrt{p}(\overline{\alpha}_n/\alpha_0)^{2\nu+d} \log^2 n \\
&\leq 42 p n^{(2\nu+d)\overline{\kappa}} \log^2 n < n^{1/2-\tau}/128. \qquad \text{(S.17)}
\end{aligned}
$$

We now apply Lemma S.21 to (S.15) with $Z = Z_n(\overline{\alpha}_n)$, $z = 16 \log^2 n$, and $\Sigma = \Lambda_{\overline{\alpha}_n}^{-1/2} H_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1/2}$ to obtain that for all sufficiently large n,

$$
\begin{aligned}
&\Pr\left(\sqrt{n}\widetilde{\theta}_{\overline{\alpha}_n}^{(2)} > \frac{\theta_0}{16} n^{-\tau}\right) \\
\leq\ &\Pr\left(Z_n(\overline{\alpha}_n)^\top \Lambda_{\overline{\alpha}_n}^{-1/2} H_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1/2} Z_n(\overline{\alpha}_n) > \frac{1}{32} n^{1/2-\tau}\right)
\end{aligned}
$$

$$\leq \Pr\left(Z_n(\overline{\alpha}_n)^\top \Lambda_{\overline{\alpha}_n}^{-1/2} H_{\overline{\alpha}_n} \Lambda_{\overline{\alpha}_n}^{-1/2} Z_n(\overline{\alpha}_n) > \mathrm{tr}(\Sigma) + 2\sqrt{\mathrm{tr}(\Sigma^2) z} + 2\|\Sigma\|_{\mathrm{op}} z\right)$$
$$\leq \exp(-z) = \exp(-16\log^2 n). \tag{S.18}$$

This proves (S.14).

The proof of (S.12) is similar to the proof of (S.14) above. (S.15) still holds by replacing all $\overline{\alpha}_n$ with α_0. We notice that $\Lambda_{\alpha_0} = I_n$, $\lambda_{\max}(\Lambda_{\alpha_0}^{-1}) = 1$, so the three upper bounds in (S.16) become $p, p, \sqrt{p}$, respectively. With $\overline{\alpha}_n$ replaced by α_0, the left-hand side (S.17) is upper bounded by $p+8\sqrt{p}\log n+32\sqrt{p}\log^2 n$, which is smaller than $n^{1/2-\tau}/32$ for all sufficiently large n. Hence (S.18) holds with $\overline{\alpha}_n$ replaced by α_0. This proves (S.12).

The proof of (S.13) is also similar to the proof of (S.14) above. (S.15) still holds by replacing all $\overline{\alpha}_n$ with $\underline{\alpha}_n$. We notice that from S.15, $\lambda_{i,n}(\underline{\alpha}_n) \geq 1$ for all $i = 1, \ldots, n$, so $\lambda_{\max}\left(\Lambda_{\underline{\alpha}_n}^{-1}\right) \leq 1$. As a result, the three upper bounds in (S.16) become $p, p, \sqrt{p}$, respectively. With $\overline{\alpha}_n$ replaced by $\underline{\alpha}_n$, the left-hand side (S.17) is upper bounded by $p + 8\sqrt{p}\log n + 32\sqrt{p}\log^2 n$, which is smaller than $n^{1/2-\tau}/32$ for all sufficiently large n. Hence (S.18) holds with $\overline{\alpha}_n$ replaced by $\underline{\alpha}_n$. This proves (S.13). □

Lemma S.7. *For $\widetilde{\theta}_\alpha^{(3)}$ defined in (S.10), for $d \in \mathbb{Z}^+$ and $\nu \in \mathbb{R}^+$, there exists a large integer N_2' that only depends on $\nu, d, T, \beta_0, \theta_0, \alpha_0$ and the $\mathcal{W}_2^{\nu+d/2}(\mathcal{S})$ norms of $\mathrm{m}_1(\cdot), \ldots, \mathrm{m}_p(\cdot)$, such that for all $n > N_2'$ and $\underline{\alpha}_n, \overline{\alpha}_n, \tau$ defined in (S.9),*

$$\Pr\left(\sqrt{n}\widetilde{\theta}_{\alpha_0}^{(3)} > \theta_0 n^{-\tau}/16\right) \leq \exp(-16\log^2 n), \tag{S.19}$$
$$\Pr\left(\sqrt{n}\widetilde{\theta}_{\underline{\alpha}_n}^{(3)} > \theta_0 n^{-\tau}/16\right) \leq \exp(-16\log^2 n), \tag{S.20}$$
$$\Pr\left(\sqrt{n}\widetilde{\theta}_{\overline{\alpha}_n}^{(3)} > \theta_0 n^{-\tau}/16\right) \leq \exp(-16\log^2 n). \tag{S.21}$$

Proof of Lemma S.7. We first prove (S.21). The proofs of (S.19) and (S.20) follow similarly.

In the definition of $\widetilde{\theta}_{\overline{\alpha}_n}^{(3)}$ in (S.10), we directly drop the positive semidefinite matrix $\big(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n + \Omega_\beta\big)^{-1}$ in the middle bracket, and obtain that

$$\begin{aligned}
\widetilde{\theta}_{\overline{\alpha}_n}^{(3)} &= \frac{\overline{\alpha}_n^{2\nu} Y_n^\top R_{\overline{\alpha}_n}^{-1} M_n \left[\big(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big)^{-1} - \big(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n + \Omega_\beta\big)^{-1}\right] M_n^\top R_{\overline{\alpha}_n}^{-1} Y_n}{n-p} \\
&\leq \frac{\overline{\alpha}_n^{2\nu} Y_n^\top R_{\overline{\alpha}_n}^{-1} M_n \big(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big)^{-1} M_n^\top R_{\overline{\alpha}_n}^{-1} Y_n}{n-p} \\
&= \frac{\overline{\alpha}_n^{2\nu} (M_n\beta_0 + X_n)^\top R_{\overline{\alpha}_n}^{-1} M_n \big(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big)^{-1} M_n^\top R_{\overline{\alpha}_n}^{-1} (M_n\beta_0 + X_n)}{n-p}.
\end{aligned} \tag{S.22}$$

For two vectors $u, v \in \mathbb{R}^n$ and an $n \times n$ symmetric positive definite matrix Σ, we have the following inequality:

$$\begin{aligned}
&(u+v)^\top \Sigma (u+v) = u^\top \Sigma u + v^\top \Sigma v + 2u^\top \Sigma v \\
&= u^\top \Sigma u + v^\top \Sigma v + 2(\Sigma^{1/2} u)^\top (\Sigma^{1/2} v) \leq u^\top \Sigma u + v^\top \Sigma v + u^\top \Sigma u + v^\top \Sigma v \\
&= 2\left(u^\top \Sigma u + v^\top \Sigma v\right).
\end{aligned} \tag{S.23}$$

We apply (S.23) to the right-hand side of (S.22), with $u = M_n\beta_0$, $v = X_n$, and $\Sigma = R_{\overline{\alpha}_n}^{-1} M_n \big(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big)^{-1} M_n^\top R_{\overline{\alpha}_n}^{-1}$ to obtain that

$$\widetilde{\theta}_{\overline{\alpha}_n}^{(3)} \leq \frac{\overline{\alpha}_n^{2\nu} (M_n\beta_0 + X_n)^\top R_{\overline{\alpha}_n}^{-1} M_n \big(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big)^{-1} M_n^\top R_{\overline{\alpha}_n}^{-1} (M_n\beta_0 + X_n)}{n-p}$$

$$
\begin{aligned}
&\leq \frac{2\overline{\alpha}_n^{2\nu}}{n-p}\left[\beta_0^\top M_n^\top R_{\overline{\alpha}_n}^{-1} M_n\beta_0 + X_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big)^{-1} M_n^\top R_{\overline{\alpha}_n}^{-1} X_n\right] \\
&\leq \frac{2\theta_0}{n-p}\beta_0^\top M_n^\top\left[(\theta_0/\overline{\alpha}_n^{2\nu})R_{\overline{\alpha}_n}\right]^{-1} M_n\beta_0 + \frac{2\overline{\alpha}_n^{2\nu}}{n-p} X_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big)^{-1} M_n^\top R_{\overline{\alpha}_n}^{-1} X_n.
\end{aligned}
\tag{S.24}
$$

We bound the two terms in (S.24). Because $\mathrm{m}_1,\ldots,\mathrm{m}_p \in \mathcal{W}_2^{\nu+d/2}(\mathcal{S})$ by Assumption (A.1), Lemma S.11 implies that $\mathrm{m}_1,\ldots,\mathrm{m}_p \in \mathcal{H}_{\sigma_0^2 K_{\alpha_0,\nu}}$, the RKHS of Matérn kernel $\sigma^2 K_{\alpha,\nu}$ for any $(\sigma^2,\alpha)\in\mathbb{R}^+\times\mathbb{R}^+$. Let $\mathrm{m}_{j,n} = (\mathrm{m}_j(s_1),\ldots,\mathrm{m}_j(s_n))^\top \in \mathbb{R}^n$ for $j=1,\ldots,p$. Then we can apply Lemma S.12, Lemma S.13, and Lemma S.11 to the first term in (S.24) and obtain that

$$
\begin{aligned}
&\frac{2\theta_0}{n-p}\beta_0^\top M_n^\top\left[(\theta_0/\overline{\alpha}_n^{2\nu})R_{\overline{\alpha}_n}\right]^{-1} M_n\beta_0 \\
&\leq \frac{2\theta_0}{n-p}\beta_0^\top\beta_0\cdot\lambda_{\max}\left(M_n^\top\left[(\theta_0/\overline{\alpha}_n^{2\nu})R_{\overline{\alpha}_n}\right]^{-1} M_n\right) \\
&\leq \frac{2\theta_0}{n-p}\|\beta_0\|^2\cdot\mathrm{tr}\left(M_n^\top\left[(\theta_0/\overline{\alpha}_n^{2\nu})R_{\overline{\alpha}_n}\right]^{-1} M_n\right) \\
&= \frac{2\theta_0}{n-p}\|\beta_0\|^2\cdot\sum_{j=1}^p \mathrm{m}_{j,n}^\top\left[(\theta_0/\overline{\alpha}_n^{2\nu})R_{\overline{\alpha}_n}\right]^{-1}\mathrm{m}_{j,n} \\
&\overset{(i)}{\leq} \frac{2\theta_0}{n-p}\|\beta_0\|^2\cdot\sum_{j=1}^p \|\mathrm{m}_j\|^2_{\mathcal{H}_{(\theta_0/\overline{\alpha}_n^{2\nu})K_{\overline{\alpha}_n,\nu}}} \\
&\overset{(ii)}{\leq} \frac{2\theta_0}{n-p}\|\beta_0\|^2\cdot\sum_{j=1}^p (\overline{\alpha}_n/\alpha_0)^{2\nu+d}\|\mathrm{m}_j\|^2_{\mathcal{H}_{\sigma_0^2 K_{\alpha_0,\nu}}} \\
&\overset{(iii)}{\leq} \frac{2\theta_0}{n-p}\|\beta_0\|^2(\overline{\alpha}_n/\alpha_0)^{2\nu+d}\cdot c_2(\sigma_0,\alpha_0)^2\sum_{j=1}^p\|\mathrm{m}_j\|^2_{\mathcal{W}_2^{\nu+d/2}(\mathcal{S})} \\
&\leq \frac{4\|\beta_0\|^2}{\alpha_0^{2\nu+d}}c_2(\sigma_0,\alpha_0)^2\left(\sum_{j=1}^p\|\mathrm{m}_j\|^2_{\mathcal{W}_2^{\nu+d/2}(\mathcal{S})}\right)\cdot\theta_0 n^{(2\nu+d)\overline{\kappa}-1} \overset{(iv)}{\leq} \theta_0 n^{-\tau-1/2}/32,
\end{aligned}
\tag{S.25}
$$

for all sufficiently large n, where (i) follows by applying Lemma S.12 to each $\mathrm{m}_1(\cdot),\ldots,\mathrm{m}_p(\cdot)$ with the covariance kernel $(\theta_0/\overline{\alpha}_n^{2\nu})K_{\overline{\alpha}_n,\nu}$, (ii) follows from Lemma S.13, (iii) follows from Lemma S.11 with the constant $c_2(\sigma_0,\alpha_0)$ defined in Lemma S.11, and (iv) follows from the definition of τ in (S.9) and $\tau < 1/2-(2\nu+d)\overline{\kappa}$.

For the second term in (S.24), we notice that the exact term

$$
\frac{\overline{\alpha}_n^{2\nu}}{n-p} X_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big)^{-1} M_n^\top R_{\overline{\alpha}_n}^{-1} X_n
$$

shows up as an upper bound for $\widetilde{\theta}^{(2)}_{\overline{\alpha}_n}$ in (S.15) in the proof of Lemma S.6. Therefore, we can directly make use of the inequalities in (S.15), (S.17), and (S.18) to conclude that for all sufficiently large n,

$$
\begin{aligned}
&\Pr\left(\frac{2\overline{\alpha}_n^{2\nu}}{n-p} X_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n\big)^{-1} M_n^\top R_{\overline{\alpha}_n}^{-1} X_n > \theta_0 n^{-\tau-1/2}/32\right) \\
&\leq \Pr\left(\frac{4\theta_0}{n} Z_n(\overline{\alpha}_n)^\top \Lambda_{\overline{\alpha}_n}^{-1/2} H_{\overline{\alpha}_n}\Lambda_{\overline{\alpha}_n}^{-1/2} Z_n(\overline{\alpha}_n) > \theta_0 n^{-\tau-1/2}/32\right) \\
&= \Pr\left(Z_n(\overline{\alpha}_n)^\top \Lambda_{\overline{\alpha}_n}^{-1/2} H_{\overline{\alpha}_n}\Lambda_{\overline{\alpha}_n}^{-1/2} Z_n(\overline{\alpha}_n) > n^{1/2-\tau}/128\right) \\
&\leq \exp(-16\log^2 n).
\end{aligned}
\tag{S.26}
$$

Therefore, we can combine (S.24), (S.25), and (S.26) together to conclude that for all sufficiently large n,

$$
\begin{aligned}
&\Pr\left(\sqrt{n}\widetilde{\theta}^{(3)}_{\overline{\alpha}_n} > \theta_0 n^{-\tau}/16\right)\\
&\leq \Pr\left(\sqrt{n}\cdot \frac{2\theta_0}{n-p}\beta_0^\top M_n^\top \left[(\theta_0/\overline{\alpha}_n^{2\nu})R_{\overline{\alpha}_n}\right]^{-1} M_n\beta_0 > \theta_0 n^{-\tau}/32\right)\\
&\quad + \Pr\left(\sqrt{n}\cdot \frac{2\overline{\alpha}_n^{2\nu}}{n-p} X_n^\top R_{\overline{\alpha}_n}^{-1} M_n \left(M_n^\top R_{\overline{\alpha}_n}^{-1} M_n\right)^{-1} M_n^\top R_{\overline{\alpha}_n}^{-1} X_n > \theta_0 n^{-\tau}/32\right)\\
&\leq 0 + \exp(-16\log^2 n) = \exp(-16\log^2 n). \qquad\qquad (S.27)
\end{aligned}
$$

This proves (S.21).

For the proofs of (S.19) and (S.20), we only need to modify the proof above for (S.21) for a looser upper bound. In particular, the relation (S.24) still holds with $\overline{\alpha}_n$ replaced by both α_0 and $\underline{\alpha}_n$; in the inequality (S.25), $(\overline{\alpha}_n/\alpha_0)^{2\nu+d}$ in step (ii) will be replaced by 1 if $\overline{\alpha}_n$ is replaced by both α_0 and $\underline{\alpha}_n$, such that $n^{(2\nu+d)\overline{\kappa}-1}$ before the last step of (S.25) is replaced by the smaller n^{-1}, which means that (S.25) remains true if $\overline{\alpha}_n$ is replaced by both α_0 and $\underline{\alpha}_n$. Given Lemma S.6, (S.26) still holds true if $\overline{\alpha}_n$ is replaced by both α_0 and $\underline{\alpha}_n$. Therefore, (S.27) holds for both α_0 and $\underline{\alpha}_n$. This completes the proof. □

Lemma S.8. *For $\widetilde{\theta}_\alpha$ defined in* (S.1)*, for $d \in \{1,2,3\}$ and $\nu \in \mathbb{R}^+$, there exists a large integer N_3' that only depends on $\nu, d, T, \beta_0, \theta_0, \alpha_0$ and the $\mathcal{W}_2^{\nu+d/2}(\mathcal{S})$ norms of $\mathrm{m}_1(\cdot), \ldots, \mathrm{m}_p(\cdot)$, such that for all $n > N_3'$,*

$$
\Pr\left(0 \leq \sqrt{n}\left(\widetilde{\theta}_{\overline{\alpha}_n} - \widetilde{\theta}_{\alpha_0}\right) \leq \frac{\theta_0}{2} n^{-\tau}\right) \geq 1 - 2\exp(-4\log^2 n), \qquad (S.28)
$$

$$
\Pr\left(0 \leq \sqrt{n}\left(\widetilde{\theta}^{(1)}_{\overline{\alpha}_n} - \widetilde{\theta}^{(1)}_{\alpha_0}\right) \leq \frac{\theta_0}{4} n^{-\tau}\right) \geq 1 - \exp(-4\log^2 n), \qquad (S.29)
$$

$$
\Pr\left(0 \leq \sqrt{n}\left(\widetilde{\theta}_{\alpha_0} - \widetilde{\theta}_{\underline{\alpha}_n}\right) \leq \frac{\theta_0}{2} n^{-\tau}\right) \geq 1 - 2\exp(-4\log^2 n), \qquad (S.30)
$$

$$
\Pr\left(0 \leq \sqrt{n}\left(\widetilde{\theta}^{(1)}_{\alpha_0} - \widetilde{\theta}^{(1)}_{\underline{\alpha}_n}\right) \leq \frac{\theta_0}{4} n^{-\tau}\right) \geq 1 - \exp(-4\log^2 n), \qquad (S.31)
$$

where $\tau, \underline{\alpha}_n, \overline{\alpha}_n$ are as defined the same as in (S.9).

Proof of Lemma S.8. Proof of (S.28) and (S.29) (for the case of $\overline{\alpha}_n = n^{\overline{\kappa}}$).

Since $\overline{\kappa} > 0$, $\overline{\alpha}_n = n^{\overline{\kappa}} > \alpha_0$ for all sufficiently large n. By Lemma S.4, we have $\widetilde{\theta}_{\overline{\alpha}_n} \geq \widetilde{\theta}_{\alpha_0}$ and $\widetilde{\theta}^{(1)}_{\overline{\alpha}_n} \geq \widetilde{\theta}^{(1)}_{\alpha_0}$. By the decomposition of $\widetilde{\theta}_\alpha$ in (S.10) of Lemma S.5 and the fact that $0 \leq \widetilde{\theta}^{(2)}_\alpha \leq \widetilde{\theta}^{(1)}_\alpha, \widetilde{\theta}^{(3)}_\alpha \geq 0$, we can rewrite the difference inside the probability in (S.28) as

$$
\begin{aligned}
0 &\leq \widetilde{\theta}_{\overline{\alpha}_n} - \widetilde{\theta}_{\alpha_0}\\
&= \widetilde{\theta}^{(1)}_{\overline{\alpha}_n} - \widetilde{\theta}^{(2)}_{\overline{\alpha}_n} + \widetilde{\theta}^{(3)}_{\overline{\alpha}_n} - \widetilde{\theta}^{(1)}_{\alpha_0} + \widetilde{\theta}^{(2)}_{\alpha_0} - \widetilde{\theta}^{(3)}_{\alpha_0}\\
&\leq \widetilde{\theta}^{(1)}_{\overline{\alpha}_n} - \widetilde{\theta}^{(1)}_{\alpha_0} + \widetilde{\theta}^{(3)}_{\overline{\alpha}_n} + \widetilde{\theta}^{(2)}_{\alpha_0}. \qquad (S.32)
\end{aligned}
$$

According to the definition of $\lambda_{k,n}(\alpha)$ $(k = 1, \ldots, n)$ in (S.54), we have that for any $\alpha > 0$,

$$
\alpha^{2\nu} R_\alpha^{-1} = \theta_0 \sigma^{-2} R_\alpha^{-1} = \theta_0 U_\alpha \Lambda_\alpha^{-1} U_\alpha^\top, \quad \alpha_0^{2\nu} R_{\alpha_0}^{-1} = \theta_0 \sigma_0^{-2} R_{\alpha_0}^{-1} = \theta_0 U_\alpha U_\alpha^\top, \qquad (S.33)
$$

where $\Lambda_\alpha = \mathrm{diag}\{\lambda_{k,n}(\alpha) : k = 1, \ldots, n\}$. Similar to the proof of Lemma S.6, for any $\alpha > 0$, we define $Z_n(\alpha) = (Z_{1,n}(\alpha), \ldots, Z_{n,n}(\alpha))^\top = U_\alpha^\top X_n$. Since $X_n \sim \mathcal{N}(0_n, \sigma_0^2 R_{\alpha_0})$, we have $Z_n(\alpha) \sim \mathcal{N}(0_n, I_n)$ for any $\alpha > 0$.

Then it follows that for n sufficiently large,

$$\begin{aligned}
\sqrt{n}\left(\widetilde{\theta}^{(1)}_{\overline{\alpha}_n} - \widetilde{\theta}^{(1)}_{\alpha_0}\right) &= \frac{\sqrt{n}}{n-p} X_n^\top \left(\overline{\alpha}_n^{2\nu} R_{\overline{\alpha}_n}^{-1} - \alpha_0^{2\nu} R_{\alpha_0}^{-1}\right) X_n \\
&= \frac{\sqrt{n}\theta_0}{n-p} X_n^\top U_{\overline{\alpha}_n} \left(\Lambda_{\overline{\alpha}_n}^{-1} - I_n\right) U_{\overline{\alpha}_n}^\top X_n \\
&= \frac{\sqrt{n}\theta_0}{n-p} \sum_{i=1}^n \left|\lambda_{i,n}(\overline{\alpha}_n)^{-1} - 1\right| Z_{i,n}(\overline{\alpha}_n)^2. \qquad \text{(S.34)}
\end{aligned}$$

$\sqrt{n}/(n-p) \leq 2n^{-1/2}$ for all large n. Now we apply Lemma S.19 and Lemma S.20 to (S.34), with $z = 4\log^2 n$, $Z_i = Z_{i,n}(\overline{\alpha}_n)$, $w_i = w_i(\overline{\alpha}_n) = \left|\lambda_{i,n}(\overline{\alpha}_n)^{-1} - 1\right|/\sqrt{n}$, to obtain that

$$\begin{aligned}
&\Pr\left(\sqrt{n}\left|\widetilde{\theta}^{(1)}_{\overline{\alpha}_n} - \widetilde{\theta}^{(1)}_{\alpha_0}\right| > \frac{\theta_0}{4} n^{-\tau}\right) \\
\leq\ &\Pr\Big(\frac{1}{\sqrt{n}} \sum_{i=1}^n \left|\lambda_{i,n}(\overline{\alpha}_n)^{-1} - 1\right| Z_{i,n}(\overline{\alpha}_n)^2 \\
&\qquad > \|w(\overline{\alpha}_n)\|_1 + 4\|w(\overline{\alpha}_n)\| \log n + 8\|w(\overline{\alpha}_n)\|_\infty \log^2 n\Big) \\
\leq\ &\exp(-4\log^2 n). \qquad \text{(S.35)}
\end{aligned}$$

This proves (S.29).

We combine (S.32), (S.35) with (S.21) from Lemma S.7 and (S.12) from Lemma S.6 to obtain that for all sufficiently large n,

$$\begin{aligned}
&\Pr\left(\sqrt{n}\left(\widetilde{\theta}_{\overline{\alpha}_n} - \widetilde{\theta}_{\alpha_0}\right) > \frac{\theta_0}{2} n^{-\tau}\right) \\
\leq\ &\Pr\left(\sqrt{n}\left(\widetilde{\theta}^{(1)}_{\overline{\alpha}_n} - \widetilde{\theta}^{(1)}_{\alpha_0} + \widetilde{\theta}^{(3)}_{\overline{\alpha}_n} + \widetilde{\theta}^{(2)}_{\alpha_0}\right) > \frac{\theta_0}{2} n^{-\tau}\right) \\
\leq\ &\Pr\left(\sqrt{n}\left(\widetilde{\theta}^{(1)}_{\overline{\alpha}_n} - \widetilde{\theta}^{(1)}_{\alpha_0}\right) > \frac{\theta_0}{4} n^{-\tau}\right) + \Pr\left(\sqrt{n}\widetilde{\theta}^{(3)}_{\overline{\alpha}_n} > \frac{\theta_0}{16} n^{-\tau}\right) + \Pr\left(\sqrt{n}\widetilde{\theta}^{(2)}_{\alpha_0} > \frac{\theta_0}{16} n^{-\tau}\right) \\
\leq\ &\exp(-4\log^2 n) + 2\exp(-16\log^2 n) < 2\exp(-4\log^2 n), \qquad \text{(S.36)}
\end{aligned}$$

which proves (S.28).

Proof of (S.30) and (S.31) (for the case of $\underline{\alpha}_n = n^{\underline{\kappa}}$).

The proof for the case of $\underline{\alpha}_n = n^{\underline{\kappa}}$ is similar to the previous case of $\overline{\alpha}_n = n^{\overline{\kappa}}$. First by Lemma S.4, we have $\widetilde{\theta}_{\underline{\alpha}_n} \leq \widetilde{\theta}_{\alpha_0}$ and $\widetilde{\theta}^{(1)}_{\underline{\alpha}_n} \leq \widetilde{\theta}^{(1)}_{\alpha_0}$ for large n. By the decomposition of $\widetilde{\theta}_\alpha$ in (S.10) of Lemma S.5 and the fact that $0 \leq \widetilde{\theta}^{(2)}_\alpha \leq \widetilde{\theta}^{(1)}_\alpha, \widetilde{\theta}^{(3)}_\alpha \geq 0$, we can rewrite the difference inside the probability in (S.30) as

$$\begin{aligned}
0 &\leq \widetilde{\theta}_{\alpha_0} - \widetilde{\theta}_{\underline{\alpha}_n} \\
&= \widetilde{\theta}^{(1)}_{\alpha_0} - \widetilde{\theta}^{(2)}_{\alpha_0} + \widetilde{\theta}^{(3)}_{\alpha_0} - \widetilde{\theta}^{(1)}_{\underline{\alpha}_n} + \widetilde{\theta}^{(2)}_{\underline{\alpha}_n} - \widetilde{\theta}^{(3)}_{\underline{\alpha}_n} \\
&\leq \widetilde{\theta}^{(1)}_{\alpha_0} - \widetilde{\theta}^{(1)}_{\underline{\alpha}_n} + \widetilde{\theta}^{(3)}_{\alpha_0} + \widetilde{\theta}^{(2)}_{\underline{\alpha}_n}. \qquad \text{(S.37)}
\end{aligned}$$

Using Lemma S.19 and Lemma S.20 with $z = 4\log^2 n$, $Z_i = Z_{i,n}(\underline{\alpha}_n)$, $w_i = w_i(\underline{\alpha}_n) = \left|\lambda_{i,n}(\underline{\alpha}_n)^{-1} - 1\right|/\sqrt{n}$, we have that

$$\Pr\left(\sqrt{n}\left|\widetilde{\theta}^{(1)}_{\underline{\alpha}_n} - \widetilde{\theta}^{(1)}_{\alpha_0}\right| > \frac{\theta_0}{2} n^{-\tau}\right)$$

$$
\begin{aligned}
&\leq \Pr\Big(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left| \lambda_{i,n}(\underline{\alpha}_n)^{-1} - 1 \right| Z_{i,n}(\overline{\alpha}_n)^2 \\
&\qquad > \|w(\underline{\alpha}_n)\|_1 + 4\|w(\underline{\alpha}_n)\| \log n + 8\|w(\underline{\alpha}_n)\|_\infty \log^2 n \Big) \\
&\leq \exp(-4\log^2 n). \qquad \text{(S.38)}
\end{aligned}
$$

This proves (S.31).

We then combine (S.37), (S.38) with (S.19) from Lemma S.7 and (S.13) from Lemma S.6 to obtain that for all sufficiently large n,

$$
\begin{aligned}
&\Pr\left(\sqrt{n}\left(\widetilde{\theta}_{\alpha_0} - \widetilde{\theta}_{\underline{\alpha}_n} \right) > \frac{\theta_0}{2} n^{-\tau} \right) \\
\leq\ &\Pr\left(\sqrt{n}\left(\widetilde{\theta}^{(1)}_{\alpha_0} - \widetilde{\theta}^{(1)}_{\underline{\alpha}_n} + \widetilde{\theta}^{(3)}_{\alpha_0} + \widetilde{\theta}^{(2)}_{\underline{\alpha}_n} \right) > \frac{\theta_0}{2} n^{-\tau} \right) \\
\leq\ &\Pr\left(\sqrt{n}\left(\widetilde{\theta}^{(1)}_{\alpha_0} - \widetilde{\theta}^{(1)}_{\underline{\alpha}_n} \right) > \frac{\theta_0}{4} n^{-\tau} \right) + \Pr\left(\sqrt{n}\widetilde{\theta}^{(3)}_{\alpha_0} > \frac{\theta_0}{16} n^{-\tau} \right) + \Pr\left(\sqrt{n}\widetilde{\theta}^{(2)}_{\underline{\alpha}_n} > \frac{\theta_0}{16} n^{-\tau} \right) \\
\leq\ &\exp(-4\log^2 n) + 2\exp(-16\log^2 n) < 2\exp(-4\log^2 n), \qquad \text{(S.39)}
\end{aligned}
$$

which proves (S.30). □

We restate and strengthen the uniform convergence in Part (ii) of Lemma 2.2 in the main text as the following lemma. The inequality in Part (ii) of Lemma 2.2 is implied by (S.40) below.

Lemma S.9 (Uniform Convergence of $\widetilde{\theta}_\alpha$ in Lemma 2.2 in the Main Text)**.** *Suppose that $d \in \{1, 2, 3\}$. For $\widetilde{\theta}_\alpha$ defined in* (S.1) *and N_3' defined in Lemma S.8, for all $n > N_3'$,*

$$
\Pr\left(\sup_{\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]} \sqrt{n} \left| \widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0} \right| \leq \frac{\theta_0}{2} n^{-\tau} \right) \geq 1 - 4\exp(-4\log^2 n), \qquad \text{(S.40)}
$$

$$
\Pr\left(\sup_{\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]} \sqrt{n} \left| \widetilde{\theta}^{(1)}_\alpha - \widetilde{\theta}^{(1)}_{\alpha_0} \right| \leq \frac{\theta_0}{4} n^{-\tau} \right) \geq 1 - 2\exp(-4\log^2 n), \qquad \text{(S.41)}
$$

where $\tau, \underline{\alpha}_n, \overline{\alpha}_n$ are as defined the same as in (S.9).

Proof of Lemma S.9. From Lemma S.4, we have that $\widetilde{\theta}_\alpha$ and $\widetilde{\theta}^{(1)}_\alpha$ are both non-decreasing in α. Therefore,

$$
\begin{aligned}
\sup_{\alpha \in [\underline{\alpha}_n, \alpha_0]} \left| \widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0} \right| &= \widetilde{\theta}_{\alpha_0} - \widetilde{\theta}_{\underline{\alpha}_n}, \\
\sup_{\alpha \in [\alpha_0, \overline{\alpha}_n]} \left| \widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0} \right| &= \widetilde{\theta}_{\overline{\alpha}_n} - \widetilde{\theta}_{\alpha_0}, \\
\sup_{\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]} \left| \widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0} \right| &= \max\left(\widetilde{\theta}_{\alpha_0} - \widetilde{\theta}_{\underline{\alpha}_n}, \widetilde{\theta}_{\overline{\alpha}_n} - \widetilde{\theta}_{\alpha_0} \right), \\
\sup_{\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]} \left| \widetilde{\theta}^{(1)}_\alpha - \widetilde{\theta}^{(1)}_{\alpha_0} \right| &= \max\left(\widetilde{\theta}^{(1)}_{\alpha_0} - \widetilde{\theta}^{(1)}_{\underline{\alpha}_n}, \widetilde{\theta}^{(1)}_{\overline{\alpha}_n} - \widetilde{\theta}^{(1)}_{\alpha_0} \right).
\end{aligned}
$$

We can then combine (S.28) and (S.30) from Lemma S.8 to obtain that for all $n > N_3'$,

$$
\begin{aligned}
&\Pr\left(\sup_{\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]} \sqrt{n} \left| \widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0} \right| > \frac{\theta_0}{2} n^{-\tau} \right) \\
= &\Pr\left(\sqrt{n}\left(\widetilde{\theta}_{\alpha_0} - \widetilde{\theta}_{\underline{\alpha}_n} \right) > \frac{\theta_0}{2} n^{-\tau} \text{ or } \sqrt{n}\left(\widetilde{\theta}_{\overline{\alpha}_n} - \widetilde{\theta}_{\alpha_0} \right) > \frac{\theta_0}{2} n^{-\tau} \right)
\end{aligned}
$$

$$\le \Pr\left(\sqrt{n}\left(\widetilde{\theta}_{\alpha_0} - \widetilde{\theta}_{\underline{\alpha}_n}\right) > \frac{\theta_0}{2} n^{-\tau}\right) + \Pr\left(\sqrt{n}\left(\widetilde{\theta}_{\overline{\alpha}_n} - \widetilde{\theta}_{\alpha_0}\right) > \frac{\theta_0}{2} n^{-\tau}\right)$$
$$\le 4\mathrm{e}^{-4\log^2 n}.$$

The inequality (S.41) follows similarly using a union bound from (S.31) and (S.29) in Lemma S.8. □

In the next lemma, we prove the asymptotic normality of $\widetilde{\theta}_\alpha$ for a fixed $\alpha > 0$ in Theorem 2.1 in the main text. We also bound the tail probability of $\left|\widetilde{\theta}_{\alpha_0} - \theta_0\right|$.

Lemma S.10. *For $d \in \mathbb{Z}^+$ and $\nu \in \mathbb{R}^+$, there exists a large integer N_4' that only depends on $\nu, d, T, \beta_0, \theta_0, \alpha_0$ and the $\mathcal{W}_2^{\nu+d/2}(\mathcal{S})$ norms of $\mathrm{m}_1(\cdot), \ldots, \mathrm{m}_p(\cdot)$, such that for all $n > N_4'$,*

$$\Pr\left(\sqrt{n}\left|\widetilde{\theta}_{\alpha_0} - \theta_0\right| \le 5\theta_0 \log n\right) \ge 1 - 3\exp(-4\log^2 n). \tag{S.42}$$

Furthermore, for $d \in \{1, 2, 3\}$ and $\nu \in \mathbb{R}^+$, for any fixed $\alpha > 0$, as $n \to \infty$,

$$\sqrt{n}\left(\widetilde{\theta}_\alpha - \theta_0\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, 2\theta_0^2\right). \tag{S.43}$$

Proof of Lemma S.10. Let $W_n = (W_{1,n}, \ldots, W_{n,n})^\top = \sigma_0^{-1} R_{\alpha_0}^{-1/2} X_n \sim \mathcal{N}(0_n, I_n)$. Using the decomposition in (S.10), we have

$$\sqrt{n}\left(\widetilde{\theta}_{\alpha_0} - \theta_0\right) = \sqrt{n}\left(\widetilde{\theta}_{\alpha_0}^{(1)} - \theta_0\right) - \sqrt{n}\widetilde{\theta}_{\alpha_0}^{(2)} + \sqrt{n}\widetilde{\theta}_{\alpha_0}^{(3)}.$$

Since $\widetilde{\theta}_{\alpha_0}^{(1)} = \alpha_0^{2\nu} X_n^\top R_{\alpha_0}^{-1} X_n/(n-p) = \theta_0 W_n^\top W_n/(n-p)$, by the central limit theorem for χ_1^2 random variables, we have that as $n \to \infty$,

$$\sqrt{n}\left(\widetilde{\theta}_{\alpha_0}^{(1)} - \theta_0\right) = \sqrt{n}\theta_0\left(\frac{W_n^\top W_n}{n-p} - 1\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 2\theta_0^2). \tag{S.44}$$

The first inequality in Lemma S.19 with $Z_i = W_{i,n}$, $w_i = 1$ for $i = 1, \ldots, n$ and $z = 4\log^2 n$ implies that for all sufficiently large n,

$$\Pr\left(\sqrt{n}\left(\widetilde{\theta}_{\alpha_0}^{(1)} - \theta_0\right) > 4.5\theta_0 \log n\right) = \Pr\left(W_n^\top W_n > n - p + \frac{4.5(n-p)\log n}{\sqrt{n}}\right)$$
$$\le \Pr\left(W_n^\top W_n > n + 4\sqrt{n}\log n + 8\log^2 n\right) \le \exp(-4\log^2 n). \tag{S.45}$$

The second inequality in Lemma S.19 with $Z_i = W_{i,n}$, $w_i = 1$ for $i = 1, \ldots, n$ and $z = 4\log^2 n$ implies that for all sufficiently large n,

$$\Pr\left(\sqrt{n}\left(\widetilde{\theta}_{\alpha_0}^{(1)} - \theta_0\right) < -4.5\theta_0 \log n\right) = \Pr\left(W_n^\top W_n < n - p - \frac{4.5(n-p)\log n}{\sqrt{n}}\right)$$
$$\le \Pr\left(W_n^\top W_n < n - 4\sqrt{n}\log n\right) \le \exp(-4\log^2 n). \tag{S.46}$$

We combine (S.45), (S.46), (S.12) from Lemma S.6 and (S.19) from Lemma S.7 to obtain that for all sufficiently large n,

$$\Pr\left(\sqrt{n}\left|\widetilde{\theta}_{\alpha_0} - \theta_0\right| > 5\theta_0 \log n\right)$$
$$\le \Pr\left(\sqrt{n}\left|\widetilde{\theta}_{\alpha_0}^{(1)} - \theta_0\right| > 4.5\theta_0 \log n\right) + \Pr\left(\sqrt{n}\widetilde{\theta}_{\alpha_0}^{(2)} > \frac{\theta_0}{4}\log n\right) + \Pr\left(\sqrt{n}\widetilde{\theta}_{\alpha_0}^{(3)} > \frac{\theta_0}{4}\log n\right)$$

$$\leq \Pr\left(\sqrt{n}\left(\widetilde{\theta}^{(1)}_{\alpha_0} - \theta_0\right) > 4.5\theta_0 \log n\right) + \Pr\left(\sqrt{n}\left(\widetilde{\theta}^{(1)}_{\alpha_0} - \theta_0\right) < -4.5\theta_0 \log n\right)$$
$$+ \Pr\left(\sqrt{n}\widetilde{\theta}^{(2)}_{\alpha_0} > \theta_0 n^{-\tau}/16\right) + \Pr\left(\sqrt{n}\widetilde{\theta}^{(3)}_{\alpha_0} > \theta_0 n^{-\tau}/16\right)$$
$$\leq 2\exp(-4\log^2 n) + 2\exp(-16\log^2 n) < 3\exp(-4\log^2 n),$$

which has proved (S.42).

Now for (S.43), we notice that (S.12) from Lemma S.6 and (S.19) from Lemma S.7 imply that both $\sqrt{n}\widetilde{\theta}^{(2)}_{\alpha_0}$ and $\sqrt{n}\widetilde{\theta}^{(3)}_{\alpha_0}$ converge to zero in $P_{(\beta_0,\sigma_0^2,\alpha_0)}$-probability as $n \to \infty$. Therefore, we combine this with (S.44) and apply the Slutsky's theorem to obtain that as $n \to \infty$,

$$\sqrt{n}\left(\widetilde{\theta}_{\alpha_0} - \theta_0\right) = \sqrt{n}\left(\widetilde{\theta}^{(1)}_{\alpha_0} - \theta_0\right) - \sqrt{n}\widetilde{\theta}^{(2)}_{\alpha_0} + \sqrt{n}\widetilde{\theta}^{(3)}_{\alpha_0} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 2\theta_0^2). \tag{S.47}$$

Since $\alpha > 0$ is fixed, it will be eventually covered by the interval $[\underline{\alpha}_n, \overline{\alpha}_n]$ as $n \to \infty$. Therefore, by Lemma S.9, for any fixed $\alpha > 0$, $\sqrt{n}\left|\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}\right| \to 0$ in $P_{(\beta_0,\sigma_0^2,\alpha_0)}$-probability as $n \to \infty$. We combine this with (S.47) and apply the Slutsky's theorem again to conclude that as $n \to \infty$,

$$\sqrt{n}\left(\widetilde{\theta}_\alpha - \theta_0\right) = \sqrt{n}\left(\widetilde{\theta}_{\alpha_0} - \theta_0\right) + \sqrt{n}\left(\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 2\theta_0^2). \tag{S.48}$$

This completes the proof. □

S1.3 Auxiliary RKHS Theory

In this subsection, we present some auxiliary technical results on the reproducing kernel Hilbert space (RKHS) of Matérn kernels that are used to handle the regression functions $\mathrm{m}_1, \ldots, \mathrm{m}_p$. We define some concepts for a generic positive definite covariance function $K(\cdot,\cdot)$ on a fixed domain $\mathcal{S} = [0,T]^d$. Let $L_2(\mathcal{S})$ be the space of square integrable functions on $\mathcal{S}$, and $\mathcal{C}(\mathcal{S})$ be the space of continuous functions on $\mathcal{S}$. We assume that $K(\cdot,\cdot)$ is symmetric with $K(s,t) = K(t,s)$ for any $s,t \in \mathcal{S}$. The reproducing kernel Hilbert space (RKHS) associated with K, denoted by $\mathcal{H}_K$ (suppressing its dependence on the domain $\mathcal{S}$), can be defined to be the space endowed with an inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}_K}$ such that: (i) $K(s,\cdot) \in \mathcal{H}_K$ for each $s \in \mathcal{S}$; (ii) reproducing property: for any $f \in \mathcal{H}_K$, $\langle f, K(\cdot,s)\rangle_{\mathcal{H}_K} = f(s)$ for all $s \in \mathcal{S}$ (see Definition 6.1 of [Rasmussen and Williams, 2006]).

For shift-invariant kernels (including the isotropic Matérn in this paper), an alternative and equivalent definition of the RKHS norm is based on the spectral density of the kernel. Details can be found in [Wendland, 2005]. Let $\imath = \sqrt{-1}$ and $\mathcal{F}[f](\omega) = (2\pi)^{-d/2}\int_{\mathcal{S}} f(x)\mathrm{e}^{-\imath x^\top \omega}\mathrm{d}\omega$ for any $\omega \in \mathbb{R}^d$. If $K(\cdot,\cdot)$ is a shift-invariant kernel on $\mathcal{S}$, with $\Phi(s-s') \equiv K(s,s')$ for any $s,s' \in \mathcal{S}$, then Theorem 10.12 of [Wendland, 2005] has shown that the RKHS associated with K can be written as

$$\mathcal{H}_K = \Big\{f \in L_2(\mathcal{S}) \cap \mathcal{C}(\mathcal{S}) : \exists g \in L_2(\mathbb{R}^d) \cap \mathcal{C}(\mathbb{R}^d) \text{ such that } g\big|_{\mathcal{S}} = f,$$
$$\|f\|^2_{\mathcal{H}_K} = \|g\|^2_{\mathcal{H}_K} = (2\pi)^{-d/2}\int_{\mathbb{R}^d} \frac{|\mathcal{F}[g](\omega)|^2}{\mathcal{F}[\Phi](\omega)}\mathrm{d}\omega < \infty\Big\}, \tag{S.49}$$

where $g\big|_{\mathcal{S}}$ is the restriction of g to the domain $\mathcal{S}$. For ease of notation, we suppress the dependence on $\mathcal{S}$ in the notation $\mathcal{H}_K$.

In particular, for the isotropic Matérn covariance function $\sigma^2 K_{\alpha,\nu}$ as defined in (2) of the main text, we know that $\mathcal{F}[\sigma^2 K_{\alpha,\nu}](\omega) = \frac{2^{d/2}\Gamma(\nu+d/2)}{\Gamma(\nu)}\frac{\sigma^2\alpha^{2\nu}}{(\alpha^2+\|\omega\|^2)^{\nu+d/2}}$. So the RKHS associated with $\sigma^2 K_{\alpha,\nu}$ can be written as

$$\mathcal{H}_{\sigma^2 K_{\alpha,\nu}} = \Big\{f \in L_2(\mathcal{S}) \cap \mathcal{C}(\mathcal{S}) : \exists g \in L_2(\mathbb{R}^d) \cap \mathcal{C}(\mathbb{R}^d) \text{ such that } g\big|_{\mathcal{S}} = f,$$

$$\|f\|^2_{\mathcal{H}_{\sigma^2 K_{\alpha,\nu}}} = \frac{\Gamma(\nu)}{2^d\pi^{d/2}\Gamma(\nu+d/2)\sigma^2\alpha^{2\nu}} \int_{\mathbb{R}^d} (\alpha^2+\|\omega\|^2)^{\nu+d/2} |\mathcal{F}[g](\omega)|^2 \mathrm{d}\omega < \infty \Big\}, \tag{S.50}$$

Lemma S.11. *([Wendland, 2005] Corollary 10.48) For any fixed $(\sigma^2, \alpha) \in \mathbb{R}^+ \times \mathbb{R}^+$, $d \in \mathbb{Z}^+$, $\nu \in \mathbb{R}^+$, $\mathcal{H}_{\sigma^2 K_{\alpha,\nu}}$ is norm equivalent to the Sobolev space $\mathcal{W}_2^{\nu+d/2}(\mathcal{S})$. In other words, there exist constants $0 < c_1(\sigma,\alpha) \le c_2(\sigma,\alpha) < \infty$, such that for any $f \in \mathcal{H}_{\sigma^2 K_{\alpha,\nu}}$,*

$$c_1(\sigma,\alpha)\|f\|_{\mathcal{W}_2^{\nu+d/2}(\mathcal{S})} \le \|f\|_{\mathcal{H}_{\sigma^2 K_{\alpha,\nu}}} \le c_2(\sigma,\alpha)\|f\|_{\mathcal{W}_2^{\nu+d/2}(\mathcal{S})}.$$

Lemma S.12. *Suppose that $f \in \mathcal{H}_K$ for a covariance function $K(\cdot,\cdot)$ defined on the fixed domain $\mathcal{S} = [0,T]^d$. Let $\mathcal{S}_n = \{s_1,\ldots,s_n\}$ be a set of distinct points in $\mathcal{S}$, $f_n = (f(s_1),\ldots,f(s_n))^\top$, and $K(\mathcal{S}_n,\mathcal{S}_n)$ be the matrix with (i,j)-entry equal to $K(s_i,s_j)$, for $i,j = 1,\ldots,n$. Then $f_n^\top K(\mathcal{S}_n,\mathcal{S}_n)^{-1} f_n \le \|f\|^2_{\mathcal{H}_K}$.*

Proof of Lemma S.12. We denote the (i,j)-entry of the matrix $K(\mathcal{S}_n,\mathcal{S}_n)^{-1}$ by $\{K(\mathcal{S}_n,\mathcal{S}_n)^{-1}\}_{ij}$, for $i,j = 1,\ldots,n$. Let $K(\mathcal{S}_n,s) = (K(s_1,s),\ldots,K(s_n,s))^\top$ for any $s \in \mathcal{S}$. Because the function $K(s,\cdot) \in \mathcal{H}_K$ for any $s \in \mathcal{S}$ by the definition of RKHS, we have that the function $f_n^\top K(\mathcal{S}_n,\mathcal{S}_n)^{-1}K(\mathcal{S}_n,\cdot) \in \mathcal{H}_K$. For any $a = (a_1,\ldots,a_n)^\top \in \mathbb{R}^n$, the RKHS norm of the function $a^\top K(\mathcal{S}_n,\cdot)$ is

$$\left\|a^\top K(\mathcal{S}_n,\cdot)\right\|^2_{\mathcal{H}_K} = \sum_{i=1}^n\sum_{j=1}^n a_i a_j K(s_i,s_j).$$

Therefore, the RKHS norm of $f_n^\top K(\mathcal{S}_n,\mathcal{S}_n)^{-1}K(\mathcal{S}_n,\cdot)$ is

$$\begin{aligned}
&\left\|f_n^\top K(\mathcal{S}_n,\mathcal{S}_n)^{-1}K(\mathcal{S}_n,\cdot)\right\|^2_{\mathcal{H}_K} \\
&= \sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sum_{l=1}^n \big\{K(\mathcal{S}_n,\mathcal{S}_n)^{-1}\big\}_{ij}\big\{K(\mathcal{S}_n,\mathcal{S}_n)^{-1}\big\}_{kl}\cdot f(s_i)f(s_k)K(s_j,s_l) \\
&= \sum_{i=1}^n\sum_{k=1}^n f(s_i)f(s_k)\left\{\sum_{j=1}^n\sum_{l=1}^n \big\{K(\mathcal{S}_n,\mathcal{S}_n)^{-1}\big\}_{ij}\big\{K(\mathcal{S}_n,\mathcal{S}_n)^{-1}\big\}_{kl}K(s_j,s_l)\right\} \\
&\overset{(i)}{=} \sum_{i=1}^n\sum_{k=1}^n f(s_i)f(s_k)\big\{K(\mathcal{S}_n,\mathcal{S}_n)^{-1}\big\}_{ik} = f_n^\top K(\mathcal{S}_n,\mathcal{S}_n)^{-1}f_n,
\end{aligned} \tag{S.51}$$

where the equality (i) follows from the expression of (i,k)-entry in the matrix multiplication $K(\mathcal{S}_n,\mathcal{S}_n)^{-1}K(\mathcal{S}_n,\mathcal{S}_n)K(\mathcal{S}_n,\mathcal{S}_n)^{-1}$.

On the other hand, using the RKHS inner product and the fact that $f \in \mathcal{H}_K$, we have

$$\begin{aligned}
&f_n^\top K(\mathcal{S}_n,\mathcal{S}_n)^{-1}f_n = \sum_{i=1}^n\sum_{k=1}^n f(s_i)f(s_k)\big\{K(\mathcal{S}_n,\mathcal{S}_n)^{-1}\big\}_{ik} \\
&= \left\langle \sum_{i=1}^n\sum_{k=1}^n f(s_k)\big\{K(\mathcal{S}_n,\mathcal{S}_n)^{-1}\big\}_{ik}K(s_i,\cdot),\ \ f(\cdot)\right\rangle_{\mathcal{H}_K} \\
&\overset{(i)}{\le} \left\|\sum_{i=1}^n\sum_{k=1}^n f(s_k)\big\{K(\mathcal{S}_n,\mathcal{S}_n)^{-1}\big\}_{ik}K(s_i,\cdot)\right\|_{\mathcal{H}_k}\cdot\|f\|_{\mathcal{H}_K} \\
&= \left\|f_n^\top K(\mathcal{S}_n,\mathcal{S}_n)^{-1}K(\mathcal{S}_n,\cdot)\right\|_{\mathcal{H}_k}\cdot\|f\|_{\mathcal{H}_K} \\
&\overset{(ii)}{=} \sqrt{f_n^\top K(\mathcal{S}_n,\mathcal{S}_n)^{-1}f_n}\cdot\|f\|_{\mathcal{H}_K},
\end{aligned} \tag{S.52}$$

where the inequality (i) follows from $\langle f_1, f_2\rangle_{\mathcal{H}_K} \leq \|f_1\|_{\mathcal{H}_K}\|f_2\|_{\mathcal{H}_K}$ for any $f_1, f_2 \in \mathcal{H}_K$, and the equality (ii) follows from (S.51). Therefore, we conclude from the left-hand side and the right-hand side of (S.52) that $\sqrt{f_n^\top K(\mathcal{S}_n,\mathcal{S}_n)^{-1} f_n} \leq \|f\|_{\mathcal{H}_K}$, or $f_n^\top K(\mathcal{S}_n,\mathcal{S}_n)^{-1} f_n \leq \|f\|^2_{\mathcal{H}_K}$. □

Lemma S.13. *For any $f \in \mathcal{W}_2^{\nu+d/2}(\mathcal{S})$, any $d \in \mathbb{Z}^+$, $\nu \in \mathbb{R}^+$, $\alpha \in \mathbb{R}^+$,*

$$\|f\|_{\mathcal{H}_{(\theta_0/\alpha^{2\nu})K_{\alpha,\nu}}} \leq \max\left\{\left(\frac{\alpha}{\alpha_0}\right)^{\nu+d/2}, 1\right\} \cdot \|f\|_{\mathcal{H}_{\sigma_0^2 K_{\alpha_0,\nu}}}. \tag{S.53}$$

Proof of Lemma S.13. From (S.50), one can see that for any function $f \in \mathcal{W}_2^{\nu+d/2}(\mathcal{S})$, for any $\alpha > 0$,

$$\begin{aligned}
&\|f\|^2_{\mathcal{H}_{(\theta_0/\alpha^{2\nu})K_{\alpha,\nu}}} \\
&= \frac{\Gamma(\nu)}{2^d\pi^{d/2}\Gamma(\nu+d/2)\theta_0}\int_{\mathbb{R}^d}(\alpha^2+\|\omega\|^2)^{\nu+d/2}|\mathcal{F}[f](\omega)|^2 \mathrm{d}\omega \\
&\leq \sup_{\omega\in\mathbb{R}^d}\left(\frac{\alpha^2+\|\omega\|^2}{\alpha_0^2+\|\omega\|^2}\right)^{\nu+d/2} \cdot \frac{\Gamma(\nu)}{2^d\pi^{d/2}\Gamma(\nu+d/2)\theta_0}\int_{\mathbb{R}^d}(\alpha_0^2+\|\omega\|^2)^{\nu+d/2}|\mathcal{F}[f](\omega)|^2 \mathrm{d}\omega \\
&\leq \sup_{\omega\in\mathbb{R}^d}\left(\frac{\alpha^2+\|\omega\|^2}{\alpha_0^2+\|\omega\|^2}\right)^{\nu+d/2} \cdot \|f\|^2_{\mathcal{H}_{(\theta_0/\alpha_0^{2\nu})K_{\alpha_0,\nu}}} \\
&\leq \max\left\{\left(\frac{\alpha}{\alpha_0}\right)^{2(\nu+d/2)}, 1\right\} \cdot \|f\|^2_{\mathcal{H}_{\sigma_0^2 K_{\alpha_0,\nu}}}.
\end{aligned}$$

Hence the conclusion follows. □

S1.4 Auxiliary Results on Spectral Analysis of Matérn Covariance Functions

In this subsection, we present a series of technical lemmas on the spectral analysis of Matérn covariance functions. For a detailed background theory on the equivalence of Gaussian measures on Hilbert spaces, we refer the interested readers to Chapter III of [Ibragimov and Rozanov, 1978] and Chapter 4 of [Stein, 1999]. Our Lemmas S.16, S.17, and S.18 below will use similar techniques in Section 4 of [Wang and Loh, 2011]. The key difference is that the theory of [Wang and Loh, 2011] only works for a *fixed and known* value of range parameter α. As a result, all those probabilistic error bounds in [Wang and Loh, 2011] do not depend on α and cannot be directly applied to varying values of α drawn from a posterior distribution. In contrast, our lemmas below will make all error bounds explicitly dependent on the value of α. This is made possible by using our new results on Matérn spectral densities in Lemma S.15, which is not shown in [Wang and Loh, 2011]. These lemmas will be used for showing the *uniform convergence* of $\left|\widehat{\theta}_\alpha - \widehat{\theta}_{\alpha_0}\right|$ over a large range of values of α as proved in Lemma S.9, which is fundamental for deriving the limiting joint posterior distribution of (θ, α).

We first consider the case when $\sigma^2\alpha^{2\nu} = \theta_0 = \sigma_0^2\alpha_0^{2\nu}$. If $d \in \{1,2,3\}$, then the two Gaussian measures $\mathrm{GP}(0, \sigma^2 K_{\alpha,\nu})$ and $\mathrm{GP}(0, \sigma_0^2 K_{\alpha_0,\nu})$ are equivalent ([Zhang, 2004]). For a generic $\alpha > 0$, we consider the two Matérn covariance matrices $\sigma_0^2 R_{\alpha_0}$ and $\sigma^2 R_\alpha$. We have the following lemma.

Lemma S.14. *For any pair $(\sigma, \alpha) \in \mathbb{R}^+ \times \mathbb{R}^+$ that satisfies $\sigma^2\alpha^{2\nu} = \theta_0 = \sigma_0^2\alpha_0^{2\nu}$, for all $d \in \mathbb{Z}^+$ and $\nu \in \mathbb{R}^+$, there exists an $n \times n$ invertible matrix U_α that depends on $\alpha, \alpha_0, \sigma_0^2, \nu$, such that*

$$\sigma_0^2 U_\alpha^\top R_{\alpha_0} U_\alpha = I_n, \qquad \sigma^2 U_\alpha^\top R_\alpha U_\alpha = \mathrm{diag}\{\lambda_{k,n}(\alpha): k = 1,\ldots,n\} \equiv \Lambda_\alpha, \tag{S.54}$$

where I_n is the $n \times n$ identity matrix, and $\{\lambda_{k,n}(\alpha), k = 1,\ldots,n\}$ are the positive diagonal entries of the diagonal matrix Λ_α.

Proof of Lemma S.14. The existence of such an invertible U_α is guaranteed by Theorem 7.6.4 and Corollary 7.6.5 on page 465–466 of [Horn and Johnson, 1985]. For completeness, we directly prove the existence of such an invertible matrix in the following general claim.

Claim: Suppose that A and B are two generic $n \times n$ symmetric positive definite matrices. Then there always exists an invertible matrix U, such that

$$U^\top A U = I_n, \qquad U^\top B U = \Lambda, \tag{S.55}$$

where I_n is the $n \times n$ identity matrix and Λ is an $n \times n$ diagonal matrix whose diagonal entries are all positive.

Proof of the Claim: Since B is symmetric positive definite, let $B = LL^\top$ be the Cholesky decomposition of B, where L is an $n \times n$ lower triangular matrix with all positive diagonal entries and L is invertible. Let $G = L^{-1} A L^{-\top}$. Then obviously G is also a symmetric positive definite matrix with $G^\top = G$. Suppose that G has the spectral decomposition $G = PDP^{-1}$ where P is an $n \times n$ orthogonal matrix ($P^{-1} = P^\top$) and D is a $n \times n$ diagonal matrix whose diagonal entries are all eigenvalues of G and they are all positive. Then $P^\top G P = D$. We let $U = L^{-\top} P D^{-1/2}$. It follows that

$$\begin{aligned} U^\top A U &= D^{-1/2} P^\top L^{-1} A L^{-\top} P D^{-1/2} \\ &= D^{-1/2} P^\top G P D^{-1/2} = D^{-1/2} D D^{-1/2} = I_n, \\ U^\top B U &= D^{-1/2} P^\top L^{-1} B L^{-\top} P D^{-1/2} \\ &= D^{-1/2} P^\top L^{-1} L L^\top L^{-\top} P D^{-1/2} = D^{-1/2} P^\top P D^{-1/2} = D^{-1}. \end{aligned}$$

We set $\Lambda = D^{-1}$ which is an $n \times n$ diagonal matrix whose diagonal entries are all positive. This proves the claim.

Based on the claim, if we set $A = \sigma_0^2 R_{\alpha_0}$ and $B = \sigma^2 R_\alpha$, then we can find an invertible matrix U such that (S.55) holds. Because $\sigma^2\alpha^{2\nu} = \theta_0 = \sigma_0^2 \alpha_0^{2\nu}$, and $\sigma_0^2, \alpha_0, \nu$ are assumed to be fixed numbers, we can see that U only changes with α and we can write it as U_α. Similarly, we write Λ_α to highlight its dependence on α. Correspondingly, we have $\sigma_0^2 U_\alpha^\top R_{\alpha_0} U_\alpha = I_n$ and $\sigma^2 U_\alpha^\top R_\alpha U_\alpha = \mathrm{diag}\{\lambda_{k,n}(\alpha) : k = 1, \ldots, n\} \equiv \Lambda_\alpha$. This proves Lemma S.14. $\square$

Let $\imath = \sqrt{-1}$. For $\omega \in \mathbb{R}^d$, let

$$\begin{aligned} f_{\sigma,\alpha}(\omega) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \mathrm{e}^{-\imath \omega^\top x} \sigma^2 K_{\alpha,\nu}(x) \mathrm{d}x \\ &= \frac{\Gamma(\nu + d/2)}{\Gamma(\nu)} \cdot \frac{\sigma^2 \alpha^{2\nu}}{\pi^{d/2} \left(\alpha^2 + \|\omega\|^2\right)^{\nu + d/2}}, \end{aligned} \tag{S.56}$$

be the isotropic spectral density of the Gaussian process with isotropic Matérn covariance function defined in (2) of the main text. For any given pair (σ, α), let $\|\psi\|_{f_{\sigma,\alpha}}^2 = \langle \psi, \psi \rangle_{f_{\sigma,\alpha}} = \int_{\mathbb{R}^d} |\psi(\omega)|^2 f_{\sigma,\alpha}(\omega) \mathrm{d}\omega$ be the norm of a generic function ψ in the Hilbert space $L_2(f_{\sigma,\alpha})$, with inner product $\langle \psi_1, \psi_2 \rangle_{f_{\sigma,\alpha}} = \int_{\mathbb{R}^d} \psi_1(\omega) \overline{\psi_2(\omega)} f_{\sigma,\alpha}(\omega) \mathrm{d}\omega$ for any $\psi_1, \psi_2 \in L_2(f_{\sigma,\alpha})$.

According to the spectral analysis in Section 4 of [Wang and Loh, 2011], using the same notation as theirs, for any given pair (σ, α) that satisfies $\sigma^2 \alpha^{2\nu} = \theta_0 = \sigma_0^2 \alpha_0^{2\nu}$, there exist orthonormal basis functions $\psi_1, \ldots, \psi_n \in L_2(f_{\sigma_0,\alpha_0})$ such that for any $j, k \in \{1, \ldots, n\}$,

$$\langle \psi_j, \psi_k \rangle_{f_{\sigma_0,\alpha_0}} = \mathcal{I}(j = k), \qquad \langle \psi_j, \psi_k \rangle_{f_{\sigma,\alpha}} = \lambda_{j,n}(\alpha)\, \mathcal{I}(j = k), \tag{S.57}$$

where $\mathcal{I}(\cdot)$ is the indicator function.

We prove the following lemma for the spectral density $f_{\sigma,\alpha}$ and the sequence $\{\lambda_{k,n}(\alpha), k = 1, \ldots, n\}$.

Lemma S.15. *Suppose that $d \in \mathbb{Z}^+$ and $\nu \in \mathbb{R}^+$. For any pair $(\sigma, \alpha) \in \mathbb{R}^+ \times \mathbb{R}^+$ that satisfies $\sigma^2\alpha^{2\nu} = \theta_0 = \sigma_0^2\alpha_0^{2\nu}$, and for all $\omega \in \mathbb{R}^d$, the following relations hold:*

$$\min\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d}, 1\right\} \le \frac{f_{\sigma,\alpha}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)} \le \max\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d}, 1\right\}, \tag{S.58}$$

$$\left|\frac{f_{\sigma,\alpha}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)} - 1\right| \le \frac{(2\nu+d)\max(\alpha_0^2,\alpha^2)\max\left(\alpha_0^{2\nu+d-2}, \alpha^{2\nu+d-2}\right)}{\alpha^{2\nu+d-2}(\alpha^2+\|\omega\|^2)}, \tag{S.59}$$

$$\lambda_{k,n}(\alpha) \le \max\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d}, 1\right\}, \tag{S.60}$$

$$\lambda_{k,n}(\alpha) \ge \min\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d}, 1\right\}, \tag{S.61}$$

for all $k = 1, \dots, n$.

Proof of Lemma S.15. For (S.58), when $\sigma^2\alpha^{2\nu} = \theta_0$, we have that

$$\frac{f_{\sigma,\alpha}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)} = \left(\frac{\alpha_0^2+\|\omega\|^2}{\alpha^2+\|\omega\|^2}\right)^{\nu+d/2}.$$

If $\alpha \ge \alpha_0$, then this ratio is an increasing function in $\|\omega\|$, which implies that $f_{\sigma,\alpha}(\omega)/f_{\sigma_0,\alpha_0}(\omega) \le 1$ (attained when $\|\omega\| \to +\infty$), and $f_{\sigma,\alpha}(\omega)/f_{\sigma_0,\alpha_0}(\omega) \ge (\alpha_0/\alpha)^{2\nu+d}$ (attained when $\|\omega\| \to 0$). The case of $\alpha < \alpha_0$ follows similarly. (S.58) summarizes the two cases.

For (S.59), if $\nu + d/2 \ge 1$, then using a first order Taylor expansion, we have that

$$\begin{aligned}
&\left|\frac{f_{\sigma,\alpha}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)} - 1\right| = \left|\frac{\left(\alpha_0^2+\|\omega\|^2\right)^{\nu+d/2}}{\left(\alpha^2+\|\omega\|^2\right)^{\nu+d/2}} - 1\right| \\
&\le \frac{(\nu+d/2)(\alpha_1^2+\|\omega\|^2)^{\nu+d/2-1}\cdot 2\alpha_1\cdot|\alpha-\alpha_0|}{\left(\alpha^2+\|\omega\|^2\right)^{\nu+d/2}} \\
&\le (2\nu+d)\max(\alpha_0^2,\alpha^2)\left(\frac{\max(\alpha_0,\alpha)^2+\|\omega\|^2}{\alpha^2+\|\omega\|^2}\right)^{\nu+d/2-1}\cdot\frac{1}{\alpha^2+\|\omega\|^2} \\
&\le \frac{(2\nu+d)\max(\alpha_0^2,\alpha^2)\max\left(\alpha_0^{2\nu+d-2},\alpha^{2\nu+d-2}\right)}{\alpha^{2\nu+d-2}(\alpha^2+\|\omega\|^2)},
\end{aligned} \tag{S.62}$$

where α_1 is a value between α_0 and α.

If $\nu + d/2 < 1$, then we have that

$$\begin{aligned}
&\left|\frac{f_{\sigma,\alpha}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)} - 1\right| = \left|\frac{\left(\alpha_0^2+\|\omega\|^2\right)^{\nu+d/2}}{\left(\alpha^2+\|\omega\|^2\right)^{\nu+d/2}} - 1\right| \\
&\le \frac{(\nu+d/2)(\alpha_1^2+\|\omega\|^2)^{\nu+d/2-1}\cdot 2\alpha_1\cdot|\alpha-\alpha_0|}{\left(\alpha^2+\|\omega\|^2\right)^{\nu+d/2}} \\
&\le (2\nu+d)\max(\alpha_0^2,\alpha^2)\left(\frac{\alpha^2+\|\omega\|^2}{\alpha_1^2+\|\omega\|^2}\right)^{1-(\nu+d/2)}\cdot\frac{1}{\alpha^2+\|\omega\|^2}.
\end{aligned} \tag{S.63}$$

In (S.63), if $\alpha \ge \alpha_1 \ge \alpha_0$, then the function $\left(\frac{\alpha^2+\|\omega\|^2}{\alpha_1^2+\|\omega\|^2}\right)^{1-(\nu+d/2)}$ is decreasing in $\|\omega\|^2$, so

$$\left(\frac{\alpha^2+\|\omega\|^2}{\alpha_1^2+\|\omega\|^2}\right)^{1-(\nu+d/2)} \le \left(\frac{\alpha}{\alpha_1}\right)^{2-(2\nu+d)} = \left(\frac{\alpha_1}{\alpha}\right)^{2\nu+d-2} \le \left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d-2}.$$

If $\alpha \le \alpha_1 \le \alpha_0$, then the function $\left(\frac{\alpha^2+\|\omega\|^2}{\alpha_1^2+\|\omega\|^2}\right)^{1-(\nu+d/2)}$ is increasing in $\|\omega\|^2$, so

$$\left(\frac{\alpha^2+\|\omega\|^2}{\alpha_1^2+\|\omega\|^2}\right)^{1-(\nu+d/2)} \le 1.$$

Considering both cases, then from (S.59), we can derive that

$$\begin{aligned}\left|\frac{f_{\sigma,\alpha}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)}-1\right| &\le (2\nu+d)\max(\alpha_0^2,\alpha^2)\left(\frac{\alpha^2+\|\omega\|^2}{\alpha_1^2+\|\omega\|^2}\right)^{1-(\nu+d/2)}\cdot\frac{1}{\alpha^2+\|\omega\|^2}\\ &\le \frac{(2\nu+d)\max(\alpha_0^2,\alpha^2)}{\alpha^2+\|\omega\|^2}\max\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d-2},1\right\}\\ &\le \frac{(2\nu+d)\max(\alpha_0^2,\alpha^2)\max\left(\alpha_0^{2\nu+d-2},\alpha^{2\nu+d-2}\right)}{\alpha^{2\nu+d-2}(\alpha^2+\|\omega\|^2)}. \end{aligned}\tag{S.64}$$

(S.62) for $\nu+d/2\ge 1$ and (S.64) for $\nu+d/2<1$ lead to (S.59).

For (S.60) and (S.61), we use the relation $\lambda_{k,n}(\alpha)=\int_{\mathbb{R}^d}|\psi_k(\omega)|^2 f_{\sigma_0,\alpha_0}(\omega)\cdot\frac{f_{\sigma,\alpha}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)}\mathrm{d}\omega$ for $k=1,\ldots,n$ and the bounds in (S.58) to obtain that

$$\begin{aligned}\lambda_{k,n}(\alpha) &\le \sup_{\omega\in\mathbb{R}^d}\frac{f_{\sigma,\alpha}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)}\cdot\int_{\mathbb{R}^d}|\psi_k(\omega)|^2 f_{\sigma_0,\alpha_0}(\omega)\mathrm{d}\omega \le \max\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d},1\right\},\\ \lambda_{k,n}(\alpha) &\ge \inf_{\omega\in\mathbb{R}^d}\frac{f_{\sigma,\alpha}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)}\cdot\int_{\mathbb{R}^d}|\psi_k(\omega)|^2 f_{\sigma_0,\alpha_0}(\omega)\mathrm{d}\omega \ge \min\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d},1\right\}. \end{aligned}\tag{S.65}$$

□

In the rest of this subsection, we focus exclusively on the case of $d\in\{1,2,3\}$. For any $a>0$, define $m_a=\lfloor a+d/2\rfloor+1$. For $\omega\in\mathbb{R}^d$, let

$$c_0(x)=\|x\|^{\frac{\nu+d/2}{2m_\nu}-d}\mathcal{I}(\|x\|\le 1), \tag{S.66}$$

$$\xi_0(\omega)=\int_{\mathbb{R}^d}\mathrm{e}^{-\imath x^\top w}c_0(x)\mathrm{d}x, \tag{S.67}$$

and $\xi_1(\omega)=\xi_0(\omega)^{2m_\nu}$ for all $\omega\in\mathbb{R}^d$. If $c_1=c_0*\ldots*c_0$ is the $2m_\nu$-fold convolution of the function c_0 with itself, then $\xi_1(\omega)$ is the Fourier transform of $c_1(x)$. Then Lemma 6 in [Wang and Loh, 2011] has proved that for $d=1,2,3$, $\xi_0(\omega)\asymp\|\omega\|^{-\frac{\nu+d/2}{2m_\nu}}$ as $\|\omega\|\to\infty$, which means that $\xi_1(\omega)\asymp\|\omega\|^{-(\nu+d/2)}$. This implies that if $\sigma^2\alpha^{2\nu}=\theta_0$, then $f_{\sigma,\alpha}(\omega)/\xi_1(\omega)\asymp 1$ as $\|\omega\|\to\infty$. In fact, using Lemma 6 in [Wang and Loh, 2011], we can prove the following lower and upper bound for his ratio.

Lemma S.16. *Suppose that $d\in\{1,2,3\}$ and $\nu\in\mathbb{R}^+$. For any pair $(\sigma,\alpha)\in\mathbb{R}^+\times\mathbb{R}^+$, the following holds for all $\omega\in\mathbb{R}^d$:*

$$\underline{c}_\xi\sigma^2\alpha^{2\nu}\min\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d},1\right\}\le\frac{f_{\sigma,\alpha}(\omega)}{\xi_1(\omega)^2}\le\overline{c}_\xi\sigma^2\alpha^{2\nu}\max\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d},1\right\}, \tag{S.68}$$

where $\underline{c}_\xi$ and $\overline{c}_\xi$ are two positive constants that only depend on d, ν and α_0.

Proof of Lemma S.16. Lemma 6 in [Wang and Loh, 2011] has proved that for $d=1,2,3$, $\xi_0(\omega)\asymp\|\omega\|^{-\frac{\nu+d/2}{2m_\nu}}$ as $\|\omega\|\to\infty$. This implies that there exists two positive absolute constants $\underline{c}_{\xi 0}$ and $\overline{c}_{\xi 0}$ that only depend on d, ν and α_0, such that

$$\underline{c}_{\xi 0}\le(\alpha_0^2+\|\omega\|^2)^{\frac{\nu+d/2}{4m_\nu}}\xi_0(\omega)\le\overline{c}_{\xi 0},$$

for all $\omega \in \mathbb{R}^d$. According to the definition of $\xi_1(\omega)$, this implies that

$$\underline{c}_{\xi 0}^{2m_\nu} \le (\alpha_0^2 + \|\omega\|^2)^{\frac{\nu+d/2}{2}} \xi_1(\omega) \le \overline{c}_{\xi 0}^{2m_\nu}, \tag{S.69}$$

for all $\omega \in \mathbb{R}^d$. Now, from the definition of $f_{\sigma,\alpha}$ in (S.56), we have that

$$\frac{f_{\sigma,\alpha}(\omega)}{\xi_1(\omega)^2} = \frac{\sigma^2\alpha^{2\nu}(\alpha_0^2 + \|\omega\|^2)^{\nu+d/2}}{\pi^{d/2}\,(\alpha^2 + \|\omega\|^2)^{\nu+d/2}} \cdot \frac{1}{(\alpha_0^2 + \|\omega\|^2)^{\nu+d/2}\xi_1(\omega)^2}. \tag{S.70}$$

Since

$$\min\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d}, 1\right\} \le \left(\frac{\alpha_0^2 + \|\omega\|^2}{\alpha^2 + \|\omega\|^2}\right)^{\nu+d/2} \le \max\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d}, 1\right\},$$

we have from (S.69) and (S.70) that

$$\frac{f_{\sigma,\alpha}(\omega)}{\xi_1(\omega)^2} \ge \frac{\sigma^2\alpha^{2\nu}}{\pi^{d/2}\overline{c}_{\xi 0}^{4m_\nu}} \min\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d}, 1\right\},$$
$$\frac{f_{\sigma,\alpha}(\omega)}{\xi_1(\omega)^2} \le \frac{\sigma^2\alpha^{2\nu}}{\pi^{d/2}\underline{c}_{\xi 0}^{4m_\nu}} \max\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d}, 1\right\}.$$

Finally, we let $\underline{c}_\xi = 1/(\pi^{d/2}\overline{c}_{\xi 0}^{4m_\nu})$ and $\overline{c}_\xi = 1/(\pi^{d/2}\underline{c}_{\xi 0}^{4m_\nu})$ and the conclusion follows. □

Now to proceed, we define the function

$$\eta(\omega) = \frac{f_{\sigma,\alpha}(\omega) - f_{\sigma_0,\alpha_0}(\omega)}{\xi_1(\omega)^2}, \quad \forall \omega \in \mathbb{R}^d. \tag{S.71}$$

Note that η depends on (σ, α), but we suppress the dependence for the ease of notation.

For any given pair $(\sigma, \alpha) \in \mathbb{R}^+ \times \mathbb{R}^+$, from (S.59) in Lemma S.15 and (S.68) in Lemma S.16, we have that

$$\begin{aligned}
&\int_{\mathbb{R}^d} \eta_n(\omega)^2 \mathrm{d}\omega = \int_{\mathbb{R}^d} \left\{\frac{f_{\sigma,\alpha}(\omega) - f_{\sigma_0,\alpha_0}(\omega)}{\xi_1(\omega)^2}\right\}^2 \mathrm{d}\omega \\
&= \int_{\mathbb{R}^d} \left\{\frac{f_{\sigma,\alpha}(\omega) - f_{\sigma_0,\alpha_0}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)}\right\}^2 \cdot \left(\frac{f_{\sigma_0,\alpha_0}(\omega)}{\xi_1(\omega)^2}\right)^2 \mathrm{d}\omega \\
&\le \sup_{\omega \in \mathbb{R}^d} \left(\frac{f_{\sigma_0,\alpha_0}(\omega)}{\xi_1(\omega)^2}\right)^2 \cdot \int_{\mathbb{R}^d} \left|\frac{f_{\sigma,\alpha}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)} - 1\right|^2 \mathrm{d}\omega \\
&\le \overline{c}_\xi^2 \theta_0^2 \cdot \int_{\mathbb{R}^d} \left\{\frac{(2\nu + d)\max(\alpha_0^2, \alpha^2)\max\left(\alpha_0^{2\nu+d-2}, \alpha^{2\nu+d-2}\right)}{\alpha^{2\nu+d-2}(\alpha^2 + \|\omega\|^2)}\right\}^2 \mathrm{d}\omega \\
&= \frac{\overline{c}_\xi^2 \theta_0^2 (2\nu + d)^2 \max(\alpha_0^4, \alpha^4)\max\left\{\alpha_0^{2(2\nu+d-2)}, \alpha^{2(2\nu+d-2)}\right\}}{\alpha^{2(2\nu+d-2)}} \\
&\quad \times \int_0^\infty \frac{r^{d-1}}{(\alpha^2 + r^2)^2} \mathrm{d}r < \infty,
\end{aligned} \tag{S.72}$$

where the last integral is finite because $\alpha > 0$ and $4 - (d - 1) \ge 2$ for $d = 1, 2, 3$. Therefore, we have shown that $\eta(\omega)$ is a square-integrable function of w. From the theory of Fourier transforms of $L_2(\mathbb{R}^d)$, there exists a square-integrable function $g : \mathbb{R}^d \to \mathbb{R}$ such that

$$\int_{\mathbb{R}^d} \{\eta(\omega) - \hat{g}_k(\omega)\}^2 \mathrm{d}\omega \to 0, \text{ as } k \to \infty,$$

where

$$\hat{g}_k(\omega) = \int_{\mathbb{R}^d} \mathrm{e}^{-\imath\omega^\top x} g(x)\,\mathcal{I}(\|x\|_\infty \le k)\mathrm{d}x. \tag{S.73}$$

Furthermore, for any fixed number $a > 0$ and $0 < b < \min(4-d, 2)$, we define the sequence $\varepsilon_n = n^{-1/(4a+2d+b)}$, such that $\varepsilon_n \to 0$ as $n \to \infty$. We define the following functions similar to Equations (35) and (36) in [Wang and Loh, 2011]. Let

$$\tilde{c}_0(x) = \|x\|^{\frac{a+d/2}{2m_a} - d}\mathcal{I}(\|x\| \le 1), \quad \forall x \in \mathbb{R}^d,$$

and $\tilde{c}_1(x) = c_0 * \ldots * c_0(x)$ be the $2m_a$-fold convolution of c_0 with itself. Let $C_q = \int_{\mathbb{R}^d} \tilde{c}_1(x)\mathrm{d}x$. Define the following functions

$$\begin{aligned}
\tilde{\xi}_0(\omega) &= \int_{\mathbb{R}^d} \mathrm{e}^{-\imath x^\top w}\tilde{c}_0(x)\mathrm{d}x, \quad \forall \omega \in \mathbb{R}^d, \\
\tilde{\xi}_1(\omega) &= \int_{\mathbb{R}^d} \mathrm{e}^{-\imath x^\top w}\tilde{c}_1(x)\mathrm{d}x = \tilde{\xi}_0(\omega)^{2m_a}, \quad \forall \omega \in \mathbb{R}^d, \\
q_n(x) &= \frac{1}{C_q \varepsilon_n^d}\tilde{c}_1\left(\frac{x}{\varepsilon_n}\right), \quad \forall x \in \mathbb{R}^d, \\
\hat{q}_n(\omega) &= \int_{\mathbb{R}^d} \mathrm{e}^{-\imath\omega^\top x} q_n(x)\mathrm{d}x = \frac{1}{C_q}\int_{\mathbb{R}^d} \mathrm{e}^{-\imath\varepsilon_n\omega^\top x}\tilde{c}_1(x)\mathrm{d}x = \frac{\tilde{\xi}_1(\varepsilon_n w)}{C_q}, \quad \forall \omega \in \mathbb{R}^d.
\end{aligned} \tag{S.74}$$

Then using Lemma 6 of [Wang and Loh, 2011], there exists a finite positive constant $C_{\hat{q}}$ that only depends on d, ν, a, b, such that

$$|\hat{q}_n(\omega)| \le \frac{C_{\hat{q}}}{(1 + \varepsilon_n\|\omega\|)^{a+d/2}}, \quad \forall \omega \in \mathbb{R}^d. \tag{S.75}$$

Lemma S.17. *Suppose that $d \in \{1, 2, 3\}$ and $\nu \in \mathbb{R}^+$. Let $a > 0$ and $0 < b < \min(4-d, 2)$ be fixed constants. Let $\varepsilon_n = n^{-1/(4a+2d+b)}$. For the g function in* (S.73) *and the q_n function in* (S.74)*, there exists a positive constant $C_{g,q}$ that depends only on d, ν, α_0, a, b, such that*

$$\left\{\int_{\mathbb{R}^d} |q_n * g(x) - g(x)|^2\,\mathrm{d}x\right\}^{1/2} \le C_{g,q}\frac{\max(\alpha_0^4, \alpha^4)\max\left\{\alpha_0^{2(2\nu+d-2)}, \alpha^{2(2\nu+d-2)}\right\}}{\alpha^{4\nu+3d/2-b/2}}\varepsilon_n^{b/2},$$

where $q_n * g(x) = \int_{\mathbb{R}^d} q_n(y)g(x-y)\mathrm{d}y$ for any $x \in \mathbb{R}^d$.

Proof of Lemma S.17. We have the following derivation:

$$\begin{aligned}
&\left\{\int_{\mathbb{R}^d} |q_n * g(x) - g(x)|^2\,\mathrm{d}x\right\}^{1/2} \\
=\ &\left[\int_{\mathbb{R}^d} \left|\int_{\|y\|\le 2m_a\varepsilon_n} \{g(x-y) - g(x)\}q_n(y)\mathrm{d}y\right|^2\right]^{1/2} \\
\overset{(i)}{\le}\ &\int_{\|y\|\le 2m_a\varepsilon_n} \left[\int_{\mathbb{R}^d} |g(x-y) - g(x)|^2\mathrm{d}x\right]^{1/2} q_n(y)\mathrm{d}y \\
\overset{(ii)}{=}\ &\int_{\|y\|\le 2m_a\varepsilon_n} \left[\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} |(\mathrm{e}^{-\imath\omega^\top y} - 1)\eta(\omega)|^2\mathrm{d}\omega\right]^{1/2} q_n(y)\mathrm{d}y \\
\overset{(iii)}{=}\ &\int_{\|y\|\le 2m_a\varepsilon_n} \left[\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \left|(\mathrm{e}^{-\imath\omega^\top y} - 1)\cdot\frac{f_{\sigma,\alpha}(\omega) - f_{\sigma_0,\alpha_0}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)}\cdot\frac{f_{\sigma_0,\alpha_0}(\omega)}{\xi_1(\omega)^2}\right|^2\mathrm{d}\omega\right]^{1/2}
\end{aligned}$$

$$
\begin{aligned}
&\cdot q_n(y)\mathrm{d}y \\
&\leq \frac{1}{(2\pi)^{d/2}} \sup_{\omega\in\mathbb{R}^d} \frac{f_{\sigma_0,\alpha_0}(\omega)}{\xi_1(\omega)^2} \cdot \\
&\quad \int_{\|y\|\leq 2m_a\varepsilon_n} \left[\int_{\mathbb{R}^d} \left|(\mathrm{e}^{-\imath\omega^\top y}-1)\cdot\left\{\frac{f_{\sigma,\alpha}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)}-1\right\}\right|^2 \mathrm{d}\omega\right]^{1/2} q_n(y)\mathrm{d}y \\
&\overset{(iv)}{\leq} \frac{1}{(2\pi)^{d/2}} \sup_{\omega\in\mathbb{R}^d} \frac{f_{\sigma_0,\alpha_0}(\omega)}{\xi_1(\omega)^2} \cdot \\
&\quad 2^{1-b/2}\int_{\|y\|\leq 2m_a\varepsilon_n} \left[\int_{\mathbb{R}^d} \|\omega\|^b \left|\left\{\frac{f_{\sigma,\alpha}(\omega)}{f_{\sigma_0,\alpha_0}(\omega)}-1\right\}\right|^2 \mathrm{d}\omega\right]^{1/2} \|y\|^{b/2} q_n(y)\mathrm{d}y \\
&\overset{(v)}{\leq} \frac{2^{1-b/2}}{(2\pi)^{d/2}} \sup_{\omega\in\mathbb{R}^d} \frac{f_{\sigma_0,\alpha_0}(\omega)}{\xi_1(\omega)^2} \\
&\quad \cdot \left[\int_{\mathbb{R}^d} \left\{\frac{(2\nu+d)\max(\alpha_0^2,\alpha^2)\max\left(\alpha_0^{2\nu+d-2},\alpha^{2\nu+d-2}\right)}{\alpha^{2\nu+d-2}}\right\}^2 \frac{\|\omega\|^b}{(\alpha^2+\|\omega\|^2)^2}\mathrm{d}\omega\right]^{1/2} \\
&\quad \cdot \int_{\|y\|\leq 2m_a\varepsilon_n} \|y\|^{b/2} q_n(y)\mathrm{d}y \\
&\overset{(vi)}{\leq} \frac{2^{1-b/2}\theta_0}{(2\pi)^{d/2}} \cdot \bar{c}_\xi \sigma^2\alpha^{2\nu} \max\left\{\left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d}, 1\right\} \\
&\quad \cdot \frac{(2\nu+d)\max(\alpha_0^2,\alpha^2)\max\left(\alpha_0^{2\nu+d-2},\alpha^{2\nu+d-2}\right)}{\alpha^{2\nu+d-2}} \\
&\quad \cdot \alpha^{b/2+d/2-2} \cdot \left[\int_0^\infty \frac{r^{b+d-1}}{(1+r^2)^2}\mathrm{d}r\right]^{1/2} \cdot (2m_a\varepsilon_n)^{b/2} \\
&\leq \left[\int_0^\infty \frac{r^{b+d-1}}{(1+r^2)^2}\mathrm{d}r\right]^{1/2} \\
&\quad \cdot \frac{2\bar{c}_\xi\theta_0(2\nu+d)m_a^{b/2}\max(\alpha_0^4,\alpha^4)\max\left(\alpha_0^{2(2\nu+d-2)},\alpha^{2(2\nu+d-2)}\right)}{(2\pi)^{d/2}\alpha^{4\nu+3d/2-b/2}} \cdot \varepsilon_n^{b/2}. \qquad \text{(S.76)}
\end{aligned}
$$

In the derivations above: (i) follows from the Minkowski's integral inequality; (ii) follows from the Plancherel's theorem; (iii) is based on the definition of $\eta(\omega)$ in (S.71); (iv) uses the fact that $|\mathrm{e}^{\imath a}-1|^2 = 4\sin^2(a/2) \leq 2^{2-b}|a|^b$ for any $a\in\mathbb{R}$ and all $0<b<2$; (v) follows from (S.59) in Lemma S.15. (vi) follows from (S.68) in Lemma S.16. Since $b<4-d$, the integral in the last display exists and hence the conclusion follows. □

Lemma S.18. *Suppose that $d\in\{1,2,3\}$ and $\nu\in\mathbb{R}^+$. Let $(\sigma,\alpha)\in\mathbb{R}^+\times\mathbb{R}^+$ satisfy $\sigma^2\alpha^{2\nu} = \theta_0 = \sigma_0^2\alpha_0^{2\nu}$. Let $a>0$ and $0<b<\min(4-d,2)$ be fixed constants. Let $\varepsilon_n = n^{-1/(4a+2d+b)}$. For the $\lambda_{k,n}(\alpha)$ in (S.57), for any $\alpha>0$, there exist positive constants $C_1^\dagger, C_1^\ddagger, C_2^\ddagger$ that depend only on d,ν,T,α_0,a,b, such that*

$$
\begin{aligned}
&\sum_{k=1}^n |\lambda_{k,n}(\alpha)-1| \\
&\leq C_1^\dagger \frac{\max(\alpha_0^6,\alpha^6)\max\left\{\alpha_0^{3(2\nu+d-2)},\alpha^{3(2\nu+d-2)}\right\}\sqrt{n}\varepsilon_n^{b/2}}{\alpha^{4\nu+3d/2-b/2}}
\end{aligned}
$$

$$+ C_1^{\ddagger} \frac{[\max(\alpha_0, \alpha)]^{2\nu+d}}{\varepsilon_n^{2a+d}} + C_2^{\ddagger} \frac{\max(\alpha_0^6, \alpha^6) \max\left\{\alpha_0^{3(2\nu+d-2)}, \alpha^{3(2\nu+d-2)}\right\}}{\alpha^{2(3\nu+d)}}. \tag{S.77}$$

Proof of Lemma S.18. For any $x, y \in \mathcal{S}$, let $b(x,y) = \mathrm{E}_{(\sigma,\alpha)}\{X(x)X(y)\} - \mathrm{E}_{(\sigma_0,\alpha_0)}\{X(x)X(y)\}$. Then using the definition of $c_0(x)$ in (S.66) and $c_1(x)$ with the support of c_1 in $[-2m_\nu, 2m_\nu]^d$, the derivation after Equation (39) of [Wang and Loh, 2011] has shown that for $s, t \in \mathcal{S}$,

$$\begin{aligned} b(x,y) &= (2\pi)^d \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} g(s-t)c_1(x-s)c_1(y-t)\mathrm{d}s\mathrm{d}t \\ &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} \mathrm{e}^{\imath(\omega^\top x - v^\top y)}\eta_n^*\left(\frac{w+v}{2}\right)\vartheta\left(\frac{w-v}{2}\right)\xi_1(\omega)\xi_1(v)\mathrm{d}\omega\mathrm{d}v \\ &\quad + \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} \mathrm{e}^{\imath(\omega^\top x - v^\top y)}\xi_1(\omega)\xi_1(v)\Bigg\{\int_{\|u\|_\infty \le 2m_\nu + 2m_a + T} \mathrm{e}^{-\imath(\omega^\top u - v^\top u)} \\ &\quad \times \hat{q}_n(\omega)\eta(v)\mathrm{d}u\Bigg\}\mathrm{d}v\mathrm{d}\omega, \end{aligned} \tag{S.78}$$

where $\eta_n^* : \mathbb{R}^d \to \mathbb{C}$ is the Fourier transform of $g - q_n * g$ for g defined in (S.73) and q_n in defined in (S.74), such that $\int_{\mathbb{R}^d} |\eta_n^*(\omega)|^2\,\mathrm{d}\omega = \int_{\mathbb{R}^d} |q_n * g(x) - g(x)|^2\,\mathrm{d}x$ which can be upper bounded by Lemma S.17; $\vartheta(\omega)$ in (S.78) is defined in the same way as Equation (23) of [Wang and Loh, 2011]:

$$\vartheta(\omega) = \frac{1}{2^d}\int_{\mathbb{R}^d} e^{-\imath t^\top w}\mathcal{I}\left(\|t\|_\infty \le 4m_\nu + 2T\right)\mathrm{d}t, \qquad \text{for all } \omega \in \mathbb{R}^d. \tag{S.79}$$

Lemma 3 of [Wang and Loh, 2011] has proved that $\int_{\mathbb{R}^d}\vartheta(\omega)^2\mathrm{d}\omega < \infty$ and its value only depends on d, ν, T.

Note that by the definition of covariance function,

$$\begin{aligned} b(x,y) &= \mathrm{E}_{(\sigma,\alpha)}\{X(x)X(y)\} - \mathrm{E}_{(\sigma_0,\alpha_0)}\{X(x)X(y)\} \\ &= \int_{\mathbb{R}^d} \mathrm{e}^{\imath(x-y)^\top w}\left\{f_{\sigma,\alpha}(\omega) - f_{\sigma_0,\alpha_0}(\omega)\right\}\mathrm{d}\omega. \end{aligned} \tag{S.80}$$

Hence, for any pair (σ, α) that satisfies $\sigma^2\alpha^{2\nu} = \theta_0 = \sigma_0^2\alpha_0^{2\nu}$, for the $\{\psi_k : k = 1\ldots, n\}$ functions in (S.57), we have that for $k = 1, \ldots, n$,

$$\lambda_{k,n}(\alpha) - 1 = \langle\psi_k, \psi_k\rangle_{f_{\sigma,\alpha}} - \langle\psi_k, \psi_k\rangle_{f_{\sigma_0,\alpha_0}} := \zeta_{k,n}^\dagger + \zeta_{k,n}^\ddagger, \tag{S.81}$$

where

$$\begin{aligned} \zeta_{k,n}^\dagger &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} \psi_k(\omega)\overline{\psi_k(v)}\eta_n^*\left(\frac{w+v}{2}\right)\vartheta\left(\frac{w-v}{2}\right)\xi_1(\omega)\xi_1(v)\mathrm{d}\omega\mathrm{d}v, \\ \zeta_{k,n}^\ddagger &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} \psi_k(\omega)\overline{\psi_k(v)}\xi_1(\omega)\xi_1(v)\hat{q}_n(\omega)\eta(v) \\ &\quad \times \left\{\int_{\|u\|_\infty \le 2m_\nu + 2m_a + T} \mathrm{e}^{-\imath(\omega^\top u - v^\top u)}\mathrm{d}u\right\}\mathrm{d}\omega\mathrm{d}v. \end{aligned} \tag{S.82}$$

We follow the derivations on page 258-259 of [Wang and Loh, 2011]. By the Bessel's inequality, we have that

$$\sum_{k=1}^n \left|\zeta_{k,n}^\dagger\right|^2 = \sum_{k=1}^n \left\{\frac{1}{(2\pi)^d}\int_{\mathbb{R}^{2d}} \psi_k(\omega)\overline{\psi_k(\omega)}\eta_n^*\left(\frac{w+v}{2}\right)\vartheta\left(\frac{w-v}{2}\right)\xi_1(\omega)\xi_1(v)\mathrm{d}\omega\mathrm{d}v\right\}^2$$

$$
\begin{aligned}
&\leq \frac{1}{(2\pi)^{2d}} \int_{\mathbb{R}^{2d}} \left| \eta_n^*\left(\frac{w+v}{2}\right) \vartheta\left(\frac{w-v}{2}\right) \right|^2 \frac{\xi_1(\omega)^2}{f_{\sigma,\alpha}(\omega)} \frac{\xi_1(v)^2}{f_{\sigma,\alpha}(v)} \mathrm{d}\omega \mathrm{d}v \\
&\overset{(i)}{\leq} \frac{1}{2^d \pi^{2d}} \left\{ \sup_{\omega\in\mathbb{R}^d} \frac{\xi_1(\omega)^2}{f_{\sigma,\alpha}(\omega)} \right\}^2 \int_{\mathbb{R}^d} |\vartheta(v)|^2 \mathrm{d}v \int_{\mathbb{R}^d} |\eta_n^*(\omega)|^2 \mathrm{d}\omega \\
&\overset{(ii)}{\leq} \frac{1}{2^d \pi^{2d}} \cdot \left\{ \frac{\max\left\{(\alpha/\alpha_0)^{2\nu+d}, 1\right\}}{\underline{c}_\xi \theta_0} \right\}^2 \cdot \int_{\mathbb{R}^d} |\vartheta(v)|^2 \mathrm{d}v \\
&\quad \times C_{g,q}^2 \left[\frac{\max(\alpha_0^4, \alpha^4) \max\left\{\alpha_0^{2(2\nu+d-2)}, \alpha^{2(2\nu+d-2)}\right\}}{\alpha^{4\nu+3d/2-b/2}} \right]^2 \cdot \varepsilon_n^b \\
&\leq (C_1^\dagger)^2 \frac{\max(\alpha_0^{12}, \alpha^{12}) \max\left\{\alpha_0^{6(2\nu+d-2)}, \alpha^{6(2\nu+d-2)}\right\}}{\alpha^{2(4\nu+3d/2-b/2)}} \varepsilon_n^b,
\end{aligned} \tag{S.83}
$$

where (i) follows from the Cauchy-Schwarz inequality; (ii) follows from Lemma S.16 and Lemma S.17, and $C_1^\dagger$ is a positive constant that depends only on $d, \nu, T, \alpha_0, a, b$.

For $\zeta_{k,n}^\ddagger$, we apply the Bessel's inequality to obtain that

$$
\begin{aligned}
&\sum_{k=1}^n \left| \zeta_{k,n}^\ddagger \right| \\
&\leq \frac{1}{(2\pi)^d} \sum_{k=1}^n \int_{\|u\|_\infty \leq 2m_\nu+2m_a+T} \left| \int_{\mathbb{R}^d} \mathrm{e}^{-\imath\omega^\top u} \psi_k(\omega) \xi_1(\omega) \hat{q}_n(\omega) \mathrm{d}\omega \right| \\
&\quad \times \left| \int_{\mathbb{R}^d} \mathrm{e}^{\imath v^\top u} \overline{\psi}_k(v) \xi_1(v) \eta(v) \mathrm{d}v \right| \mathrm{d}u \\
&\leq \frac{1}{2(2\pi)^d} \int_{\|u\|_\infty \leq 2m_\nu+2m_a+T} \sum_{k=1}^n \left\{ \left| \int_{\mathbb{R}^d} \mathrm{e}^{-\imath\omega^\top u} \psi_k(\omega) \frac{\xi_1(\omega)}{f_{\sigma,\alpha}(\omega)} \hat{q}_n(\omega) f_{\sigma,\alpha}(\omega) \mathrm{d}\omega \right|^2 \right. \\
&\quad \left. + \left| \int_{\mathbb{R}^d} \mathrm{e}^{-\imath v^\top u} \overline{\psi}_k(v) \frac{\xi_1(v)}{f_{\sigma,\alpha}(v)} \eta(v) f_{\sigma,\alpha}(v) \mathrm{d}v \right|^2 \right\} \mathrm{d}u \\
&\leq \frac{1}{2(2\pi)^d} \int_{\|u\|_\infty \leq 2m_\nu+2m_a+T} \left\{ \sup_{\omega\in\mathbb{R}^d} \frac{\xi_1(\omega)^2}{f_{\sigma,\alpha}(\omega)} \int_{\mathbb{R}^d} |\hat{q}_n(\omega)|^2 \mathrm{d}\omega \right. \\
&\quad \left. + \sup_{\omega\in\mathbb{R}^d} \frac{f_{\sigma_0,\alpha_0}(\omega)}{\xi_1(\omega)^2} \int_{\mathbb{R}^d} \left| \frac{f_{\sigma,\alpha}(v)}{f_{\sigma_0,\alpha_0}(v)} - 1 \right|^2 \mathrm{d}v \right\} \mathrm{d}u \\
&\overset{(i)}{\leq} \frac{1}{2(2\pi)^d} \cdot (4m_\nu + 4m_a + 2T)^d \cdot \left\{ \frac{\max\left\{(\alpha/\alpha_0)^{2\nu+d}, 1\right\}}{\underline{c}_\xi \theta_0} \right\} \\
&\quad \times \int_{\mathbb{R}^d} \frac{C_{\hat{q}}^2}{(1+\varepsilon_n\|\omega\|)^{2a+d}} \mathrm{d}\omega \\
&\quad + \frac{1}{2(2\pi)^d} \cdot (4m_\nu + 4m_a + 2T)^d \cdot \overline{c}_\xi \theta_0 \max\left\{(\alpha_0/\alpha)^{2\nu+d}, 1\right\} \\
&\quad \times \int_{\mathbb{R}^d} \left\{ \frac{(2\nu+d)\max(\alpha_0^2, \alpha^2) \max\left(\alpha_0^{2\nu+d-2}, \alpha^{2\nu+d-2}\right)}{\alpha^{2\nu+d-2}} \right\}^2 \frac{1}{(\alpha^2+\|v\|^2)^2} \mathrm{d}v \\
&\leq \frac{(4m_\nu + 4m_a + 2T)^d}{2(2\pi)^d} \cdot \frac{C_{\hat{q}}^2 [\max(\alpha_0, \alpha)]^{2\nu+d}}{\underline{c}_\xi \theta_0 \alpha_0^{2\nu+d} \varepsilon_n^{2a+d}} \left\{ \int_0^\infty \frac{r^{d-1}}{(1+r)^{2a+d}} \mathrm{d}r \right\}
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{(4m_\nu+4m_a+2T)^d\overline{c}_\xi\theta_0}{2(2\pi)^d}\cdot\frac{(2\nu+d)^2\max(\alpha_0^6,\alpha^6)\max\left\{\alpha_0^{3(2\nu+d-2)},\alpha^{3(2\nu+d-2)}\right\}}{\alpha^{3(2\nu+d)-4}}\\
&\times\alpha^{d-4}\left\{\int_0^\infty\frac{r^{d-1}}{(1+r^2)^2}\mathrm{d}r\right\}\\
\le\,& C_1^\ddagger\frac{[\max(\alpha_0,\alpha)]^{2\nu+d}}{\varepsilon_n^{2a+d}}+C_2^\ddagger\frac{\max(\alpha_0^6,\alpha^6)\max\left\{\alpha_0^{3(2\nu+d-2)},\alpha^{3(2\nu+d-2)}\right\}}{\alpha^{2(3\nu+d)}},
\end{aligned}\tag{S.84}
$$

where (i) follows from Lemma S.15, Lemma S.16, and the inequality (S.75), and $C_1^\ddagger, C_2^\ddagger$ are positive constants that depend only on d,ν,T,α_0,a,b.

Finally, we combine (S.83) and (S.84) to conclude that for any pair (σ,α) that satisfies $\sigma^2\alpha^{2\nu}=\theta_0=\sigma_0^2\alpha_0^{2\nu}$,

$$
\begin{aligned}
&\sum_{k=1}^n|\lambda_{k,n}(\alpha)-1|\le\sum_{k=1}^n\left(\left|\zeta_{k,n}^\dagger\right|+\left|\zeta_{k,n}^\ddagger\right|\right)\le\left(n\sum_{k=1}^n\left|\zeta_{k,n}^\dagger\right|^2\right)^{1/2}+\sum_{k=1}^n\left|\zeta_{k,n}^\ddagger\right|\\
&\le C_1^\dagger\frac{\max(\alpha_0^6,\alpha^6)\max\left\{\alpha_0^{3(2\nu+d-2)},\alpha^{3(2\nu+d-2)}\right\}\sqrt{n}\varepsilon_n^{b/2}}{\alpha^{4\nu+3d/2-b/2}}\\
&+C_1^\ddagger\frac{[\max(\alpha_0,\alpha)]^{2\nu+d}}{\varepsilon_n^{2a+d}}+C_2^\ddagger\frac{\max(\alpha_0^6,\alpha^6)\max\left\{\alpha_0^{3(2\nu+d-2)},\alpha^{3(2\nu+d-2)}\right\}}{\alpha^{2(3\nu+d)}}.
\end{aligned}
$$

□

Lemma S.19. *([Laurent and Massart, 2000] Lemma 1) Let $Z_1,\ldots,Z_n$ be i.i.d. $\mathcal{N}(0,1)$ random variables. Let $\{w_i: i=1\ldots,n\}$ be nonnegative constants. Let $\|w\|_\infty=\max_{1\le i\le n}w_i$, $\|w\|_1=\sum_{i=1}^n w_i$, and $\|w\|^2=\sum_{i=1}^n w_i^2$. Then for any positive $z>0$,*

$$
\begin{aligned}
&\Pr\left\{\sum_{i=1}^n w_iZ_i^2\ge\|w\|_1+2\|w\|\sqrt{z}+2\|w\|_\infty z\right\}\le \mathrm{e}^{-z},\\
&\Pr\left\{\sum_{i=1}^n w_iZ_i^2\le\|w\|_1-2\|w\|\sqrt{z}\right\}\le \mathrm{e}^{-z}.
\end{aligned}
$$

Lemma S.20. *Suppose that $d\in\{1,2,3\}$ and $\nu\in\mathbb{R}^+$. For any $\alpha>0$, we define $w_i(\alpha)=\left|\lambda_{i,n}(\alpha)^{-1}-1\right|/\sqrt{n}$ for $i=1,\ldots,n$ and $w(\alpha)=(w_1(\alpha),\ldots,w_n(\alpha))^\top$, where $\lambda_{i,n}(\alpha)$'s are as defined in* (S.54) *and* (S.57)*. Then there exists a large integer N_5' that only depends on ν,d,T,α_0, such that for all $n>N_5'$, for $\tau,\underline{\alpha}_n,\overline{\alpha}_n$ defined in* (S.9),

$$
\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]}\left\{\|w(\alpha)\|_1+4\|w(\alpha)\|\log n+8\|w(\alpha)\|_\infty\log^2 n\right\}\le n^{-\tau}/8,
$$

where $\|w(\alpha)\|_1=\sum_{i=1}^n|w_i(\alpha)|$, $\|w(\alpha)\|=\left(\sum_{i=1}^n w_i(\alpha)^2\right)^{1/2}$, and $\|w(\alpha)\|_\infty=\max_{1\le i\le n}|w_i(\alpha)|$.

Proof of Lemma S.20. For abbreviation, we use Γ to denote the right-hand side of Equation S.77 in Lemma S.18. From Lemma S.15 and Lemma S.18, we can obtain that

$$
\begin{aligned}
\|w(\alpha)\|_1&=\sum_{i=1}^n w_i(\alpha)=\frac{1}{\sqrt{n}}\sum_{i=1}^n\left|\lambda_{i,n}(\alpha)^{-1}-1\right|\\
&\le\frac{1}{\sqrt{n}\min_{1\le i\le n}\lambda_{i,n}(\alpha)}\sum_{i=1}^n|\lambda_{i,n}(\alpha)-1|\le\frac{\{\max(\alpha_0,\alpha)\}^{2\nu+d}}{\sqrt{n}\alpha_0^{2\nu+d}}\times\Gamma,
\end{aligned}\tag{S.85}
$$

$$
\begin{aligned}
\|w(\alpha)\|^2 &= \sum_{i=1}^n w_i^2 = \frac{1}{n}\sum_{i=1}^n \left|\lambda_{i,n}(\alpha)^{-1} - 1\right|^2 \\
&\le \frac{1}{n\left\{\min_{1\le i\le n}\lambda_{i,n}(\alpha)\right\}^2}\sum_{i=1}^n |\lambda_{i,n}(\alpha) - 1|^2 \\
&\le \frac{1}{n\left\{\min_{1\le i\le n}\lambda_{i,n}(\alpha)\right\}^2}\left(\sum_{i=1}^n |\lambda_{i,n}(\alpha) - 1|\right)^2 \le \frac{\{\max(\alpha_0,\alpha)\}^{2(2\nu+d)}}{n\alpha_0^{2(2\nu+d)}} \times \Gamma^2. \qquad \text{(S.86)}
\end{aligned}
$$

We can see the upper bound in (S.86) is exactly the square of the upper bound in (S.85).

$$
\begin{aligned}
\|w(\alpha)\|_\infty &= \max_{1\le i\le n} w_i = \frac{1}{\sqrt{n}}\max_{1\le i\le n}\left|\lambda_{i,n}(\alpha)^{-1} - 1\right| \\
&\le \frac{1}{\sqrt{n}\min_{1\le i\le n}\lambda_{i,n}(\alpha)}\max_{1\le i\le n}|\lambda_{i,n}(\alpha) - 1| \le \frac{\max_{1\le i\le n}\lambda_{i,n}(\alpha) + 1}{\sqrt{n}\min_{1\le i\le n}\lambda_{i,n}(\alpha)} \\
&\le \frac{\max\left\{(\alpha_0/\alpha)^{2\nu+d}, 1\right\} + 1}{\sqrt{n}\min\left\{(\alpha_0/\alpha)^{2\nu+d}, 1\right\}} \le \frac{\left[\{\max(\alpha_0,\alpha)\}^{2\nu+d} + \alpha^{2\nu+d}\right]\{\max(\alpha_0,\alpha)\}^{2\nu+d}}{\sqrt{n}\alpha_0^{2\nu+d}\alpha^{2\nu+d}} \\
&\le \frac{2\{\max(\alpha_0,\alpha)\}^{2(2\nu+d)}}{\sqrt{n}\alpha_0^{2\nu+d}\alpha^{2\nu+d}}. \qquad \text{(S.87)}
\end{aligned}
$$

Since $\varepsilon_n = n^{-1/(4a+2d+b)}$ in Lemma S.17 and Lemma S.18, we have $\sqrt{n}\varepsilon_n^{b/2} = 1/\varepsilon_n^{2a+d} = n^{(2a+d)/(4a+2d+b)}$. Let $z = 4\log^2 n$ in Lemma S.19. In the following, we analyze the necessary condition for $\underline{\alpha}_n$ and $\overline{\alpha}_n$ such that $\|w(\alpha)\|_1 + 4\|w(\alpha)\|\sqrt{z} + 8\|w(\alpha)\|_\infty z = o(1)$ for any $\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]$ as $n \to \infty$. We consider two situations according to the value of α, each of which has two further sub-cases according to the sign of $2\nu + d - 2$.

(1) When $\alpha \in [\alpha_0, \overline{\alpha}_n]$ and possibly $\alpha \to +\infty$ as $n \to \infty$:

In this case, in the upper bounds of (S.85) and (S.86), since $\alpha \ge \alpha_0$, we have that $\max\{\alpha_0^{3(2\nu+d-2)}, \alpha^{3(2\nu+d-2)}\} \preceq \alpha^{3(2\nu+d-2)}$ if $2\nu + d - 2 \ge 0$, and that $\max\{\alpha_0^{3(2\nu+d-2)}, \alpha^{3(2\nu+d-2)}\} \preceq 1$ if $-1 < 2\nu + d - 2 < 0$. We discuss the two sub-cases respectively:

(1)-(i) When $2\nu + d - 2 \ge 0$, we have $\max\{\alpha_0^{3(2\nu+d-2)}, \alpha^{3(2\nu+d-2)}\} \preceq \alpha^{3(2\nu+d-2)}$. Using (S.85), (S.86), and (S.87), we can see that (neglecting all multiplicative constants by using the order relation $\preceq$):

$$
\begin{aligned}
&\|w(\alpha)\|_1 + 2\|w(\alpha)\|\sqrt{z} + 2\|w(\alpha)\|_\infty z \\
&\preceq \frac{\alpha^{2\nu+d}\log n}{\sqrt{n}}\left(\alpha^{2\nu+3d/2+b/2}\sqrt{n}\varepsilon_n^{b/2} + \frac{\alpha^{2\nu+d}}{\varepsilon_n^{2a+d}} + \alpha^d\right) + \frac{\alpha^{2\nu+d}\log^2 n}{\sqrt{n}} \\
&\preceq \frac{\overline{\alpha}_n^{2\nu+d}\log n}{\sqrt{n}} \cdot \overline{\alpha}_n^{2\nu+3d/2+b/2} n^{(2a+d)/(4a+2d+b)} + \frac{\overline{\alpha}_n^{2\nu+d}\log^2 n}{\sqrt{n}} \\
&= \frac{\overline{\alpha}_n^{4\nu+5d/2+b/2}\log n}{n^{b/(8a+4d+2b)}} + \frac{\overline{\alpha}_n^{2\nu+d}\log^2 n}{\sqrt{n}}. \qquad \text{(S.88)}
\end{aligned}
$$

In order to make the last upper bound $o(1)$, given that $\overline{\alpha}_n \succ 1$, we further need

$$
\overline{\alpha}_n \prec n^{\frac{b}{(4a+2d+b)(8\nu+5d+b)}}(\log n)^{-\frac{2}{8\nu+5d+b}}, \quad \overline{\alpha}_n \prec n^{\frac{1}{2(2\nu+d)}}(\log n)^{-\frac{2}{2\nu+d}}, \qquad \text{(S.89)}
$$

which holds as long as

$$
\overline{\kappa} < \frac{b}{(4a+2d+b)(8\nu+5d+b)}, \quad \overline{\kappa} < \frac{1}{2(2\nu+d)}. \qquad \text{(S.90)}
$$

(1)-(ii) When $-1 < 2\nu + d - 2 < 0$, we have $\max\{\alpha_0^{3(2\nu+d-2)}, \alpha^{3(2\nu+d-2)}\} \preceq 1$. Note that this special case can only happen when $d = 1$ and $\nu \in (0, 1/2)$. Using (S.85), (S.86), and (S.87), we can see that:

$$\begin{aligned}
&\|w(\alpha)\|_1 + 2\|w(\alpha)\|\sqrt{z} + 2\|w(\alpha)\|_\infty z \\
&\preceq \frac{\alpha^{2\nu+d}\log n}{\sqrt{n}}\Big\{\alpha^{6-4\nu-3d/2+b/2}n^{(2a+d)/(4a+2d+b)} + \alpha^{2\nu+d}n^{(2a+d)/(4a+2d+b)} \\
&\quad + \alpha^{6-6\nu-2d}\Big\} + \frac{\alpha^{2\nu+d}\log^2 n}{\sqrt{n}}. \qquad \text{(S.91)}
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\|w(\alpha)\|_1 + 2\|w(\alpha)\|\sqrt{z} + 2\|w(\alpha)\|_\infty z \\
&\preceq \frac{\overline{\alpha}_n^{2\nu+d}\log n}{\sqrt{n}}\Big\{\overline{\alpha}_n^{6-4\nu-3d/2+b/2}n^{(2a+d)/(4a+2d+b)} + \overline{\alpha}_n^{6-6\nu-2d}\Big\} + \frac{\overline{\alpha}_n^{2\nu+d}\log^2 n}{\sqrt{n}} \\
&\preceq \frac{\overline{\alpha}_n^{6-2\nu-d/2+b/2}\log n}{n^{b/(8a+4d+2b)}} + \frac{\overline{\alpha}_n^{6-4\nu-d}\log n}{\sqrt{n}} + \frac{\overline{\alpha}_n^{2\nu+d}\log^2 n}{\sqrt{n}}. \qquad \text{(S.92)}
\end{aligned}$$

In order to make the last upper bound $o(1)$, given that $\overline{\alpha}_n \succ 1$ and $d = 1$, we need

$$\begin{aligned}
&\overline{\alpha}_n \prec n^{\frac{b}{(4a+2+b)(11-4\nu+b)}}(\log n)^{-\frac{2}{11-4\nu+b}}, \\
&\overline{\alpha}_n \prec n^{\frac{1}{2(5-4\nu)}}(\log n)^{-\frac{1}{5-4\nu}}, \quad \overline{\alpha}_n \prec n^{\frac{1}{2(2\nu+1)}}(\log n)^{-\frac{2}{2\nu+1}}. \qquad \text{(S.93)}
\end{aligned}$$

Since $d = 1$ and $\nu \in (0, 1/2)$ in this case, $11 - 4\nu + b > 0$ and $10 > 2(5 - 4\nu) > 2(2\nu + 1)$. We need that

$$\overline{\kappa} < \frac{b}{(4a+2+b)(11-4\nu+b)}, \quad \overline{\kappa} < \frac{1}{10}. \qquad \text{(S.94)}$$

(2) When $\alpha \in [\underline{\alpha}_n, \alpha_0]$ and possibly $\alpha \to 0+$ as $n \to \infty$:

In this case, in the upper bounds of (S.85) and (S.86), since $\alpha \le \alpha_0$, we have that $\max\{\alpha_0^{3(2\nu+d-2)}, \alpha^{3(2\nu+d-2)}\} \preceq 1$ if $2\nu + d - 2 \ge 0$, and that $\max\{\alpha_0^{3(2\nu+d-2)}, \alpha^{3(2\nu+d-2)}\} \preceq \alpha^{3(2\nu+d-2)}$ if $-1 < 2\nu + d - 2 < 0$. We discuss the two sub-cases respectively:

(2)-(i) When $2\nu + d - 2 \ge 0$, we have $\max\{\alpha_0^{3(2\nu+d-2)}, \alpha^{3(2\nu+d-2)}\} \preceq 1$ and $\max(\alpha_0, \alpha) \preceq 1$. Using (S.85), (S.86), and (S.87), we can see that in this case:

$$\begin{aligned}
&\|w(\alpha)\|_1 + 2\|w(\alpha)\|\sqrt{z} + 2\|w(\alpha)\|_\infty z \\
&\preceq \frac{\log n}{\sqrt{n}}\left(\frac{\sqrt{n}\varepsilon_n^{b/2}}{\alpha^{4\nu+3d/2-b/2}} + \frac{1}{\varepsilon_n^{2a+d}} + \frac{1}{\alpha^{2(3\nu+d)}}\right) + \frac{\log^2 n}{\sqrt{n}\alpha^{2\nu+d}} \\
&\preceq \frac{\log n}{\underline{\alpha}_n^{4\nu+3d/2-b/2}n^{\frac{b}{2(4a+2d+b)}}} + \frac{\log n}{\sqrt{n}\underline{\alpha}_n^{2(3\nu+d)}} + \frac{\log^2 n}{\sqrt{n}\underline{\alpha}_n^{2\nu+d}}. \qquad \text{(S.95)}
\end{aligned}$$

In order to make the last upper bound $o(1)$, given that $\underline{\alpha}_n \prec 1$, we need that

$$\begin{aligned}
&\underline{\alpha}_n \succ n^{-\frac{b}{(4a+2d+b)(8\nu+3d-b)}}(\log n)^{\frac{2}{8\nu+3d-b}}, \\
&\underline{\alpha}_n \succ n^{-\frac{1}{4(3\nu+d)}}(\log n)^{\frac{1}{2(3\nu+d)}}, \quad \underline{\alpha}_n \succ n^{-\frac{1}{2(2\nu+d)}}(\log n)^{\frac{2}{2\nu+d}}. \qquad \text{(S.96)}
\end{aligned}$$

Since $4(3\nu + d) > 2(2\nu + d)$, we only need

$$\underline{\kappa} < \frac{b}{(4a + 2d + b)(8\nu + 3d - b)}, \quad \underline{\kappa} < \frac{1}{4(3\nu + d)}. \tag{S.97}$$

(2)-(ii) When $-1 < 2\nu + d - 2 < 0$, we have $\max\{\alpha_0^{3(2\nu+d-2)}, \alpha^{3(2\nu+d-2)}\} \preceq \alpha^{3(2\nu+d-2)}$ and $\max(\alpha_0, \alpha) \preceq \alpha^{3(2\nu+d-2)}$. Note that this special case can only happen when $d = 1$ and $\nu \in (0, 1/2)$. Using (S.85), (S.86), and (S.87), we can see that in this case:

$$\begin{aligned}
&\|w(\alpha)\|_1 + 2\|w(\alpha)\|\sqrt{z} + 2\|w(\alpha)\|_\infty z \\
\preceq\ & \frac{\log n}{\sqrt{n}} \left(\frac{\sqrt{n}\varepsilon_n^{b/2}}{\alpha^{6-2\nu-3d/2-b/2}} + \frac{1}{\varepsilon_n^{2a+d}} + \frac{1}{\alpha^{6-d}} \right) + \frac{\log^2 n}{\sqrt{n}\alpha^{2\nu+d}} \\
\preceq\ & \frac{\log n}{\underline{\alpha}_n^{6-2\nu-3d/2-b/2} n^{\frac{b}{2(4a+2d+b)}}} + \frac{\log n}{\sqrt{n}\underline{\alpha}_n^{6-d}} + \frac{\log^2 n}{\sqrt{n}\underline{\alpha}_n^{2\nu+d}}.
\end{aligned} \tag{S.98}$$

In order to make the last upper bound $o(1)$, given that $\underline{\alpha}_n \prec 1$ and $d = 1$, we only need that

$$\begin{aligned}
&\underline{\alpha}_n \succ n^{-\frac{b}{(4a+2+b)(9-4\nu-b)}} (\log n)^{\frac{2}{9-4\nu-b}}, \\
&\underline{\alpha}_n \succ n^{-\frac{1}{10}} (\log n)^{\frac{1}{5}}, \quad \underline{\alpha}_n \succ n^{-\frac{1}{2(2\nu+1)}} (\log n)^{\frac{2}{2\nu+1}}.
\end{aligned} \tag{S.99}$$

Note that since $d = 1$ and $\nu \in (0, 1/2)$ in this case, $9 - 4\nu - b > 0$ and $2(2\nu + d) < 10$. Therefore, we only need

$$\underline{\kappa} < \frac{b}{(4a + 2 + b)(9 - 4\nu - b)}, \quad \underline{\kappa} < \frac{1}{10}. \tag{S.100}$$

Since all the right-hand sides of (S.90), (S.94), (S.97), and (S.100) are positive, we choose $a = 0.01$ and $b = 0.9$ such that $a > 0$ and $0 < b < \min(4 - d, 2)$ with $d \in \{1, 2, 3\}$ are both satisfied. Then the choice of $\overline{\kappa}$ and $\underline{\kappa}$ in (S.9) satisfy (S.90), (S.94), (S.97), and (S.100). Furthermore, for τ defined in (S.9), $n^{-\tau}/8$ is strictly larger in order than the maximum of the right-hand sides of (S.88), (S.92), (S.95), and (S.98).

With this τ and $z = 4\log^2 n$, we have shown that uniformly for all $\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]$, there exists a large integer N_5' that depends only on ν, d, T, α_0, such that for all $n > N_5'$,

$$\|w(\alpha)\|_1 + 4\|w(\alpha)\| \log n + 8\|w(\alpha)\|_\infty \log^2 n \le n^{-\tau}/8.$$

□

Lemma S.21. *([Hsu et al., 2012] Proposition 1.1) Let $Z_1, \ldots, Z_n$ be i.i.d. $\mathcal{N}(0, 1)$ random variables and $Z = (Z_1, \ldots, Z_n)^\top$. Let Σ be an $n \times n$ symmetric positive semidefinite matrix. Then for any positive $z > 0$,*

$$\Pr\left\{ Z^\top \Sigma Z \ge \operatorname{tr}(\Sigma) + 2\sqrt{\operatorname{tr}(\Sigma^2) z} + 2\|\Sigma\|_{\rm op} z \right\} \le e^{-z}.$$

S2 Technical Lemmas for Profile Restricted Log-Likelihood

In this section, we derive some useful results for the profile restricted log-likelihood $\widetilde{\mathcal{L}}_n(\alpha)$ defined in (8) of the main text. In particular, we show Lemma S.22, Lemma S.24, Lemma S.25, and Lemma S.27. These four lemmas play key roles in controlling the tail part of the posterior of α, and will be used in the proof of Theorem 2.3. Finally, Lemma S.28 proves the existence of the profile posterior $\widetilde{\pi}(\alpha|Y_n)$ as stated in Theorem 2.3.

We recall from the main text that the profile restricted log-likelihood $\widetilde{\mathcal{L}}_n(\alpha)$ defined in (8) of the main text is

$$\begin{aligned}\widetilde{\mathcal{L}}_n(\alpha) &\equiv \mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha) \\ &= -\frac{n-p}{2}\log\frac{Y_n^\top\left[R_\alpha^{-1} - R_\alpha^{-1}M_n\left(M_n^\top R_\alpha^{-1}M_n + \Omega_\beta\right)^{-1}M_n^\top R_\alpha^{-1}\right]Y_n}{n-p} \\ &\quad - \frac{1}{2}\log|R_\alpha| - \frac{1}{2}\log\left|M_n^\top R_\alpha^{-1}M_n + \Omega_\beta\right| - \frac{n-p}{2}. \end{aligned} \tag{S.101}$$

Lemma S.22. *Suppose that $d \in \mathbb{Z}^+$ and $\nu \in \mathbb{R}^+$. The profile restricted log-likelihood function defined in* (S.101) *satisfies that for any $0 < \alpha_1 < \alpha_2 < \infty$, for all possible value of $Y_n \in \mathbb{R}^n$,*

$$\left(\frac{\alpha_1}{\alpha_2}\right)^{n(\nu+d/2)} < \exp\left\{\widetilde{\mathcal{L}}_n(\alpha_2) - \widetilde{\mathcal{L}}_n(\alpha_1)\right\} < \left(\frac{\alpha_2}{\alpha_1}\right)^{n(\nu+d/2)}.$$

Proof of Lemma S.22. From the expression (S.101), we have that for any $0 < \alpha_1 < \alpha_2 < \infty$,

$$\begin{aligned}&\widetilde{\mathcal{L}}_n(\alpha_2) - \widetilde{\mathcal{L}}_n(\alpha_1) \\ &= -\frac{n-p}{2}\log\frac{Y_n^\top\left[R_{\alpha_2}^{-1} - R_{\alpha_2}^{-1}M_n\left(M_n^\top R_{\alpha_2}^{-1}M_n + \Omega_\beta\right)^{-1}M_n^\top R_{\alpha_2}^{-1}\right]Y_n}{Y_n^\top\left[R_{\alpha_1}^{-1} - R_{\alpha_1}^{-1}M_n\left(M_n^\top R_{\alpha_1}^{-1}M_n + \Omega_\beta\right)^{-1}M_n^\top R_{\alpha_1}^{-1}\right]Y_n} \\ &\quad - \frac{1}{2}\log\frac{|R_{\alpha_2}|}{|R_{\alpha_1}|} - \frac{1}{2}\log\frac{\left|M_n^\top R_{\alpha_2}^{-1}M_n + \Omega_\beta\right|}{\left|M_n^\top R_{\alpha_1}^{-1}M_n + \Omega_\beta\right|}. \end{aligned} \tag{S.102}$$

From (S.7) in the proof of Lemma S.4, we have that for any value of $Y_n \in \mathbb{R}^n$,

$$\frac{Y_n^\top\left[R_{\alpha_2}^{-1} - R_{\alpha_2}^{-1}M_n\left(M_n^\top R_{\alpha_2}^{-1}M_n + \Omega_\beta\right)^{-1}M_n^\top R_{\alpha_2}^{-1}\right]Y_n}{Y_n^\top\left[R_{\alpha_1}^{-1} - R_{\alpha_1}^{-1}M_n\left(M_n^\top R_{\alpha_1}^{-1}M_n + \Omega_\beta\right)^{-1}M_n^\top R_{\alpha_1}^{-1}\right]Y_n} \geq \left(\frac{\alpha_1}{\alpha_2}\right)^{2\nu}. \tag{S.103}$$

Similar to the proof of (S.7), now we notice that the second relation in Lemma S.3 implies that $\alpha_1^{-d}R_{\alpha_1}^{-1} > \alpha_2^{-d}R_{\alpha_2}^{-1}$ for any $0 < \alpha_1 < \alpha_2 < \infty$. Therefore, we apply Lemma S.2 with $A_1 = \alpha_2^{-d}R_{\alpha_2}^{-1}$, $A_2 = \alpha_1^{-d}R_{\alpha_1}^{-1}$, $G = M_n$, and $\Omega = \alpha_2^{-d}\Omega_\beta$ to obtain that

$$\begin{aligned}0_{n\times n} &\overset{(i)}{\leq} \left[\alpha_1^{-d}R_{\alpha_1}^{-1} - \alpha_1^{-d}R_{\alpha_1}^{-1}M_n\left(\alpha_1^{-d}M_n^\top R_{\alpha_1}^{-1}M_n + \alpha_2^{-d}\Omega_\beta\right)^{-1}M_n^\top\left(\alpha_1^{-d}R_{\alpha_1}^{-1}\right)\right] \\ &\quad - \left[\alpha_2^{-d}R_{\alpha_2}^{-1} - \alpha_2^{-d}R_{\alpha_2}^{-1}M_n\left(\alpha_2^{-d}M_n^\top R_{\alpha_2}^{-1}M_n + \alpha_2^{-d}\Omega_\beta\right)^{-1}M_n^\top\left(\alpha_2^{-d}R_{\alpha_2}^{-1}\right)\right] \\ &\overset{(ii)}{\leq} \left[\alpha_1^{-d}R_{\alpha_1}^{-1} - \alpha_1^{-d}R_{\alpha_1}^{-1}M_n\left(\alpha_1^{-d}M_n^\top R_{\alpha_1}^{-1}M_n + \alpha_1^{-d}\Omega_\beta\right)^{-1}M_n^\top\left(\alpha_1^{-d}R_{\alpha_1}^{-1}\right)\right] \\ &\quad - \left[\alpha_2^{-d}R_{\alpha_2}^{-1} - \alpha_2^{-d}R_{\alpha_2}^{-1}M_n\left(\alpha_2^{-d}M_n^\top R_{\alpha_2}^{-1}M_n + \alpha_2^{-d}\Omega_\beta\right)^{-1}M_n^\top\left(\alpha_2^{-d}R_{\alpha_2}^{-1}\right)\right] \\ &= \alpha_1^{-d}\left[R_{\alpha_1}^{-1} - R_{\alpha_1}^{-1}M_n\left(M_n^\top R_{\alpha_1}^{-1}M_n + \Omega_\beta\right)^{-1}M_n^\top R_{\alpha_1}^{-1}\right] \\ &\quad - \alpha_2^{-d}\left[R_{\alpha_2}^{-1} - R_{\alpha_2}^{-1}M_n\left(M_n^\top R_{\alpha_2}^{-1}M_n + \Omega_\beta\right)^{-1}M_n^\top R_{\alpha_2}^{-1}\right], \end{aligned} \tag{S.104}$$

where (i) follows from the conclusion of Lemma S.2 and (ii) follows from replacing $\alpha_2^{-d}\Omega_\beta$ inside the first inverse by $\alpha_1^{-d}\Omega_\beta$. This implies that the right-hand side of (S.104) is positive semidefinite. Therefore, we have that if $\alpha_1 < \alpha_2$, then for any value of $Y_n \in \mathbb{R}^n$,

$$\frac{Y_n^\top \left[R_{\alpha_2}^{-1} - R_{\alpha_2}^{-1} M_n \left(M_n^\top R_{\alpha_2}^{-1} M_n + \Omega_\beta\right)^{-1} M_n^\top R_{\alpha_2}^{-1}\right] Y_n}{Y_n^\top \left[R_{\alpha_1}^{-1} - R_{\alpha_1}^{-1} M_n \left(M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right)^{-1} M_n^\top R_{\alpha_1}^{-1}\right] Y_n} \leq \left(\frac{\alpha_2}{\alpha_1}\right)^d. \tag{S.105}$$

Using Lemma S.3 again, we can see that $\alpha_2^{2\nu} R_{\alpha_2}^{-1} > \alpha_1^{2\nu} R_{\alpha_1}^{-1}$ and $\alpha_2^d R_{\alpha_2} > \alpha_1^d R_{\alpha_1}$ imply

$$\left(\frac{\alpha_1}{\alpha_2}\right)^{nd} \leq \frac{|R_{\alpha_2}|}{|R_{\alpha_1}|} \leq \left(\frac{\alpha_2}{\alpha_1}\right)^{2n\nu}. \tag{S.106}$$

Next we find upper and lower bounds for the last term in (S.102) involving $\left|M_n^\top R_\alpha^{-1} M_n + \Omega_\beta\right|$. We first notice that

$$\frac{\left|M_n^\top R_{\alpha_2}^{-1} M_n + \Omega_\beta\right|}{\left|M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right|} = \left|\left(M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right)^{-1} \left(M_n^\top R_{\alpha_2}^{-1} M_n + \Omega_\beta\right)\right|. \tag{S.107}$$

For a lower bound of this ratio, we use the result of Lemma S.3 that $\alpha_2^{2\nu} R_{\alpha_2}^{-1} > \alpha_1^{2\nu} R_{\alpha_1}^{-1}$ if $\alpha_1 < \alpha_2$ and derive that

$$\begin{aligned}
&\frac{\left|M_n^\top R_{\alpha_2}^{-1} M_n + \Omega_\beta\right|}{\left|M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right|} = \left|\left(M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right)^{-1} \left(M_n^\top R_{\alpha_2}^{-1} M_n + \Omega_\beta\right)\right| \\
&= \left|\left(M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right)^{-1} \left[\alpha_2^{-2\nu}\left(\alpha_2^{2\nu} M_n^\top R_{\alpha_2}^{-1} M_n - \alpha_1^{2\nu} M_n^\top R_{\alpha_1}^{-1} M_n + \alpha_1^{2\nu} M_n^\top R_{\alpha_1}^{-1} M_n\right) + \Omega_\beta\right]\right| \\
&\overset{(i)}{\geq} \left|\left(M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right)^{-1} \left[\left(\frac{\alpha_1}{\alpha_2}\right)^{2\nu} M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right]\right| \\
&\overset{(ii)}{\geq} \left|\left(M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right)^{-1} \left(\frac{\alpha_1}{\alpha_2}\right)^{2\nu} \left[M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right]\right| = \left(\frac{\alpha_1}{\alpha_2}\right)^{2p\nu},
\end{aligned} \tag{S.108}$$

where (i) follows from that $\alpha_2^{2\nu} M_n^\top R_{\alpha_2}^{-1} M_n - \alpha_1^{2\nu} M_n^\top R_{\alpha_1}^{-1} M_n$ is positive semidefinite and that the determinant $|A + B| \geq |B|$ if both A and B are positive semidefinite matrices, and (ii) follows from $(\alpha_1/\alpha_2)^{2\nu} < 1$ and that the matrix inside the determinant is $p \times p$.

Similarly, we have the upper bound from Lemma S.3 that $\alpha_2^{-d} R_{\alpha_2}^{-1} < \alpha_1^{-d} R_{\alpha_1}^{-1}$ if $\alpha_1 < \alpha_2$:

$$\begin{aligned}
&\frac{\left|M_n^\top R_{\alpha_2}^{-1} M_n + \Omega_\beta\right|}{\left|M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right|} = \left|\left(M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right)^{-1} \left(M_n^\top R_{\alpha_2}^{-1} M_n + \Omega_\beta\right)\right| \\
&= \left|\left(M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right)^{-1} \left[\alpha_2^{d}\left(\alpha_2^{-d} M_n^\top R_{\alpha_2}^{-1} M_n - \alpha_1^{-d} M_n^\top R_{\alpha_1}^{-1} M_n + \alpha_1^{-d} M_n^\top R_{\alpha_1}^{-1} M_n\right) + \Omega_\beta\right]\right| \\
&\leq \left|\left(M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right)^{-1} \left[\left(\frac{\alpha_2}{\alpha_1}\right)^{d} M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right]\right| \\
&\leq \left|\left(M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right)^{-1} \left(\frac{\alpha_2}{\alpha_1}\right)^{d} \left[M_n^\top R_{\alpha_1}^{-1} M_n + \Omega_\beta\right]\right| = \left(\frac{\alpha_2}{\alpha_1}\right)^{pd}.
\end{aligned} \tag{S.109}$$

Therefore, we can combine the inequalities in (S.103), (S.105), (S.106), (S.108), and (S.109) with (S.102) to conclude that for any $0 < \alpha_1 < \alpha_2 < \infty$,

$$\begin{aligned}
&\widetilde{\mathcal{L}}_n(\alpha_2) - \widetilde{\mathcal{L}}_n(\alpha_1) \\
&\geq -\frac{n-p}{2} \log\left(\frac{\alpha_2}{\alpha_1}\right)^d - \frac{1}{2}\log\left(\frac{\alpha_2}{\alpha_1}\right)^{2n\nu} - \frac{1}{2}\log\left(\frac{\alpha_2}{\alpha_1}\right)^{pd} = n(\nu + d/2)\log\left(\frac{\alpha_1}{\alpha_2}\right),
\end{aligned}$$

$$\begin{aligned}&\widetilde{\mathcal{L}}_n(\alpha_2) - \widetilde{\mathcal{L}}_n(\alpha_1)\\ \le& -\frac{n-p}{2}\log\left(\frac{\alpha_1}{\alpha_2}\right)^{2\nu} - \frac{1}{2}\log\left(\frac{\alpha_1}{\alpha_2}\right)^{nd} - \frac{1}{2}\log\left(\frac{\alpha_1}{\alpha_2}\right)^{2p\nu} = n(\nu+d/2)\log\left(\frac{\alpha_2}{\alpha_1}\right).\end{aligned}$$

Exponentiating both sides leads to the conclusion. □

The following lemma is a consequence of Lemmas S.5, S.6, S.7, S.8, S.9 in Section S1. It will be used in proving Lemma S.24, Lemma S.25 and Lemma S.27 below.

Lemma S.23. *For $\tau, \underline{\alpha}_n, \overline{\alpha}_n$ defined in* (S.9) *and $\widetilde{\theta}_\alpha, \widetilde{\theta}_\alpha^{(1)}$ defined in* (S.10)*, for $d \in \{1,2,3\}$ and $\nu \in \mathbb{R}^+$, there exists a large integer $N'_{6,1}$ that only depends on $\nu, d, T, \beta_0, \theta_0, \alpha_0$ and the $\mathcal{W}_2^{\nu+d/2}(\mathcal{S})$ norms of $\mathrm{m}_1(\cdot), \ldots, \mathrm{m}_p(\cdot)$, such that for all $n > N'_{6,1}$,*

$$\Pr\left(\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \frac{\left|\widetilde{\theta}_\alpha - \widetilde{\theta}_\alpha^{(1)}\right|}{\widetilde{\theta}_\alpha^{(1)}} \le 2n^{-1/2-\tau}\right) \ge 1 - 10\exp(-4\log^2 n). \tag{S.110}$$

Furthermore, for any given $c \ge 1/(2\nu+d)$, for all $d \in \mathbb{Z}^+$ and $\nu \in \mathbb{R}^+$, there exists a large integer $N'_{6,2}$ that only depends on $c, \nu, d, T, \beta_0, \theta_0, \alpha_0$ and the $\mathcal{W}_2^{\nu+d/2}(\mathcal{S})$ norms of $\mathrm{m}_1(\cdot), \ldots, \mathrm{m}_p(\cdot)$, such that for all $n > N'_{6,2}$,

$$\begin{aligned}&\Pr\left(\sup_{\alpha\in\left[(1-n^{-c})\alpha_0,(1+n^{-c})\alpha_0\right]} \frac{\left|\widetilde{\theta}_\alpha - \widetilde{\theta}_\alpha^{(1)}\right|}{\widetilde{\theta}_\alpha^{(1)}} \le n^{-1}\log^4 n\right) \ge 1 - 8\exp(-4\log^2 n),\\ &\Pr\left(\sup_{\alpha\in\left[(1-n^{-c})\alpha_0,(1+n^{-c})\alpha_0\right]} \left|\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}\right| \le 10\theta_0 n^{-(2\nu+d)c}\right) \ge 1 - 8\exp(-4\log^2 n).\end{aligned} \tag{S.111}$$

Proof of Lemma S.23. <u>Proof of (S.110):</u>

We consider the case of $d \in \{1,2,3\}$. From the inequalities (S.12), (S.19), (S.40), (S.41) and (S.42), a simple union bound shows that for all sufficiently large n,

$$\begin{aligned}&\Pr\Big(\sqrt{n}\widetilde{\theta}_{\alpha_0}^{(2)} \le \frac{\theta_0}{16}n^{-\tau},\ \sqrt{n}\widetilde{\theta}_{\alpha_0}^{(3)} \le \frac{\theta_0}{16}n^{-\tau},\ \sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]}\sqrt{n}\left|\widetilde{\theta}_\alpha^{(1)} - \widetilde{\theta}_{\alpha_0}^{(1)}\right| \le \frac{\theta_0}{4}n^{-\tau},\\ &\quad \sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]}\sqrt{n}\left|\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}\right| \le \frac{\theta_0}{2}n^{-\tau},\ \sqrt{n}\left|\widetilde{\theta}_{\alpha_0} - \theta_0\right| \le 5\theta_0\log n\Big)\\ &\ge 1 - \exp(-16\log^2 n) - \exp(-16\log^2 n) - 2\exp(-4\log^2 n)\\ &\quad - 4\exp(-4\log^2 n) - 3\exp(-4\log^2 n) > 1 - 10\exp(-4\log^2 n).\end{aligned} \tag{S.112}$$

From Lemma S.5, we have $\widetilde{\theta}_\alpha = \widetilde{\theta}_\alpha^{(1)} - \widetilde{\theta}_\alpha^{(2)} + \widetilde{\theta}_\alpha^{(3)}$, $\widetilde{\theta}_\alpha^{(1)} \ge \widetilde{\theta}_\alpha^{(2)} \ge 0$, and $\widetilde{\theta}_\alpha^{(3)} \ge 0$ for all $\alpha \in \mathbb{R}^+$. Therefore, with probability at least $1 - 10\exp(-4\log^2 n)$, uniformly over all $\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]$,

$$\begin{aligned}\frac{\left|\widetilde{\theta}_\alpha - \widetilde{\theta}_\alpha^{(1)}\right|}{\widetilde{\theta}_\alpha^{(1)}} &= \frac{\left|\left(\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}\right) - \left(\widetilde{\theta}_\alpha^{(1)} - \widetilde{\theta}_{\alpha_0}^{(1)}\right) - \widetilde{\theta}_{\alpha_0}^{(2)} + \widetilde{\theta}_{\alpha_0}^{(3)}\right|}{\left(\widetilde{\theta}_\alpha^{(1)} - \widetilde{\theta}_{\alpha_0}^{(1)}\right) + \widetilde{\theta}_{\alpha_0}^{(2)} - \widetilde{\theta}_{\alpha_0}^{(3)} + \left(\widetilde{\theta}_{\alpha_0} - \theta_0\right) + \theta_0}\\ &\le \frac{\left|\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}\right| + \left|\widetilde{\theta}_\alpha^{(1)} - \widetilde{\theta}_{\alpha_0}^{(1)}\right| + \widetilde{\theta}_{\alpha_0}^{(2)} + \widetilde{\theta}_{\alpha_0}^{(3)}}{\theta_0 - \left|\widetilde{\theta}_\alpha^{(1)} - \widetilde{\theta}_{\alpha_0}^{(1)}\right| - \left|\widetilde{\theta}_{\alpha_0} - \theta_0\right| - \widetilde{\theta}_{\alpha_0}^{(2)} - \widetilde{\theta}_{\alpha_0}^{(3)}}\\ &\le \frac{(\theta_0/2)n^{-\frac{1}{2}-\tau} + (\theta_0/4)n^{-\frac{1}{2}-\tau} + (\theta_0/16)n^{-\frac{1}{2}-\tau} + (\theta_0/16)n^{-\frac{1}{2}-\tau}}{\theta_0 - (\theta_0/4)n^{-\frac{1}{2}-\tau} - 5\theta_0 n^{-1/2}\log n - (\theta_0/16)n^{-\frac{1}{2}-\tau} - (\theta_0/16)n^{-\frac{1}{2}-\tau}}\end{aligned}$$

$$\leq 2n^{-1/2-\tau}.$$

Proof of (S.111):

Now we consider the case of $d \in \mathbb{Z}^+$ and change the interval of supremum to $[(1-n^{-c})\alpha_0, (1+n^{-c})\alpha_0]$. According to (S.60) and (S.61) in Lemma S.15, if $\alpha \in [\alpha_0, (1+n^{-c})\alpha_0]$, then for all $k = 1, \ldots, n$ and all sufficiently large n,

$$1 \geq \lambda_{k,n}(\alpha) \geq \left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d} \geq (1+n^{-c})^{-(2\nu+d)} \geq 1 - 2n^{-(2\nu+d)c} > \frac{1}{2}. \tag{S.113}$$

If $\alpha \in [(1-n^{-c})\alpha_0, \alpha_0]$, then for all $k = 1, \ldots, n$ and all sufficiently large n,

$$1 \leq \lambda_{k,n}(\alpha) \leq \left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d} \leq (1-n^{-c})^{-(2\nu+d)} \leq 1 + 2n^{-(2\nu+d)c} < 2. \tag{S.114}$$

For short, we let $\alpha_{1n} = (1-n^{-c})\alpha_0$ and $\alpha_{2n} = (1+n^{-c})\alpha_0$. Following a similar argument to the proof of Lemmas S.6 and S.7, we can show that for all sufficiently large n, with probability $1 - 6\exp(-16\log^2 n)$,

$$\begin{aligned}
&\widetilde{\theta}^{(2)}_{\alpha_0} \leq (\theta_0/16)n^{-1}\log^3 n, \qquad \widetilde{\theta}^{(3)}_{\alpha_0} \leq (\theta_0/16)n^{-1}\log^3 n, \\
&\widetilde{\theta}^{(2)}_{\alpha_{1n}} \leq (\theta_0/16)n^{-1}\log^3 n, \quad \text{and } \widetilde{\theta}^{(3)}_{\alpha_{1n}} \leq (\theta_0/16)n^{-1}\log^3 n \\
&\widetilde{\theta}^{(2)}_{\alpha_{2n}} \leq (\theta_0/16)n^{-1}\log^3 n, \quad \text{and } \widetilde{\theta}^{(3)}_{\alpha_{2n}} \leq (\theta_0/16)n^{-1}\log^3 n.
\end{aligned} \tag{S.115}$$

For $\widetilde{\theta}^{(1)}_{\alpha_{2n}}$, we first notice that by Lemma S.4, $\widetilde{\theta}^{(1)}_{\alpha_0} \leq \widetilde{\theta}^{(1)}_{\alpha} \leq \widetilde{\theta}^{(1)}_{\alpha_{2n}}$ for all $\alpha \in [\alpha_0, (1+n^{-c})\alpha_0]$. Similar to (S.34) in the proof of Lemma S.8, we have that

$$\widetilde{\theta}^{(1)}_{\alpha_{2n}} - \widetilde{\theta}^{(1)}_{\alpha_0} = \frac{\theta_0}{n-p}\sum_{i=1}^{n}\left\{\lambda_{i,n}(\alpha_{2n})^{-1} - 1\right\} Z_{i,n}(\alpha_{2n})^2, \tag{S.116}$$

where $Z_n(\alpha) = (Z_{1,n}(\alpha), \ldots, Z_{n,n}(\alpha))^\top = U_\alpha^\top X_n$ with U_α given in Lemma S.14. We let $w = (w_1, \ldots, w_n)^\top$ with $w_i = \frac{\theta_0}{n-p}\left|\lambda_{i,n}(\alpha_{2n})^{-1} - 1\right|$ for $i = 1, \ldots, n$. Then by (S.114), we have

$$\begin{aligned}
\|w\|_1 &\leq \frac{2\theta_0}{n}\frac{\sum_{i=1}^n [1 - \lambda_{i,n}(\alpha_{2n})]}{\min_{1\leq i\leq n}\lambda_{i,n}(\alpha_{2n})} \leq \frac{8\theta_0 n \cdot n^{-(2\nu+d)c}}{n} = 8\theta_0 n^{-(2\nu+d)c}, \\
\|w\| &\leq \frac{2\theta_0}{n}\frac{\left\{\sum_{i=1}^n [1 - \lambda_{i,n}(\alpha_{2n})]^2\right\}^{1/2}}{\min_{1\leq i\leq n}\lambda_{i,n}(\alpha_{2n})} \leq \frac{4\theta_0\left(n\cdot 4n^{-2(2\nu+d)c}\right)^{1/2}}{n} = 8\theta_0 n^{-1/2-(2\nu+d)c}, \\
\|w\|_\infty &\leq \frac{2\theta_0}{n}\frac{\max_{1\leq i\leq n}[1 - \lambda_{i,n}(\alpha_{2n})]}{\min_{1\leq i\leq n}\lambda_{i,n}(\alpha_{2n})} \leq 8\theta_0 n^{-1-(2\nu+d)c}.
\end{aligned}$$

Therefore, if we apply the first inequality in Lemma S.19 with $z = 16\log^2 n$ and w_i's given as above, we obtain that for all sufficiently large n,

$$\begin{aligned}
&\Pr\left(\sup_{\alpha\in[\alpha_0,(1+n^{-c})\alpha_0]}\left(\widetilde{\theta}^{(1)}_{\alpha} - \widetilde{\theta}^{(1)}_{\alpha_0}\right) > 9\theta_0 n^{-(2\nu+d)c}\right) \\
&= \Pr\left(\widetilde{\theta}^{(1)}_{\alpha_{2n}} - \widetilde{\theta}^{(1)}_{\alpha_0} > 9\theta_0 n^{-(2\nu+d)c}\right) \\
&\leq \Pr\left(\widetilde{\theta}^{(1)}_{\alpha_{2n}} - \widetilde{\theta}^{(1)}_{\alpha_0} > \|w\|_1 + 8\|w\|\log n + 32\|w\|_\infty\log^2 n\right) \\
&\leq \exp(-16\log^2 n).
\end{aligned} \tag{S.117}$$

Similarly we can show from (S.113) that

$$\Pr\left(\sup_{\alpha\in[(1-n^{-c})\alpha_0,\alpha_0]}\left(\widetilde{\theta}^{(1)}_{\alpha} - \widetilde{\theta}^{(1)}_{\alpha_0}\right) > 9\theta_0 n^{-(2\nu+d)c}\right) \leq \exp(-16\log^2 n). \tag{S.118}$$

(S.117) and (S.118) together imply that for all sufficiently large n,

$$\Pr\left(\sup_{\alpha\in\left[(1-n^{-c})\alpha_0,(1+n^{-c})\alpha_0\right]}\left|\widetilde{\theta}_{\alpha}^{(1)}-\widetilde{\theta}_{\alpha_0}^{(1)}\right|>9\theta_0 n^{-(2\nu+d)c}\right)\leq 2\exp(-16\log^2 n). \qquad \text{(S.119)}$$

Finally, from Lemma S.5, Lemma S.10, (S.115), (S.117) and (S.118), we obtain that for all sufficiently large n, with probability at least $1-8\exp(-4\log^2 n)$, uniformly over all $\alpha\in[\alpha_0,(1+n^{-c})\alpha_0]$,

$$\begin{aligned}
\frac{\left|\widetilde{\theta}_{\alpha}-\widetilde{\theta}_{\alpha}^{(1)}\right|}{\widetilde{\theta}_{\alpha}^{(1)}} &\leq \frac{\left(\widetilde{\theta}_{\alpha_{2n}}-\widetilde{\theta}_{\alpha_0}\right)+\left(\widetilde{\theta}_{\alpha_{2n}}^{(1)}-\widetilde{\theta}_{\alpha_0}^{(1)}\right)+\widetilde{\theta}_{\alpha_0}^{(2)}+\widetilde{\theta}_{\alpha_0}^{(3)}}{\theta_0-\left(\widetilde{\theta}_{\alpha_{2n}}^{(1)}-\widetilde{\theta}_{\alpha_0}^{(1)}\right)-\left|\widetilde{\theta}_{\alpha_0}-\theta_0\right|-\widetilde{\theta}_{\alpha_0}^{(2)}-\widetilde{\theta}_{\alpha_0}^{(3)}} \\
&\leq \frac{2\left(\widetilde{\theta}_{\alpha_{2n}}^{(1)}-\widetilde{\theta}_{\alpha_0}^{(1)}\right)+2\widetilde{\theta}_{\alpha_0}^{(2)}+2\widetilde{\theta}_{\alpha_0}^{(3)}+\widetilde{\theta}_{\alpha_{2n}}^{(2)}+\widetilde{\theta}_{\alpha_{2n}}^{(3)}}{\theta_0-\left(\widetilde{\theta}_{\alpha_{2n}}^{(1)}-\widetilde{\theta}_{\alpha_0}^{(1)}\right)-\left|\widetilde{\theta}_{\alpha_0}-\theta_0\right|-\widetilde{\theta}_{\alpha_0}^{(2)}-\widetilde{\theta}_{\alpha_0}^{(3)}} \\
&\leq \frac{18\theta_0 n^{-(2\nu+d)c}+(\theta_0/4)n^{-1}\log^3 n+(\theta_0/8)n^{-1}\log^3 n}{\theta_0-9\theta_0 n^{-(2\nu+d)c}-5\theta_0 n^{-1/2}\log n-(\theta_0/8)n^{-1}\log^3 n} \\
&\leq n^{-\min\{(2\nu+d)c,1\}}\log^4 n \overset{(i)}{=} n^{-1}\log^4 n,
\end{aligned}$$

and similarly for all $\alpha\in[(1-n^{-c})\alpha_0,\alpha_0]$, $\left|\widetilde{\theta}_{\alpha}-\widetilde{\theta}_{\alpha}^{(1)}\right|/\widetilde{\theta}_{\alpha}^{(1)}\leq n^{-1}\log^4 n$. The step (i) follows from our condition $c\geq 1/(2\nu+d)$. This proves the first inequality in (S.111). The second inequality in (S.111) follows from combining the first inequality with (S.119). □

Lemma S.24. *For $\tau,\underline{\alpha}_n,\overline{\alpha}_n$ defined in (S.9), for all $d\in\mathbb{Z}^+,\nu\in\mathbb{R}^+$, for any $c>1/(2\nu+d)$, there exists a large integer N_7' that only depends on $c,\nu,d,T,\beta_0,\theta_0,\alpha_0$ and the $\mathcal{W}_2^{\nu+d/2}(\mathcal{S})$ norms of $\mathrm{m}_1(\cdot),\ldots,\mathrm{m}_p(\cdot)$, such that with probability at least $1-9\exp(-4\log^2 n)$, for all $n>N_7'$,*

$$\inf_{\alpha\in[\alpha_0,(1+n^{-c})\alpha_0]}\exp\left\{\widetilde{\mathcal{L}}_n(\alpha)-\widetilde{\mathcal{L}}_n(\alpha_0)\right\}\geq\exp\left(-3\log^4 n\right). \qquad \text{(S.120)}$$

Proof of Lemma S.24. Let $\overline{\lambda}_n(\alpha)=\{\prod_{k=1}^n\lambda_{k,n}(\alpha)\}^{1/n}$. (S.113) implies that for all $\alpha\in[\alpha_0,(1+n^{-c})\alpha_0]$, $\overline{\lambda}_n(\alpha)\leq 1$. Let $Z_n(\alpha)=U_\alpha^\top X_n=(Z_{1,n}(\alpha),\ldots,Z_{n,n}(\alpha))^\top\sim\mathcal{N}(0_n,I_n)$ for any given $\alpha>0$, where U_α is given in (S.54) of Lemma S.14. Then using (S.54) in Lemma S.14 and the definition $\widetilde{\theta}_{\alpha}^{(1)}$ in (S.10) in Lemma S.5, we have that

$$\begin{aligned}
& -\frac{n-p}{2}\log\frac{\alpha^{-2\nu}\widetilde{\theta}_{\alpha}^{(1)}}{\alpha_0^{-2\nu}\widetilde{\theta}_{\alpha_0}^{(1)}}-\frac{1}{2}\log\frac{|R_\alpha|}{|R_{\alpha_0}|} \\
={}& -\frac{n-p}{2}\log\frac{\alpha^{-2\nu}X_n^\top U_\alpha\Lambda_\alpha^{-1}U_\alpha^\top X_n}{\alpha_0^{-2\nu}X_n^\top U_\alpha U_\alpha^\top X_n}-\frac{1}{2}\log\frac{\alpha^{2\nu n}\prod_{k=1}^n\lambda_{k,n}(\alpha)}{|U_\alpha|^2}+\frac{1}{2}\log\frac{\alpha_0^{2\nu n}}{|U_\alpha|^2} \\
={}& -\frac{n-p}{2}\log\frac{\sum_{k=1}^n\lambda_{k,n}(\alpha)^{-1}Z_{k,n}(\alpha)^2}{\sum_{k=1}^n Z_{k,n}(\alpha)^2}-\frac{1}{2}\sum_{k=1}^n\log\lambda_{k,n}(\alpha)-p\nu\log\frac{\alpha}{\alpha_0}. \qquad \text{(S.121)}
\end{aligned}$$

Denote the event on the left-hand side of the first inequality in (S.111) in Lemma S.23 as $\mathcal{A}_{1n}$ such that $\Pr(\mathcal{A}_{1n})\geq 1-8\exp(-4\log^2 n)$ given the condition $c>1/(2\nu+d)$. Then from the expression (S.101) and the relation (S.54), we have that on the event $\mathcal{A}_{1n}$, uniformly over all $\alpha\in[\alpha_0,(1+n^{-c})\alpha_0]$,

$$\widetilde{\mathcal{L}}_n(\alpha)-\widetilde{\mathcal{L}}_n(\alpha_0)$$

$$
\begin{aligned}
&= -\frac{n-p}{2}\log \frac{Y_n^\top \left[R_\alpha^{-1} - R_\alpha^{-1}M_n\big(M_n^\top R_\alpha^{-1}M_n + \Omega_\beta\big)^{-1}M_n^\top R_\alpha^{-1}\right]Y_n}{Y_n^\top \left[R_{\alpha_0}^{-1} - R_{\alpha_0}^{-1}M_n\big(M_n^\top R_{\alpha_0}^{-1}M_n + \Omega_\beta\big)^{-1}M_n^\top R_{\alpha_0}^{-1}\right]Y_n} \\
&\quad - \frac{1}{2}\log\frac{|R_\alpha|}{|R_{\alpha_0}|} - \frac{1}{2}\log\frac{\left|M_n^\top R_{\alpha_0}^{-1}M_n + \Omega_\beta\right|}{\left|M_n^\top R_{\alpha_0}^{-1}M_n + \Omega_\beta\right|} \\
&= -\frac{n-p}{2}\log\frac{\alpha^{-2\nu}\widetilde{\theta}_\alpha}{\alpha_0^{-2\nu}\widetilde{\theta}_{\alpha_0}} - \frac{1}{2}\log\frac{|R_\alpha|}{|R_{\alpha_0}|} - \frac{1}{2}\log\frac{\left|M_n^\top R_{\alpha_0}^{-1}M_n + \Omega_\beta\right|}{\left|M_n^\top R_{\alpha_0}^{-1}M_n + \Omega_\beta\right|} \\
&\overset{(i)}{\geq} -\frac{n-p}{2}\log\frac{\alpha^{-2\nu}\widetilde{\theta}_\alpha^{(1)}\left(1+n^{-1}\log^4 n\right)}{\alpha_0^{-2\nu}\widetilde{\theta}_{\alpha_0}^{(1)}\left(1-n^{-1}\log^4 n\right)} - \frac{1}{2}\log\frac{|R_\alpha|}{|R_{\alpha_0}|} - \frac{1}{2}\log\frac{\left|M_n^\top R_\alpha^{-1}M_n + \Omega_\beta\right|}{\left|M_n^\top R_{\alpha_0}^{-1}M_n + \Omega_\beta\right|} \\
&\overset{(ii)}{=} -\frac{n-p}{2}\log\frac{\sum_{k=1}^n \lambda_{k,n}(\alpha)^{-1}Z_{k,n}(\alpha)^2}{\sum_{k=1}^n Z_{k,n}(\alpha)^2} - \frac{1}{2}\sum_{k=1}^n \log\lambda_{k,n}(\alpha) - p\nu\log\frac{\alpha}{\alpha_0} \\
&\quad + \frac{n-p}{2}\log\frac{1-n^{-1}\log^4 n}{1+n^{-1}\log^4 n} - \frac{1}{2}\log\frac{\left|M_n^\top R_\alpha^{-1}M_n + \Omega_\beta\right|}{\left|M_n^\top R_{\alpha_0}^{-1}M_n + \Omega_\beta\right|} \\
&\overset{(iii)}{\geq} -\frac{n-p}{2}\log\frac{\sum_{k=1}^n \lambda_{k,n}(\alpha)^{-1}Z_{k,n}(\alpha)^2}{\sum_{k=1}^n Z_{k,n}(\alpha)^2} - \frac{1}{2}\sum_{k=1}^n \log\lambda_{k,n}(\alpha) - \frac{p\nu}{2\nu+d}\log 2 \\
&\quad - 2\log^4 n - \frac{pd\log 2}{2(2\nu+d)} \\
&= -\frac{n-p}{2}\log\frac{\sum_{k=1}^n \lambda_{k,n}(\alpha)^{-1}Z_{k,n}(\alpha)^2}{\sum_{k=1}^n Z_{k,n}(\alpha)^2} - \frac{1}{2}\sum_{k=1}^n \log\lambda_{k,n}(\alpha) - 2\log^4 n - \frac{p}{2}\log 2, \qquad \text{(S.122)}
\end{aligned}
$$

where (i) follows from (S.111) in Lemma S.23; (ii) follows from (S.121); to derive (iii), we first apply

$$
\frac{n-p}{2}\log\frac{1-n^{-1}\log^4 n}{1+n^{-1}\log^4 n} \geq \frac{n}{2}\cdot\left(-3n^{-1}\log^4 n\right) = -2\log^4 n, \qquad \text{(S.123)}
$$

for all sufficiently large n, then notice that $p\nu\log(\alpha/\alpha_0) \leq \frac{p\nu}{2\nu+d}\log 2$ for all $\alpha \in [\alpha_0, (1+n^{-c})\alpha_0] \subseteq [\alpha_0, 2^{1/(2\nu+d)}\alpha_0]$, and finally apply (S.109) to obtain that

$$
-\frac{1}{2}\log\frac{\left|M_n^\top R_\alpha^{-1}M_n + \Omega_\beta\right|}{\left|M_n^\top R_{\alpha_0}^{-1}M_n + \Omega_\beta\right|} \geq -\frac{1}{2}\log\left(\frac{\alpha}{\alpha_0}\right)^{pd} \geq -\frac{1}{2}\log 2^{pd/(2\nu+d)} = \frac{pd\log 2}{2(2\nu+d)},
$$

for all $\alpha \in [\alpha_0, (1+n^{-c})\alpha_0] \subseteq [\alpha_0, 2^{1/(2\nu+d)}\alpha_0]$.

Now we further control the first two terms on the right-hand side of (S.122). Since $\overline{\lambda}_n(\alpha) \leq 1$ for all $\alpha \in [\alpha_0, (1+n^{-c})\alpha_0]$, we have that

$$
\begin{aligned}
&\exp\left\{-\frac{n-p}{2}\log\frac{\sum_{k=1}^n \lambda_{k,n}(\alpha)^{-1}Z_{k,n}(\alpha)^2}{\sum_{k=1}^n Z_{k,n}(\alpha)^2} - \frac{1}{2}\sum_{k=1}^n \log\lambda_{k,n}(\alpha)\right\} \\
&= \left[\frac{\sum_{k=1}^n \lambda_{k,n}(\alpha)^{-1}Z_{k,n}(\alpha)^2}{\sum_{k=1}^n Z_{k,n}(\alpha)^2}\cdot\left\{\prod_{k=1}^n \lambda_{k,n}(\alpha)\right\}^{1/(n-p)}\right]^{-(n-p)/2} \\
&\geq \left[\frac{\sum_{k=1}^n \lambda_{k,n}(\alpha)^{-1}Z_{k,n}(\alpha)^2}{\sum_{k=1}^n Z_{k,n}(\alpha)^2}\right]^{-(n-p)/2} \\
&= \left[1 + \frac{\sum_{k=1}^n \left\{\lambda_{k,n}(\alpha)^{-1} - 1\right\}Z_{k,n}(\alpha)^2}{\sum_{k=1}^n Z_{k,n}(\alpha)^2}\right]^{-(n-p)/2}. \qquad \text{(S.124)}
\end{aligned}
$$

By (S.116) and (S.117) in the proof of Lemma S.23, we have that on the event $\mathcal{A}_{1n}$,

$$\sup_{\alpha\in[\alpha_0,(1+n^{-c})\alpha_0]} \sum_{k=1}^n \left\{\lambda_{k,n}(\alpha)^{-1}-1\right\} Z_{k,n}(\alpha)^2 \\ \leq \sup_{\alpha\in[\alpha_0,(1+n^{-c})\alpha_0]} (n-p)\left(\widetilde{\theta}^{(1)}_{\alpha}-\widetilde{\theta}^{(1)}_{\alpha_0}\right)/\theta_0 \leq 9n^{1-(2\nu+d)c}. \tag{S.125}$$

On the other hand, for any $\alpha>0$,

$$\sum_{k=1}^n Z_{k,n}(\alpha)^2 = Z_n(\alpha)^\top Z_n(\alpha) = X_n^\top U_\alpha U_\alpha^\top X_n = X_n^\top(\sigma_0^2 R_{\alpha_0})^{-1}X_n \\ = W_n^\top W_n = \sum_{k=1}^n W_{k,n}^2, \tag{S.126}$$

where $W_n=(W_{1,n},\ldots,W_{n,n})^\top = \sigma_0^{-1}R_{\alpha_0}^{-1/2}X_n \sim \mathcal{N}(0_n,I_n)$. Therefore, we apply the second inequality in Lemma S.19 directly to the χ^2_1 random variables of $\{W_{k,n}^2 : k=1,\ldots,n\}$ with $z=4\log^2 n$ and obtain that for all sufficiently large n,

$$\Pr\left(\inf_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]}\sum_{k=1}^n Z_{k,n}(\alpha)^2 \leq n-4\sqrt{n}\log n\right) \\ = \Pr\left(\sum_{k=1}^n W_{k,n}^2 \leq n-4\sqrt{n}\log n\right) \leq \exp(-4\log^2 n). \tag{S.127}$$

We combine (S.122), (S.124), (S.125) and (S.127) to obtain that with probability at least $1-9\exp(-4\log^2 n)$, uniformly for all $\alpha\in[\alpha_0,(1+n^{-c})\alpha_0]$ and for all sufficiently large n,

$$\inf_{\alpha\in[\alpha_0,(1+n^{-c})\alpha_0]} \exp\left\{\widetilde{\mathcal{L}}_n(\alpha)-\widetilde{\mathcal{L}}_n(\alpha_0)\right\} \\ \geq \left[1+\frac{\sup_{\alpha\in[\alpha_0,(1+n^{-c})\alpha_0]}\sum_{k=1}^n\left\{\lambda_{k,n}(\alpha)^{-1}-1\right\}Z_{k,n}(\alpha)^2}{\inf_{\alpha\in[\alpha_0,(1+n^{-c})\alpha_0]}\sum_{k=1}^n Z_{k,n}(\alpha)^2}\right]^{-(n-p)/2} \\ \times \exp\left\{-2\log^4 n-\frac{p}{2}\log 2\right\} \\ \geq \left(1+\frac{9n^{1-(2\nu+d)c}}{n-4\sqrt{n}\log n}\right)^{-(n-p)/2}\cdot\exp\left\{-2\log^4 n-\frac{p}{2}\log 2\right\} \\ \geq \left(1+\frac{10}{n^{(2\nu+d)c}}\right)^{-(n-p)/2}\cdot\exp\left\{-2\log^4 n-\frac{p}{2}\log 2\right\} \\ \overset{(i)}{\geq} \exp\left\{-10n^{1-(2\nu+d)c}-2\log^4 n-\frac{p}{2}\log 2\right\} \\ \geq \exp\left(-3\log^4 n\right), \tag{S.128}$$

where in (i), we apply the relation $(1+x^{-1})^x\leq\exp(1)$ for all $x>0$ and the condition $c>1/(2\nu+d)$. □

Lemma S.25. *For $\tau,\underline{\alpha}_n,\overline{\alpha}_n$ defined in (S.9), for $d\in\{1,2,3\}$ and $\nu\in\mathbb{R}^+$, there exists a large integer N_8' that only depends on $\nu,d,T,\beta_0,\theta_0,\alpha_0$ and the $\mathcal{W}_2^{\nu+d/2}(\mathcal{S})$ norms of $\mathrm{m}_1(\cdot),\ldots,\mathrm{m}_p(\cdot)$, such that with probability at least $1-10\exp(-4\log^2 n)$, for all $n>N_8'$,*

$$\sup_{\alpha\in[\underline{\alpha}_n,\alpha_0]}\exp\left\{\widetilde{\mathcal{L}}_n(\alpha)-\widetilde{\mathcal{L}}_n(\alpha_0)\right\} < \exp\left(3n^{1/2-\tau}\right). \tag{S.129}$$

Proof of Lemma S.25. According to (S.60) and (S.61) in Lemma S.15, we have that for all $k = 1, \ldots, n$ and all $\alpha \in [\underline{\alpha}_n, \alpha_0]$,

$$1 \le \lambda_{k,n}(\alpha) \le \left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d} \le \left(\frac{\alpha_0}{\underline{\alpha}_n}\right)^{2\nu+d}. \tag{S.130}$$

Let $\overline{\lambda}_n(\alpha) = \{\prod_{k=1}^n \lambda_{k,n}(\alpha)\}^{1/n}$. (S.130) implies that $\overline{\lambda}_n(\alpha) \ge 1$. For any $\alpha > 0$, let $Z_n(\alpha) = U_\alpha^\top X_n = (Z_{1,n}(\alpha), \ldots, Z_{n,n}(\alpha))^\top$ with U_α given in (S.54).

Denote the event on the left-hand side of (S.110) in Lemma S.23 as $\mathcal{A}_{2n}$ such that $\Pr(\mathcal{A}_{2n}) \ge 1 - 10\exp(-4\log^2 n)$. Then using the relation (S.121), we have that on the event $\mathcal{A}_{2n}$,

$$\begin{aligned}
&\exp\left\{\mathcal{L}_n(\alpha) - \widetilde{\mathcal{L}}_n(\alpha_0)\right\} \\
&\le \exp\left\{-\frac{n-p}{2}\log\frac{\alpha^{-2\nu}\widetilde{\theta}_\alpha^{(1)}\left(1-2n^{-1/2-\tau}\right)}{\alpha_0^{-2\nu}\widetilde{\theta}_{\alpha_0}^{(1)}\left(1+2n^{-1/2-\tau}\right)} - \frac{1}{2}\log\frac{|R_\alpha|}{|R_{\alpha_0}|} - \frac{1}{2}\log\frac{\left|M_n^\top R_\alpha^{-1}M_n + \Omega_\beta\right|}{\left|M_n^\top R_{\alpha_0}^{-1}M_n + \Omega_\beta\right|}\right\} \\
&\overset{(i)}{\le} \left[\frac{\sum_{k=1}^n \lambda_{k,n}(\alpha)^{-1}Z_{k,n}(\alpha)^2}{\sum_{k=1}^n Z_{k,n}(\alpha)^2}\cdot\left\{\prod_{k=1}^n \lambda_{k,n}(\alpha)\right\}^{1/(n-p)}\right]^{-(n-p)/2} \\
&\quad \times \exp\left\{2n^{1/2-\tau} - p\nu\log\frac{\alpha}{\alpha_0} + \frac{1}{2}\log\left(\frac{\alpha_0}{\alpha}\right)^{pd}\right\} \\
&\overset{(ii)}{\le} \left[\frac{\sum_{k=1}^n \lambda_{k,n}(\alpha)^{-1}Z_{k,n}(\alpha)^2}{\sum_{k=1}^n Z_{k,n}(\alpha)^2}\right]^{-(n-p)/2}\cdot\exp\left\{2n^{1/2-\tau} + \frac{p(2\nu+d)}{2}\log\left(\frac{\alpha_0}{\alpha}\right)\right\} \\
&= \left[1 + \frac{\sum_{k=1}^n\left\{\lambda_{k,n}(\alpha)^{-1}-1\right\}Z_{k,n}(\alpha)^2}{\sum_{k=1}^n Z_{k,n}(\alpha)^2}\right]^{-(n-p)/2}\cdot\exp\left\{2n^{1/2-\tau} + \frac{p(2\nu+d)}{2}\log\left(\frac{\alpha_0}{\alpha}\right)\right\},
\end{aligned} \tag{S.131}$$

where in (i), we use the inequality

$$\frac{n-p}{2}\log\frac{1+2n^{-1/2-\tau}}{1-2n^{-1/2-\tau}} \le \frac{n}{2}\cdot\left(4n^{-1/2-\tau}\right) = 2n^{1/2-\tau}, \tag{S.132}$$

for all sufficiently large n and (S.109) similar to the derivation of (S.122); in (ii) we use the fact that $\overline{\lambda}_n(\alpha) \ge 1$.

Notice that $\lambda_{k,n}^{-1}(\alpha) - 1 \le 0$ for all $k = 1, \ldots, n$ for all $\alpha \in [\underline{\alpha}_n, \alpha_0]$. Then using the relation (S.34) in the proof of Lemma S.8, on the event $\mathcal{A}_{2n}$, uniformly for all $\alpha \in [\underline{\alpha}_n, \alpha_0]$ and for all sufficiently large n,

$$\begin{aligned}
&\inf_{\alpha\in[\underline{\alpha}_n,\alpha_0]}\sum_{k=1}^n\left\{\lambda_{k,n}(\alpha)^{-1}-1\right\}Z_{k,n}(\alpha)^2 \\
&= \inf_{\alpha\in[\underline{\alpha}_n,\alpha_0]}\frac{(n-p)\left(\widetilde{\theta}_\alpha^{(1)} - \widetilde{\theta}_{\alpha_0}^{(1)}\right)}{\theta_0} \ge -n^{1/2-\tau}/4.
\end{aligned} \tag{S.133}$$

We combine (S.131), (S.133), and (S.127) together to derive that uniformly for all all $\alpha \in [\underline{\alpha}_n, \alpha_0]$, for all sufficiently large n, with probability at least $1 - 10\exp(-4\log^2 n)$,

$$\begin{aligned}
&\exp\left\{\widetilde{\mathcal{L}}_n(\alpha) - \widetilde{\mathcal{L}}_n(\alpha_0)\right\} \\
&\le \left[1 - \frac{n^{1/2-\tau}/4}{\inf_{\alpha\in[\underline{\alpha}_n,\alpha_0]}\sum_{k=1}^n Z_{k,n}(\alpha)^2}\right]^{-(n-p)/2}\cdot\exp\left\{2n^{1/2-\tau} + \inf_{\alpha\in[\underline{\alpha}_n,\alpha_0]}\frac{p(2\nu+d)}{2}\log\left(\frac{\alpha_0}{\alpha}\right)\right\}
\end{aligned}$$

$$\overset{(i)}{\le} \left(1 - \frac{n^{1/2-\tau}/4}{n - 4\sqrt{n}\log n}\right)^{-(n-p)/2} \cdot \exp\left\{2n^{1/2-\tau} + \frac{p(2\nu+d)}{2}\left(\log\alpha_0 + \underline{\kappa}\log n\right)\right\}$$
$$\le \left(1 - \frac{1}{2n^{1/2+\tau}}\right)^{-n/2} \cdot \exp\left\{2n^{1/2-\tau} + \frac{p(2\nu+d)}{2}\left(\log\alpha_0 + \underline{\kappa}\log n\right)\right\}$$
$$= \left\{\left(1 - \frac{1}{2n^{1/2+\tau}}\right)^{2n^{1/2+\tau}}\right\}^{-n^{1/2-\tau}/4} \cdot \exp\left\{2n^{1/2-\tau} + \frac{p(2\nu+d)}{2}\left(\log\alpha_0 + \underline{\kappa}\log n\right)\right\}$$
$$\overset{(ii)}{<} \exp\left(n^{1/2-\tau}/2\right) \cdot \exp\left\{2n^{1/2-\tau} + \frac{p(2\nu+d)}{2}\left(\log\alpha_0 + \underline{\kappa}\log n\right)\right\} < \exp\left(3n^{1/2-\tau}\right), \quad \text{(S.134)}$$

where (i) follows from (S.127), and for (ii), we use the fact that the function $(1 - x^{-1})^x$ is continuous and monotonically increasing to $1/\mathrm{e}$ for $x > 1$, so $(1 - x^{-1})^x > 1/\mathrm{e}^2$ for $x = n^{1/2+\tau}$ given that n is sufficiently large. □

Lemma S.26. *Suppose that the sequence $\{w_i : i = 1, \ldots, n\}$ satisfies $\sum_{i=1}^n w_i \ge n - c_1 n^{b_1}$, $\max_{1\le i\le n} w_i \le 1$ and $\min_{1\le i\le n} w_i \ge c_2 n^{-b_2}$, where $0 < b_2 < b_1 < 1$, $c_1 > 0$, and $c_2 > 0$ are all constants. Then $\prod_{i=1}^n w_i \ge \exp\left(-4b_2 c_1 n^{b_1}\log n\right)$ for all $n > \max\left\{c_2^{-1/b_2}, (2c_2)^{1/b_2}\right\}$.*

Proof of Lemma S.26. Given the constraints in the lemma, minimizing $\prod_{i=1}^n w_i$ is equivalent to choosing as many w_i's to reach the lower bound of $c_2 n^{-b_2}$ as possible. On the other hand, the constraints $\sum_{i=1}^n w_i \ge n - c_1 n^{b_1}$ and $\max_{1\le i\le n} w_i \le 1$ imply that the number of w_i's that attain the lower bound cannot be too large. Suppose that out of n terms of w_i's, $w_1 = \ldots = w_k = c_2 n^{-b_2}$, where k is an integer between 1 and n. Then k must satisfy the relation (since all w_i's satisfy $w_i \le 1$):

$$k c_2 n^{-b_2} + (n-k)\cdot 1 \ge n - c_1 n^{b_1},$$

which implies that $k \le c_1 n^{b_1}/(1 - c_2 n^{-b_2})$. Therefore,

$$\prod_{i=1}^n w_i \ge (c_2 n^{-b_2})^k \cdot 1^{n-k} \ge (c_2 n^{-b_2})^{\frac{c_1 n^{b_1}}{1 - c_2 n^{-b_2}}}.$$

Finally, for all $n > \max\left\{c_2^{-1/b_2}, (2c_2)^{1/b_2}\right\}$, we have that $c_2 > n^{-b_2}$ and $1 - c_2 n^{-b_2} < 1/2$. Hence the conclusion follows. □

Lemma S.27. *For $\tau, \underline{\alpha}_n, \overline{\alpha}_n$ defined in* (S.9)*, for $d \in \{1, 2, 3\}$ and $\nu \in \mathbb{R}^+$, there exist constants $\kappa_1 \in (1/2 - \tau, 1)$, $C_{p,1} > 0$, and a large integer N_9' that only depend on $\nu, d, T, \beta_0, \theta_0, \alpha_0$ and the $\mathcal{W}_2^{\nu+d/2}(\mathcal{S})$ norms of $\mathrm{m}_1(\cdot), \ldots, \mathrm{m}_p(\cdot)$, such that with probability at least $1 - 10\exp(-4\log^2 n)$, for all $n > N_9'$,*

$$\sup_{\alpha\in[\alpha_0,\overline{\alpha}_n]} \exp\left\{\widetilde{\mathcal{L}}_n(\alpha) - \widetilde{\mathcal{L}}_n(\alpha_0)\right\} < \exp\left(C_{p,1} n^{\kappa_1}\log n\right). \quad \text{(S.135)}$$

Proof of Lemma S.27. According to (S.60) and (S.61) in Lemma S.15, we have that for all $k = 1, \ldots, n$ and all $\alpha \in [\alpha_0, \overline{\alpha}_n]$,

$$1 \ge \lambda_{k,n}(\alpha) \ge \left(\frac{\alpha_0}{\alpha}\right)^{2\nu+d} \ge \left(\frac{\alpha_0}{\overline{\alpha}_n}\right)^{2\nu+d} = \frac{\alpha_0^{2\nu+d}}{n^{(2\nu+d)\overline{\kappa}}}. \quad \text{(S.136)}$$

Let $\overline{\lambda}_n(\alpha) = \{\prod_{k=1}^n \lambda_{k,n}(\alpha)\}^{1/n}$.

If $2\nu + d - 2 \geq 0$, then by (S.77) of Lemma S.18, for all $\alpha \in [\alpha_0, \overline{\alpha}_n]$, and for all sufficiently large n,

$$\sum_{k=1}^{n} \{1 - \lambda_{k,n}(\alpha)\} \preceq n^{(2\nu+3d/2+b/2)\overline{\kappa}} \cdot n^{(2a+d)/(4a+2d+b)} + n^{(2\nu+d)\overline{\kappa}} \cdot n^{(2a+d)/(4a+2d+b)} + n^{d\overline{\kappa}}. \tag{S.137}$$

Given the definition of $\overline{\kappa}$ in (S.9) and $d \geq 1$, with the choice $a = 0.01$ and $b = 0.9$,

$$(2\nu + 3d/2 + b/2)\overline{\kappa} + \frac{2a + d}{4a + 2d + b} < 1,$$
$$(2\nu + d)\overline{\kappa} + \frac{2a + d}{4a + 2d + b} < 1, \quad d\overline{\kappa} < 1.$$

Therefore, (S.137) implies that there exist constants $\kappa_1 \in (0, 1)$ (κ_1 can be chosen close to 1) and $C_1 > 0$, such that $\sum_{k=1}^{n} \{1 - \lambda_{k,n}(\alpha)\} < C_1 n^{\kappa_1}$.

If $-1 < 2\nu + d - 2 < 0$ ($d = 1$ and $\nu \in (0, 1/2)$), then for all $\alpha \in [\alpha_0, \overline{\alpha}_n]$, and for all sufficiently large n, (S.77) of Lemma S.18 implies that

$$\sum_{k=1}^{n} \{1 - \lambda_{k,n}(\alpha)\} \preceq n^{(6-4\nu-3d/2+b/2)\overline{\kappa}} \cdot n^{(2a+d)/(4a+2d+b)} + n^{(2\nu+d)\overline{\kappa}} \cdot n^{(2a+d)/(4a+2d+b)} + n^{d\overline{\kappa}}. \tag{S.138}$$

Again given $\overline{\kappa}$ in (S.9) and the choice $a = 0.01$, $b = 0.9$, we have that

$$(6 - 4\nu - 3d/2 + b/2)\overline{\kappa} + \frac{2a + d}{4a + 2d + b} < 1,$$
$$(2\nu + d)\overline{\kappa} + \frac{2a + d}{4a + 2d + b} < 1, \quad d\overline{\kappa} < 1.$$

Therefore, (S.138) also implies that there exist constants $\kappa_1 \in (0, 1)$ (κ_1 can be chosen close to 1) and $C_1 > 0$, such that $\sum_{k=1}^{n} \{1 - \lambda_{k,n}(\alpha)\} < C_1 n^{\kappa_1}$. Combining (S.137) and (S.138), we have that for all sufficiently large n,

$$\sum_{k=1}^{n} \{1 - \lambda_{k,n}(\alpha)\} \leq C_1 n^{\kappa_1}, \quad \text{or} \sum_{k=1}^{n} \lambda_{k,n}(\alpha) \geq n - C_1 n^{\kappa_1}. \tag{S.139}$$

Now in Lemma S.26, we set $w_i = \lambda_{i,n}$, $c_1 = C_1$, $b_1 = \kappa_1$, $c_2 = \alpha_0^{2\nu+d}$, $b_2 = (2\nu + d)\overline{\kappa}$, and use (S.136) and (S.139) to obtain that for all sufficiently large n,

$$\inf_{\alpha \in [\alpha_0, \overline{\alpha}_n]} \overline{\lambda}_n(\alpha) = \left(\inf_{\alpha \in [\alpha_0, \overline{\alpha}_n]} \prod_{k=1}^{n} \lambda_{k,n}(\alpha) \right)^{1/n} \geq \exp\left\{-4C_1(2\nu + d)\overline{\kappa} n^{\kappa_1 - 1} \log n\right\}. \tag{S.140}$$

On the other hand, (S.136) implies that

$$\sum_{k=1}^{n} \left\{\lambda_{k,n}(\alpha)^{-1} - 1\right\} Y_{k,n}(\alpha)^2 \geq 0. \tag{S.141}$$

Therefore, on the event $\mathcal{A}_{2n}$ (the event on the left-hand side of (S.110) in Lemma S.23, where for any $\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]$, $\left|\widetilde{\theta}_\alpha - \widetilde{\theta}_\alpha^{(1)}\right| / \widetilde{\theta}_\alpha^{(1)} \leq 2n^{-1/2-\tau}$), we have that for all $\alpha \in [\alpha_0, \overline{\alpha}_n]$, for all sufficiently large n,

$$\exp\left\{\widetilde{\mathcal{L}}_n(\alpha) - \widetilde{\mathcal{L}}_n(\alpha_0)\right\}$$

$$
\begin{aligned}
&\le \exp\left\{ -\frac{n-p}{2}\log\frac{\alpha^{-2\nu}\widetilde{\theta}_{\alpha}^{(1)}\left(1-2n^{-1/2-\tau}\right)}{\alpha_0^{-2\nu}\widetilde{\theta}_{\alpha_0}^{(1)}\left(1+2n^{-1/2-\tau}\right)} - \frac{1}{2}\log\frac{|R_\alpha|}{|R_{\alpha_0}|} - \frac{1}{2}\log\frac{\left|M_n^\top R_\alpha^{-1}M_n+\Omega_\beta\right|}{\left|M_n^\top R_{\alpha_0}^{-1}M_n+\Omega_\beta\right|}\right\} \\
&\overset{(i)}{\le} \overline{\lambda}_n(\alpha)^{-(n-p)/2}\left[1+\frac{\sum_{k=1}^n\left\{\lambda_{k,n}(\alpha)^{-1}-1\right\}Z_{k,n}(\alpha)^2}{\sum_{k=1}^n Z_{k,n}(\alpha)^2}\right]^{-(n-p)/2} \\
&\quad\times\exp\left\{2n^{1/2-\tau}-p\nu\log\frac{\alpha}{\alpha_0}-\frac{1}{2}\log\frac{\left|M_n^\top R_\alpha^{-1}M_n+\Omega_\beta\right|}{\left|M_n^\top R_{\alpha_0}^{-1}M_n+\Omega_\beta\right|}\right\} \\
&\overset{(ii)}{\le} \overline{\lambda}_n(\alpha)^{-(n-p)/2}\cdot 1^{-(n-p)/2}\cdot\exp\left\{2n^{1/2-\tau}-p\nu\log\frac{\alpha}{\alpha_0}-\frac{1}{2}\log\left(\frac{\alpha_0}{\alpha}\right)^{2p\nu}\right\} \\
&\overset{(iii)}{\le} \exp\left\{2C_1(2\nu+d)\overline{\kappa}n^{\kappa_1}\log n\right\}\cdot\exp\left(2n^{1/2-\tau}\right) \\
&\overset{(iv)}{\le} \exp\left\{3C_1(2\nu+d)\overline{\kappa}n^{\kappa_1}\log n\right\}, \qquad \text{(S.142)}
\end{aligned}
$$

where (i) follows from (S.121) and (S.132); (ii) follows from (S.108) and (S.141); (iii) follows from (S.140); (iv) follows since we can choose $\kappa_1 \in (1/2-\tau, 1)$. The conclusion follows by taking $C_{p,1} = 3C_1(2\nu+d)\overline{\kappa}$. □

Lemma S.28. *Suppose that Assumptions (A.1), (A.2) and (A.4) hold. Then for all $d \in \mathbb{Z}^+$ and $\nu \in \mathbb{R}^+$, the profile posterior distribution of α given by $\widetilde{\pi}(\alpha|Y_n)$ in* (19) *is a proper posterior almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$ for any given $n \ge p$.*

Proof of Lemma S.28. We consider a fixed $n \ge p$. Since the Matérn covariance function is continuous in $\alpha \in \mathbb{R}^+$, R_α is also continuous in $\alpha \in \mathbb{R}^+$, and so is the profile restricted likelihood $\exp\{\widetilde{\mathcal{L}}_n(\alpha)\}$. Furthermore, both $\pi(\theta_0|\alpha)$ and $\pi(\alpha)$ are continuous functions in $\alpha \in \mathbb{R}^+$ by Assumptions (A.2) and (A.4). As a result, the profile posterior in (19) is well defined as long as the function $\exp\{\widetilde{\mathcal{L}}_n(\alpha)\}\pi(\theta_0|\alpha)\pi(\alpha)$ is integrable as $\alpha \to +\infty$ and $\alpha \to 0+$.

As $\alpha \to +\infty$, $R_\alpha \to I_n$ elementwise. Since M_n is rank-p for all $n \ge p$ by Assumption (A.1), $M_n^\top M_n$ is invertible for each fixed n and $\mathcal{S}_n$. Therefore, as $\alpha \to +\infty$, the profile restricted likelihood $\exp\{\widetilde{\mathcal{L}}_n(\alpha)\}$ becomes proportional to

$$
\begin{aligned}
&\exp\left\{-\frac{n-p}{2}\log\frac{Y_n^\top\left[I_n-M_n(M_n^\top M_n+\Omega_\beta)^{-1}M_n^\top\right]Y_n}{n}-\frac{1}{2}\log\left|M_n^\top M_n+\Omega_\beta\right|\right\} \\
&=\left(\frac{Y_n^\top\left[I_n-M_n(M_n^\top M_n+\Omega_\beta)^{-1}M_n^\top\right]Y_n}{n}\right)^{-(n-p)/2}\cdot\left|M_n^\top M_n+\Omega_\beta\right|^{-1/2},
\end{aligned}
$$

which is a finite positive number almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$ for any given $n \ge p$. Since Assumption (A.4) says that $\int_0^\infty \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha < \infty$, and $\exp\{\widetilde{\mathcal{L}}_n(\alpha)\}$ is a continuous function in α, it follows that the integral of $\exp\{\widetilde{\mathcal{L}}_n(\alpha)\}\pi(\theta_0|\alpha)\pi(\alpha)$ on $\alpha \in [1, +\infty)$ is finite.

Then we consider the case when $\alpha \to 0+$. The property of the Matérn covariance function as $\alpha \to 0+$ has been analyzed in detail in [Berger et al., 2001] and [Gu et al., 2018]. Lemma 3.3 of [Gu et al., 2018] has shown that for given n, M_n and Y_n, the profile restricted likelihood function converges to zero as $\alpha \to 0+$ with the following rates:

$$
\exp\{\widetilde{\mathcal{L}}_n(\alpha)\} \le \begin{cases} C(n,M_n,Y_n)\alpha^\nu, & \text{if } \nu \in (0,1), \\ C(n,M_n,Y_n)\alpha\left\{\log(1/\alpha)\right\}^{1/2}, & \text{if } \nu = 1, \\ C(n,M_n,Y_n)\alpha, & \text{if } \nu > 1, \end{cases}
$$

where $C(n,M_n,Y_n)$ is a finite positive number that depends on d, n, M_n and Y_n but not α. In all three cases, $\exp\{\widetilde{\mathcal{L}}_n(\alpha)\} \to 0$ as $\alpha \to 0+$. Together with $\int_0^\infty \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha < \infty$ from

Assumption (A.4), we conclude that the integral of $\exp\{\widetilde{\mathcal{L}}_n(\alpha)\}\pi(\theta_0|\alpha)\pi(\alpha)$ on $\alpha \in (0,1)$ is also finite. Therefore, $\int_0^\infty \exp\{\widetilde{\mathcal{L}}_n(\alpha)\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha < \infty$, and the profile posterior defined in (19) is a proper posterior almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$ for any given $n \geq p$. □

S3 Proof of Theorems 2.1 and 2.3

In this section, we provide the proof of Theorems 2.1 and 2.3 in the main text. We first prove a useful Lemma S.29 that establishes the local asymptotic normality (LAN) condition for the microergodic parameter θ for a given α. This lemma is essential for showing the limiting normal posterior for θ. In Section S3.4, we present the theory on the limiting posterior distribution of (θ, α) for the case of $d \geq 5$.

S3.1 Proof of Lemma S.29

For a given $\alpha > 0$, let $t = \sqrt{n-p}(\theta - \widetilde{\theta}_\alpha)$ be the local parameter. We define the following function:

$$\varrho_n(t;\alpha) = \exp\left\{\mathcal{L}_n(\alpha^{-2\nu}(\widetilde{\theta}_\alpha + \tfrac{t}{\sqrt{n-p}}),\alpha) - \mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)\right\} \cdot \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\middle|\; \alpha\right)}{\pi(\theta_0|\alpha)} - \exp\left(-\frac{t^2}{4\theta_0^2}\right). \tag{S.143}$$

Lemma S.29. *Suppose that Assumption (A.1) and (A.2) hold. Then for all $d \in \mathbb{Z}^+, \nu \in \mathbb{R}^+$, for any fixed $\alpha > 0$, for any positive sequences $\epsilon_{1n} \to 0$ as $n \to \infty$ and $1 \preceq s_n \prec \min\left(n^{1/6}, \epsilon_{1n}^{-1/2}\right)$ that do not depend on α, for all sufficiently large n, the ϱ_n function in* (S.143) *satisfies the following upper bound on the event $\mathcal{E}_1(\epsilon_{1n},\alpha) = \{|\widetilde{\theta}_\alpha - \theta_0| < \epsilon_{1n}\}$:*

$$\int_{\mathbb{R}} |\varrho_n(t;\alpha)|\mathrm{d}t \leq B_n(\alpha), \tag{S.144}$$

where

$$\begin{aligned}
B_n(\alpha) \equiv\ & 4\theta_0 \exp\left(-\frac{n-p}{64}\right) + \frac{\sqrt{n-p}}{\pi(\theta_0|\alpha)}\exp\{-0.007(n-p)\} \\
& + 10\theta_0 \exp\left(-\frac{4s_n^2}{125\theta_0^2}\right) \cdot \sup_{\theta\in\left(\frac{1}{2}\theta_0,\frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} + 4\theta_0\exp\left(-\frac{s_n^2}{4\theta_0^2}\right) \\
& + \frac{8}{\theta_0^2}\left(s_n^2\epsilon_{1n} + \frac{2s_n^3}{\sqrt{n-p}}\right) \cdot \sup_{\theta\in\left(\frac{3}{4}\theta_0,\frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \\
& + 4\theta_0 \sup_{\theta\in\left(\frac{3}{4}\theta_0,\frac{3}{2}\theta_0\right)} \left|\frac{\partial \log\pi(\theta|\alpha)}{\partial\theta}\right| \sup_{\theta\in\left(\frac{3}{4}\theta_0,\frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \cdot \left(\epsilon_{1n} + \frac{s_n}{\sqrt{n-p}}\right).
\end{aligned} \tag{S.145}$$

Proof of Lemma S.29. we first take the difference of the log-likelihood in (6) and the profile restricted log-likelihood in (8) of the main text, and use the definition of $\widetilde{\theta}_\alpha$ in (7) of the main text to obtain that

$$\mathcal{L}_n(\alpha^{-2\nu}\theta,\alpha) - \mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha) = -\frac{n-p}{2}\log\frac{\theta}{\widetilde{\theta}_\alpha} + \frac{(n-p)(\theta-\widetilde{\theta}_\alpha)}{2\theta} \tag{S.146}$$

$$= -\frac{n-p}{2}\log\left(1 + \frac{t}{\sqrt{n-p}\cdot\widetilde{\theta}_\alpha}\right) + \frac{\sqrt{n-p}\cdot t}{2\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}}\right)} \tag{S.147}$$

We decompose the integral in (S.144) into three parts:

$$\int_{\mathbb{R}} |\varrho_n(t;\alpha)|\mathrm{d}t = \int_{A_1} |\varrho_n(t;\alpha)|\mathrm{d}t + \int_{A_2} |\varrho_n(t;\alpha)|\mathrm{d}t + \int_{A_3} |\varrho_n(t;\alpha)|\mathrm{d}t, \tag{S.148}$$

where $A_1 = \{t \in \mathbb{R} : |t| \geq (\theta_0/4)\sqrt{n-p}\}$, $A_2 = \{t \in \mathbb{R} : s_n \leq |t| < (\theta_0/4)\sqrt{n-p}\}$, and $A_3 = \{t \in \mathbb{R} : |t| < s_n\}$, with the sequence s_n as specified in the lemma.

Bound the first term in (S.148): We have

$$\begin{aligned}\int_{A_1} |\varrho_n(t;\alpha)|\mathrm{d}t \leq& \int_{A_1} \exp\left\{\mathcal{L}_n(\alpha^{-2\nu}\theta,\alpha) - \mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)\right\} \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\middle|\; \alpha\right)}{\pi(\theta_0|\alpha)}\mathrm{d}t \\ &+ \int_{A_1} \mathrm{e}^{-\frac{t^2}{4\theta_0^2}}\mathrm{d}t.\end{aligned} \tag{S.149}$$

The second term in (S.149) can be bounded by

$$\begin{aligned}\int_{A_1} \mathrm{e}^{-\frac{t^2}{4\theta_0^2}}\mathrm{d}t &\leq 2\sqrt{\pi}\theta_0 \cdot \int_{|t|\geq(\theta_0/4)\sqrt{n-p}} \frac{1}{\sqrt{2\pi\cdot 2\theta_0^2}}\mathrm{e}^{-\frac{t^2}{4\theta_0^2}}\mathrm{d}t \\ &\leq 2\sqrt{\pi}\theta_0 \exp\left\{-\frac{(n-p)(\theta_0/4)^2}{4\theta_0^2}\right\} = 2\sqrt{\pi}\theta_0 \exp\left(-\frac{n-p}{64}\right),\end{aligned} \tag{S.150}$$

where the last inequality follows from the tail bounds for a normal random variable: if $Z \sim \mathcal{N}(0,1)$, then for any $z > 0$,

$$\Pr(|Z| > z) \leq \mathrm{e}^{-z^2/2}. \tag{S.151}$$

For the first term in (S.149), we note that θ is a linear transformation of t. We use the relation (S.146) and obtain that

$$\begin{aligned}&\int_{A_1} \exp\left\{\mathcal{L}_n(\alpha^{-2\nu}\theta,\alpha) - \mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)\right\} \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\middle|\; \alpha\right)}{\pi(\theta_0|\alpha)}\mathrm{d}t \\ &= \int_{|t|\geq(\theta_0/4)\sqrt{n-p}} \exp\left\{-\frac{n-p}{2}\log\frac{\theta}{\widetilde{\theta}_\alpha} + \frac{(n-p)(\theta-\widetilde{\theta}_\alpha)}{2\theta}\right\} \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\middle|\; \alpha\right)}{\pi(\theta_0|\alpha)}\mathrm{d}t \\ &\leq \sqrt{n-p}\int_{|\theta-\widetilde{\theta}_\alpha|\geq\theta_0/4} \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\middle|\; \alpha\right)}{\pi(\theta_0|\alpha)} \cdot \exp\left\{-\frac{n-p}{2}\varphi\left(\frac{\widetilde{\theta}_\alpha}{\theta}\right)\right\}\mathrm{d}\theta.\end{aligned} \tag{S.152}$$

For any constant $\epsilon > 0$, define the event $\mathcal{E}_1(\epsilon,\alpha) = \{|\widetilde{\theta}_\alpha - \theta_0| < \epsilon\}$. Let $0 < \epsilon_{1n} < \theta_0/4$, where $\epsilon_{1n} \to 0$ as $n \to \infty$ and its order will be determined later. Then, on the event $\mathcal{E}_1(\epsilon_{1n},\alpha)$ and $\{|\theta - \widetilde{\theta}_\alpha| \geq \theta_0/4\}$, we consider two cases: If $\theta > \widetilde{\theta}_\alpha + \theta_0/4$, then

$$1 - \frac{\widetilde{\theta}_\alpha}{\theta} = 1 - \frac{\widetilde{\theta}_\alpha}{\theta - \widetilde{\theta}_\alpha + \widetilde{\theta}_\alpha} \geq 1 - \frac{\widetilde{\theta}_\alpha}{\theta_0/4 + \widetilde{\theta}_\alpha} = \frac{\theta_0/4}{\theta_0/4 + \widetilde{\theta}_\alpha} > \frac{\theta_0/4}{\theta_0/4 + \theta_0 + \epsilon_{1n}} > \frac{1}{6}.$$

If $\theta < \widetilde{\theta}_\alpha - \theta_0/4$, then

$$\frac{\widetilde{\theta}_\alpha}{\theta} - 1 = \frac{\widetilde{\theta}_\alpha}{\theta - \widetilde{\theta}_\alpha + \widetilde{\theta}_\alpha} - 1 \geq \frac{\widetilde{\theta}_\alpha}{-\theta_0/4 + \widetilde{\theta}_\alpha} - 1 = \frac{\theta_0/4}{\widetilde{\theta}_\alpha - \theta_0/4} > \frac{\theta_0/4}{\theta_0 + \epsilon_{1n} - \theta_0/4} > \frac{1}{4}.$$

This implies that on the event $\mathcal{E}_1(\epsilon_{1n},\alpha)$ and $\{|\theta - \widetilde{\theta}_\alpha| \geq \theta_0/4\}$, we must have either $\widetilde{\theta}_\alpha/\theta < \frac{5}{6}$ or $\widetilde{\theta}_\alpha/\theta > \frac{5}{4}$. Since the function $\varphi(u) = u - \log u - 1$ is monotonically decreasing on $(0,1)$ and monotonically increasing on $[1,+\infty)$, we have that on the event $\mathcal{E}_1(\epsilon_{1n},\alpha)$ and $\{|\theta-\widetilde{\theta}_\alpha|\geq\theta_0/4\}$, either $\varphi(\widetilde{\theta}_\alpha/\theta)>\min\{\varphi(5/6),\varphi(5/4)\}>0.015$. Therefore, from (S.152), we obtain that on the event $\mathcal{E}_1(\epsilon_{1n},\alpha)$,

$$\begin{aligned}
&\int_{A_1}\exp\left\{\mathcal{L}_n(\alpha^{-2\nu}\theta,\alpha)-\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)\right\}\frac{\pi\left(\widetilde{\theta}_\alpha+\frac{t}{\sqrt{n-p}}\;\middle|\;\alpha\right)}{\pi(\theta_0|\alpha)}\mathrm{d}t\\
&\leq\sqrt{n-p}\int_{|\theta-\widetilde{\theta}_\alpha|\geq\theta_0/4}\frac{\pi\left(\theta|\alpha\right)}{\pi(\theta_0|\alpha)}\cdot\exp\left\{-\frac{0.015(n-p)}{2}\right\}\mathrm{d}\theta\\
&<\frac{\sqrt{n-p}}{\pi(\theta_0|\alpha)}\exp\{-0.007(n-p)\}, \qquad\qquad (S.153)
\end{aligned}$$

where in the last inequality, we use the fact that $\pi(\theta|\alpha)$ is a proper prior density. Thus, combining (S.149), (S.150) and (S.153) yields that on the event $\mathcal{E}_1(\epsilon_{1n},\alpha)$,

$$\int_{A_1}|\varrho_n(t;\alpha)|\mathrm{d}t\leq 2\sqrt{\pi}\theta_0\exp\left(-\frac{n-p}{64}\right)+\frac{\sqrt{n-p}}{\pi(\theta_0|\alpha)}\exp\{-0.007(n-p)\}. \qquad (S.154)$$

Bound the second term in (S.148): On the event $\mathcal{E}_1(\epsilon_{1n},\alpha)$ and $\{|\theta-\widetilde{\theta}_\alpha|<\theta_0/4\}$ with $0<\epsilon_{1n}<\theta_0/4$, if $\theta\geq\widetilde{\theta}_\alpha$, then

$$1-\frac{\widetilde{\theta}_\alpha}{\theta}=1-\frac{\widetilde{\theta}_\alpha}{\theta-\widetilde{\theta}_\alpha+\widetilde{\theta}_\alpha}<1-\frac{\widetilde{\theta}_\alpha}{\theta_0/4+\widetilde{\theta}_\alpha}=\frac{\theta_0/4}{\theta_0/4+\widetilde{\theta}_\alpha}\leq\frac{\theta_0/4}{\theta_0/4+\theta_0-\epsilon_{1n}}<\frac{1}{4}.$$

If $\theta<\widetilde{\theta}_\alpha$, then

$$\frac{\widetilde{\theta}_\alpha}{\theta}-1=\frac{\widetilde{\theta}_\alpha}{\theta-\widetilde{\theta}_\alpha+\widetilde{\theta}_\alpha}-1<\frac{\widetilde{\theta}_\alpha}{-\theta_0/4+\widetilde{\theta}_\alpha}-1=\frac{\theta_0/4}{\widetilde{\theta}_\alpha-\theta_0/4}<\frac{\theta_0/4}{\theta_0-\epsilon_{1n}-\theta_0/4}<\frac{1}{2}.$$

Hence on the event $\mathcal{E}_1(\epsilon_{1n},\alpha)$ and $\{|\theta-\widetilde{\theta}_\alpha|<\theta_0/4\}$, $\widetilde{\theta}_\alpha/\theta\in(\frac{3}{4},\frac{3}{2})$. For any $u\in(\frac{3}{4},\frac{3}{2})$, by simple calculus, we have

$$\left|\varphi(u)-\frac{1}{2}\left(\frac{1}{u}-1\right)^2\right|\leq\frac{6}{5}\left|\frac{1}{u}-1\right|^3. \qquad (S.155)$$

Let

$$\begin{aligned}
g_n(t)&=\frac{1}{n-p}\left[\mathcal{L}_n(\alpha^{-2\nu}(\widetilde{\theta}_\alpha+\tfrac{t}{\sqrt{n-p}}),\alpha)-\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)\right]-\frac{t^2}{2(n-p)\widetilde{\theta}_\alpha^2}\\
&=\varphi\left(\left[1+\frac{t}{\sqrt{n-p}\cdot\widetilde{\theta}_\alpha}\right]^{-1}\right)-\frac{t^2}{2(n-p)\widetilde{\theta}_\alpha^2}.
\end{aligned}$$

In (S.155), if we set $u=\widetilde{\theta}_\alpha/\theta$, then $\frac{1}{2}\left(\frac{1}{u}-1\right)^2=t^2/[2(n-p)\widetilde{\theta}_\alpha^2]$. Thus, we can obtain that on the event $\mathcal{E}_1(\epsilon_{1n},\alpha)$ and $t\in A_2$ (so that $|\theta-\widetilde{\theta}_\alpha|<\theta_0/4$),

$$\begin{aligned}
|g_n(t)|&=\left|\varphi\left(\left[1+\frac{t}{\sqrt{n-p}\cdot\widetilde{\theta}_\alpha}\right]^{-1}\right)-\frac{t^2}{2(n-p)\widetilde{\theta}_\alpha^2}\right|\leq\frac{6|t|^3}{5(n-p)^{3/2}\widetilde{\theta}_\alpha^3}=\frac{6|\theta-\widetilde{\theta}_\alpha|^3}{5\widetilde{\theta}_\alpha^3}\\
&\leq\frac{12|\theta-\widetilde{\theta}_\alpha|}{5\widetilde{\theta}_\alpha}\cdot\frac{|\theta-\widetilde{\theta}_\alpha|^2}{2\widetilde{\theta}_\alpha^2}\leq\frac{4}{5}\cdot\frac{|\theta-\widetilde{\theta}_\alpha|^2}{2\widetilde{\theta}_\alpha^2}=\frac{2t^2}{5(n-p)\widetilde{\theta}_\alpha^2}. \qquad (S.156)
\end{aligned}$$

Therefore, on the event $\mathcal{E}_1(\epsilon_{1n}, \alpha)$ with $0 < \epsilon_{1n} < \theta_0/4$,

$$
\begin{aligned}
&\int_{A_2} |\varrho_n(t;\alpha)|\mathrm{d}t \le \int_{A_2} \exp\left\{-\frac{n-p}{2}\varphi(\widetilde{\theta}_\alpha/\theta)\right\} \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\Big|\; \alpha\right)}{\pi(\theta_0|\alpha)}\mathrm{d}t + \int_{A_2} \mathrm{e}^{-\frac{t^2}{4\theta_0^2}}\mathrm{d}t \\
&\le \int_{A_2} \exp\left\{-\frac{t^2}{4\widetilde{\theta}_\alpha^2} + \frac{n-p}{2}|g_n(t)|\right\} \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\Big|\; \alpha\right)}{\pi(\theta_0|\alpha)} + \int_{A_2} \mathrm{e}^{-\frac{t^2}{4\theta_0^2}}\mathrm{d}t \\
&\overset{(i)}{\le} \int_{A_2} \exp\left\{-\frac{t^2}{20\widetilde{\theta}_\alpha^2}\right\} \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\Big|\; \alpha\right)}{\pi(\theta_0|\alpha)}\mathrm{d}t + \int_{A_2} \mathrm{e}^{-\frac{t^2}{4\theta_0^2}}\mathrm{d}t \\
&\le \sup_{|\theta-\widetilde{\theta}_\alpha|<\theta_0/4} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \cdot \int_{A_2} \exp\left\{-\frac{t^2}{20\widetilde{\theta}_\alpha^2}\right\}\mathrm{d}t + \int_{A_2} \mathrm{e}^{-\frac{t^2}{4\theta_0^2}}\mathrm{d}t \\
&\overset{(ii)}{\le} \sup_{\theta\in\left(\frac{1}{2}\theta_0, \frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \cdot \int_{|t|>s_n} \exp\left\{-\frac{t^2}{20\widetilde{\theta}_\alpha^2}\right\}\mathrm{d}t + \int_{|t|>s_n} \mathrm{e}^{-\frac{t^2}{4\theta_0^2}}\mathrm{d}t \\
&\overset{(iii)}{\le} \sup_{\theta\in\left(\frac{1}{2}\theta_0, \frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \cdot 2\sqrt{5\pi}\widetilde{\theta}_\alpha \exp\left(-\frac{s_n^2}{20\widetilde{\theta}_\alpha^2}\right) + 2\sqrt{\pi}\theta_0 \exp\left(-\frac{s_n^2}{4\theta_0^2}\right) \\
&\overset{(iv)}{\le} \sup_{\theta\in\left(\frac{1}{2}\theta_0, \frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \cdot \frac{5}{2}\sqrt{5\pi}\theta_0 \exp\left(-\frac{4s_n^2}{125\theta_0^2}\right) + 2\sqrt{\pi}\theta_0 \exp\left(-\frac{s_n^2}{4\theta_0^2}\right),
\end{aligned}
\tag{S.157}
$$

where (i) is from the upper bound of $g_n(t)$ in (S.156); (ii) is based on the relation $|\theta - \theta_0| \le |\theta - \widetilde{\theta}_\alpha| + |\widetilde{\theta}_\alpha - \theta_0| < \theta_0/4 + \epsilon_{1n} < \theta_0/2$; (iii) follows from the normal tail inequality (S.151); (iv) is based on the relation $\widetilde{\theta}_\alpha \le \theta_0 + \epsilon_{1n} < \theta_0 + \theta_0/4 < 5\theta_0/4$.

Bound the third term in (S.148): We continue to use the bound in (S.155) and (S.156) for $t \in A_3$ on the event $\mathcal{E}_1(\epsilon_{1n}, \alpha)$ and obtain that

$$
|g_n(t)| \le \frac{6|t|^3}{5(n-p)^{3/2}\widetilde{\theta}_\alpha^3} \le \frac{6s_n^3}{5(n-p)^{3/2}\widetilde{\theta}_\alpha^3}. \tag{S.158}
$$

Therefore,

$$
\begin{aligned}
&\int_{A_3} |\varrho_n(t;\alpha)|\mathrm{d}t \\
&= \int_{A_3} \left| \exp\left\{-\frac{n-p}{2}\varphi(\widetilde{\theta}_\alpha/\theta)\right\} \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\Big|\; \alpha\right)}{\pi(\theta_0|\alpha)} - \mathrm{e}^{-\frac{t^2}{4\theta_0^2}} \right| \mathrm{d}t \\
&= \int_{A_3} \left| \exp\Big\{-\frac{t^2}{4\widetilde{\theta}_\alpha^2} - \frac{n-p}{2}g_n(t)\Big\} \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\Big|\; \alpha\right)}{\pi(\theta_0|\alpha)} - \mathrm{e}^{-\frac{t^2}{4\theta_0^2}} \right| \mathrm{d}t \\
&\le \int_{A_3} \left| \exp\Big\{-\frac{t^2}{4\widetilde{\theta}_\alpha^2} - \frac{n-p}{2}g_n(t)\Big\} - \exp\Big(-\frac{t^2}{4\theta_0^2}\Big) \right| \cdot \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\Big|\; \alpha\right)}{\pi(\theta_0|\alpha)}\mathrm{d}t \\
&\quad + \int_{A_3} \mathrm{e}^{-\frac{t^2}{4\theta_0^2}} \left| \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\Big|\; \alpha\right)}{\pi(\theta_0|\alpha)} - 1 \right| \mathrm{d}t \\
&\le \sup_{|t|<s_n} \left| \exp\left\{\frac{t^2}{4}\left(\theta_0^{-2} - \widetilde{\theta}_\alpha^{-2}\right) - \frac{n-p}{2}g_n(t)\right\} - 1 \right| \cdot \sup_{|t|<s_n} \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\Big|\; \alpha\right)}{\pi(\theta_0|\alpha)} \cdot \int_{|t|<s_n} \mathrm{e}^{-\frac{t^2}{4\theta_0^2}}\mathrm{d}t
\end{aligned}
$$

$$
+ \sup_{|t|<s_n} \left| \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\middle|\; \alpha\right)}{\pi(\theta_0|\alpha)} - 1 \right| \times \int_{|t|<s_n} \mathrm{e}^{-\frac{t^2}{4\theta_0^2}} \mathrm{d}t
$$

$$
\leq 2\sqrt{\pi}\theta_0 \cdot \sup_{|t|<s_n} \left| \exp\left\{ \frac{t^2}{4}\left(\theta_0^{-2} - \widetilde{\theta}_\alpha^{-2}\right) - \frac{n-p}{2} g_n(t) \right\} - 1 \right| \cdot \sup_{|t|<s_n} \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\middle|\; \alpha\right)}{\pi(\theta_0|\alpha)}
$$

$$
+ 2\sqrt{\pi}\theta_0 \cdot \sup_{|t|<s_n} \left| \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\middle|\; \alpha\right)}{\pi(\theta_0|\alpha)} - 1 \right|. \tag{S.159}
$$

For the first term in (S.159), we can choose $\epsilon_{1n} \to 0$ as $n \to \infty$ and $\epsilon_{1n} < \theta_0/4$, such that on the event $\mathcal{E}_1(\epsilon_{1n}, \alpha)$, for all $|t| < s_n$, using (S.158), we have

$$
\begin{aligned}
&\left| \frac{t^2}{4}\left(\theta_0^{-2} - \widetilde{\theta}_\alpha^{-2}\right) - \frac{n-p}{2} g_n(t) \right| \leq \frac{s_n^2}{4} \frac{\left|\widetilde{\theta}_\alpha^2 - \theta_0^2\right|}{\widetilde{\theta}_\alpha^2 \theta_0^2} + \left| \frac{n-p}{2} g_n(t) \right| \\
&\leq \frac{s_n^2 \epsilon_{1n}}{4} \frac{\left|\widetilde{\theta}_\alpha + \theta_0\right|}{\widetilde{\theta}_\alpha^2 \theta_0^2} + \left| \frac{n-p}{2} g_n(t) \right| \leq \frac{s_n^2 \epsilon_{1n}}{4} \frac{(2\theta_0 + \epsilon_{1n})}{(\theta_0 - \epsilon_{1n})^2 \theta_0^2} + \frac{3 s_n^3}{5\sqrt{n-p}\,\widetilde{\theta}_\alpha^3} \\
&< \frac{s_n^2 \epsilon_{1n}}{\theta_0^3} + \frac{2 s_n^3}{\sqrt{n-p}\,\theta_0^3}.
\end{aligned} \tag{S.160}
$$

We choose sufficiently large n that satisfies $\epsilon_{1n} \leq \frac{\theta_0^3}{2s_n^2}$ and $n \geq \frac{16 s_n^6}{\theta_0^6} + p$, such that the upper bound in (S.160) is smaller than 1. Then we can apply the inequality $|\mathrm{e}^u - 1| \leq 2|u|$ for all $|u| \leq 1$ and obtain that

$$
\begin{aligned}
&\sup_{|t|<s_n} \left| \exp\left\{ \frac{t^2}{4}\left(\theta_0^{-2} - \widetilde{\theta}_\alpha^{-2}\right) - \frac{n-p}{2} g_n(t) \right\} - 1 \right| \\
&\leq 2 \left| \frac{t^2}{4}\left(\theta_0^{-2} - \widetilde{\theta}_\alpha^{-2}\right) - \frac{n-p}{2} g_n(t) \right| < \frac{2 s_n^2 \epsilon_{1n}}{\theta_0^3} + \frac{4 s_n^3}{\sqrt{n-p}\,\theta_0^3}.
\end{aligned} \tag{S.161}
$$

Furthermore, we can choose $n \geq \frac{16 s_n^2}{\theta_0^2} + p$ such that for all $|t| < s_n$, on the event $\mathcal{E}_1(\epsilon_{1n}, \alpha)$, $\widetilde{\theta}_\alpha + t/\sqrt{n-p} \leq \theta_0 + \epsilon_{1n} + s_n/\sqrt{n-p} < \frac{3}{2}\theta_0$ and $\widetilde{\theta}_\alpha + t/\sqrt{n-p} > \theta_0 - \epsilon_{1n} > \frac{3}{4}\theta_0$. Then from Assumption (A.2) (ii), we have that on the interval $(\frac{3}{4}\theta_0, \frac{3}{2}\theta_0)$,

$$
\sup_{|t|<s_n} \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\middle|\; \alpha\right)}{\pi(\theta_0|\alpha)} \leq \sup_{\theta \in \left(\frac{3}{4}\theta_0, \frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)}. \tag{S.162}
$$

For the second term in (S.160), by Assumption (A.2) and the fact that $\epsilon_{1n} \to 0, s_n/\sqrt{n-p} \to 0$, we have that on the event $\mathcal{E}_1(\epsilon_{1n}, \alpha)$, for all sufficiently large n,

$$
\begin{aligned}
&\sup_{|t|<s_n} \left| \frac{\pi\left(\widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} \;\middle|\; \alpha\right)}{\pi(\theta_0|\alpha)} - 1 \right| \leq \sup_{\theta \in (3\theta_0/4, 3\theta_0/2)} \left| \frac{\pi(\theta|\alpha) - \pi(\theta_0|\alpha)}{\pi(\theta_0|\alpha)} \right| \\
&\leq \sup_{\theta \in (3\theta_0/4, 3\theta_0/2)} \left| \frac{\partial \log \pi(\theta|\alpha)}{\partial \theta} \right| \cdot \sup_{\theta \in (3\theta_0/4, 3\theta_0/2)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \cdot \sup_{|t|<s_n} \left| \widetilde{\theta}_\alpha + \frac{t}{\sqrt{n-p}} - \theta_0 \right| \\
&\leq \sup_{\theta \in \left(\frac{3}{4}\theta_0, \frac{3}{2}\theta_0\right)} \left| \frac{\partial \log \pi(\theta|\alpha)}{\partial \theta} \right| \sup_{\theta \in \left(\frac{3}{4}\theta_0, \frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \cdot \left(\epsilon_{1n} + \frac{s_n}{\sqrt{n-p}} \right).
\end{aligned} \tag{S.163}
$$

Therefore, (S.159), (S.161), (S.162), and (S.163) together yield that on the event $\mathcal{E}_1(\epsilon_{1n}, \alpha)$, with $\epsilon_{1n} \leq \min\left(\frac{\theta_0^3}{2s_n^2}, \frac{\theta_0}{4}\right)$ and $n \geq \max\left(\frac{16s_n^6}{\theta_0^6}, \frac{16s_n^2}{\theta_0^2}\right) + p$,

$$\begin{aligned}
&\int_{A_3} |\varrho_n(t;\alpha)|\mathrm{d}t \\
\leq\ & \frac{4\sqrt{\pi}}{\theta_0^2}\left(s_n^2\epsilon_{1n} + \frac{2s_n^3}{\sqrt{n-p}}\right) \cdot \sup_{\theta\in\left(\frac{3}{4}\theta_0, \frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \\
&+ 2\sqrt{\pi}\theta_0 \sup_{\theta\in\left(\frac{3}{4}\theta_0, \frac{3}{2}\theta_0\right)} \left|\frac{\partial \log \pi(\theta|\alpha)}{\partial\theta}\right| \sup_{\theta\in\left(\frac{3}{4}\theta_0, \frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \cdot \left(\epsilon_{1n} + \frac{s_n}{\sqrt{n-p}}\right). \qquad \text{(S.164)}
\end{aligned}$$

Finally, we combine (S.154), (S.157), and (S.164) to conclude that on the event $\mathcal{E}_1(\epsilon_{1n}, \alpha)$ with $\epsilon_{1n} \leq \min\left(\frac{\theta_0^3}{2s_n^2}, \frac{\theta_0}{4}\right)$ and $n \geq \max\left(\frac{16s_n^6}{\theta_0^6}, \frac{16s_n^2}{\theta_0^2}\right) + p$,

$$\begin{aligned}
\int_{\mathbb{R}} |\varrho_n(t;\alpha)|\mathrm{d}t &\leq 2\sqrt{\pi}\theta_0 \exp\{-(n-p)/64\} + \frac{\sqrt{n-p}}{\pi(\theta_0|\alpha)} \exp\{-0.007(n-p)\} \\
&+ \sup_{\theta\in\left(\frac{1}{2}\theta_0, \frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \cdot \frac{5}{2}\sqrt{5\pi}\theta_0 \exp\left(-\frac{4s_n^2}{125\theta_0^2}\right) + 2\sqrt{\pi}\theta_0 \exp\left(-\frac{s_n^2}{4\theta_0^2}\right) \\
&+ \frac{4\sqrt{\pi}}{\theta_0^2}\left(s_n^2\epsilon_{1n} + \frac{2s_n^3}{\sqrt{n-p}}\right) \cdot \sup_{\theta\in\left(\frac{3}{4}\theta_0, \frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \\
&+ 2\sqrt{\pi}\theta_0 \sup_{\theta\in\left(\frac{3}{4}\theta_0, \frac{3}{2}\theta_0\right)} \left|\frac{\partial \log \pi(\theta|\alpha)}{\partial\theta}\right| \sup_{\theta\in\left(\frac{3}{4}\theta_0, \frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \cdot \left(\epsilon_{1n} + \frac{s_n}{\sqrt{n-p}}\right). \qquad \text{(S.165)}
\end{aligned}$$

By adjusting the constants to be slightly larger, we obtain the bound in (S.145). □

The proof of Theorem 2.1 has used on the following lemmas.

Lemma S.30. *For two nonnegative functions f and g, if their integrals are $F = \int f$ and $G = \int g$, then*

$$\int \left|\frac{f}{\int f} - \frac{g}{\int g}\right| \leq \frac{2\int |f-g|}{G}.$$

Proof of Lemma S.30.

$$\begin{aligned}
\int \left|\frac{f}{\int f} - \frac{g}{\int g}\right| &= \int \frac{|fG - gF|}{FG} \leq \int \frac{f|G-F| + F|g-f|}{FG} \\
&= \frac{|G-F|\int f + F\int |f-g|}{FG} \leq \frac{F\int |f-g| + F\int |f-g|}{FG} = \frac{2\int |f-g|}{G}.
\end{aligned}$$

□

Lemma S.31. *For two univariate normal distributions $\mathcal{N}(\mu_1, \sigma^2)$ and $\mathcal{N}(\mu_2, \sigma^2)$ on $\mathbb{R}$, their total variation distance is given by*

$$\left\|\mathcal{N}(\mu_1, \sigma^2) - \mathcal{N}(\mu_2, \sigma^2)\right\|_{\mathrm{TV}} = 2\Phi\left(\frac{|\mu_1 - \mu_2|}{2\sigma}\right) - 1,$$

where $\Phi(x) = \int_{-\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-z^2/2}\mathrm{d}z$ is the standard normal cdf.

Proof of Lemma S.31. Let $f_i(x)$ be the normal density of $\mathcal{N}(\mu_i, \sigma^2)$, $i = 1, 2$. Suppose that $\mu_1 < \mu_2$ without loss of generality. Then it is clear that $f_1(x) > f_2(x)$ if $x < (\mu_1 + \mu_2)/2$ and $f_1(x) < f_2(x)$ if $x > (\mu_1 + \mu_2)/2$. Therefore,

$$\begin{aligned}
&\left\|\mathcal{N}(\mu_1, \sigma^2) - \mathcal{N}(\mu_2, \sigma^2)\right\|_{\mathrm{TV}} \\
&= \frac{1}{2}\int_{-\infty}^{+\infty} |f_1(x) - f_2(x)| \mathrm{d}x \\
&= \frac{1}{2}\int_{-\infty}^{(\mu_1+\mu_2)/2} \{f_1(x) - f_2(x)\} \mathrm{d}x + \frac{1}{2}\int_{(\mu_1+\mu_2)/2}^{+\infty} \{f_2(x) - f_1(x)\} \mathrm{d}x \\
&= \frac{1}{2}\left[\Phi\left(\frac{\mu_2 - \mu_1}{2\sigma}\right) - \Phi\left(\frac{\mu_1 - \mu_2}{2\sigma}\right) + 1 - \Phi\left(\frac{\mu_1 - \mu_2}{2\sigma}\right) - \left\{1 - \Phi\left(\frac{\mu_2 - \mu_1}{2\sigma}\right)\right\}\right] \\
&= 2\Phi\left(\frac{\mu_2 - \mu_1}{2\sigma}\right) - 1.
\end{aligned}$$

□

S3.2 Proof of Theorem 2.1

Proof of Theorem 2.1. The asymptotic normality of $\widetilde{\theta}_\alpha$, i.e., $\sqrt{n}(\widetilde{\theta}_\alpha - \theta_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 2\theta_0^2)$ as $n \to \infty$, has already been proved in Lemma S.10. In the following, we focus on proving the normal limit for the conditional posterior of θ.

From (11), the posterior density of θ can be written as

$$\pi(\theta|Y_n, \alpha) = \frac{\mathrm{e}^{\mathcal{L}_n\left(\alpha^{-2\nu}\theta, \alpha\right)} \pi\left(\theta|\alpha\right)}{\int_0^\infty \mathrm{e}^{\mathcal{L}_n\left(\alpha^{-2\nu}\theta, \alpha\right)} \pi\left(\theta|\alpha\right) \mathrm{d}\theta} = \frac{\mathrm{e}^{\mathcal{L}_n\left(\alpha^{-2\nu}\theta, \alpha\right) - \mathcal{L}_n\left(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)}}{\int_0^\infty \mathrm{e}^{\mathcal{L}_n\left(\alpha^{-2\nu}\theta, \alpha\right) - \mathcal{L}_n\left(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \mathrm{d}\theta}. \quad \text{(S.166)}$$

We can rewrite (S.144) in Lemma S.29 in terms of $\theta = \widetilde{\theta}_\alpha + (n-p)^{-1/2} t$:

$$\int_{\mathbb{R}} \left| \mathrm{e}^{\mathcal{L}_n\left(\alpha^{-2\nu}\theta, \alpha\right) - \mathcal{L}_n\left(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha\right)} \frac{\pi\left(\theta|\alpha\right)}{\pi(\theta_0|\alpha)} - \mathrm{e}^{-\frac{(n-p)(\theta - \widetilde{\theta}_\alpha)^2}{4\theta_0^2}} \right| \mathrm{d}\theta \le \frac{B_n(\alpha)}{\sqrt{n-p}}. \quad \text{(S.167)}$$

For the fixed $\alpha > 0$, define the events $\mathcal{E}_1'(\epsilon, \alpha) = \{|\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}| < \epsilon\}$ and $\mathcal{E}_1''(\epsilon) = \{|\widetilde{\theta}_{\alpha_0} - \theta_0| < \epsilon\}$ for any $\epsilon > 0$. From Lemma S.9, $\Pr\left\{\mathcal{E}_1'(\theta_0 n^{-1/2-\tau}/2, \alpha)\right\} \ge 1 - 4\exp(-4\log^2 n)$ for all sufficiently large n. From Lemma S.10, $\Pr\left\{\mathcal{E}_1''(5\theta_0 n^{-1/2}\log n)\right\} \ge 1 - 3\exp(-4\log^2 n)$ for all sufficiently large n. Since when n is sufficiently large,

$$\mathcal{E}_1'(\theta_0 n^{-1/2-\tau}/2, \alpha) \cap \mathcal{E}_1''(5\theta_0 n^{-1/2}\log n, \alpha) \subseteq \mathcal{E}_1(6\theta_0 n^{-1/2}\log n, \alpha),$$

we have that $\Pr\left\{\mathcal{E}_1(6\theta_0 n^{-1/2}\log n, \alpha)\right\} \ge 1 - 7\exp(-4\log^2 n)$. In the expression of $B_n(\alpha)$ in (S.145), we set $\epsilon_{1n} = 6\theta_0 n^{-1/2}\log n$ and $s_n = \log n$ which satisfies the conditions in Lemma S.29. By Assumption (A.2), for a fixed $\alpha > 0$, there exists some finite constant $C_1 > 0$ that depends on α, such that

$$\sup_{\theta \in \left(\frac{1}{2}\theta_0, \frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \le C_1, \qquad \sup_{\theta \in \left(\frac{3}{4}\theta_0, \frac{3}{2}\theta_0\right)} \left|\frac{\partial \log \pi(\theta|\alpha)}{\partial \theta}\right| \le C_1. \quad \text{(S.168)}$$

Hence, on the event $\mathcal{E}_1(6\theta_0 n^{-1/2}\log n, \alpha)$, the order of $B_n(\alpha)$ can be quantified from (S.145) in Lemma S.29:

$$B_n(\alpha) \le 4\theta_0 \exp\left(-\frac{n-p}{64}\right) + \frac{\sqrt{n}}{\pi(\theta_0|\alpha)} \exp\left\{-0.007(n-p)\right\}$$

$$
\begin{aligned}
&+10C_1\theta_0 \exp\left(-\frac{4\log^2 n}{125\theta_0^2}\right) + 4\theta_0 \exp\left(-\frac{\log^2 n}{4\theta_0^2}\right) \\
&+ \frac{8C_1}{\theta_0^2}\left(6\theta_0 n^{-1/2}\log^3 n + 2(n-p)^{-1/2}\log^3 n\right) + 4C_1^2\theta_0\left(6\theta_0+1\right)(n-p)^{-1/2}\log n \\
&\le C_2 n^{-1/2}\log^3 n \to 0, \ \text{ as } n\to\infty, \qquad \text{(S.169)}
\end{aligned}
$$

for some constant $C_2 > 0$ that depends on $\theta_0, p, \pi(\theta_0|\alpha)$ and C_1 in (S.168). This together with (S.167) implies that on the event $\mathcal{E}_1(6\theta_0 n^{-1/2}\log n, \alpha)$, the denominator of (S.166) converges to

$$
\int_{\mathbb{R}} \exp\left\{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}\right\} \mathrm{d}\theta = 2\theta_0\sqrt{\pi/(n-p)}.
$$

Now in Lemma S.30, we set f to be the numerator of (S.166) and g to be $\exp\left\{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}\right\}$. Using (S.169), we obtain that on the event $\mathcal{E}_1(6\theta_0 n^{-1/2}\log n, \alpha)$, as $n\to\infty$,

$$
\begin{aligned}
&\int_{\mathbb{R}} \left|\pi(\theta|Y_n,\alpha) - \frac{1}{2\sqrt{\pi/(n-p)}\theta_0}\exp\left\{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}\right\}\right| \mathrm{d}\theta \\
&\le \frac{2\int_{\mathbb{R}}\left|\mathrm{e}^{\mathcal{L}_n\left(\alpha^{-2\nu}\theta,\alpha\right)-\mathcal{L}_n\left(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha\right)}\frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} - \exp\left\{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}\right\}\right|\mathrm{d}\theta}{2\theta_0\sqrt{\pi/(n-p)}} \\
&\le \frac{B_n(\alpha)/\sqrt{n-p}}{\theta_0\sqrt{\pi/(n-p)}} = \frac{B_n(\alpha)}{\theta_0\sqrt{\pi}} \le C_3 n^{-1/2}\log^3 n \to 0, \qquad \text{(S.170)}
\end{aligned}
$$

for some constant $C_3 > 0$ that depends on $\theta_0, p, \pi(\theta_0|\alpha)$ and C_1 in (S.168).

Since $\Pr\left(\left\{\mathcal{E}_1(6\theta_0 n^{-1/2}\log n,\alpha)\right\}^c\right) \le 7\exp(-4\log^2 n)$ and $\sum_{n=1}^\infty 7\exp(-4\log^2 n) < \infty$, by the Borel-Cantelli lemma, we have shown that as $n\to\infty$ almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$,

$$
\left\|\Pi(\mathrm{d}\theta|Y_n,\alpha) - \mathcal{N}\left(\mathrm{d}\theta\middle|\widetilde{\theta}_\alpha, 2\theta_0^2/(n-p)\right)\right\|_{\mathrm{TV}} \le \frac{B_n(\alpha)}{2\theta_0\sqrt{\pi}} \to 0. \qquad \text{(S.171)}
$$

On the other hand, Theorem 1.3 of [Devroye et al., 2018] implies that

$$
\begin{aligned}
&\left\|\mathcal{N}\left(\mathrm{d}\theta\middle|\widetilde{\theta}_\alpha, 2\theta_0^2/n\right) - \mathcal{N}\left(\mathrm{d}\theta\middle|\widetilde{\theta}_\alpha, 2\theta_0^2/(n-p)\right)\right\|_{\mathrm{TV}} \\
&\le \frac{3}{2}\cdot\frac{2\theta_0^2/(n-p) - 2\theta_0^2/n}{2\theta_0^2/n} = \frac{3p}{2(n-p)} \to 0, \quad \text{as } n\to\infty. \qquad \text{(S.172)}
\end{aligned}
$$

Therefore, by (S.171), (S.172), and the triangle inequality, we have

$$
\left\|\Pi(\mathrm{d}\theta|Y_n,\alpha) - \mathcal{N}\left(\mathrm{d}\theta\middle|\widetilde{\theta}_\alpha, 2\theta_0^2/n\right)\right\|_{\mathrm{TV}} \le C_3 n^{-1/2}\log^3 n + \frac{3p}{2(n-p)} \le C_4 n^{-1/2}\log^3 n \to 0,
$$

as $n\to\infty$ almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$, for some constant $C_4 > 0$ that depends on $\theta_0, p, \pi(\theta_0|\alpha)$ and C_1 in (S.168). This completes the proof of Theorem 2.1. □

S3.3 Proof of Theorem 2.3

Proof of Theorem 2.3. It has been proved in Lemma S.28 that the profile posterior density (19) is well defined almost surely for every $n \ge p$. The convergence in total variation norm for the marginal posterior distributions of θ and α will follow trivially once the convergence for the joint posterior is proved. The convergence in total variation norm for the joint posterior (18) is implied by adding the following relations using a triangle inequality:

$$\int_0^\infty \int_{\mathbb{R}} \left| \pi(\theta, \alpha|Y_n) - \frac{\sqrt{n-p}}{2\sqrt{\pi}\theta_0} \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} \cdot \widetilde{\pi}(\alpha|Y_n) \right| \mathrm{d}\theta \mathrm{d}\alpha \to 0, \tag{S.173}$$

$$\int_0^\infty \int_{\mathbb{R}} \left| \frac{\sqrt{n-p}}{2\sqrt{\pi}\theta_0} \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} - \frac{\sqrt{n}}{2\sqrt{\pi}\theta_0} \mathrm{e}^{-\frac{n(\theta-\widetilde{\theta}_{\alpha_0})^2}{4\theta_0^2}} \right| \cdot \widetilde{\pi}(\alpha|Y_n) \mathrm{d}\theta \mathrm{d}\alpha \to 0, \tag{S.174}$$

as $n \to \infty$ almost surely $P_{(\beta_0, \sigma_0^2, \alpha_0)}$. We prove (S.173) and (S.174) respectively.

<u>Proof of (S.173):</u>

In Lemma S.30, we take

$$\begin{aligned} f &= \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\theta, \alpha) - \mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta|\alpha) \cdot \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\alpha), \\ g &= \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} \pi(\theta_0|\alpha) \cdot \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\alpha), \end{aligned}$$

such that by applying Lemma S.30, we can obtain that

$$\begin{aligned} &\int_0^\infty \int_{\mathbb{R}} \left| \pi(\theta, \alpha|Y_n) - \frac{\sqrt{n-p}}{2\sqrt{\pi}\theta_0} \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} \cdot \widetilde{\pi}(\alpha|Y_n) \right| \mathrm{d}\theta \mathrm{d}\alpha \\ =& \int_0^\infty \int_{\mathbb{R}} \left| \frac{\mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\theta, \alpha) - \mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \cdot \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta|\alpha)\pi(\alpha)}{\int_0^\infty \int_0^\infty \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\theta, \alpha) - \mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \cdot \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta|\alpha)\pi(\alpha) \mathrm{d}\theta \mathrm{d}\alpha} \right. \\ &- \left. \frac{\mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} \cdot \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta_0|\alpha)\pi(\alpha)}{\int_0^\infty \int_{\mathbb{R}} \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} \cdot \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta_0|\alpha)\pi(\alpha) \mathrm{d}\theta \mathrm{d}\alpha} \right| \mathrm{d}\theta \mathrm{d}\alpha \le \frac{\mathsf{N}}{\mathsf{D}}, \end{aligned} \tag{S.175}$$

where (with $\varrho_n(t;\alpha)$ defined in (S.143))

$$\begin{aligned} \mathsf{N} &= 2\int_0^\infty \int_{\mathbb{R}} \left| \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\theta, \alpha) - \mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} - \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} \right| \\ &\quad \times \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta_0|\alpha)\pi(\alpha) \mathrm{d}\theta \mathrm{d}\alpha \\ &= 2\int_0^\infty \int_{\mathbb{R}} \left| \varrho_n(\sqrt{n-p}(\theta - \widetilde{\theta}_\alpha); \alpha) \right| \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta_0|\alpha)\pi(\alpha) \mathrm{d}\theta \mathrm{d}\alpha, \end{aligned} \tag{S.176}$$

$$\begin{aligned} \mathsf{D} &= \int_0^\infty \int_{\mathbb{R}} \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} \cdot \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta_0|\alpha)\pi(\alpha) \mathrm{d}\theta \mathrm{d}\alpha \\ &= \frac{2\theta_0\sqrt{\pi}}{\sqrt{n-p}} \int_0^\infty \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta_0|\alpha)\pi(\alpha) \mathrm{d}\alpha, \end{aligned} \tag{S.177}$$

We decompose the numerator in (S.176) into three terms:

$$\begin{aligned} \mathsf{N} &= \mathsf{N}_1 + \mathsf{N}_2 + \mathsf{N}_3, \\ \mathsf{N}_1 &= 2\int_{\underline{\alpha}_n}^{\overline{\alpha}_n} \int_{\mathbb{R}} \left| \varrho_n(\sqrt{n-p}(\theta - \widetilde{\theta}_\alpha); \alpha) \right| \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta_0|\alpha)\pi(\alpha) \mathrm{d}\theta \mathrm{d}\alpha, \\ \mathsf{N}_2 &= 2\int_0^{\underline{\alpha}_n} \int_{\mathbb{R}} \left| \varrho_n(\sqrt{n-p}(\theta - \widetilde{\theta}_\alpha); \alpha) \right| \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta_0|\alpha)\pi(\alpha) \mathrm{d}\theta \mathrm{d}\alpha, \\ \mathsf{N}_3 &= 2\int_{\overline{\alpha}_n}^\infty \int_{\mathbb{R}} \left| \varrho_n(\sqrt{n-p}(\theta - \widetilde{\theta}_\alpha); \alpha) \right| \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta_0|\alpha)\pi(\alpha) \mathrm{d}\theta \mathrm{d}\alpha, \end{aligned} \tag{S.178}$$

To show (S.173), from (S.175) and (S.178), it suffices to show that $\mathsf{N}_j/\mathsf{D} \to 0$ for $j = 1, 2, 3$ as $n \to \infty$ almost surely $P_{(\beta_0, \sigma_0^2, \alpha_0)}$.

Proof of $\mathsf{N}_1/\mathsf{D} \to 0$:

We consider all $\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]$. For any $\epsilon > 0$, define three events

$$\mathcal{E}_2(\epsilon) = \Big\{ \sup_{\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]} |\widetilde{\theta}_\alpha - \theta_0| < \epsilon \Big\}, \quad \mathcal{E}_3(\epsilon) = \Big\{ \sup_{\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]} |\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}| < \epsilon \Big\},$$
$$\mathcal{E}_4(\epsilon) = \Big\{ |\widetilde{\theta}_{\alpha_0} - \theta_0| < \epsilon \Big\}. \tag{S.179}$$

For sufficiently large n, Lemma S.9 shows that $\Pr\{\mathcal{E}_3(\theta_0 n^{-1/2-\tau}/2)\} \geq 1 - 4\exp(-4\log^2 n)$ for some constant $\tau \in (0, 1/2)$. Lemma S.10 shows that $\Pr\{\mathcal{E}_4(5\theta_0 n^{-1/2}\log n)\} \geq 1 - 3\exp(-4\log^2 n)$. By the triangle inequality, for sufficiently large n,

$$\mathcal{E}_2(6\theta_0 n^{-1/2}\log n) \supseteq \mathcal{E}_3(\theta_0 n^{-1/2-\tau}/2) \cap \mathcal{E}_4(5\theta_0 n^{-1/2}\log n),$$

it follows that $\Pr\{\mathcal{E}_2(6\theta_0 n^{-1/2}\log n)\} \geq 1 - 7\exp(-4\log^2 n)$.

We again use the inequality (S.167) from Lemma S.29, with $B_n(\alpha)$ defined in (S.145) with $\epsilon_{1n} = 6\theta_0 n^{-1/2}\log n$ and $s_n = \log n$. Since $\mathcal{E}_1(6\theta_0 n^{-1/2}\log n, \alpha) \supseteq \mathcal{E}_2(6\theta_0 n^{-1/2}\log n)$ for every $\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]$, Lemma S.29 can be applied to all $\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]$ with $\epsilon_{1n} = 6\theta_0 n^{-1/2}\log n$ and $s_n = \log n$. Therefore, (S.167) holds uniformly for all $\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]$ on the event $\mathcal{E}_2(6\theta_0 n^{-1/2}\log n)$, such that $\Pr\{\mathcal{E}_2(6\theta_0 n^{-1/2}\log n)\} \geq 1 - 7\exp(-4\log^2 n)$.

Integrating (S.167) over the interval $[\underline{\alpha}_n, \overline{\alpha}_n]$ gives that

$$\begin{aligned}
&\int_{\underline{\alpha}_n}^{\overline{\alpha}_n} \int_{\mathbb{R}} \left| \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\theta,\alpha) - \mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} - \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} \right| \\
&\quad \times \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\theta\mathrm{d}\alpha \\
&\leq \int_{\underline{\alpha}_n}^{\overline{\alpha}_n} \frac{B_n(\alpha)}{\sqrt{n-p}} \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \\
&\leq \frac{\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} B_n(\alpha)}{\sqrt{n-p}} \int_{\underline{\alpha}_n}^{\overline{\alpha}_n} \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha.
\end{aligned} \tag{S.180}$$

According to Assumption (A.3), with $\epsilon_{1n} = 6\theta_0 n^{-1/2}\log n$ and $s_n = \log n$, $B_n(\alpha)$ as defined in (S.145) satisfies that for all sufficiently large n,

$$\begin{aligned}
&\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} B_n(\alpha) \\
&\leq 4\theta_0 \exp\left(-\frac{n-p}{64}\right) + \frac{\sqrt{n-p}}{\inf_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \pi(\theta_0|\alpha)} \exp\{-0.007(n-p)\} \\
&+ \sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \sup_{\theta\in\left(\frac{1}{2}\theta_0,\frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \cdot 10\theta_0 \exp\left(-\frac{4\log^2 n}{125\theta_0^2}\right) + 4\theta_0 \exp\left(-\frac{\log^2 n}{4\theta_0^2}\right) \\
&+ \frac{8}{\theta_0^2}\left(\frac{6\theta_0\log^3 n}{\sqrt{n}} + \frac{2\log^3 n}{\sqrt{n-p}}\right) \cdot \sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \sup_{\theta\in\left(\frac{3}{4}\theta_0,\frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \\
&+ 4\theta_0 \sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \sup_{\theta\in\left(\frac{3}{4}\theta_0,\frac{3}{2}\theta_0\right)} \left| \frac{\partial \log\pi(\theta|\alpha)}{\partial\theta} \right| \\
&\times \sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \sup_{\theta\in\left(\frac{3}{4}\theta_0,\frac{3}{2}\theta_0\right)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \left(\frac{6\theta_0\log n}{\sqrt{n}} + \frac{\log n}{\sqrt{n-p}}\right)
\end{aligned}$$

$$
\leq 4\theta_0 \exp\left(-\frac{n-p}{64}\right) + \exp\left(n^{C_{\pi,3}}\right) \cdot \sqrt{n}\exp\{-0.007(n-p)\} + n^{C_{\pi,2}} \cdot 10\theta_0 \exp\left(-\frac{4\log^2 n}{125\theta_0^2}\right)
$$

$$
+ 4\theta_0 \exp\left(-\frac{\log^2 n}{4\theta_0^2}\right) + \frac{8(6\theta_0+2)}{\theta_0^2}\frac{\log^3 n}{\sqrt{n-p}} \cdot n^{C_{\pi,2}} + 4(6\theta_0+1)\theta_0 n^{C_{\pi,1}+C_{\pi,2}} \cdot \frac{\log n}{\sqrt{n-p}}
$$

$$
\to 0, \text{ as } n \to \infty, \tag{S.181}
$$

where in the last step, we have used the fact that $C_{\pi,3} < 1$ and $C_{\pi,1} + C_{\pi,2} < 1/2$ according to Assumption (A.3).

Therefore, (S.180), (S.181), (S.178), and (S.177) together imply that on the event $\mathcal{E}_2(6\theta_0 n^{-1/2}\log n)$,

$$
\begin{aligned}
\frac{\mathsf{N}_1}{\mathsf{D}} &= \frac{2\int_{\underline{\alpha}_n}^{\overline{\alpha}_n}\int_{\mathbb{R}} \left|\varrho_n(\sqrt{n-p}(\theta-\widetilde{\theta}_\alpha);\alpha)\right| \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\theta\mathrm{d}\alpha}{\frac{2\theta_0\sqrt{\pi}}{\sqrt{n-p}}\int_0^\infty \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha} \\
&\leq \frac{\frac{\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} B_n(\alpha)}{\sqrt{n-p}} \int_{\underline{\alpha}_n}^{\overline{\alpha}_n} \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}{\frac{\theta_0\sqrt{\pi}}{\sqrt{n-p}}\int_{\underline{\alpha}_n}^{\overline{\alpha}_n} \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha} \\
&\leq \frac{\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} B_n(\alpha)}{\theta_0\sqrt{\pi}} \to 0,
\end{aligned} \tag{S.182}
$$

as $n \to \infty$. Since $\Pr\{\mathcal{E}_2(6\theta_0 n^{-1/2}\log n)^c\} \leq 7\exp(-4\log^2 n)$ and $\sum_{n=1}^\infty 7\exp(-4\log^2 n) < \infty$, by the Borel-Cantelli lemma, we have shown that $\mathsf{N}_1/\mathsf{D} \to 0$ as $n \to \infty$ almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$.

Proof of $\mathsf{N}_2/\mathsf{D} \to 0$:

We start with an upper bound for N_2:

$$
\begin{aligned}
\mathsf{N}_2 &= 2\int_0^{\underline{\alpha}_n}\int_{\mathbb{R}} \left| \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\theta,\alpha)-\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\theta|\alpha) - \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}}\pi(\theta_0|\alpha)\right| \\
&\quad \times \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\alpha)\mathrm{d}\theta\mathrm{d}\alpha, \\
&\leq 2\int_0^{\underline{\alpha}_n}\int_{\mathbb{R}} \left(\mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\theta,\alpha)-\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\theta|\alpha) + \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}}\pi(\theta_0|\alpha)\right) \\
&\quad \times \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\alpha)\mathrm{d}\theta\mathrm{d}\alpha, \\
&\overset{(i)}{\leq} 2\int_0^{\underline{\alpha}_n}\left\{\int_0^\infty \pi(\theta|\alpha)\mathrm{d}\theta\right\}\mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\alpha)\mathrm{d}\alpha \\
&\quad + 2\int_0^{\underline{\alpha}_n}\left\{\int_{\mathbb{R}} \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}}\mathrm{d}\theta\right\}\mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \\
&\leq 2\int_0^{\underline{\alpha}_n}\mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\alpha)\mathrm{d}\alpha + \frac{4\theta_0\sqrt{\pi}}{\sqrt{n-p}}\int_0^{\underline{\alpha}_n}\mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha,
\end{aligned} \tag{S.183}
$$

where (i) follows from the fact that $\mathcal{L}_n(\alpha^{-2\nu}\theta,\alpha) \leq \mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)$ as $\widetilde{\theta}_\alpha$ is the maximizer of $\mathcal{L}_n(\alpha^{-2\nu}\theta,\alpha)$ given α.

On the other hand, since $2\nu+d > 1$, we choose $c = 1 > 1/(2\nu+d)$ in Lemma S.24, and define $\mathcal{E}_5$ to be the event that (S.120) in Lemma S.24 happens, such that $\Pr(\mathcal{E}_5) \geq 1-9\exp(-4\log^2 n)$. Then on the event $\mathcal{E}_5$, the denominator (S.177) can be lower bounded by

$$
\mathsf{D} \geq \frac{2\theta_0\sqrt{\pi}}{\sqrt{n-p}}\mathrm{e}^{\widetilde{\mathcal{L}}_n(\alpha_0)}\int_{\alpha_0}^{(1+n^{-1})\alpha_0}\mathrm{e}^{\widetilde{\mathcal{L}}_n(\alpha)-\widetilde{\mathcal{L}}_n(\alpha_0)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha
$$

$$
\begin{aligned}
&\geq \frac{2\theta_0\sqrt{\pi}}{\sqrt{n-p}}\exp\left\{\widetilde{\mathcal{L}}_n(\alpha_0)-3\log^4 n\right\}\int_{\alpha_0}^{(1+n^{-1})\alpha_0}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \\
&\overset{(i)}{\geq} \frac{2\theta_0\sqrt{\pi}c_{\pi,0}}{n\sqrt{n-p}}\exp\left\{\widetilde{\mathcal{L}}_n(\alpha_0)-3\log^4 n\right\}, \qquad \text{(S.184)}
\end{aligned}
$$

where $c_{\pi,0}=\pi(\theta_0|\alpha_0)\pi(\alpha_0)\cdot\alpha_0/4$, and the inequality (i) holds because by Assumptions (A.2) and (A.4), $\pi(\theta_0|\alpha)>\pi(\theta_0|\alpha_0)/2>0$ and $\pi(\alpha)>\pi(\alpha_0)/2>0$ for all $\alpha\in[\alpha_0,(1+n^{-1})\alpha_0]$ and sufficiently large n, such that $\int_{\alpha_0}^{(1+n^{-1})\alpha_0}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha\geq n^{-1}\alpha_0\cdot\pi(\theta_0|\alpha_0)\pi(\alpha_0)/4=c_{\pi,0}n^{-1}$.

We combine (S.183) and (S.184) to obtain that

$$
\begin{aligned}
\frac{\mathsf{N}_2}{\mathsf{D}} &\leq \frac{n^{3/2}}{\theta_0\sqrt{\pi}c_{\pi,0}}\exp\left(3\log^4 n\right)\int_0^{\underline{\alpha}_n}\mathrm{e}^{\widetilde{\mathcal{L}}_n(\alpha)-\widetilde{\mathcal{L}}_n(\alpha_0)}\pi(\alpha)\mathrm{d}\alpha \\
&\quad+\frac{2n}{c_{\pi,0}}\exp\left(3\log^4 n\right)\int_0^{\underline{\alpha}_n}\mathrm{e}^{\widetilde{\mathcal{L}}_n(\alpha)-\widetilde{\mathcal{L}}_n(\alpha_0)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha. \qquad \text{(S.185)}
\end{aligned}
$$

To upper bound the two terms in (S.185), we first derive a simple relation for the part $\exp\{\widetilde{\mathcal{L}}_n(\alpha)-\widetilde{\mathcal{L}}_n(\alpha_0)\}$. Let $\mathcal{E}_6$ be the event on which (S.129) in Lemma S.25 happens, such that $\Pr(\mathcal{E}_6)\geq 1-10\exp(-4\log^2 n)$ for sufficiently large n. On the event $\mathcal{E}_6$, the monotonicity bound from Lemma S.22 and the upper bound from Lemma S.25 imply that for any $\alpha\in(0,\underline{\alpha}_n)$,

$$
\begin{aligned}
&\exp\left\{\widetilde{\mathcal{L}}_n(\alpha)-\widetilde{\mathcal{L}}_n(\alpha_0)\right\} \\
=\ &\exp\left\{\widetilde{\mathcal{L}}_n(\alpha)-\widetilde{\mathcal{L}}_n(\underline{\alpha}_n)\right\}\cdot\exp\left\{\widetilde{\mathcal{L}}_n(\underline{\alpha}_n)-\widetilde{\mathcal{L}}_n(\alpha_0)\right\} \\
<\ &\left(\frac{\underline{\alpha}_n}{\alpha}\right)^{n(\nu+d/2)}\exp\left(3n^{1/2-\tau}\right) \\
=\ &\alpha^{-n(\nu+d/2)}\exp\left\{-(\nu+d/2)\underline{\kappa}n\log n+3n^{1/2-\tau}\right\}, \qquad \text{(S.186)}
\end{aligned}
$$

where $\tau\in(0,1/2)$ and $\underline{\kappa}\in(0,1/2)$ are defined in (S.9). Since $3\log^4 n/n^{1/2-\tau}\to 0$ as $n\to\infty$, we now plug (S.186) in (S.185) and use Assumption (A.4) to obtain that on the event $\mathcal{E}_5\cap\mathcal{E}_6$,

$$
\begin{aligned}
\frac{\mathsf{N}_2}{\mathsf{D}} &\leq \frac{n^{3/2}}{\theta_0\sqrt{\pi}c_{\pi,0}}\exp\left\{-(\nu+d/2)\underline{\kappa}n\log n+4n^{1/2-\tau}\right\}\int_0^{\underline{\alpha}_n}\alpha^{-n(\nu+d/2)}\pi(\alpha)\mathrm{d}\alpha \\
&\quad+\frac{2n}{c_{\pi,0}}\exp\left\{-(\nu+d/2)\underline{\kappa}n\log n+4n^{1/2-\tau}\right\}\int_0^{\underline{\alpha}_n}\alpha^{-n(\nu+d/2)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \\
&\leq \frac{n^{3/2}}{\theta_0\sqrt{\pi}c_{\pi,0}}\exp\left\{-(\nu+d/2)\underline{\kappa}n\log n+4n^{1/2-\tau}+\underline{c_\pi}n\log n\right\} \\
&\quad+\frac{2n}{c_{\pi,0}}\exp\left\{-(\nu+d/2)\underline{\kappa}n\log n+4n^{1/2-\tau}+\underline{c_\pi}n\log n\right\} \\
&\to 0,\ \text{as } n\to\infty, \qquad \text{(S.187)}
\end{aligned}
$$

where the last step follows because $\underline{c_\pi}<(\nu+d/2)\underline{\kappa}$ by Assumption (A.4) and $\tau\in(0,1/2)$. Since $\Pr\{(\mathcal{E}_5\cap\mathcal{E}_6)^c\}\leq 20\exp(-4\log^2 n)$ and $\sum_{n=1}^\infty 20\exp(-4\log^2 n)<\infty$, by the Borel-Cantelli lemma, we have shown that $\mathsf{N}_2/\mathsf{D}\to 0$ as $n\to\infty$ almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$.

Proof of $\mathsf{N}_3/\mathsf{D}\to 0$:

Similar to the derivation of (S.183), we have the following upper bound for N_3:

$$
\mathsf{N}_3\leq 2\int_{\overline{\alpha}_n}^\infty \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\alpha)\mathrm{d}\alpha+\frac{4\theta_0\sqrt{\pi}}{\sqrt{n-p}}\int_{\overline{\alpha}_n}^\infty \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha. \qquad \text{(S.188)}
$$

(S.184) and (S.188) imply that on the event $\mathcal{E}_5$,

$$\frac{\mathsf{N}_3}{\mathsf{D}} \leq \frac{n^{3/2}}{\theta_0\sqrt{\pi}c_{\pi,0}} \exp\left(3\log^4 n\right) \int_{\overline{\alpha}_n}^{\infty} \mathrm{e}^{\widetilde{\mathcal{L}}_n(\alpha)-\widetilde{\mathcal{L}}_n(\alpha_0)}\pi(\alpha)\mathrm{d}\alpha$$
$$+ \frac{2n}{c_{\pi,0}} \exp\left(3\log^4 n\right) \int_{\overline{\alpha}_n}^{\infty} \mathrm{e}^{\widetilde{\mathcal{L}}_n(\alpha)-\widetilde{\mathcal{L}}_n(\alpha_0)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha. \tag{S.189}$$

Let $\mathcal{E}_7$ be the event on which (S.135) in Lemma S.27 happens, such that $\Pr(\mathcal{E}_7) \geq 1 - 10\exp(-4\log^2 n)$ for sufficiently large n. Similar to the proof of $\mathsf{N}_2/\mathsf{D} \to 0$, on the event $\mathcal{E}_7$, we use Lemma S.22 and Lemma S.27 to obtain that for any $\alpha \in (\overline{\alpha}_n, +\infty)$,

$$\begin{aligned}
&\exp\left\{\widetilde{\mathcal{L}}_n(\alpha) - \widetilde{\mathcal{L}}_n(\alpha_0)\right\} \\
&= \exp\left\{\widetilde{\mathcal{L}}_n(\alpha) - \widetilde{\mathcal{L}}_n(\overline{\alpha}_n)\right\} \cdot \exp\left\{\widetilde{\mathcal{L}}_n(\overline{\alpha}_n) - \widetilde{\mathcal{L}}_n(\alpha_0)\right\} \\
&< \left(\frac{\alpha}{\overline{\alpha}_n}\right)^{n(\nu+d/2)} \exp\left(C_{p,1}n^{\kappa_1}\log n\right) \\
&= \alpha^{n(\nu+d/2)} \exp\left\{-(\nu+d/2)\overline{\kappa} n\log n + C_{p,1}n^{\kappa_1}\log n\right\},
\end{aligned} \tag{S.190}$$

where $C_{p,1} > 0$ and $\kappa_1 \in (1/2-\tau, 1)$ are given in Lemma S.27, and $\overline{\kappa} \in (0, 1/2)$ is given in (S.9). Since $3\log^4 n/(C_{p,1}n^{\kappa_1}\log n) \to 0$ as $n \to \infty$, we now plug (S.190) in (S.189) and use Assumption (A.4) to obtain that on the event $\mathcal{E}_5 \cap \mathcal{E}_7$,

$$\begin{aligned}
\frac{\mathsf{N}_3}{\mathsf{D}} &\leq \frac{n^{3/2}}{\theta_0\sqrt{\pi}c_{\pi,0}} \exp\left\{-(\nu+d/2)\overline{\kappa} n\log n + 2C_{p,1}n^{\kappa_1}\log n\right\} \times \int_{\overline{\alpha}_n}^{\infty} \alpha^{n(\nu+d/2)}\pi(\alpha)\mathrm{d}\alpha \\
&\quad + \frac{2n}{c_{\pi,0}} \exp\left\{-(\nu+d/2)\overline{\kappa} n\log n + 2C_{p,1}n^{\kappa_1}\log n\right\} \times \int_{\overline{\alpha}_n}^{\infty} \alpha^{n(\nu+d/2)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \\
&\leq \frac{n^{3/2}}{\theta_0\sqrt{\pi}c_{\pi,0}} \exp\left\{-(\nu+d/2)\overline{\kappa} n\log n + 2C_{p,1}n^{\kappa_1}\log n + \overline{c_\pi} n\log n\right\} \\
&\quad + \frac{2n}{c_{\pi,0}} \exp\left\{-(\nu+d/2)\overline{\kappa} n\log n + 2C_{p,1}n^{\kappa_1}\log n + \overline{c_\pi} n\log n\right\} \\
&\to 0, \text{ as } n \to \infty,
\end{aligned} \tag{S.191}$$

where the last step follows because $\overline{c_\pi} < (\nu+d/2)\overline{\kappa}$ by Assumption (A.4) and $\kappa_1 \in (1/2 - \tau, 1)$. Since $\Pr\{(\mathcal{E}_5 \cap \mathcal{E}_7)^c\} \leq 20\exp(-4\log^2 n)$ and $\sum_{n=1}^{\infty} 20\exp(-4\log^2 n) < \infty$, by the Borel-Cantelli lemma, we have shown that $\mathsf{N}_3/\mathsf{D} \to 0$ as $n \to \infty$ almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$.

Proof of (S.174):

We use Lemma S.31 and obtain that

$$\begin{aligned}
&\int_0^{\infty}\int_{\mathbb{R}} \left|\frac{\sqrt{n-p}}{2\sqrt{\pi}\theta_0}\mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} - \frac{\sqrt{n}}{2\sqrt{\pi}\theta_0}\mathrm{e}^{-\frac{n(\theta-\widetilde{\theta}_{\alpha_0})^2}{4\theta_0^2}}\right| \cdot \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\theta\mathrm{d}\alpha \\
&= \int_0^{\infty} \left\|\mathcal{N}(\widetilde{\theta}_\alpha, 2\theta_0^2/(n-p)) - \mathcal{N}(\widetilde{\theta}_{\alpha_0}, 2\theta_0^2/n)\right\|_{\mathrm{TV}} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha \\
&\overset{(i)}{\leq} \int_0^{\infty} \left\|\mathcal{N}(\widetilde{\theta}_\alpha, 2\theta_0^2/(n-p)) - \mathcal{N}(\widetilde{\theta}_{\alpha_0}, 2\theta_0^2/(n-p))\right\|_{\mathrm{TV}} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha \\
&\quad + \int_0^{\infty} \left\|\mathcal{N}(\widetilde{\theta}_{\alpha_0}, 2\theta_0^2/(n-p)) - \mathcal{N}(\widetilde{\theta}_{\alpha_0}, 2\theta_0^2/n)\right\|_{\mathrm{TV}} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha \\
&\overset{(ii)}{\leq} \int_0^{\infty} \left\{2\Phi\left(\frac{(n-p)^{1/2}|\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}|}{2\sqrt{2}\theta_0}\right) - 1\right\} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha
\end{aligned}$$

$$
\begin{aligned}
&+ \int_0^{\infty} \frac{3}{2} \cdot \frac{2\theta_0^2/(n-p) - 2\theta_0^2/n}{2\theta_0^2/n} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha \\
\overset{(iii)}{\leq} & \int_{\underline{\alpha}_n}^{\overline{\alpha}_n} \frac{(n-p)^{1/2}}{2\sqrt{\pi}\theta_0} |\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}| \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha \\
&+ \int_0^{\underline{\alpha}_n} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha + \int_{\overline{\alpha}_n}^{\infty} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha + \frac{3p}{2(n-p)}, \qquad \text{(S.192)}
\end{aligned}
$$

where (i) follows from the triangle inequality of total variation distance; (ii) follows from Lemma S.31 and Theorem 1.3 of [Devroye et al., 2018]; for (iii), we use the relation $\Phi(x) - 0.5 = \Phi(x) - \Phi(0) \leq \phi(0)x = x/\sqrt{2\pi}$ for all $x \geq 0$ (where $\phi(x)$ is the standard normal density), and the direct bound $|2\Phi(x) - 1| \leq 1$ for all $x \in \mathbb{R}$.

On the event $\mathcal{E}_3(\theta_0 n^{-1/2-\tau}/2)$, we have that $n^{1/2}|\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}| \leq \theta_0 n^{-\tau}/2$ uniformly for all $\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]$. Together with the fact that $\widetilde{\pi}(\alpha|Y_n)$ is almost surely a proper probability density from Lemma S.28, we can derive from (S.192) that on the event $\mathcal{E}_3(\theta_0 n^{-1/2-\tau}/2)$,

$$
\int_{\underline{\alpha}_n}^{\overline{\alpha}_n} \frac{(n-p)^{1/2}}{2\sqrt{\pi}\theta_0} |\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}| \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha \leq \frac{n^{-\tau}}{4\sqrt{\pi}} \int_0^{\infty} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha \leq \frac{n^{-\tau}}{4\sqrt{\pi}} \to 0, \qquad \text{(S.193)}
$$

as $n \to \infty$. Since $\Pr\left\{\mathcal{E}_3(\theta_0 n^{-1/2-\tau}/2)^c\right\} \leq 4\exp(-4\log^2 n)$ and $\sum_{n=1}^{\infty} 4\exp(-4\log^2 n) < \infty$, by the Borel-Cantelli lemma, we have shown that (S.193) holds as $n \to \infty$ almost surely $P_{(\beta_0, \sigma_0^2, \alpha_0)}$.

For the second term on the right-hand side of (S.192), we have that by the definition (19),

$$
\int_0^{\underline{\alpha}_n} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha \leq \frac{\int_0^{\underline{\alpha}_n} \mathrm{e}^{\widetilde{\mathcal{L}}_n(\alpha) - \widetilde{\mathcal{L}}_n(\alpha_0)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}{\int_{\alpha_0}^{(1+n^{-1})\alpha_0} \mathrm{e}^{\widetilde{\mathcal{L}}_n(\alpha) - \widetilde{\mathcal{L}}_n(\alpha_0)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}.
$$

The denominator is lower bounded by $c_{\pi,0} n^{-1} \exp(-3\log^4 n)$ on the event $\mathcal{E}_5$, similar to the proof of (S.184). The numerator can be upper bounded on the event $\mathcal{E}_6$, using the same derivation as in (S.185) and (S.186). As a result, on the event $\mathcal{E}_5 \cap \mathcal{E}_6$, using $\underline{c_\pi} < (\nu + d/2)\underline{\kappa}$ in Assumption (A.4), we have that

$$
\begin{aligned}
\int_0^{\underline{\alpha}_n} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha &\leq \frac{\exp\left\{-(\nu + d/2)\underline{\kappa} n \log n + 3n^{1/2-\tau}\right\} \int_0^{\underline{\alpha}_n} \alpha^{-n(\nu+d/2)} \pi(\alpha)\mathrm{d}\alpha}{c_{\pi,0} n^{-1} \exp(-3\log^4 n)} \\
&\leq \frac{n}{c_{\pi,0}} \exp\left\{-(\nu + d/2)\underline{\kappa} n \log n + 3n^{1/2-\tau} + \underline{c_\pi} n \log n + 3\log^4 n\right\} \\
&\to 0, \text{ as } n \to \infty. \qquad \text{(S.194)}
\end{aligned}
$$

(S.194) holds as $n \to \infty$ almost surely $P_{(\beta_0, \sigma_0^2, \alpha_0)}$ since $\Pr\{(\mathcal{E}_5 \cap \mathcal{E}_6)^c\} \leq 20\exp(-4\log^2 n)$ and $\sum_{n=1}^{\infty} 20\exp(-4\log^2 n) < \infty$.

Similarly, for the third term on the right-hand side of (S.192), we have that by the definition (19),

$$
\int_{\overline{\alpha}_n}^{\infty} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha \leq \frac{\int_{\overline{\alpha}_n}^{\infty} \mathrm{e}^{\widetilde{\mathcal{L}}_n(\alpha) - \widetilde{\mathcal{L}}_n(\alpha_0)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}{\int_{\alpha_0}^{(1+n^{-1})\alpha_0} \mathrm{e}^{\widetilde{\mathcal{L}}_n(\alpha) - \widetilde{\mathcal{L}}_n(\alpha_0)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}.
$$

On the event $\mathcal{E}_5 \cap \mathcal{E}_7$, the denominator is lower bounded by $c_{\pi,0} n^{-1} \exp(-3\log^4 n)$, and the numerator can be upper bounded using the same derivation as in (S.190) and (S.191). As a result, using $\overline{c_\pi} < (\nu + d/2)\overline{\kappa}$ in Assumption (A.4), we have that on $\mathcal{E}_5 \cap \mathcal{E}_7$,

$$
\int_{\overline{\alpha}_n}^{\infty} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha
$$

$$
\begin{aligned}
&\leq \frac{\exp\left\{-(\nu+d/2)\overline{\kappa}n\log n + C_{p,1}n^{\kappa_1}\log n\right\}\int_{\overline{\alpha}_n}^{\infty}\alpha^{n(\nu+d/2)}\pi(\alpha)\mathrm{d}\alpha}{c_{\pi,0}n^{-1}\exp(-3\log^4 n)} \\
&\leq \frac{n}{c_{\pi,0}}\exp\left\{-(\nu+d/2)\overline{\kappa}n\log n + C_{p,1}n^{\kappa_1}\log n + \overline{c_\pi}n\log n + 3\log^4 n\right\} \\
&\to 0, \text{ as } n\to\infty. \qquad \text{(S.195)}
\end{aligned}
$$

(S.195) holds as $n\to\infty$ almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$ since $\Pr\{(\mathcal{E}_5\cap\mathcal{E}_7)^c\}\leq 20\exp(-4\log^2 n)$ and $\sum_{n=1}^{\infty}20\exp(-4\log^2 n)<\infty$.

Finally, (S.193), (S.194), and (S.195) together imply that the right-hand side of (S.192) converges to zero as $n\to\infty$ almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$. This has proved (S.174), and hence has completed the proof of Theorem 2.3. □

S3.4 Limiting Posterior Distribution When $d\geq 5$

We present a theorem for the limiting posterior distribution of (θ,α) when the domain dimension $d\geq 5$ in the universal kriging model (1) with the isotropic Matérn covariance function (2). The theorem is similar to Theorem 2.3 for the case of $d\in\{1,2,3\}$ but requires more assumptions and has some important difference in its proof from that of Theorem 2.3, mainly because that the range parameter α can be consistently estimated for $d\geq 5$ ([Anderes, 2010]).

For any $\epsilon_1>0,\epsilon_2>0$, we define the set

$$
\mathcal{B}_0(\epsilon_1,\epsilon_2)=\Big\{(\beta,\theta,\alpha)\in\mathbb{R}^p\times\mathbb{R}^+\times\mathbb{R}^+ : |\theta/\theta_0-1|<\epsilon_1, |\alpha/\alpha_0-1|<\epsilon_2\Big\}. \qquad \text{(S.196)}
$$

This set can be viewed as a neighborhood of (θ_0,α_0). For the case of $d\geq 5$, the following assumptions will replace Assumption (A.4) in the main text for the case of $d\in\{1,2,3\}$.

(S.1) For the model (1) with isotropic Matérn covariance function in (2) with $d\geq 5$, there exist constants $0<\kappa_1'\leq 1/2$, $1/(2\nu+d)<\kappa_2'\leq 1/2$, $c_5>0$ and consistent estimators $\widehat{\theta}_n$ for θ and $\widehat{\alpha}_n$ for α based on (Y_n,M_n), such that for any $\epsilon_1>0$ and $\epsilon_2>0$,

$$
\begin{aligned}
P_{(\beta_0,\sigma_0^2,\alpha_0)}\left(\left|\widehat{\theta}_n/\theta_0-1\right|\geq\epsilon_1/2\right)&\leq\exp\left\{-c_5\min\left[n^{\kappa_1'}\epsilon_1,(n^{\kappa_1'}\epsilon_1)^2\right]\right\}, \\
P_{(\beta_0,\sigma_0^2,\alpha_0)}\left(|\widehat{\alpha}_n/\alpha_0-1|\geq\epsilon_2/2\right)&\leq\exp\left\{-c_5\min\left[n^{\kappa_2'}\epsilon_2,(n^{\kappa_2'}\epsilon_2)^2\right]\right\}, \\
\sup_{\mathcal{B}_0(\epsilon_1,\epsilon_2)^c\cap\mathcal{F}_n}P_{(\beta,\theta/\alpha^{2\nu},\alpha)}\left(\left|\widehat{\theta}_n/\theta_0-1\right|\leq\epsilon_1/2\right)&\leq\exp\left\{-c_5\min\left[n^{\kappa_1'}\epsilon_1,(n^{\kappa_1'}\epsilon_1)^2\right]\right\}, \\
\sup_{\mathcal{B}_0(\epsilon_1,\epsilon_2)^c\cap\mathcal{F}_n}P_{(\beta,\theta/\alpha^{2\nu},\alpha)}\left(|\widehat{\alpha}_n/\alpha_0-1|\leq\epsilon_2/2\right)&\leq\exp\left\{-c_5\min\left[n^{\kappa_2'}\epsilon_2,(n^{\kappa_2'}\epsilon_2)^2\right]\right\},
\end{aligned}
\qquad \text{(S.197)}
$$

where the sieve $\mathcal{F}_n\subseteq\{(\beta,\theta,\alpha)\in\mathbb{R}^p\times\mathbb{R}^+\times\mathbb{R}^+\}$, such that the prior satisfies $\Pi(\mathcal{F}_n^c)\leq n^{-(3p+6)}$ for all sufficiently large n.

Assumption (S.1) requires the existence of consistent estimators $\widehat{\theta}_n$ and $\widehat{\alpha}_n$. The exponentially small tail bounds in the inequalities in (S.197) imply the convergence rates of $O(n^{-\kappa_1'})$ and $O(n^{-\kappa_2'})$ for $\widehat{\theta}_n$ and $\widehat{\alpha}_n$, respectively. The inequalities in (S.197) will be used to construct exponentially consistent tests for θ and α, which are commonly used for showing the posterior consistency and posterior contraction rates in the Bayesian nonparametrics literature; see for example, Sections 6.4 and 8.2 in [Ghosal and van der Vaart, 2017].

Since Assumption (S.1) is a high level condition, we explain why such estimators $\widehat{\theta}_n$ and $\widehat{\alpha}_n$ exist for the isotropic Matérn covariance function with $d\geq 5$. To the best of our knowledge, [Anderes, 2010] is the only work that has systematically studied the fixed-domain asymptotics for the isotropic Matérn covariance function with domain dimension $d\geq 5$. [Anderes, 2010] has considered a special case of our model (1), in which (i) $Y(\cdot)$ is a GP with mean zero and no regression terms $\mathrm{m}(\cdot)^\top\beta$, and (ii) the sampling location set $\mathcal{S}_n$ consists of equispaced grids in a fixed domain. For this special case, [Anderes, 2010] proposed consistent moment estimators for both θ and α when $d \geq 5$ if we set their M matrix to be the identity matrix; see their Theorem 1, Theorem 2, and the discussion after the two theorems. The proofs of Theorems 1 and 2 in [Anderes, 2010] have derived tail bound inequalities similar to (S.197), where both κ_1' and κ_2' can be taken as 1/2, which satisfies our condition $0 < \kappa_1' \leq 1/2$ and $1/(2\nu+d) < \kappa_2' \leq 1/2$ since $1/(2\nu+d) < 1/5$ when $d \geq 5$.

The supremum in the inequalities of (S.197) can often be established using a union bound argument over the set $\mathcal{B}_0(\epsilon_1,\epsilon_2)^c \cap \mathcal{F}_n$. The parameter set $\mathcal{F}_n$ in Assumption (S.1) is typically a bounded set whose radius increases slowly with n, such that it is a sieve to the whole parameter space of $\{(\beta,\theta,\alpha) \in \mathbb{R}^p \times \mathbb{R}^+ \times \mathbb{R}^+\}$. The supremum inequalities and the sieve are also commonly used in Bayesian nonparametrics for showing posterior consistency and contraction rates; see for example, Theorem 6.17, Theorem 8.9 and their proofs in [Ghosal and van der Vaart, 2017]. We assume that the prior mass outside the sieve $\mathcal{F}_n$ is polynomially small, which is usually satisfied if β is assigned a normal prior and (θ,α) are assigned the priors described in Section 2.3. In Bayesian nonparametrics, it is often assumed that $\Pi(\mathcal{F}_n^c)$ is exponentially small in n, so our assumption is weaker in comparison.

Although Assumption (S.1) is currently verifiable only for the special case considered in [Anderes, 2010], we expect that the inequalities in (S.197) continue to hold for more general sampling designs and the model with regression terms in the case of $d \geq 5$, where the two constants κ_1' and κ_2' can be possibly smaller than 1/2 depending on the sampling designs. Detailed construction of such consistent estimators $\widehat{\theta}_n$ and $\widehat{\alpha}_n$ for $d \geq 5$ in the general universal kriging model (1) can be based on the recently proposed higher-order quadratic variation techniques in [Loh, 2015] and [Loh et al., 2021] and will be left for future investigation.

Before stating the main theorem for $d \geq 5$, we first prove two technical lemmas. Lemma S.32 can be used to show a theoretical lower bound of the denominator in the posterior distribution for $d \geq 5$. Lemma S.33 proves the posterior contraction for (θ,α) for $d \geq 5$. This will be used later for truncating the posterior to a shrinking neighborhood of (θ_0,α_0), which will be important for deriving the limiting posterior distribution for $d \geq 5$.

Lemma S.32. *Suppose that Assumptions (A.1) holds for $d \geq 5$ and $\nu \in \mathbb{R}^+$. Let*

$$\mathcal{A}_n^\dagger = \big\{(\beta,\theta,\alpha) \in \mathbb{R}^p \times \mathbb{R}^+ \times \mathbb{R}^+ : \ \|\beta-\beta_0\| \leq n^{-3}, \\ \theta_0 \leq \theta < \theta_0(1+n^{-2}),\ \alpha_0(1-n^{-2}) \leq \alpha \leq \alpha_0\big\}.$$

Then $\inf_{\mathcal{A}_n^\dagger}\left\{\mathcal{L}_n(\beta,\theta/\alpha^{2\nu},\alpha) - \mathcal{L}_n(\beta_0,\sigma_0^2,\alpha_0)\right\} \geq c_{5L}n^{-1}$ *with probability at least* $1-\exp(-16\log^2 n)$ *for all sufficiently large n, where $c_{5L}>0$ is a constant that depends on $\nu, d, T, \beta_0, \sigma_0^2, \alpha_0$ and the $\mathcal{W}_2^{\nu+d/2}(\mathcal{S})$ norms of $\mathrm{m}_1(\cdot),\ldots,\mathrm{m}_p(\cdot)$.*

Proof of Lemma S.32. By definition of the log-likelihood function $\mathcal{L}_n(\beta,\sigma^2,\alpha)$ in (3) and the true model $Y_n = M_n\beta_0 + X_n$, we have

$$\begin{aligned}
\mathcal{L}_n(\beta,\theta/\alpha^{2\nu},\alpha) - \mathcal{L}_n(\beta_0,\sigma_0^2,\alpha_0) &= -\frac{n}{2}\log\frac{\theta}{\theta_0} + \nu n \log\frac{\alpha}{\alpha_0} - \frac{1}{2}\log\frac{|R_\alpha|}{|R_{\alpha_0}|} \\
&- (Y_n - M_n\beta)^\top\left(\frac{\alpha^{2\nu}R_\alpha^{-1}}{2\theta} - \frac{\alpha_0^{2\nu}R_{\alpha_0}^{-1}}{2\theta_0}\right)(Y_n - M_n\beta) \\
&+ (\beta-\beta_0)^\top M_n^\top \frac{\alpha_0^{2\nu}R_{\alpha_0}^{-1}}{\theta_0}X_n - (\beta-\beta_0)^\top M_n^\top \frac{\alpha_0^{2\nu}R_{\alpha_0}^{-1}}{2\theta_0}M_n(\beta-\beta_0). \qquad \text{(S.198)}
\end{aligned}$$

On the right-hand side of (S.198), using Lemma S.14 and Lemma S.15, the first line can be lower bounded as follows in the set $\mathcal{A}_n^\dagger$ for all sufficiently large n:

$$-\frac{n}{2}\log\frac{\theta}{\theta_0} + \nu n \log\frac{\alpha}{\alpha_0} - \frac{1}{2}\log\frac{|R_\alpha|}{|R_{\alpha_0}|}$$

$$
\begin{aligned}
&\overset{(i)}{\geq} -\frac{n}{2}\log(1+n^{-2}) + \nu n \log\frac{\alpha}{\alpha_0} - \frac{1}{2}\sum_{k=1}^{n}\log\lambda_{k,n}(\alpha) - \nu p \log\frac{\alpha}{\alpha_0} \\
&\overset{(ii)}{\geq} -\frac{n}{2}\log(1+n^{-2}) + \nu(n-p)\log(1-n^{-2}) + \frac{2\nu+d}{2}\sum_{k=1}^{n}\log\left(1-n^{-2}\right) \\
&\overset{(iii)}{\geq} -\frac{1}{2n} - \frac{2\nu}{n} - \frac{2\nu+d}{n},
\end{aligned}
\tag{S.199}
$$

where (i) follows from Lemma S.14, (ii) follows from (S.60) in Lemma S.15 given that $\alpha \leq \alpha_0$ on $\mathcal{A}_n^\dagger$, and (iii) follows from $\log(1+x) \leq x$ and $\log(1-x) \geq -2x$ for $x \in (0, 1/2)$.

By Lemma S.3, in the set $\mathcal{A}_n^\dagger$, $\frac{\alpha^{2\nu}R_\alpha^{-1}}{2\theta} - \frac{\alpha_0^{2\nu}R_{\alpha_0}^{-1}}{2\theta_0}$ is negative definite. Therefore,

$$
-(Y_n - M_n\beta)^\top \left(\frac{\alpha^{2\nu}R_\alpha^{-1}}{2\theta} - \frac{\alpha_0^{2\nu}R_{\alpha_0}^{-1}}{2\theta_0}\right)(Y_n - M_n\beta) \geq 0. \tag{S.200}
$$

For the third line of (S.198), since $X_n \sim \mathcal{N}(0, \sigma_0^2 R_{\alpha_0})$, by Lemma S.19, $\Pr(\|\sigma_0^{-1}R_{\alpha_0}^{-1/2}X_n\|^2 \leq n + 8\log n + 16\log^2 n) \geq 1 - \exp(-16\log^2 n)$. By Assumption (A.1), using similar derivation to (S.25) based on Lemmas S.11, S.12 and S.13, we have that on the set $\mathcal{A}_n^\dagger$,

$$
\left\|\sigma_0^{-1}R_{\alpha_0}^{-1/2}M_n(\beta-\beta_0)\right\|^2 \leq \sum_{j=1}^{p}\|\mathrm{m}_j\|^2_{\mathcal{H}_{\sigma_0^2 K_{\alpha_0,\nu}}} \cdot \|\beta-\beta_0\|^2 \leq c_2(\sigma_0,\alpha_0)^2 \sum_{j=1}^{p}\|\mathrm{m}_j\|^2_{\mathcal{W}_2^{\nu+d/2}(\mathcal{S})} \cdot n^{-6}.
$$

Therefore, by Cauchy-Schwarz inequality, with probability at least $1 - \exp(-16\log^2 n)$,

$$
\begin{aligned}
&(\beta-\beta_0)^\top M_n^\top \frac{\alpha_0^{2\nu}R_{\alpha_0}^{-1}}{\theta_0}X_n - (\beta-\beta_0)^\top M_n^\top \frac{\alpha_0^{2\nu}R_{\alpha_0}^{-1}}{2\theta_0}M_n(\beta-\beta_0) \\
&\geq -(n + 8\log n + 16\log^2 n)^{1/2} c_2(\sigma_0,\alpha_0)\left(\sum_{j=1}^{p}\|\mathrm{m}_j\|^2_{\mathcal{W}_2^{\nu+d/2}(\mathcal{S})}\right)^{1/2} n^{-3} \\
&\quad - \frac{c_2(\sigma_0,\alpha_0)^2}{2}\sum_{j=1}^{p}\|\mathrm{m}_j\|^2_{\mathcal{W}_2^{\nu+d/2}(\mathcal{S})} \cdot n^{-6} \\
&\geq -c_1' n^{-2},
\end{aligned}
\tag{S.201}
$$

where $c_1' > 0$ is a constant dependent on the $\mathcal{W}_2^{\nu+d/2}(\mathcal{S})$ norms of $\mathrm{m}_1(\cdot), \ldots, \mathrm{m}_p(\cdot)$.

Finally, we combine (S.198), (S.199), (S.200) and (S.201) to obtain that on the set $\mathcal{A}_n^\dagger$, with probability at least $1 - \exp(-16\log^2 n)$,

$$
\mathcal{L}_n(\beta, \theta/\alpha^{2\nu}, \alpha) - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0) \geq -\frac{1}{2n} - \frac{2\nu}{n} - \frac{2\nu+d}{n} - c_1' n^{-2} \geq c_{5L} n^{-1},
$$

for some constant $c_{5L} > 0$. This completes the proof. □

Lemma S.33. *Suppose that Assumptions (A.1), (A.2) and (S.1) hold for $d \geq 5$ and $\nu \in \mathbb{R}^+$. Then the profile posterior distribution satisfies*

$$
\Pi\left(|\theta/\theta_0 - 1| \leq n^{-\kappa_1'}\log^2 n,\ |\alpha/\alpha_0 - 1| \leq n^{-\kappa_2'}\log^2 n \;\middle|\; Y_n\right) \to 1,
$$

as $n \to \infty$ almost surely $P_{(\beta_0, \sigma_0^2, \alpha_0)}$.

Proof of Lemma S.33. The proof proceeds in a similar way to that of the Schwartz's theorem for posterior consistency ([Schwartz, 1965]); see for example, Theorem 6.17 and its proof in

[Ghosal and van der Vaart, 2017]. Let $\epsilon_{1n}' = n^{-\kappa_1'} \log^2 n$ and $\epsilon_{2n}' = n^{-\kappa_2'} \log^2 n$. Define the testing function (indicator function):

$$T_n = \mathcal{I}\left(|\widehat{\theta}_n/\theta_0 - 1| \geq \epsilon_{1n}'/2, \text{ or } |\widehat{\alpha}_n/\alpha_0 - 1| \geq \epsilon_{2n}'/2 \right), \tag{S.202}$$

where $\widehat{\theta}_n$ and $\widehat{\alpha}_n$ are the consistent estimators of θ and α from Assumption (S.1). Recall that the log-likelihood function $\mathcal{L}_n(\beta, \theta/\alpha^{2\nu}, \alpha)$ is defined in (3) of the main text. We have the following decomposition:

$$\begin{aligned}
&\Pi\left(\mathcal{B}_0(\epsilon_{1n}', \epsilon_{2n}')^c \mid Y_n\right) \\
&= \frac{(T_n + 1 - T_n) \int_{\mathcal{B}_0(\epsilon_{1n}', \epsilon_{2n}')^c} \exp\left\{\mathcal{L}_n(\beta, \theta/\alpha^{2\nu}, \alpha)\right\} \pi(\beta, \theta, \alpha) \mathrm{d}\beta \mathrm{d}\theta \mathrm{d}\alpha}{\int_{\mathbb{R}^p \times \mathbb{R}^+ \times \mathbb{R}^+} \exp\left\{\mathcal{L}_n(\beta', \theta'/\alpha'^{2\nu}, \alpha')\right\} \pi(\beta', \theta', \alpha') \mathrm{d}\beta' \mathrm{d}\theta' \mathrm{d}\alpha'} \\
&\leq T_n + \frac{(1 - T_n) \int_{\mathcal{B}_0(\epsilon_{1n}', \epsilon_{2n}')^c \cap \mathcal{F}_n} \exp\left\{\mathcal{L}_n(\beta, \theta/\alpha^{2\nu}, \alpha) - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta, \theta, \alpha) \mathrm{d}\beta \mathrm{d}\theta \mathrm{d}\alpha}{\int_{\mathbb{R}^p \times \mathbb{R}^+ \times \mathbb{R}^+} \exp\left\{\mathcal{L}_n(\beta', \theta'/\alpha'^{2\nu}, \alpha') - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta', \theta', \alpha') \mathrm{d}\beta' \mathrm{d}\theta' \mathrm{d}\alpha'} \\
&\quad + \frac{\int_{\mathcal{F}_n^c} \exp\left\{\mathcal{L}_n(\beta, \theta/\alpha^{2\nu}, \alpha) - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta, \theta, \alpha) \mathrm{d}\beta \mathrm{d}\theta \mathrm{d}\alpha}{\int_{\mathbb{R}^p \times \mathbb{R}^+ \times \mathbb{R}^+} \exp\left\{\mathcal{L}_n(\beta', \theta'/\alpha'^{2\nu}, \alpha') - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta', \theta', \alpha') \mathrm{d}\beta' \mathrm{d}\theta' \mathrm{d}\alpha'}.
\end{aligned} \tag{S.203}$$

By Assumption (S.1), we have that as $n \to \infty$,

$$\begin{aligned}
\mathrm{E}_{(\beta_0, \sigma_0^2, \alpha_0)}(T_n) &\leq P_{(\beta_0, \sigma_0^2, \alpha_0)}\left(\left|\widehat{\theta}_n/\theta_0 - 1\right| \geq \epsilon_{1n}'/2\right) + P_{(\beta_0, \sigma_0^2, \alpha_0)}\left(\left|\widehat{\alpha}_n/\alpha_0 - 1\right| \geq \epsilon_{2n}'/2\right) \\
&\leq 2 \exp\left(-c_5 \log^2 n\right).
\end{aligned} \tag{S.204}$$

For the second term in (S.203), we use the same proof technique as the Schwartz's theorem for posterior consistency. By Assumption (S.1) and the Fubini's theorem, its numerator has expectation upper bounded by

$$\begin{aligned}
&\mathrm{E}_{(\beta_0, \sigma_0^2, \alpha_0)}(1 - T_n) \int_{\mathcal{B}_0(\epsilon_{1n}', \epsilon_{2n}')^c \cap \mathcal{F}_n} \exp\left\{\mathcal{L}_n(\beta, \theta/\alpha^{2\nu}, \alpha) - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta, \theta, \alpha) \mathrm{d}\beta \mathrm{d}\theta \mathrm{d}\alpha \\
&= \int_{\mathcal{B}_0(\epsilon_{1n}', \epsilon_{2n}')^c \cap \mathcal{F}_n} \mathrm{E}_{(\beta_0, \sigma_0^2, \alpha_0)}(1 - T_n) \exp\left\{\mathcal{L}_n(\beta, \theta/\alpha^{2\nu}, \alpha) - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta, \theta, \alpha) \mathrm{d}\beta \mathrm{d}\theta \mathrm{d}\alpha \\
&= \int_{\mathcal{B}_0(\epsilon_{1n}', \epsilon_{2n}')^c \cap \mathcal{F}_n} \mathrm{E}_{(\beta, \theta/\alpha^{2\nu}, \alpha)}(1 - T_n) \pi(\beta, \theta, \alpha) \mathrm{d}\beta \mathrm{d}\theta \mathrm{d}\alpha \\
&\leq \sup_{\mathcal{B}_0(\epsilon_{1n}', \epsilon_{2n}')^c \cap \mathcal{F}_n} \mathrm{E}_{(\beta, \theta/\alpha^{2\nu}, \alpha)} \left\{(1 - T_n) \int_{\mathcal{B}_0(\epsilon_{1n}', \epsilon_{2n}')^c \cap \mathcal{F}_n} \pi(\beta, \theta, \alpha) \mathrm{d}\beta \mathrm{d}\theta \mathrm{d}\alpha\right\} \\
&\leq \sup_{\mathcal{B}_0(\epsilon_{1n}', \epsilon_{2n}')^c \cap \mathcal{F}_n} \mathrm{E}_{(\beta, \theta/\alpha^{2\nu}, \alpha)}(1 - T_n) \\
&\leq \sup_{\mathcal{B}_0(\epsilon_{1n}', \epsilon_{2n}')^c \cap \mathcal{F}_n} P_{(\beta, \theta/\alpha^{2\nu}, \alpha)}\left(\left|\widehat{\theta}_n/\theta_0 - 1\right| \leq \epsilon_{1n}'/2\right) \\
&\quad + \sup_{\mathcal{B}_0(\epsilon_{1n}', \epsilon_{2n}')^c \cap \mathcal{F}_n} P_{(\beta, \theta/\alpha^{2\nu}, \alpha)}\left(\left|\widehat{\alpha}_n/\alpha_0 - 1\right| \leq \epsilon_{2n}'/2\right) \\
&\leq 2 \exp\left(-c_5 \log^2 n\right).
\end{aligned} \tag{S.205}$$

Since $\sum_{n=1}^{\infty} 2 \exp\left(-c_5 \log^2 n\right) < \infty$, by applying the Markov's inequality and the Borel-Cantelli lemma, the numerator of the second term in (S.203) is upper bounded by $2 \exp\left\{-(c_5/2) \log^2 n\right\}$ as $n \to \infty$ almost surely $P_{(\beta_0, \sigma_0^2, \alpha_0)}$. On the other hand, by Lemma S.32, for $d \geq 5$ and for all sufficiently large n, with probability at least $1 - \exp(-16 \log^2 n)$ the denominator of the second term in (S.203) can be lower bounded by

$$\int_{\mathbb{R}^p \times \mathbb{R}^+ \times \mathbb{R}^+} \exp\left\{\mathcal{L}_n(\beta', \theta'/\alpha'^{2\nu}, \alpha') - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta', \theta', \alpha') \mathrm{d}\beta' \mathrm{d}\theta' \mathrm{d}\alpha'$$

$$\geq \int_{\mathcal{A}_n^\dagger} \exp\left\{\mathcal{L}_n(\beta', \theta'/\alpha'^{2\nu}, \alpha') - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta', \theta', \alpha') \mathrm{d}\beta' \mathrm{d}\theta' \mathrm{d}\alpha'$$

$$\geq \exp(-c_{5L} n^{-1}) \cdot \Pi(\mathcal{A}_n^\dagger) \overset{(i)}{\geq} c_{6L} n^{-(3p+4)}, \tag{S.206}$$

for some constant $c_{6L} > 0$, where (i) follows because $\exp(-c_{5L} n^{-1}) > 1/2$ for large n, the prior density $\pi(\beta, \theta, \alpha) = \pi(\beta|\theta/\alpha^{2\nu})\pi(\theta|\alpha)\pi(\alpha)$ is lower bounded by constant in the set $\mathcal{A}_n^\dagger$ by Assumptions (A.1) and (A.2), and the set $\mathcal{A}_n^\dagger$ defined in Lemma S.32 has a volume of order $n^{-3p} \cdot n^{-2} \cdot n^{-2} = n^{-(3p+4)}$. Therefore, we combine (S.205) and (S.206) to obtain that almost surely $P_{(\beta_0, \sigma_0^2, \alpha_0)}$ as $n \to \infty$, the second term in (S.203) is upper bounded by

$$\frac{(1 - T_n) \int_{\mathcal{B}_0(\epsilon_{1n}', \epsilon_{2n}')^c \cap \mathcal{F}_n} \exp\left\{\mathcal{L}_n(\beta, \theta/\alpha^{2\nu}, \alpha) - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta, \theta, \alpha) \mathrm{d}\beta \mathrm{d}\theta \mathrm{d}\alpha}{\int_{\mathbb{R}^p \times \mathbb{R}^+ \times \mathbb{R}^+} \exp\left\{\mathcal{L}_n(\beta', \theta'/\alpha'^{2\nu}, \alpha') - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta', \theta', \alpha') \mathrm{d}\beta' \mathrm{d}\theta' \mathrm{d}\alpha'}$$

$$\leq 2 c_{6L}^{-1} n^{3p+4} \exp\left\{-(c_5/2) \log^2 n\right\} \to 0, \text{ as } n \to \infty. \tag{S.207}$$

For the third term in (S.203), similar to (S.205), by the Fubini's theorem and Assumption (S.1), we have that

$$\mathrm{E}_{(\beta_0, \sigma_0^2, \alpha_0)} \int_{\mathcal{F}_n^c} \exp\left\{\mathcal{L}_n(\beta, \theta/\alpha^{2\nu}, \alpha) - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta, \theta, \alpha) \mathrm{d}\beta \mathrm{d}\theta \mathrm{d}\alpha$$

$$= \int_{\mathcal{F}_n^c} \mathrm{E}_{(\beta_0, \sigma_0^2, \alpha_0)} \exp\left\{\mathcal{L}_n(\beta, \theta/\alpha^{2\nu}, \alpha) - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta, \theta, \alpha) \mathrm{d}\beta \mathrm{d}\theta \mathrm{d}\alpha$$

$$= \int_{\mathcal{F}_n^c} \mathrm{E}_{(\beta, \theta/\alpha^{2\nu}, \alpha)} \pi(\beta, \theta, \alpha) \mathrm{d}\beta \mathrm{d}\theta \mathrm{d}\alpha = \Pi(\mathcal{F}_n^c) \leq n^{-(3p+6)}, \tag{S.208}$$

which by the Markov's inequality and the Borel-Cantelli lemma, implies that the numerator of the second term in (S.203) is upper bounded by $n^{-(3p+6)}$ as $n \to \infty$ almost surely $P_{(\beta_0, \sigma_0^2, \alpha_0)}$. Therefore, (S.206) and (S.208) imply that the second term in (S.203) is upper bounded by

$$\frac{\int_{\mathcal{F}_n^c} \exp\left\{\mathcal{L}_n(\beta, \theta/\alpha^{2\nu}, \alpha) - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta, \theta, \alpha) \mathrm{d}\beta \mathrm{d}\theta \mathrm{d}\alpha}{\int_{\mathbb{R}^p \times \mathbb{R}^+ \times \mathbb{R}^+} \exp\left\{\mathcal{L}_n(\beta', \theta'/\alpha'^{2\nu}, \alpha') - \mathcal{L}_n(\beta_0, \sigma_0^2, \alpha_0)\right\} \pi(\beta', \theta', \alpha') \mathrm{d}\beta' \mathrm{d}\theta' \mathrm{d}\alpha'}$$

$$\leq c_{6L}^{-1} n^{3p+4} \cdot n^{-(3p+6)} = c_{6L}^{-1} n^{-2} \to 0, \text{ as } n \to \infty. \tag{S.209}$$

The conclusion follows by combining (S.203), (S.204), (S.207), and (S.209). □

We state and prove the following Theorem S.1 for the limiting posterior distribution of the covariance parameters (θ, α) for the case of $d \geq 5$. Theorem S.1 for $d \geq 5$ is a parallel to Theorem 2.3 in the main text for $d \in \{1, 2, 3\}$. We emphasize that in Theorem S.1, we only derive the asymptotic normality for the posterior of θ, since the limiting posterior distribution of the range parameter α will depend on the exact form of sampling design $\mathcal{S}_n$. Another difference in Theorem S.1 from Theorem 2.3 is that the profile posterior distribution for α will be a truncated distribution to the neighborhood $\left[(1 - n^{-\kappa_2'} \log^2 n)\alpha_0, (1 + n^{-\kappa_2'} \log^2 n)\alpha_0\right]$, given the posterior contraction result in Lemma S.33.

Theorem S.1. *Suppose that Assumptions (A.1), (A.2), (A.3) and (S.1) hold for $d \geq 5$ and $\nu \in \mathbb{R}^+$. The posterior distributions of θ and α are asymptotically independent, in the sense that the joint posterior distribution of (θ, α) satisfies*

$$\left\| \Pi(\mathrm{d}\theta, \mathrm{d}\alpha | Y_n) - \mathcal{N}\left(\mathrm{d}\theta \middle| \widetilde{\theta}_{\alpha_0}, 2\theta_0^2/n\right) \times \widetilde{\Pi}^\dagger(\mathrm{d}\alpha | Y_n) \right\|_{\mathrm{TV}} \to 0, \tag{S.210}$$

as $n \to \infty$ almost surely $P_{(\beta_0, \sigma_0^2, \alpha_0)}$, and $\widetilde{\Pi}^\dagger(\mathrm{d}\alpha | Y_n)$ is the truncated profile posterior distribution with the density

$$\widetilde{\pi}^\dagger(\alpha | Y_n) = \frac{\exp\left\{\widetilde{\mathcal{L}}_n(\alpha)\right\} \pi(\alpha | \theta_0)}{\int_{\max\left\{0, \left(1 - n^{-\kappa_2'} \log^2 n\right)\alpha_0\right\}}^{(1 + n^{-\kappa_2'} \log^2 n)\alpha_0} \exp\left\{\widetilde{\mathcal{L}}_n(\alpha')\right\} \pi(\alpha' | \theta_0) \mathrm{d}\alpha'}, \tag{S.211}$$

where the profile restricted log-likelihood $\widetilde{\mathcal{L}}_n(\alpha)$ is given in (8) *of the main text and $\pi(\alpha|\theta_0)$ is the conditional prior density of α given $\theta = \theta_0$.*

Proof of Theorem S.1. For short, let $\mathcal{B}_{0n} = \mathcal{B}_0(n^{-\kappa_1'} \log^2 n, n^{-\kappa_2'} \log^2 n)$ as defined in (S.196). For the joint posterior distribution $\Pi(\mathrm{d}\theta, \mathrm{d}\alpha|Y_n)$, we define the truncated posterior distribution $\Pi^\dagger(\mathrm{d}\theta, \mathrm{d}\alpha|Y_n) = \Pi(\mathrm{d}\theta, \mathrm{d}\alpha|Y_n) \cdot \mathcal{I}\{(\theta, \alpha) \in \mathcal{B}_{0n}\}/\Pi(\mathcal{B}_{0n}|Y_n)$ on the truncated support $\mathcal{B}_{0n}$. For all sufficiently large n, this support is a subset of $\mathbb{R}^+ \times \mathbb{R}^+$. By Lemma S.33, the posterior probability of the set $\mathcal{B}_{0n}$ converges to 1 as $n \to \infty$ almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$, which immediately implies that

$$\begin{aligned} \left\|\Pi^\dagger(\mathrm{d}\theta, \mathrm{d}\alpha|Y_n) - \Pi(\mathrm{d}\theta, \mathrm{d}\alpha|Y_n)\right\|_{\mathrm{TV}} &= \sup_{\mathcal{A}\in\mathbb{R}^+\times\mathbb{R}^+} \left|\Pi^\dagger(\mathcal{A}|Y_n) - \Pi(\mathcal{A}|Y_n)\right| \\ = \frac{\sup_{\mathcal{A}\in\mathbb{R}^+\times\mathbb{R}^+} \Pi(\mathcal{A}|Y_n)\left[1 - \Pi(\mathcal{B}_{0n}|Y_n)\right]}{\Pi(\mathcal{B}_{0n}|Y_n)} &= \frac{\Pi(\mathcal{B}_{0n}^c|Y_n)}{\Pi(\mathcal{B}_{0n}|Y_n)} \to 0, \end{aligned} \tag{S.212}$$

as $n \to \infty$ almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$. Therefore, to show (S.210), it suffices to show that as $n \to \infty$ almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$,

$$\left\|\Pi^\dagger(\mathrm{d}\theta, \mathrm{d}\alpha|Y_n) - \mathcal{N}\left(\mathrm{d}\theta\middle|\widetilde{\theta}_{\alpha_0}, 2\theta_0^2/n\right) \times \widetilde{\Pi}^\dagger(\mathrm{d}\alpha|Y_n)\right\|_{\mathrm{TV}} \to 0. \tag{S.213}$$

The rest of the proof proceeds in a similar way to the proof of Theorem 2.3, with a few key differences. Without loss of generality, we only consider those sufficiently large n such that $1 - n^{-\kappa_2'} \log^2 n > 0$. For short, let $\alpha_{1n} = \left(1 - n^{-\kappa_2'} \log^2 n\right)\alpha_0$ and $\alpha_{2n} = \left(1 + n^{-\kappa_2'} \log^2 n\right)\alpha_0$. First, (S.173) and (S.174) in the proof of Theorem 2.3 will be replaced by

$$\int_{\alpha_{1n}}^{\alpha_{2n}} \int_{\mathbb{R}} \left|\pi^\dagger(\theta, \alpha|Y_n) - \frac{\sqrt{n-p}}{2\sqrt{\pi}\theta_0} \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} \cdot \widetilde{\pi}^\dagger(\alpha|Y_n)\right| \mathrm{d}\theta \mathrm{d}\alpha \to 0, \tag{S.214}$$

$$\int_{\alpha_{1n}}^{\alpha_{2n}} \int_{\mathbb{R}} \left|\frac{\sqrt{n-p}}{2\sqrt{\pi}\theta_0} \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} - \frac{\sqrt{n}}{2\sqrt{\pi}\theta_0} \mathrm{e}^{-\frac{n(\theta-\widetilde{\theta}_{\alpha_0})^2}{4\theta_0^2}}\right| \cdot \widetilde{\pi}^\dagger(\alpha|Y_n) \mathrm{d}\theta \mathrm{d}\alpha \to 0, \tag{S.215}$$

as $n \to \infty$ almost surely $P_{(\beta_0,\sigma_0^2,\alpha_0)}$, where $\pi^\dagger(\theta, \alpha|Y_n)$ is the density of $\Pi^\dagger(\mathrm{d}\theta, \mathrm{d}\alpha|Y_n)$ and $\widetilde{\pi}^\dagger(\alpha|Y_n)$ is as defined in (S.211). The lower and upper bounds in the integrals of (S.214) and (S.215) are because the range parameter α in both $\pi^\dagger(\theta, \alpha|Y_n)$ and $\widetilde{\pi}^\dagger(\alpha|Y_n)$ is supported on $[\alpha_{1n}, \alpha_{2n}]$. Similar to (S.175), using Lemma S.30 and the definition of $\varrho_n(t; \alpha)$ in (S.143), the left-hand side of (S.214) is smaller than N'/D', where

$$\mathsf{N}' = 2 \int_{\alpha_{1n}}^{\alpha_{2n}} \int_{\mathbb{R}} \left|\varrho_n(\sqrt{n-p}(\theta - \widetilde{\theta}_\alpha); \alpha)\right| \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta_0|\alpha)\pi(\alpha) \mathrm{d}\theta \mathrm{d}\alpha, \tag{S.216}$$

$$\mathsf{D}' = \frac{2\theta_0\sqrt{\pi}}{\sqrt{n-p}} \int_{\alpha_{1n}}^{\alpha_{2n}} \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha)} \pi(\theta_0|\alpha)\pi(\alpha) \mathrm{d}\alpha, . \tag{S.217}$$

For any $\epsilon > 0$, let $\mathcal{E}_2'(\epsilon) = \left\{\sup_{\alpha\in[\alpha_{1n},\alpha_{2n}]} |\widetilde{\theta}_\alpha - \theta_0| < \epsilon\right\}$, $\mathcal{E}_3'(\epsilon) = \left\{\sup_{\alpha\in[\alpha_{1n},\alpha_{2n}]} |\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}| < \epsilon\right\}$, $\mathcal{E}_4(\epsilon) = \left\{|\widetilde{\theta}_{\alpha_0} - \theta_0| < \epsilon\right\}$. We can set $c = \kappa_2'$ in Lemma S.23, which satisfies $c = \kappa_2' > 1/(2\nu + d)$ given Assumption (S.1) and hence $[\alpha_{1n}, \alpha_{2n}] \subseteq \left[\left(1 - n^{-1/(2\nu+d)}\right)\alpha_0, \left(1 + n^{-1/(2\nu+d)}\right)\alpha_0\right]$. Thus we can apply (S.111) of Lemma S.23 to obtain that $\Pr\left\{\mathcal{E}_3'(10\theta_0 n^{-(2\nu+d)\kappa_2'} \log^4 n)\right\} \geq 1 - 8\exp(-4\log^2 n)$. Lemma S.10 implies that $\Pr\left\{\mathcal{E}_4(5\theta_0 n^{-1/2} \log n)\right\} \geq 1 - 3\exp(-4\log^2 n)$. Since $(2\nu + d)\kappa_2' > 1$ from Assumption (S.1), by the triangle inequality, for sufficiently large n,

$$\mathcal{E}_2'\left(6\theta_0 n^{-1/2} \log n\right) \supseteq \mathcal{E}_3'\left(10\theta_0 n^{-(2\nu+d)\kappa_2'} \log^4 n\right) \cap \mathcal{E}_4\left(5\theta_0 n^{-1/2} \log n\right),$$

and hence it follows that $\Pr\left\{\mathcal{E}_2'(6\theta_0 n^{-1/2} \log n)\right\} \geq 1 - 11\exp(-4\log^2 n)$.

Lemma S.29 still applies when $d \geq 5$ and $\mathcal{E}_1(6\theta_0 n^{-1/2}\log n, \alpha) \supseteq \mathcal{E}_2'(6\theta_0 n^{-1/2}\log n)$ for every $\alpha \in [\alpha_{1n}, \alpha_{2n}]$. Also, under Assumption (A.3), the inequality and convergence in (S.181) in the proof of Theorem 2.3 still holds. Since $[\alpha_{1n}, \alpha_{2n}] \subseteq [\underline{\alpha}_n, \overline{\alpha}_n]$, we apply Lemma S.29 with $\epsilon_{1n} = 6\theta_0 n^{-1/2}\log n$ and $s_n = \log n$ and obtain from (S.180) and (S.181) that

$$\int_{\alpha_{1n}}^{\alpha_{2n}} \int_{\mathbb{R}} \left| \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\theta,\alpha) - \mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)} \frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} - \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} \right| \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\theta\mathrm{d}\alpha$$
$$\leq \frac{\sup_{\alpha\in[\alpha_{1n},\alpha_{2n}]} B_n(\alpha)}{\sqrt{n-p}} \int_{\alpha_{1n}}^{\alpha_{2n}} \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \to 0, \text{ as } n\to\infty.$$

Therefore, similar to (S.182), we have that on the event $\mathcal{E}_2'(6\theta_0 n^{-1/2}\log n)$,

$$\begin{aligned}
\frac{\mathsf{N}'}{\mathsf{D}'} &\leq \frac{2\int_{\alpha_{1n}}^{\alpha_{2n}} \int_{\mathbb{R}} \left| \varrho_n(\sqrt{n-p}(\theta-\widetilde{\theta}_\alpha);\alpha) \right| \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\theta\mathrm{d}\alpha}{\frac{2\theta_0\sqrt{\pi}}{\sqrt{n-p}} \int_{\alpha_{1n}}^{\alpha_{2n}} \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha} \\
&\leq \frac{\frac{\sup_{\alpha\in[\alpha_{1n},\alpha_{2n}]} B_n(\alpha)}{\sqrt{n-p}} \int_{\alpha_{1n}}^{\alpha_{2n}} \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}{\frac{\theta_0\sqrt{\pi}}{\sqrt{n-p}} \int_{\alpha_{1n}}^{\alpha_{2n}} \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha} \\
&\leq \frac{\sup_{\alpha\in[\alpha_{1n},\alpha_{2n}]} B_n(\alpha)}{\theta_0\sqrt{\pi}} \to 0, \text{ as } n\to\infty,
\end{aligned}$$

where the last step follows from (S.181). This has proved (S.214).

For (S.215), similar to (S.192) and (S.193), on the event $\mathcal{E}_3'(10\theta_0 n^{-(2\nu+d)\kappa_2'}\log^4 n)$, we have that for all sufficiently large n,

$$\begin{aligned}
&\int_{\alpha_{1n}}^{\alpha_{2n}} \int_{\mathbb{R}} \left| \frac{\sqrt{n-p}}{2\sqrt{\pi}\theta_0} \mathrm{e}^{-\frac{(n-p)(\theta-\widetilde{\theta}_\alpha)^2}{4\theta_0^2}} - \frac{\sqrt{n}}{2\sqrt{\pi}\theta_0} \mathrm{e}^{-\frac{n(\theta-\widetilde{\theta}_{\alpha_0})^2}{4\theta_0^2}} \right| \cdot \widetilde{\pi}^\dagger(\alpha|Y_n)\mathrm{d}\theta\mathrm{d}\alpha \\
&\leq \int_{\alpha_{1n}}^{\alpha_{2n}} \frac{(n-p)^{1/2}}{2\sqrt{\pi}\theta_0} \left|\widetilde{\theta}_\alpha - \widetilde{\theta}_{\alpha_0}\right| \widetilde{\pi}^\dagger(\alpha|Y_n)\mathrm{d}\alpha \\
&\leq \frac{5}{\sqrt{\pi}} n^{-(2\nu+d)\kappa_2'+1/2} \log^4 n \int_{\alpha_{1n}}^{\alpha_{2n}} \widetilde{\pi}^\dagger(\alpha|Y_n)\mathrm{d}\alpha \\
&\overset{(i)}{\leq} \frac{5}{\sqrt{\pi}} n^{-1/2}\log^4 n \to 0, \text{ as } n\to\infty,
\end{aligned} \tag{S.218}$$

where (i) follows from the condition $\kappa_2' > 1/(2\nu+d)$. This has proved (S.215). Therefore, (S.214) and (S.215) together imply (S.213), and (S.213) together with (S.212) proves Theorem S.1. □

S4 Proof of Propositions 2.4 and 2.5

In this section, we provide the proof of Propositions 2.4 and 2.5 in the main text, which verify Assumptions (A.3) and (A.4) on the prior, respectively.

Proof of Proposition 2.4: (i) Since $\pi(\theta|\alpha) = \pi(\theta)$ and does not depend on α, we have that $\frac{\partial \log \pi(\theta|\alpha)}{\partial\theta} = \pi'(\theta)/\pi(\theta)$. Since $\pi(\theta) > 0$ and $\pi'(\theta) = \mathrm{d}\pi(\theta)/\mathrm{d}\theta$ is continuous on $\mathbb{R}^+$, (13) is satisfied for all sufficiently large n since

$$\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \sup_{\theta\in(\theta_0/2,2\theta_0)} \left| \frac{\partial \log \pi(\theta|\alpha)}{\partial\theta} \right| \leq \frac{\sup_{\theta\in(\theta_0/2,2\theta_0)} \pi'(\theta)}{\inf_{\theta\in(\theta_0/2,2\theta_0)} \pi(\theta)} < n^{C_{\pi,1}},$$

for arbitrary $C_{\pi,1} > 0$.

The prior density $\pi(\theta)$ has finite supremum and positive infimum on $(\theta_0/2, 2\theta_0)$. Hence (14) is satisfied for all sufficiently large n since

$$\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]}\sup_{\theta\in(\theta_0/2,2\theta_0)}\frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \le \frac{\sup_{\theta\in(\theta_0/2,2\theta_0)}\pi(\theta)}{\inf_{\theta\in(\theta_0/2,2\theta_0)}\pi(\theta)} < n^{C_{\pi,2}},$$

for arbitrary $C_{\pi,2} > 0$. Since $C_{\pi,1}$ and $C_{\pi,2}$ can be arbitrarily small, $C_{\pi,1}+C_{\pi,2} < 1/2$ is satisfied. Finally, (15) is satisfied for all sufficiently large n since $\pi(\theta_0) > 0$ and for all sufficiently large n,

$$\inf_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]}\log\pi(\theta_0|\alpha) = \log\pi(\theta_0) > -n^{C_{\pi,3}},$$

for arbitrarily small $C_{\pi,3} > 0$.

(ii) If $\pi(\alpha)$ is supported on a compact interval $[\alpha_1, \alpha_2]$, then all $\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]}$ can be replaced by $\sup_{\alpha\in[\alpha_1,\alpha_2]}$. Based on the conditions, for all sufficiently large n,

$$\sup_{\alpha\in[\alpha_1,\alpha_2]}\sup_{\theta\in(\theta_0/2,2\theta_0)}\left|\frac{\partial\log\pi(\theta|\alpha)}{\partial\theta}\right| < n^{C_{\pi,1}},$$

for arbitrary $C_{\pi,1} > 0$.

Since $\pi(\theta|\alpha) > 0$ for all $(\theta, \alpha) \in \mathbb{R}^+ \times \mathbb{R}^+$, for all sufficiently large n,

$$\sup_{\alpha\in[\alpha_1,\alpha_2]}\sup_{\theta\in(\theta_0/2,2\theta_0)}\left|\frac{\partial\log\pi(\theta|\alpha)}{\partial\theta}\right| < n^{C_{\pi,2}},$$

for arbitrary $C_{\pi,2} > 0$. Since $C_{\pi,1}$ and $C_{\pi,2}$ can be arbitrarily small, $C_{\pi,1} + C_{\pi,2} < 1/2$ is satisfied.

Since $\pi(\theta|\alpha) > 0$ is continuous in $\alpha \in \mathbb{R}^+$, for all sufficiently large n,

$$\inf_{\alpha\in[\alpha_1,\alpha_2]}\log\pi(\theta_0|\alpha) = \inf_{\alpha\in[\alpha_1,\alpha_2]}\log\pi(\theta_0|\alpha) > -n^{C_{\pi,3}},$$

for arbitrarily small $C_{\pi,3} > 0$.

(iii) If the prior of σ^2 is independent of α, then by the relation $\theta = \sigma^2\alpha^{2\nu}$, the prior of θ given α is $\pi(\theta|\alpha) = \pi_{\sigma^2}(\theta/\alpha^{2\nu})/\alpha^{2\nu}$, where we use $\pi_{\sigma^2}(\cdot)$ to denote the prior density of σ^2. Therefore, $\frac{\partial\log\pi(\theta|\alpha)}{\partial\theta} = \frac{\pi'_{\sigma^2}(\theta/\alpha^{2\nu})}{\alpha^{2\nu}\pi_{\sigma^2}(\theta/\alpha^{2\nu})}$. For the transformed beta family density, the derivative is

$$\pi'_{\sigma^2}(\sigma^2) = \frac{\Gamma(\gamma_1+\gamma_2)}{\Gamma(\gamma_1)\Gamma(\gamma_2)}\frac{\left(\frac{\sigma^2}{b}\right)^{\gamma_2/\gamma-2}\left[\gamma_2-\gamma-(\gamma_1+\gamma)\left(\frac{\sigma^2}{b}\right)^{1/\gamma}\right]}{(b\gamma)^2[1+(\sigma^2/b)^{1/\gamma}]^{\gamma_1+\gamma_2+1}}.$$

Therefore, for all sufficiently large n,

$$\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]}\sup_{\theta\in(\theta_0/2,2\theta_0)}\left|\frac{\partial\log\pi(\theta|\alpha)}{\partial\theta}\right| \le \sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]}\sup_{\theta\in(\theta_0/2,2\theta_0)}\frac{\left|\gamma_2-\gamma-(\gamma_1+\gamma)\left(\frac{\theta}{b\alpha^{2\nu}}\right)^{1/\gamma}\right|}{b\gamma\alpha^{2\nu}\left(\frac{\theta}{b\alpha^{2\nu}}\right)\left[1+\left(\frac{\theta}{b\alpha^{2\nu}}\right)^{1/\gamma}\right]}$$
$$\le \frac{2|\gamma_2-\gamma|}{\gamma\theta_0}+\frac{2(\gamma_1+\gamma)}{\gamma\theta_0} < n^{C_{\pi,1}},$$

for arbitrary $C_{\pi,1} > 0$.

$$\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]}\sup_{\theta\in(\theta_0/2,2\theta_0)}\frac{\pi(\theta|\alpha)}{\pi(\theta_0|\alpha)} \le \sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]}\sup_{\theta\in(\theta_0/2,2\theta_0)}\left(\frac{\theta}{\theta_0}\right)^{\gamma_2/\gamma-1}\left[\frac{b^{1/\gamma}\alpha^{2\nu/\gamma}+\theta_0^{1/\gamma}}{b^{1/\gamma}\alpha^{2\nu/\gamma}+\theta^{1/\gamma}}\right]^{\gamma_1+\gamma_2}$$

$$\leq \sup_{\theta\in(\theta_0/2,2\theta_0)} \max\left\{\left(\frac{\theta}{\theta_0}\right)^{\gamma_2/\gamma-1}, \left(\frac{\theta}{\theta_0}\right)^{-\gamma_1/\gamma-1}\right\} < n^{C_{\pi,2}},$$

for arbitrary $C_{\pi,2} > 0$. Since $C_{\pi,1}$ and $C_{\pi,2}$ can be arbitrarily small, $C_{\pi,1} + C_{\pi,2} < 1/2$ is satisfied.

$$\begin{aligned}
\inf_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \log\pi(\theta_0|\alpha) \geq& \inf_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \left\{ -2\nu\log\alpha - \log\frac{\Gamma(\gamma_1+\gamma_2)}{b\gamma\Gamma(\gamma_1)\Gamma(\gamma_2)} + \left(\frac{\gamma_2}{\gamma}-1\right)\log\frac{\theta_0}{b\alpha^{2\nu}} \right.\\
&\left. -(\gamma_1+\gamma_2)\log\left[1+\left(\frac{\theta_0}{b\alpha^{2\nu}}\right)^{1/\gamma}\right]\right\}\\
\geq& -2\nu\overline{\kappa}\log n - \log\frac{\Gamma(\gamma_1+\gamma_2)}{b\gamma\Gamma(\gamma_1)\Gamma(\gamma_2)} + \left(\frac{\gamma_2}{\gamma}-1\right)\log\frac{\theta_0}{b}\\
&-\left|\frac{\gamma_2}{\gamma}-1\right|\cdot 2\nu(\overline{\kappa}+\underline{\kappa})\log n - (\gamma_1+\gamma_2)\log\left[1+\left(\frac{\theta_0 n^{2\nu\underline{\kappa}}}{b}\right)^{1/\gamma}\right]\\
\succeq& -\log n \succ -n^{C_{\pi,3}},
\end{aligned}$$

for arbitrarily small $C_{\pi,3} > 0$. □

Proof of Proposition 2.5. We will verify only (20) with $0 < \overline{c_\pi} < (\nu+d/2)\overline{\kappa}$ for each conditions in the list. The verification of (21) with $0 < \underline{c_\pi} < (\nu+d/2)\underline{\kappa}$ is similar and omitted.

For $p(\alpha)$ that satisfies (i), we use the change of variable $u = \alpha^{1/\delta_1}$ to obtain that

$$\begin{aligned}
&\int_{\overline{\alpha}_n}^{\infty} \alpha^{n(\nu+d/2)} p(\alpha)\mathrm{d}\alpha \leq \int_{\overline{\alpha}_n}^{\infty} \alpha^{n(\nu+d/2)} \exp\left(-\alpha^{\delta_1}\right)\mathrm{d}\alpha\\
&\leq \frac{1}{\delta_1}\int_{\overline{\alpha}_n^{\delta_1}}^{\infty} u^{\{n(\nu+d/2)+1\}/\delta_1-1}\mathrm{e}^{-u}\mathrm{d}u < \frac{1}{\delta_1}\int_0^{\infty} u^{\{n(\nu+d/2)+1\}/\delta_1-1}\mathrm{e}^{-u}\mathrm{d}u\\
&= \frac{1}{\delta_1}\Gamma\left(\delta_1^{-1}\{n(\nu+d/2)+1\}\right),
\end{aligned} \tag{S.219}$$

where $\Gamma(x) = \int_0^\infty u^{x-1}\mathrm{e}^{-u}\mathrm{d}u$ is the gamma function. Using the Stirling's approximation for gamma functions ($\Gamma(x) < 2\sqrt{2\pi x}(x/\mathrm{e})^x$ for all large $x > 0$), we have that for sufficiently large n,

$$\begin{aligned}
&\Gamma\left(\delta_1^{-1}\{n(\nu+d/2)+1\}\right)\\
&< 2\sqrt{2\pi\delta_1^{-1}\{n(\nu+d/2)+1\}}\left(\mathrm{e}^{-1}\delta_1^{-1}\{n(\nu+d/2)+1\}\right)^{\delta_1^{-1}\{n(\nu+d/2)+1\}}.
\end{aligned} \tag{S.220}$$

From (S.219) and (S.220), we can see that (20) will be satisfied if for all sufficiently large n,

$$\begin{aligned}
&2\delta_1^{-1}\sqrt{2\pi\delta_1^{-1}\{n(\nu+d/2)+1\}}\left(\mathrm{e}^{-1}\delta_1^{-1}\{n(\nu+d/2)+1\}\right)^{\delta_1^{-1}\{n(\nu+d/2)+1\}}\\
&< \exp(\overline{c_\pi} n\log n).
\end{aligned}$$

A comparison of the orders in n on both sides immediately shows that this relation holds for all sufficiently large n, as long as $\delta_1^{-1}(\nu+d/2) < \overline{c_\pi}$. Since $\overline{c_\pi}$ can be chosen as any constant between 0 and $(\nu+d/2)\overline{\kappa}$, it suffices to have $\delta_1^{-1}(\nu+d/2) < (\nu+d/2)\overline{\kappa}$, or equivalently $\delta_1 > 1/\overline{\kappa}$.

For $p(\alpha)$ that satisfies (ii), we use the change of variable $u = n^{\delta_2}\alpha$ and the Stirling's approximation to obtain that

$$\int_{\overline{\alpha}_n}^{\infty} \alpha^{n(\nu+d/2)} p(\alpha)\mathrm{d}\alpha \leq \int_{\overline{\alpha}_n}^{\infty} \alpha^{n(\nu+d/2)} n^{\delta_3}\exp\left(-n^{\delta_2}\alpha\right)\mathrm{d}\alpha$$

$$
\begin{aligned}
&\le n^{\delta_3-\delta_2\{n(\nu+d/2)+1\}}\int_{n^{\delta_2}\overline{\alpha}_n}^{\infty} u^{n(\nu+d/2)}\mathrm{e}^{-u}\mathrm{d}u < n^{\delta_3-\delta_2\{n(\nu+d/2)+1\}}\int_{0}^{\infty} u^{n(\nu+d/2)}\mathrm{e}^{-u}\mathrm{d}u\\
&= n^{\delta_3-\delta_2\{n(\nu+d/2)+1\}}\Gamma\left(n(\nu+d/2)+1\right)\\
&\le n^{\delta_3-\delta_2\{n(\nu+d/2)+1\}}\cdot 2\sqrt{2\pi\{n(\nu+d/2)+1\}}\times\left(\mathrm{e}^{-1}\{n(\nu+d/2)+1\}\right)^{n(\nu+d/2)+1}.
\end{aligned}
$$

From the last display, (20) will be satisfied if for all sufficiently large n,

$$
\begin{aligned}
&n^{\delta_3-\delta_2\{n(\nu+d/2)+1\}}\cdot 2\sqrt{2\pi\{n(\nu+d/2)+1\}}\\
&\quad\times\left(\mathrm{e}^{-1}\{n(\nu+d/2)+1\}\right)^{n(\nu+d/2)+1} < \exp(\overline{c_\pi} n\log n).
\end{aligned}
$$

A comparison of the orders in n on both sides immediately shows that this relation holds for all sufficiently large n, as long as $-\delta_2(\nu+d/2)+(\nu+d/2)<\overline{c_\pi}$. Since $\overline{c_\pi}$ can be chosen as any constant between 0 and $(\nu+d/2)\overline{\kappa}$, it suffices to have $(1-\delta_2)(\nu+d/2)<(\nu+d/2)\overline{\kappa}$, or equivalently $\delta_2>1-\overline{\kappa}$. □

S5 Proof of Theorem 2.6 and Corollary 2.7

In this section, we provide the proof of Theorem 2.6 and Corollary 2.7 for the limiting distribution for 1-dimensional Ornstein-Uhlenbeck process. Before that, we first elaborate on the possible choices of prior $\pi(\alpha)$ and its hyperparameters that satisfy the relaxed Assumption (A.4') on the tails of $\pi(\alpha)$.

- If we take $\pi(\alpha)$ to be the gamma density $\pi(\alpha)=\frac{b^a}{\Gamma(a)}\alpha^{a-1}\mathrm{e}^{-b\alpha}$, then for all sufficiently large n,

$$
\begin{aligned}
\sqrt{n}\int_0^{\underline{\alpha}_n}\sqrt{\alpha}\pi(\alpha)\mathrm{d}\alpha &= \sqrt{n}\int_0^{n^{-\underline{\kappa}}}\frac{b^a}{\Gamma(a)}\alpha^{a+1/2-1}\mathrm{e}^{-b\alpha}\mathrm{d}\alpha\\
&\le \frac{\sqrt{n}b^a}{\Gamma(a)}\int_0^{n^{-\underline{\kappa}}}\alpha^{a+1/2-1}\mathrm{d}\alpha = \frac{b^a}{(a+1/2)\Gamma(a)}n^{-\underline{\kappa}(a+1/2)+1/2},\\
\text{and}\quad \sqrt{n}\int_{\overline{\alpha}_n}^{\infty}\sqrt{\alpha}\pi(\alpha)\mathrm{d}\alpha &= \sqrt{n}\int_{n^{\overline{\kappa}}}^{\infty}\frac{b^a}{\Gamma(a)}\alpha^{a+1/2-1}\mathrm{e}^{-b\alpha}\mathrm{d}\alpha\\
&\le \frac{\sqrt{n}b^a}{\Gamma(a)}\int_{n^{\overline{\kappa}}}^{\infty}\mathrm{e}^{-b\alpha/2}\mathrm{d}\alpha = \frac{2b^{a-1}}{\Gamma(a)}\sqrt{n}\exp(-bn^{\overline{\kappa}}/2).
\end{aligned}
$$

  To satisfy (24) in Assumption (A.4'), we need the condition $-\underline{\kappa}(a+1/2)+1/2<0$, or $a>(\underline{\kappa}^{-1}-1)/2$. Therefore, Assumption (A.4') holds for the gamma prior density $\pi(\alpha)$ with hyperparameters $a>(\underline{\kappa}^{-1}-1)/2$ and all $b>0$.

- If we take $\pi(\alpha)$ to be the inverse gamma density $\pi(\alpha)=\frac{b^a}{\Gamma(a)}\alpha^{-(a+1)}\mathrm{e}^{-b/\alpha}$, then similar to the derivation above, we obtain that Assumption (A.4') holds for the inverse gamma prior density $\pi(\alpha)$ with hyperparameters $a>(\overline{\kappa}^{-1}-1)/2$ and all $b>0$.

- If we take $\pi(\alpha)$ to be the inverse Gaussian density $\pi(\alpha)=\sqrt{\frac{b}{2\pi\alpha^3}}\exp\left\{-\frac{b(\alpha-a)^2}{2a^2\alpha}\right\}$ for $a>0, b>0$, then for all sufficiently large n,

$$
\begin{aligned}
\sqrt{n}\int_0^{\underline{\alpha}_n}\sqrt{\alpha}\pi(\alpha)\mathrm{d}\alpha &= \sqrt{n}\int_0^{n^{-\underline{\kappa}}}\sqrt{\frac{b}{2\pi\alpha^3}}\exp\left\{-\frac{b(\alpha-a)^2}{2a^2\alpha}\right\}\mathrm{d}\alpha\\
&\le \sqrt{n}\sqrt{\frac{b}{2\pi}}\exp(b/a)\int_{n^{\underline{\kappa}}}^{\infty}t^{-1/2}\exp\left\{-bt/(2a^2)\right\}\mathrm{d}t
\end{aligned}
$$

$$
\begin{aligned}
&\leq \sqrt{n}\sqrt{\frac{b}{2\pi}}\exp(b/a)\int_{n^{\underline{\kappa}}}^{\infty}\exp\left\{-bt/(4a^2)\right\}\mathrm{d}t \\
&= \frac{4a^2}{b}\sqrt{\frac{b}{2\pi}}\exp(b/a)\sqrt{n}\exp\left\{-bn^{\underline{\kappa}}/(4a^2)\right\} \to 0,
\end{aligned}
$$

and

$$
\begin{aligned}
\sqrt{n}\int_{\overline{\alpha}_n}^{\infty}\sqrt{\alpha}\pi(\alpha)\mathrm{d}\alpha &= \sqrt{n}\int_{n^{\overline{\kappa}}}^{\infty}\sqrt{\frac{b}{2\pi\alpha^3}}\exp\left\{-\frac{b(\alpha-a)^2}{2a^2\alpha}\right\}\mathrm{d}\alpha \\
&\leq \sqrt{n}\sqrt{\frac{b}{2\pi}}\exp(b/a)\int_{n^{\overline{\kappa}}}^{\infty}\alpha^{-3/2}\exp\left\{-b\alpha/(2a^2)\right\}\mathrm{d}\alpha \\
&\leq \sqrt{n}\sqrt{\frac{b}{2\pi}}\exp(b/a)\int_{n^{\overline{\kappa}}}^{\infty}\exp\left\{-b\alpha/(2a^2)\right\}\mathrm{d}\alpha \\
&= \frac{2a^2}{b}\sqrt{\frac{b}{2\pi}}\exp(b/a)\sqrt{n}\exp\left\{-bn^{\overline{\kappa}}/(2a^2)\right\} \to 0.
\end{aligned}
$$

Therefore, the inverse Gaussian density $\pi(\alpha)$ satisfies (24) in Assumption (A.4') for all hyperparameter values of $a > 0$ and $b > 0$.

- If we take $\pi(\alpha)$ to be the generalized beta density of the second kind:

$$
\pi(\alpha) = \frac{\Gamma(\gamma_1+\gamma_2)}{\Gamma(\gamma_1)\Gamma(\gamma_2)}\frac{(\alpha/b)^{\gamma_2/\gamma-1}}{b\gamma[1+(\alpha/b)^{1/\gamma}]^{\gamma_1+\gamma_2}}
$$

with parameters $b > 0, \gamma > 0, \gamma_1 > 0, \gamma_2 > 0$, then for all sufficiently large n,

$$
\begin{aligned}
\sqrt{n}\int_0^{\underline{\alpha}_n}\sqrt{\alpha}\pi(\alpha)\mathrm{d}\alpha &= \sqrt{n}\int_0^{n^{-\underline{\kappa}}}\frac{\Gamma(\gamma_1+\gamma_2)}{\Gamma(\gamma_1)\Gamma(\gamma_2)}\frac{(\alpha/b)^{\gamma_2/\gamma-1}}{b\gamma[1+(\alpha/b)^{1/\gamma}]^{\gamma_1+\gamma_2}}\mathrm{d}\alpha \\
&\leq \frac{\Gamma(\gamma_1+\gamma_2)}{b^{\gamma_2/\gamma}\gamma\Gamma(\gamma_1)\Gamma(\gamma_2)}\sqrt{n}\int_0^{n^{-\underline{\kappa}}}\alpha^{\gamma_2/\gamma-1}\mathrm{d}\alpha \\
&= \frac{\Gamma(\gamma_1+\gamma_2)}{b^{\gamma_2/\gamma}\gamma_2\Gamma(\gamma_1)\Gamma(\gamma_2)}n^{-\underline{\kappa}\gamma_2/\gamma+1/2},
\end{aligned}
$$

and

$$
\begin{aligned}
\sqrt{n}\int_{\overline{\alpha}_n}^{\infty}\sqrt{\alpha}\pi(\alpha)\mathrm{d}\alpha &= \sqrt{n}\int_{n^{\overline{\kappa}}}^{\infty}\frac{\Gamma(\gamma_1+\gamma_2)}{\Gamma(\gamma_1)\Gamma(\gamma_2)}\frac{(\alpha/b)^{\gamma_2/\gamma-1}}{b\gamma[1+(\alpha/b)^{1/\gamma}]^{\gamma_1+\gamma_2}}\mathrm{d}\alpha \\
&\leq \frac{\Gamma(\gamma_1+\gamma_2)}{b\gamma\Gamma(\gamma_1)\Gamma(\gamma_2)}\sqrt{n}\int_{n^{\overline{\kappa}}}^{\infty}(\alpha/b)^{-(\gamma_1/\gamma+1)}\mathrm{d}\alpha \\
&= \frac{\Gamma(\gamma_1+\gamma_2)}{b^{\gamma_1/\gamma}\gamma_1\Gamma(\gamma_1)\Gamma(\gamma_2)}n^{-\overline{\kappa}\gamma_1/\gamma+1/2}.
\end{aligned}
$$

To satisfy (24) in Assumption (A.4'), we need the conditions $-\underline{\kappa}\gamma_2/\gamma + 1/2 < 0$ and $-\overline{\kappa}\gamma_1/\gamma + 1/2 < 0$, or equivalently, $\gamma_2/\gamma > 1/(2\underline{\kappa})$ and $\gamma_1/\gamma > 1/(2\overline{\kappa})$. Therefore, if $\pi(\alpha)$ is the generalized beta density of the second kind, then it satisfies Assumption (A.4') if its hyperparameters $(b, \gamma, \gamma_1, \gamma_2)$ satisfy $\gamma_1/\gamma > 1/(2\overline{\kappa})$ and $\gamma_2/\gamma > 1/(2\underline{\kappa})$.

S5.1 Proof of Theorem 2.6

Recall that for Case (i) in Section 2.4 of the main text, we observe the 1-dimensional Ornstein-Uhlenbeck process without regression term $Y(\cdot) = X(\cdot) \sim \mathrm{GP}(0, \sigma_0^2 K_{\alpha_0,\nu})$ on the grid $s_i = i/n$, for $i = 1, \ldots, n$. Since $Y(s_i) = X(s_i)$ for all $i = 1, \ldots, n$ in this case, we have

$$
A_1 = \sum_{i=2}^{n-1} X(s_i)^2, \quad A_2 = \sum_{i=1}^{n-1} X(s_i)X(s_{i+1}), \quad A_3 = \sum_{i=1}^{n} X(s_i)^2.
$$

In the following, for any random variable Z_n, we write $Z_n \asymp 1$ to denote that Z_n is lower bounded away from zero and upper bounded from infinity as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability. The $O_p(\cdot)$ notation refers to the true probability measure $P_{(\sigma_0^2,\alpha_0)}$.

We introduce two technical Lemmas S.34 and S.35.

Lemma S.34. *Under the model setup of Theorem 2.6, we have the following results:*

(i) $A_1 + A_3 - 2A_2 > 0$ *a.s.* $P_{(\sigma_0^2,\alpha_0)}$;

(ii) $A_1 + A_3 - 2A_2 \asymp 1$ *as* $n \to \infty$ *a.s.* $P_{(\sigma_0^2,\alpha_0)}$;

(iii) $|A_1 - A_2| = O_p(1)$ *as* $n \to \infty$ *in* $P_{(\sigma_0^2,\alpha_0)}$*-probability, and* $|A_1 - A_2| \preceq \log^2 n$ *as* $n \to \infty$ *a.s.* $P_{(\sigma_0^2,\alpha_0)}$;

(iv) $A_1/n \asymp 1$ *and* $A_3/n \asymp 1$ *as* $n \to \infty$ *in* $P_{(\sigma_0^2,\alpha_0)}$*-probability;*

(v) $|u_*| = n|A_1 - A_2|/A_1 = O_p(1)$ *as* $n \to \infty$ *in* $P_{(\sigma_0^2,\alpha_0)}$*-probability;*

(vi) $v_* = n(A_1 - 2A_2 + A_3)/A_1 \asymp 1$ *as* $n \to \infty$ *in* $P_{(\sigma_0^2,\alpha_0)}$*-probability;*

(vii) $\frac{A_1A_3 - A_2^2}{A_1(A_1 - 2A_2 + A_3)} = 1 + O_p(n^{-1})$ *as* $n \to \infty$ *in* $P_{(\sigma_0^2,\alpha_0)}$*-probability;*

(viii) Uniformly over all $\alpha \in [0, n^{1/6}]$,

$$\frac{\left|\left(A_1 e^{-2\alpha/n} - 2A_2 e^{-\alpha/n} + A_3\right) - \left[A_1\left(\frac{\alpha}{n} - \frac{A_1 - A_2}{A_1}\right)^2 + \frac{A_1A_3 - A_2^2}{A_1}\right]\right|}{A_1\left(\frac{\alpha}{n} - \frac{A_1 - A_2}{A_1}\right)^2 + \frac{A_1A_3 - A_2^2}{A_1}} = O_p\left(n^{-3/2}\right),$$

as $n \to \infty$ *in* $P_{(\sigma_0^2,\alpha_0)}$*-probability;;*

(ix) Uniformly over all $\alpha \in [0, n^{1/6}]$,

$$\sqrt{1 - e^{-2\alpha/n}} = \sqrt{\frac{2\alpha}{n}}\left[1 + O(n^{-5/12})\right],$$

as $n \to \infty$.

Proof of Lemma S.34. (i) By definition, $A_1 + A_3 - 2A_2 = \sum_{i=1}^{n-1}\left[X(s_{i+1}) - X(s_i)\right]^2 > 0$ almost surely $P_{(\sigma_0^2,\alpha_0)}$.

(ii) Let $W_{i,n} = [X(s_i) - e^{-\alpha_0/n}X(s_{i-1})]/\sqrt{\sigma_0^2(1 - e^{-2\alpha_0/n})}$ for $i = 2, \ldots, n$. Then by the Markov property of Ornstein-Uhlenbeck process, $W_{i,n}$'s are i.i.d. $\mathcal{N}(0,1)$ random variables, such that $W_{i,n}$ is independent of $X(s_{i-1})$, and $X(s_i) = e^{-\alpha_0/n}X(s_{i-1}) + \sqrt{\sigma_0^2(1 - e^{-2\alpha_0/n})}W_{i,n}$, for $i = 2, \ldots, n$. We can derive that

$$\begin{aligned}
A_1 + A_3 - 2A_2 &= \sum_{i=1}^{n-1}\left[X(s_{i+1}) - e^{-\alpha_0/n}X(s_i) - (1 - e^{-\alpha_0/n})X(s_i)\right]^2 \\
&= \sum_{i=1}^{n-1}\left[X(s_{i+1}) - e^{-\alpha_0/n}X(s_i)\right]^2 + \sum_{i=1}^{n-1}(1 - e^{-\alpha_0/n})^2X(s_i)^2 \\
&\quad + 2\sum_{i=1}^{n-1}(1 - e^{-\alpha_0/n})X(s_i)\left[X(s_{i+1}) - e^{-\alpha_0/n}X(s_i)\right]. \qquad \text{(S.221)}
\end{aligned}$$

The first term in (S.221) is

$$\sum_{i=1}^{n-1}\left[X(s_{i+1})-\mathrm{e}^{-\alpha_0/n}X(s_i)\right]^2=\sum_{i=2}^{n}\sigma_0^2(1-\mathrm{e}^{-2\alpha_0/n})W_{i,n}^2=\frac{\sigma_0^2\alpha_0[1+o(1)]}{n}\sum_{i=2}^{n}W_{i,n}^2,$$

using a Taylor expansion of $1-\mathrm{e}^{-x}$ around $x=0$. Since $W_{i,n}$'s are i.i.d. $\mathcal{N}(0,1)$ random variables, we have that $n^{-1}\sum_{i=2}^{n}W_{i,n}^2\to 1$ as $n\to\infty$ almost surely $P_{(\sigma_0^2,\alpha_0)}$.

The second term in (S.221) is

$$\sum_{i=1}^{n-1}(1-\mathrm{e}^{-\alpha_0/n})^2X(s_i)^2\le\frac{\alpha_0^2}{n}\sup_{s\in[0,1]}X(s)^2.$$

For the Ornstein-Uhlenbeck process, $\sup_{s\in[0,1]}X(s)^2<\infty$ almost surely $P_{(\sigma_0^2,\alpha_0)}$. Therefore, $\sum_{i=1}^{n-1}(1-\mathrm{e}^{-\alpha_0/n})^2X(s_i)^2=O(1/n)$ almost surely $P_{(\sigma_0^2,\alpha_0)}$.

The third term in (S.221) can be upper bounded by

$$\begin{aligned}
&2\sum_{i=1}^{n-1}(1-\mathrm{e}^{-\alpha_0/n})X(s_i)\left[X(s_{i+1})-\mathrm{e}^{-\alpha_0/n}X(s_i)\right]\\
&=2\sum_{i=1}^{n-1}\sqrt{\sigma_0^2(1-\mathrm{e}^{-2\alpha_0/n})}(1-\mathrm{e}^{-\alpha_0/n})X(s_i)W_{i+1,n}\\
&=\frac{2\sqrt{2}\sigma_0\alpha_0^{3/2}[1+o(1)]}{n^{3/2}}\sum_{i=1}^{n-1}X(s_i)W_{i+1,n}\\
&\le\frac{2\sqrt{2}\sigma_0\alpha_0^{3/2}[1+o(1)]}{n}\sqrt{\sum_{i=1}^{n-1}X(s_i)^2W_{i+1,n}^2}\\
&\le\frac{2\sqrt{2}\sigma_0\alpha_0^{3/2}[1+o(1)]}{\sqrt{n}}\sqrt{\sup_{s\in[0,1]}X(s)^2\frac{1}{n}\sum_{i=2}^{n}W_{i,n}^2},
\end{aligned}\tag{S.222}$$

which shows that the third term is $O(n^{-1/2})$ almost surely $P_{(\sigma_0^2,\alpha_0)}$.

In combination with (S.221), we have shown that $A_1+A_3-A_2\to\sigma_0^2\alpha_0=\theta_0>0$ as $n\to\infty$ almost surely $P_{(\sigma_0^2,\alpha_0)}$, which means that $A_1+A_3-A_2\asymp 1$.

(iii)

$$|A_1-A_2|\le\frac{1}{2}|A_1+A_3-2A_2|+\frac{1}{2}[X(s_1)^2+X(s_n)^2].$$

Since $X(s_1)\sim\mathcal{N}(0,\sigma_0^2)$, $X(s_n)\sim\mathcal{N}(0,\sigma_0^2)$, we have $X(s_1)=O_p(1)$ and $X(s_n)=O_p(1)$. Furthermore, by the Borel-Cantelli lemma, $X(s_1)\preceq\log n$ and $X(s_n)\preceq\log n$ as $n\to\infty$ almost surely $P_{(\sigma_0^2,\alpha_0)}$. Then the conclusion follows by combining these relations with Part (ii).

(iv) First $A_3/n\le\sup_{s\in[0,1]}X(s)^2<\infty$ almost surely $P_{(\sigma_0^2,\alpha_0)}$. The expectation of A_3/n is $\mathrm{E}_{(\sigma_0^2,\alpha_0)}(A_3/n)=\sum_{i=1}^{n}\mathrm{E}_{(\sigma_0^2,\alpha_0)}(X(s_i)^2)/n=\sigma_0^2$. To calculate the variance of A_3, we let $T_{ij}=\left[X(s_i)-\mathrm{e}^{-\alpha_0|i-j|/n}X(s_j)\right]/\sqrt{\sigma_0^2\left(1-\mathrm{e}^{-2\alpha_0|i-j|/n}\right)}$ for any $i\neq j$ and $i,j=1,\ldots,n$. By the Markov property of the OU process, $T_{ij}\sim\mathcal{N}(0,1)$. Therefore, given that each $X(s_i)\sim\mathcal{N}(0,\sigma_0^2)$, $\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left[X(s_i)^2\right]=\sigma_0^2$, $\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left[X(s_i)^3\right]=0$, $\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left[X(s_i)^4\right]=3\sigma_0^4$, we have that

$$\mathrm{Var}_{(\sigma_0^2,\alpha_0)}(A_3/n)$$

$$= \frac{1}{n^2}\left\{\sum_{i=1}^{n}\mathrm{Var}_{(\sigma_0^2,\alpha_0)}(X(s_i)^2) + \sum_{i\neq j}\mathrm{Cov}_{(\sigma_0^2,\alpha_0)}\left(X(s_i)^2, X(s_j)^2\right)\right\}$$

$$= \frac{1}{n^2}\left\{2n\sigma_0^4 + \sum_{i\neq j}\left[\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left(X(s_i)^2X(s_j)^2\right) - \mathrm{E}_{(\sigma_0^2,\alpha_0)}\left(X(s_i)^2\right)\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left(X(s_j)^2\right)\right]\right\}$$

$$= \frac{1}{n^2}\left\{2n\sigma_0^4 + \sum_{i\neq j}\left[\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left(\left[\mathrm{e}^{-\alpha_0|i-j|/n}X(s_j) + \sqrt{\sigma_0^2\left(1-\mathrm{e}^{-2\alpha_0|i-j|/n}\right)}T_{ij}\right]^2 X(s_j)^2\right) - \sigma_0^4\right]\right\}$$

$$= \frac{1}{n^2}\left\{2n\sigma_0^4 + \sum_{i\neq j}\left[\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left\{\mathrm{e}^{-2\alpha_0|i-j|/n}X(s_j)^4 + \sigma_0^2\left(1-\mathrm{e}^{-2\alpha_0|i-j|/n}\right)T_{ij}^2X(s_j)^2\right.\right.\right.$$

$$\left.\left.\left. + 2X(s_j)^3\cdot\sqrt{\sigma_0^2\left(1-2\mathrm{e}^{-\alpha_0|i-j|/n}\right)}T_{ij}\right\} - \sigma_0^4\right]\right\}$$

$$= \frac{1}{n^2}\left\{2n\sigma_0^4 + \sum_{i\neq j}\left[3\sigma_0^4\mathrm{e}^{-2\alpha_0|i-j|/n} + \sigma_0^4\left(1-\mathrm{e}^{-2\alpha_0|i-j|/n}\right) + 0 - \sigma_0^4\right]\right\}$$

$$= \frac{1}{n^2}\left\{2n\sigma_0^4 + 2\sigma_0^4\sum_{i\neq j}\mathrm{e}^{-2\alpha_0|i-j|/n}\right\}$$

$$= 2\sigma_0^4\left\{\frac{1}{n} + \frac{2(n-1)\mathrm{e}^{-2\alpha_0/n} - 2n\mathrm{e}^{-4\alpha_0/n} + 2\mathrm{e}^{-2\alpha_0(n+1)/n}}{n^2\left(1-\mathrm{e}^{-2\alpha_0/n}\right)^2}\right\}.$$

Therefore, as $n\to\infty$, we have $\mathrm{Var}_{(\sigma_0^2,\alpha_0)}(A_3/n) \to \frac{\sigma_0^4\mathrm{e}^{-2\alpha_0}}{\alpha_0^2}$.

For any small number $\epsilon \in (0,1)$, we can apply the one-sided Chebyshev's inequality (or Cantelli's inequality) to obtain that as $n\to\infty$,

$$\Pr\left(A_3/n \le \mathrm{E}_{(\sigma_0^2,\alpha_0)}(A_3/n)(1-\epsilon)\right)$$
$$\le \frac{\mathrm{Var}_{(\sigma_0^2,\alpha_0)}(A_3/n)}{\mathrm{Var}_{(\sigma_0^2,\alpha_0)}(A_3/n) + \epsilon^2\left[\mathrm{E}_{(\sigma_0^2,\alpha_0)}(A_3/n)\right]^2} \to \frac{\mathrm{e}^{-2\alpha_0}/\alpha_0^2}{\mathrm{e}^{-2\alpha_0}/\alpha_0^2+\epsilon^2} < 1.$$

Therefore, for any $\epsilon\in(0,1)$, $\Pr\left(A_3/n > (1-\epsilon)\sigma_0^2\right) > 0$ for all sufficiently large n, which implies that A_3/n is lower bounded as $n\to\infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability (or equivalently, A_3/n does not converge to zero as $n\to\infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability).

Since $A_1 \le A_3$, A_1/n is also upper bounded as $n\to\infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability. Since $A_1/n = A_3/n - \left[X(s_1)^2 + X(s_n)^2\right]/n$ and $\left[X(s_1)^2 + X(s_n)^2\right]/n \to 0$ as $n\to\infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability, we can see that A_1/n is also lower bounded as $n\to\infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability.

(v) Since $u_* = (A_1-A_2)/(A_1/n)$, the conclusion follows from (iii) and (iv).

(vi) Since $v_* = (A_1-2A_2+A_3)/(A_1/n)$, the conclusion follows from (ii) and (iv).

(vii) Using the notation of u_* and v_* in Parts (v) and (vi), we have

$$1 - \frac{A_1A_3 - A_2^2}{A_1(A_1-2A_2+A_3)} = \frac{(A_1-A_2)^2}{A_1(A_1-2A_2+A_3)} = \frac{u_*^2}{nv_*}.$$

From Parts (v) and (vi), we have that $u_*^2/(nv_*) = O_p(n^{-1})$ as $n\to\infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability.

(viii) We have

$$A_1\mathrm{e}^{-2\alpha/n} - 2A_2\mathrm{e}^{-\alpha/n} + A_3$$

$$= A_1\left[(1-\mathrm{e}^{-\alpha/n}) - \frac{A_1-A_2}{A_1}\right]^2 + A_1 + A_3 - 2A_2 - \frac{(A_1-A_2)^2}{A_1}$$
$$= A_1\left[(1-\mathrm{e}^{-\alpha/n}) - \frac{A_1-A_2}{A_1}\right]^2 + \frac{A_1A_3-A_2^2}{A_1}. \tag{S.223}$$

Now if we replace $1-\mathrm{e}^{-\alpha/n}$ with α/n for all $\alpha \in [0, n^{1/6}]$, then the difference would be

$$\left|\left(A_1\mathrm{e}^{-2\alpha/n} - 2A_2\mathrm{e}^{-\alpha/n} + A_3\right) - \left[A_1\left(\frac{\alpha}{n} - \frac{A_1-A_2}{A_1}\right)^2 + \frac{A_1A_3-A_2^2}{A_1}\right]\right|$$
$$\le A_1\left|\left[(1-\mathrm{e}^{-\alpha/n}) - \frac{A_1-A_2}{A_1}\right]^2 - \left(\frac{\alpha}{n} - \frac{A_1-A_2}{A_1}\right)^2\right|$$
$$= A_1\left|1-\mathrm{e}^{-\alpha/n} + \frac{\alpha}{n} + \frac{2(A_1-A_2)}{A_1}\right| \cdot \left|1-\mathrm{e}^{-\alpha/n} - \frac{\alpha}{n}\right|$$
$$\overset{(i)}{\le} \left(A_1\frac{n^{1/6}}{n} + |A_1-A_2|\right)\frac{n^{1/3}}{n^2}, \tag{S.224}$$

where (i) follows from the fact that $1-\mathrm{e}^{-x} \le x$ and $|x-(1-\mathrm{e}^{-x})| \le x^2/2$ for all $x>0$. (S.224) implies that

$$\frac{\left|\left(A_1\mathrm{e}^{-2\alpha/n} - 2A_2\mathrm{e}^{-\alpha/n} + A_3\right) - \left[A_1\left(\frac{\alpha}{n} - \frac{A_1-A_2}{A_1}\right)^2 + \frac{A_1A_3-A_2^2}{A_1}\right]\right|}{A_1\left(\frac{\alpha}{n} - \frac{A_1-A_2}{A_1}\right)^2 + \frac{A_1A_3-A_2^2}{A_1}}$$
$$\le \frac{n^{1/3}}{n^2}\cdot\frac{\left(\frac{A_1n^{1/6}}{n} + |A_1-A_2|\right)}{\frac{A_1A_3-A_2^2}{A_1}}.$$

Using Parts (ii), (iii), (iv) and (vii) together with the definition of $\overline{\alpha}_n$, we observe that

$$\frac{n^{1/3}}{n^2}\cdot\frac{\left(\frac{A_1n^{1/6}}{n} + |A_1-A_2|\right)}{\frac{A_1A_3-A_2^2}{A_1}} \preceq \frac{n^{1/3}}{n^2}\cdot\frac{\left(n^{1/6} + \log^2 n\right)}{(A_1+A_3-2A_2)(1+O_p(n^{-1}))}$$
$$\preceq \frac{n^{1/3}}{n^2}\cdot\frac{n^{1/6}}{1+O_p(n^{-1})} = O_p\left(n^{-3/2}\right),$$

as $n\to\infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability. Hence the conclusion follows.

(ix) For $\alpha \in [0, n^{1/6}]$, $\alpha/n \le n^{-5/6} \to 0$ as $n\to\infty$. With the Taylor expansion of $1-\mathrm{e}^{-x}$ around $x=0$, as $n\to\infty$ almost surely $P_{(\sigma_0^2,\alpha_0)}$,

$$\sqrt{1-\mathrm{e}^{-2\alpha/n}} = \sqrt{\frac{2\alpha}{n}\left[1+O(n^{-5/6})\right]} = \sqrt{\frac{2\alpha}{n}}\left[1+O(n^{-5/12})\right]$$

and the $o(1)$ term is uniformly over all $\alpha \in [0, n^{1/6}]$. □

Lemma S.35. *Define a normalized log profile likelihood function*

$$\widetilde{\mathcal{L}}_*(\alpha) = \mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha, \alpha) + \frac{n}{2}\log\left(\frac{A_1A_3-A_2^2}{A_1}\right) + \frac{1}{2}\log\frac{n}{2}$$

$$= -\frac{n}{2}\log\left(A_1 e^{-2\alpha/n} - 2A_2 e^{-\alpha/n} + A_3\right) + \frac{1}{2}\log\left(1 - e^{-2\alpha/n}\right)$$
$$+ \frac{n}{2}\log\left(\frac{A_1A_3 - A_2^2}{A_1}\right) + \frac{1}{2}\log\frac{n}{2}. \tag{S.225}$$

$\widetilde{\mathcal{L}}_*(\alpha)$ in (S.225) is well defined for all sufficiently large n in $P_{(\sigma_0^2,\alpha_0)}$-probability. Then, under the model setup of Theorem 2.6 and Assumptions (A.2), (A.3), and (A.4'), the integrals

$$\int_0^\infty \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha, \quad \text{and} \quad \int_0^\infty \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\alpha)\mathrm{d}\alpha$$

are lower bounded by positive constants in $P_{(\sigma_0^2,\alpha_0)}$-probability. Furthermore, the following convergence relations hold

$$\int_0^\infty \left|\exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\} - \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2v_*}\right\}\right|\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \to 0, \tag{S.226}$$
$$\int_0^\infty \left|\exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\} - \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2v_*}\right\}\right|\pi(\alpha)\mathrm{d}\alpha \to 0, \tag{S.227}$$
$$\int_0^\infty \left|\widetilde{\pi}(\alpha|Y_n) - \pi_*(\alpha|Y_n)\right|\mathrm{d}\alpha \to 0, \tag{S.228}$$

as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability, for $\widetilde{\pi}(\alpha|Y_n)$ given in Theorem 2.3 and $\pi_*(\alpha|Y_n)$ given in Theorem 2.6.

Proof of Lemma S.35. Based on Part (vii) of Lemma S.34, $(A_1A_3 - A_2^2)/A_1 > 0$ as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability. Therefore, $\widetilde{\mathcal{L}}_*(\alpha)$ in (S.225) is well defined for all sufficiently large n in $P_{(\sigma_0^2,\alpha_0)}$-probability.

We first prove the convergence in $P_{(\sigma_0^2,\alpha_0)}$-probability in (S.226), and that the integral $\int_0^\infty \exp\{\widetilde{\mathcal{L}}_*(\alpha)\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha$ is lower bounded by positive constant in $P_{(\sigma_0^2,\alpha_0)}$-probability. Note that the only difference between (S.226) and (S.227) is that $\pi(\theta_0|\alpha)\pi(\alpha)$ is replaced by $\pi(\alpha)$. The integral condition (24) in Assumption (A.4') guarantees that in the following derivation, all $\pi(\theta_0|\alpha)\pi(\alpha)$ can be replaced by $\pi(\alpha)$. Therefore, in the derivation below, we will only prove for the integrals involving $\pi(\theta_0|\alpha)\pi(\alpha)$, and the proof of (S.227) and lower boundedness of $\int_0^\infty \exp\{\widetilde{\mathcal{L}}_*(\alpha)\}\pi(\alpha)\mathrm{d}\alpha$ follow similarly.

Proof of (S.226):

Define the following quantities

$$\widetilde{\mathsf{N}}_1 = \int_0^{n^{1/6}} \left|\exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\} - \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2v_*}\right\}\right|\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha,$$
$$\widetilde{\mathsf{N}}_2 = \int_{n^{1/6}}^\infty \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha,$$
$$\widetilde{\mathsf{N}}_3 = \int_{n^{1/6}}^\infty \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2v_*}\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha,$$
$$\widetilde{\mathsf{D}} = \int_0^\infty \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2v_*}\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha.$$

We define an auxiliary "variance" $\widetilde{v}_* = n(A_1A_3 - A_2^2)/A_1^2$ which is positive as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability given Parts (i) and (vii) of Lemma S.34. Then, we have that uniformly for all $\alpha \in [0, n^{1/6}]$, as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability,

$$\left|\exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\} - \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2\widetilde{v}_*}\right\}\right|$$

$$
\begin{aligned}
&= \left| \frac{[(A_1A_3 - A_2^2)/A_1]^{n/2}\sqrt{\frac{n}{2}(1-\mathrm{e}^{-2\alpha/n})}}{\left(A_1\mathrm{e}^{-2\alpha/n} - 2A_2\mathrm{e}^{-\alpha/n} + A_3\right)^{n/2}} - \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2\widetilde{v}_*}\right\} \right| \\
&\overset{(i)}{=} \left| \frac{[(A_1A_3 - A_2^2)/A_1]^{n/2}\sqrt{\alpha}\left[1+O(n^{-5/12})\right]}{\left[A_1\left(\frac{\alpha}{n} - \frac{A_1-A_2}{A_1}\right)^2 + \frac{A_1A_3-A_2^2}{A_1}\right]^{n/2}\left[1+O_p\left(n^{-3/2}\right)\right]^{n/2}} \right. \\
&\qquad \left. - \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2\widetilde{v}_*}\right\} \right| \\
&\overset{(ii)}{=} \left| \sqrt{\alpha}\left[1+O_p(n^{-5/12})\right]\left[1+\frac{1}{n}\frac{(\alpha-u_*)^2}{\widetilde{v}_*}\right]^{-n/2} - \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2\widetilde{v}_*}\right\} \right| \\
&\overset{(iii)}{\le} O_p(n^{-5/12})\cdot\sqrt{\alpha}\left[1+\frac{1}{n}\frac{(\alpha-u_*)^2}{\widetilde{v}_*}\right]^{-n/2} + \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2\widetilde{v}_*}\right\} \\
&\qquad \times \left| \exp\left(\frac{n}{2}\left\{\frac{(\alpha-u_*)^2}{n\widetilde{v}_*} - \log\left[1+\frac{(\alpha-u_*)^2}{n\widetilde{v}_*}\right]\right\}\right) - 1 \right| \\
&\overset{(iv)}{\le} O_p(n^{-1/3}) + \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2\widetilde{v}_*}\right\}\cdot\left[\exp\left(\frac{n}{2}\left\{\frac{(\alpha-u_*)^2}{n\widetilde{v}_*}\right\}^{11/6}\right) - 1\right] \\
&\overset{(v)}{\le} O_p(n^{-1/3}) + \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2\widetilde{v}_*}\right\}\cdot\frac{n}{2}\left[\frac{(\alpha-u_*)^2}{n\widetilde{v}_*}\right]^{11/6} \\
&\overset{(vi)}{\le} O_p(n^{-1/3}) + O_p(n^{1/12}\cdot 1\cdot n/2\cdot n^{-11/9}) = O_p(n^{-5/36}).
\end{aligned} \tag{S.229}
$$

In the derivations above, (i) follows from Lemma S.34 (vii) and (viii); (ii) follows from the fact that $[1+O_p\left(n^{-3/2}\right)]^{-n/2} = 1 + O_p(n^{-1/2})$ and the definitions of u_* and $\widetilde{v}_*$; (iii) follows from the triangle inequality; (iv) follows from Lemma S.34 (v), (vi), and the fact that $\alpha \in [0, n^{1/6}]$, hence $(\alpha-u_*)^2/(2v_*) \preceq n^{1/3}$, and the inequality $0 < x - \log(1+x) \le x^{11/6}$ for all $x > 0$; (v) follows from the inequality $\mathrm{e}^x - 1 \le 2x$ for $x \in (0,1)$ and for sufficiently large n; (vi) follows from a comparison of orders.

On the other hand, if we replace $\widetilde{v}_*$ with v_*, then Part (vii) of Lemma S.34 implies that $(\widetilde{v}_* - v_*)/v_* = O_p(n^{-1})$ as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability. Therefore, uniformly for all $\alpha \in [0, n^{1/6}]$, as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability,

$$
\left|\frac{(\alpha-u_*)^2(\widetilde{v}_* - v_*)}{v_*\widetilde{v}_*}\right| = \left|\frac{(\alpha-u_*)^2\cdot O_p(n^{-1})}{v_*\left(1+O_p(n^{-1})\right)}\right| \le O_p(n^{1/3}\cdot n^{-1}) = O_p(n^{-2/3}),
$$

and hence by $|\mathrm{e}^x - 1| \le 2|x|$ for all $|x| \le 1/2$,

$$
\begin{aligned}
&\left|\sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2\widetilde{v}_*}\right\} - \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2v_*}\right\}\right| \\
&= \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2v_*}\right\}\left|\exp\left\{-\frac{(\alpha-u_*)^2(\widetilde{v}_* - v_*)}{2v_*\widetilde{v}_*}\right\} - 1\right| \\
&\le \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2v_*}\right\}\cdot\left|\frac{(\alpha-u_*)^2(\widetilde{v}_* - v_*)}{v_*\widetilde{v}_*}\right| \\
&\le O_p\left(n^{1/12}\cdot 1\cdot n^{-2/3}\right) = O_p(n^{-1/2}).
\end{aligned} \tag{S.230}
$$

We combine (S.229) and (S.230) with the triangle inequality to conclude that uniformly for all

$\alpha \in [0, n^{1/6}]$, as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability,

$$\left| \exp\left\{ \widetilde{\mathcal{L}}_*(\alpha) \right\} - \sqrt{\alpha} \exp\left\{ -\frac{(\alpha - u_*)^2}{2v_*} \right\} \right| \leq O_p(n^{-5/36}) + O_p(n^{-1/2}) \leq O_p(n^{-5/36}). \quad \text{(S.231)}$$

As a result, we have that there exists a constant $C_1 > 0$ such that as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability,

$$\widetilde{\mathsf{N}}_1 \leq C_1 n^{-5/36} \int_0^\infty \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \to 0. \quad \text{(S.232)}$$

For $\widetilde{\mathsf{N}}_2$, since $(A_1A_3 - A_2^2)/A_1 > 0$ as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability, we have that

$$\begin{aligned}
\widetilde{\mathsf{N}}_2 &= \int_{n^{1/6}}^\infty \exp\left\{ \widetilde{\mathcal{L}}_*(\alpha) \right\} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \\
&= \int_{n^{1/6}}^\infty \frac{[(A_1A_3 - A_2^2)/A_1]^{n/2} \sqrt{\frac{n}{2}\left(1 - \mathrm{e}^{-2\alpha/n}\right)}}{\left(A_1\mathrm{e}^{-2\alpha/n} - 2A_2\mathrm{e}^{-\alpha/n} + A_3\right)^{n/2}} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \\
&= \int_{n^{1/6}}^\infty \frac{[(A_1A_3 - A_2^2)/A_1]^{n/2} \sqrt{\frac{n}{2}\left(1 - \mathrm{e}^{-2\alpha/n}\right)}}{\left\{ A_1\left[(1 - \mathrm{e}^{-\alpha/n}) - \frac{A_1 - A_2}{A_1}\right]^2 + (A_1A_3 - A_2^2)/A_1 \right\}^{n/2}} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \\
&\leq \int_{n^{1/6}}^\infty \sqrt{\frac{n}{2}\left(1 - \mathrm{e}^{-2\alpha/n}\right)} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \\
&\leq \int_{n^{1/6}}^\infty \sqrt{\alpha}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \to 0,
\end{aligned} \quad \text{(S.233)}$$

as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability according to Assumption (A.4') since $\overline{\kappa} < 1/6$.

For $\widetilde{\mathsf{N}}_3$, similarly we have that as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability,

$$\begin{aligned}
\widetilde{\mathsf{N}}_3 &= \int_{n^{1/6}}^\infty \sqrt{\alpha} \exp\left\{ -\frac{(\alpha - u_*)^2}{2v_*} \right\} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \\
&\leq \int_{n^{1/6}}^\infty \sqrt{\alpha}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \to 0.
\end{aligned} \quad \text{(S.234)}$$

Hence, (S.226) follows by combining (S.232), (S.233), and (S.234) using the triangle inequality.

<u>Proof of the lower boundedness of $\int_0^\infty \exp\left\{ \widetilde{\mathcal{L}}_*(\alpha) \right\} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha$:</u>

We first derive a lower bound for $\widetilde{\mathsf{D}}$. By Lemma S.34 (v), $|u_*| \leq C_5$ for some constant $C_5 > 0$ as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability. By Lemma S.34 (vi), $v_* > C_6$ for some constant $C_6 > 0$ as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability. By Assumptions (A.2) and (A.4'), $\inf_{\alpha \in [1,2]} \pi(\theta_0|\alpha)\pi(\alpha) \geq C_7 > 0$ for some constant C_7. This implies that there exists a constant $C_8 > 0$, such that

$$\begin{aligned}
\widetilde{\mathsf{D}} &\geq \int_1^2 \sqrt{\alpha} \exp\left\{ -\frac{(\alpha - u_*)^2}{2v_*} \right\} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \geq C_7 \int_1^2 \exp\left\{ -\frac{(\alpha - u_*)^2}{2v_*} \right\} \mathrm{d}\alpha \\
&\geq C_7 \int_1^2 \exp\left\{ -\frac{(\alpha + C_5)^2}{2C_6} \right\} \mathrm{d}\alpha \equiv C_8 > 0,
\end{aligned} \quad \text{(S.235)}$$

as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability.

Now given the convergence in (S.226), we have that as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability,

$$\left| \int_0^\infty \exp\left\{ \widetilde{\mathcal{L}}_*(\alpha) \right\} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha - \int_0^\infty \sqrt{\alpha} \exp\left\{ -\frac{(\alpha - u_*)^2}{2v_*} \right\} \pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \right| \to 0.$$

This and (S.235) together imply that

$$\begin{aligned}
&\int_0^{\infty} \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \\
&\geq \widetilde{\mathsf{D}} - \left|\int_0^{\infty} \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha - \int_0^{\infty}\sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2v_*}\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha\right| \\
&> C_8/2,
\end{aligned} \tag{S.236}$$

as $n\to\infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability, which proves the lower boundedness.

We note that as stated at the beginning of this proof, proving the convergence in (S.227) and the lower boundedness of $\int_0^\infty \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\alpha)\mathrm{d}\alpha$ follows exactly the same procedure as proving (S.226) and the lower boundedness of $\int_0^\infty \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha$ under Assumption (A.4'), and is therefore omitted.

Proof of (S.228):

Based on the definitions of $\widetilde{\pi}(\alpha|Y_n)$ and $\pi_*(\alpha|Y_n)$, by Lemma S.30, the convergence in (S.228) holds true if the following relation holds as $n\to\infty$, in $P_{(\sigma_0^2,\alpha_0)}$-probability,

$$\frac{\int_0^\infty \left|\exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\} - \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2v_*}\right\}\right|\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}{\int_0^\infty \sqrt{\alpha}\exp\left\{-\frac{(\alpha-u_*)^2}{2v_*}\right\}\mathrm{d}\alpha} \to 0, \tag{S.237}$$

which follows from (S.226) and (S.235). Hence the proof for Lemma S.30 is complete. □

Proof of Theorem 2.6

Proof of Theorem 2.6. We first prove the convergence in (25). The proof follows the same process in the proof of Theorem 2.3, with some differences due to the new Assumption (A.4'). The conclusion of Theorem 2.3 is proved by showing (S.173) and (S.174). We show them respectively under the new Assumption (A.4'). We notice that since $p=0$ in Theorem 2.6, $n-p=n$ in (S.173) and (S.174).

Proof of (S.173):

Using the same notation as in the proof of Theorem 2.3, we define N_1, N_2, N_3, and D as in (S.178) and (S.177). The first step of showing $\mathsf{N}_1/\mathsf{D}\to 0$ is exactly the same as in the proof of Theorem 2.3, since this step only relies on Assumptions (A.2) and (A.3), which are both assumed in Theorem 2.6 as well. The main differences lie in the next two steps of showing $\mathsf{N}_2/\mathsf{D}\to 0$ and $\mathsf{N}_3/\mathsf{D}\to 0$.

Proof of $\mathsf{N}_2/\mathsf{D}\to 0$:

Using the upper bound of N_2 in (S.183), together with the definition of D in (S.177), we have that

$$\begin{aligned}
\frac{\mathsf{N}_2}{\mathsf{D}} &\leq \frac{2\int_0^{\alpha_n} \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\alpha)\mathrm{d}\alpha + \frac{4\theta_0\sqrt{\pi}}{\sqrt{n}}\int_0^{\alpha_n}\mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}{\frac{2\theta_0\sqrt{\pi}}{\sqrt{n}}\int_0^\infty \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha} \\
&= \frac{\sqrt{n}\int_0^{\alpha_n}\exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\alpha)\mathrm{d}\alpha}{\theta_0\sqrt{\pi}\int_0^\infty \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha} + \frac{2\int_0^{\alpha_n}\exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}{\int_0^\infty\exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha},
\end{aligned} \tag{S.238}$$

where $\widetilde{\mathcal{L}}_*(\alpha)$ is the normalized log profile likelihood defined in (S.225).

We now show the first term in (S.238) converges to zero in probability. For the numerator, by the definition of $\widetilde{\mathcal{L}}_*(\alpha)$, since $(A_1A_3 - A_2^2)/A_1 > 0$ as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability, we have that

$$
\begin{aligned}
&\sqrt{n}\int_0^{\underline{\alpha}_n} \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\alpha)\mathrm{d}\alpha \\
&= \sqrt{n}\int_0^{\underline{\alpha}_n} \frac{[(A_1A_3 - A_2^2)/A_1]^{n/2}}{(A_1\mathrm{e}^{-2\alpha/n} - 2A_2\mathrm{e}^{-\alpha/n} + A_3)^{n/2}}\sqrt{\frac{n}{2}(1-\mathrm{e}^{-2\alpha/n})}\pi(\alpha)\mathrm{d}\alpha \\
&= \sqrt{n}\int_0^{\underline{\alpha}_n} \frac{[(A_1A_3 - A_2^2)/A_1]^{n/2}}{\left[A_1\left(1-\mathrm{e}^{-\alpha/n} - \frac{A_1-A_2}{A_1}\right)^2 + (A_1A_3-A_2^2)/A_1\right]^{n/2}}\sqrt{\frac{n}{2}(1-\mathrm{e}^{-2\alpha/n})}\pi(\alpha)\mathrm{d}\alpha \\
&\leq \sqrt{n}\int_0^{\underline{\alpha}_n} 1\cdot\sqrt{\frac{n}{2}(1-\mathrm{e}^{-2\alpha/n})}\pi(\alpha)\mathrm{d}\alpha \\
&\leq \sqrt{n}\int_0^{\underline{\alpha}_n} \sqrt{\alpha}\pi(\alpha)\mathrm{d}\alpha, \qquad \text{(S.239)}
\end{aligned}
$$

where in the last step, the first ratio in the integral is less than 1 and we have used $1-\mathrm{e}^{-x} \leq x$ for all $x > 0$. By (24) in Assumption (A.4'), we have that this upper bound goes to zero as $n \to \infty$. Therefore, $\sqrt{n}\int_0^{\underline{\alpha}_n} \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\alpha)\mathrm{d}\alpha \to 0$ as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability. Since the denominator $\theta_0\sqrt{\pi}\int_0^\infty \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha$ is lower bounded by positive constant in $P_{(\sigma_0^2,\alpha_0)}$-probability according to Lemma S.35 (in (S.236)), we have that the first term in (S.238) converges to zero as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability.

We then show the second term in (S.238) converges to zero in probability. For the numerator, similar to (S.239), we have that

$$
\int_0^{\underline{\alpha}_n} \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha \leq \int_0^{\underline{\alpha}_n} \sqrt{\alpha}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha,
$$

which converges to zero as $n \to \infty$ since $\underline{\alpha}_n \to 0$ as $n \to \infty$ and $\int_0^\infty \sqrt{\alpha}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha$ is finite according to Assumption (A.4'). Therefore, with the lower bounded denominator, the second term in (S.238) also converges to zero as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability. This together with (S.238) has shown that $\mathsf{N}_2/\mathsf{D} \to 0$ as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability.

Proof of $\mathsf{N}_3/\mathsf{D} \to 0$: Using the upper bound of N_3 in (S.188), together with the definition of D in (S.177), we have that

$$
\begin{aligned}
\frac{\mathsf{N}_3}{\mathsf{D}} &\leq \frac{2\int_{\overline{\alpha}_n}^\infty \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\alpha)\mathrm{d}\alpha + \frac{4\theta_0\sqrt{\pi}}{\sqrt{n}}\int_{\overline{\alpha}_n}^\infty \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}{\frac{2\theta_0\sqrt{\pi}}{\sqrt{n}}\int_0^\infty \mathrm{e}^{\mathcal{L}_n(\alpha^{-2\nu}\widetilde{\theta}_\alpha,\alpha)}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha} \\
&= \frac{\sqrt{n}\int_{\overline{\alpha}_n}^\infty \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\alpha)\mathrm{d}\alpha}{\theta_0\sqrt{\pi}\int_0^\infty \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha} + \frac{2\int_{\overline{\alpha}_n}^\infty \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}{\int_0^\infty \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}, \qquad \text{(S.240)}
\end{aligned}
$$

For both terms in (S.240), the denominators are lower bounded by positive constants in $P_{(\sigma_0^2,\alpha_0)}$-probability by Lemma S.35. Using the same derivation as in (S.239), the numerator in the first term of (S.240) can be upper bounded by

$$
\sqrt{n}\int_{\overline{\alpha}_n}^\infty \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\alpha)\mathrm{d}\alpha \leq \sqrt{n}\int_{\overline{\alpha}_n}^\infty \sqrt{\alpha}\pi(\alpha)\mathrm{d}\alpha,
$$

which converges to zero as $n \to \infty$ by (24) in Assumption (A.4'). The numerator in the second term of (S.240) also converges to zero since $\overline{\alpha}_n \to \infty$ as $n \to \infty$ and $\int_0^\infty \sqrt{\alpha}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha$ is finite according to Assumption (A.4'). Therefore, it follows that $\mathsf{N}_3/\mathsf{D} \to 0$ as $n \to \infty$ in

$P_{(\sigma_0^2,\alpha_0)}$-probability. Thus, the convergence in (S.173) happens as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability.

Proof of (S.174):
Compared to the proof of (S.174) in the proof of Theorem 2.3, the upper bounds in (S.192) and (S.193) still hold. We only need to show the convergence in (S.194) and (S.195) using the new Assumption (A.4'). In particular, using the definition of $\widetilde{\mathcal{L}}_*(\alpha)$ in (S.225), we have

$$\int_0^{\underline{\alpha}_n} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha = \frac{\int_0^{\underline{\alpha}_n} \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}{\int_0^{\infty} \exp\left\{\widetilde{\mathcal{L}}_*(\alpha)\right\}\pi(\theta_0|\alpha)\pi(\alpha)\mathrm{d}\alpha}$$

which converges to zero in $P_{(\sigma_0^2,\alpha_0)}$-probability as already shown above in the proof of $\mathsf{N}_2/\mathsf{D} \to 0$. Similarly, $\int_{\overline{\alpha}_n}^{\infty} \widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha \to 0$ in $P_{(\sigma_0^2,\alpha_0)}$-probability as shown in the proof of $\mathsf{N}_3/\mathsf{D} \to 0$. Therefore, the convergence in (S.174) happens as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability. This completes the proof of the convergence in (25).

For the proof of the convergence in (26), we notice that

$$\begin{aligned}
&\int_0^{\infty}\int_{\mathbb{R}} \left| \frac{\sqrt{n}}{2\sqrt{\pi}\theta_0} \mathrm{e}^{-\frac{n(\theta-\widetilde{\theta}_{\alpha_0})^2}{4\theta_0^2}} \cdot \widetilde{\pi}(\alpha|Y_n) - \frac{\sqrt{n}}{2\sqrt{\pi}\theta_0} \mathrm{e}^{-\frac{n(\theta-\widetilde{\theta}_{\alpha_0})^2}{4\theta_0^2}} \cdot \pi_*(\alpha|Y_n) \right| \mathrm{d}\theta\mathrm{d}\alpha \\
= &\int_0^{\infty} \left|\widetilde{\pi}(\alpha|Y_n) - \pi_*(\alpha|Y_n)\right| \mathrm{d}\alpha \to 0,
\end{aligned}$$

as $n \to \infty$ in $P_{(\sigma_0^2,\alpha_0)}$-probability, by (S.228) of Lemma S.35. Then (26) follows from (25) and the triangle inequality. □

S5.2 Proof of Corollary 2.7

Proof of Corollary 2.7. Recall that for Case (ii) in Section 2.4 of the main text, we observe the 1-dimensional Ornstein-Uhlenbeck process with a constant regression term $\mathrm{m}_1(\cdot) \equiv 1$, so $Y(\cdot) = \beta_0 + X(\cdot) \sim \mathrm{GP}(0, \sigma_0^2 K_{\alpha_0,\nu})$ on the grid $s_i = i/n$, for $i = 1, \ldots, n$, where β_0 denotes the true mean parameter. In Corollary 2.7, we have defined $B_1 = \sum_{i=2}^{n-1} Y(s_i)$, $B_2 = \sum_{i=1}^{n} Y(s_i)$, and A_1, A_2, A_3 as in (23).

We briefly explain the derivation of the expressions for $\widetilde{\theta}_\alpha$ and $\widetilde{\mathcal{L}}_n(\alpha)$ in Corollary 2.7. With $M_n = 1_n$, using the expression of R_α in Section 2.4, it follows that

$$M_n^\top R_\alpha^{-1} M_n = \frac{(n-2)\left(1-\mathrm{e}^{-\alpha/n}\right)+2}{1+\mathrm{e}^{-\alpha/n}}, \quad M_n^\top R_\alpha^{-1} Y_n = \frac{B_2 - B_1\mathrm{e}^{-\alpha/n}}{1+\mathrm{e}^{-\alpha/n}}.$$

We then plug in these formulas to the expression of $\widetilde{\theta}_\alpha$ in (7) and $\widetilde{\mathcal{L}}_n(\alpha)$ in (8) with $\Omega_\beta = 0_{p\times p}$ to obtain the expressions for $\widetilde{\theta}_\alpha$ and $\widetilde{\mathcal{L}}_n(\alpha)$ in Corollary 2.7. We notice that the profile restricted log-likelihood $\widetilde{\mathcal{L}}_n(\alpha)$ is defined up to an additive constant.

Similarly, we obtain the normal conditional posterior of β in (27) of the main text, by plugging the formulas above to the conditional posterior of β in (5) of the main text. The convergence in total variation norm of (28) follows directly from Theorem 2.3, under Assumptions (A.1), (A.2), (A.3), and (A.4).

Next, we prove that the posterior of β is inconsistent for β_0. We already know that the conditional posterior of β is given by $\beta|Y_n, \theta, \alpha \sim \mathcal{N}(\mu_n, v_n)$, where

$$\mu_n = \frac{B_2 - B_1\mathrm{e}^{-\alpha/n}}{(n-2)(1-\mathrm{e}^{-\alpha/n})+2}, \quad v_n = \frac{\theta\left(1+\mathrm{e}^{-\alpha/n}\right)}{\left[(n-2)(1-\mathrm{e}^{-\alpha/n})+2\right]\alpha}.$$

Let $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-z^2/2} \mathrm{d}z$ be the standard normal cumulative distribution function. For a given $\epsilon_0 > 0$ whose value will be chosen later, using the mean value theorem, we have that

$$
\begin{aligned}
&\Pi(|\beta - \beta_0| > \epsilon_0 | Y_n, \theta, \alpha) = 1 - \Pi(|\beta - \beta_0| \le \epsilon_0 | Y_n, \theta, \alpha) \\
&= 1 - \left\{ \Phi\left(\frac{\beta_0 + \epsilon_0 - \mu_n}{\sqrt{v_n}} \right) - \Phi\left(\frac{\beta_0 - \epsilon_0 - \mu_n}{\sqrt{v_n}} \right) \right\} \\
&= 1 - \frac{2\epsilon_0}{\sqrt{2\pi v_n}} \exp\left(-\frac{x_1^2}{2v_n} \right) \\
&\ge 1 - \frac{2\epsilon_0}{\sqrt{2\pi v_n}},
\end{aligned}
\tag{S.241}
$$

for some value $x_1 \in [\beta_0 - \epsilon_0 - \mu_n, \beta_0 + \epsilon_0 - \mu_n]$, where the inequality follows from the bound $\exp\left(-x_1^2/(2v_n)\right) \le 1$.

Let $\mathcal{E}_9 = \{\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]\}$. Then under Assumptions (A.1)-(A.4), Theorem 2.3 implies that $\left|\Pi(\mathcal{E}_9^c | Y_n) - \widetilde{\Pi}(\mathcal{E}_9^c | Y_n)\right| \to 0$ as $n \to \infty$ almost surely $P_{(\beta_0, \sigma_0^2, \alpha_0)}$. (S.194) and (S.195) in the proof of Theorem 2.3 imply that $\widetilde{\Pi}(\mathcal{E}_9^c | Y_n) \to 0$ as $n \to \infty$ almost surely $P_{(\beta_0, \sigma_0^2, \alpha_0)}$. Therefore, $\Pi(\mathcal{E}_9^c | Y_n) \to 0$ as $n \to \infty$ almost surely $P_{(\beta_0, \sigma_0^2, \alpha_0)}$. This implies that given any $\eta \in (0, 1/4)$, any $\delta \in (0, 1/4)$, there exist two numbers $0 < \alpha_1 < \alpha_2 < \infty$ and a sufficiently large integer $N_{10}' > 0$ ($\alpha_1, \alpha_2, N_{10}'$ are dependent on η, δ), such that for all $n > N_{10}'$, $\Pr\left(\Pi(\mathcal{A}_{3n}^c \mid Y_n) \le \delta/2\right) > 1 - \eta/2$, where we let $\mathcal{A}_{3n} = \{\alpha \in [\alpha_1, \alpha_2]\}$.

We find the limit of v_n. For the α_1, α_2 above, it is clear that using Taylor series expansion for e^{-x}, we have that as $n \to \infty$,

$$
\sup_{\alpha \in [\alpha_1, \alpha_2]} n \left| 1 - \mathrm{e}^{-\alpha/n} - \alpha/n \right| \to 0. \tag{S.242}
$$

Therefore, for a given $\theta > 0$, as $n \to \infty$,

$$
\sup_{\alpha \in [\alpha_1, \alpha_2]} \left| v_n - \frac{2\theta}{\alpha(\alpha + 2)} \right| \to 0. \tag{S.243}
$$

(S.241) and (S.243) imply that by choosing N_{10}' to be large, for all $n > N_{10}'$, on the event $\mathcal{A}_{3n}$, $v_n > \frac{\theta}{2\alpha(\alpha+2)}$, such that

$$
\begin{aligned}
&\Pi(|\beta - \beta_0| > \epsilon_0 | Y_n, \theta, \alpha) \\
&\ge 1 - \frac{2\epsilon_0}{\sqrt{2\pi v_n}} > 1 - \frac{2\epsilon_0}{\sqrt{2\pi \cdot \frac{\theta}{2\alpha(\alpha+2)}}} = 1 - 2\epsilon_0 \sqrt{\frac{\alpha_2(\alpha_2 + 2)}{\pi\theta}}.
\end{aligned}
\tag{S.244}
$$

Let $\mathcal{A}_{4n} = \left\{ |\theta - \theta_0| \le n^{-1/2} \log^2 n \right\}$. Under Assumptions (A.1)-(A.4), Theorem 2.3 and Lemma S.10 imply that $\Pi(\mathcal{A}_{4n}^c | Y_n) \to 0$ as $n \to \infty$ almost surely $P_{(\beta_0, \sigma_0^2, \alpha_0)}$. In other words, for any small $\eta \in (0, 1/4)$, any small $\delta \in (0, 1/4)$, there exists a large integer N_{11}', such that for all $n > N_{11}'$, $\Pr(\Pi(\mathcal{A}_{4n}^c | Y_n) \le \delta/2) > 1 - \eta/2$. Therefore, together with $\Pr\left(\Pi(\mathcal{A}_{3n}^c \mid Y_n) \le \delta/2\right) > 1 - \eta/2$ for all $n > N_{10}'$, we have that $\Pr(\Pi(\mathcal{A}_{3n}^c \cup \mathcal{A}_{4n}^c | Y_n) \le \delta) > 1 - \eta$ for all $n > \max(N_{10}', N_{11}')$, which implies that $\Pi(\mathcal{A}_{3n} \cap \mathcal{A}_{4n} | Y_n) > 1 - \delta$ happens with $P_{(\beta_0, \sigma_0^2, \alpha_0)}$-probability at least $1 - \eta$ for all $n > \max(N_{10}', N_{11}')$. On the event $\Pi(\mathcal{A}_{3n} \cap \mathcal{A}_{4n} | Y_n) > 1 - \delta$ for all $n > \max(N_{10}', N_{11}')$,

$$
\begin{aligned}
&\Pi(|\beta - \beta_0| > \epsilon_0 | Y_n) = \mathrm{E}_{\Pi(\mathrm{d}\theta, \mathrm{d}\alpha | Y_n)} \left[\Pi(|\beta - \beta_0| > \epsilon_0 | Y_n, \theta, \alpha) \right] \\
&= \mathrm{E}_{\Pi(\mathrm{d}\theta, \mathrm{d}\alpha | Y_n)} \left[\Pi(|\beta - \beta_0| > \epsilon_0 | Y_n, \theta, \alpha) \cdot \mathcal{I}(\mathcal{A}_{3n} \cap \mathcal{A}_{4n}) \right] \\
&\quad + \mathrm{E}_{\Pi(\mathrm{d}\theta, \mathrm{d}\alpha | Y_n)} \left[\Pi(|\beta - \beta_0| > \epsilon_0 | Y_n, \theta, \alpha) \cdot \mathcal{I}(\mathcal{A}_{3n}^c \cup \mathcal{A}_{4n}^c) \right] \\
&\overset{(i)}{\ge} \mathrm{E}_{\Pi(\mathrm{d}\theta, \mathrm{d}\alpha | Y_n)} \left[\left\{ 1 - 2\epsilon_0 \sqrt{\frac{\alpha_2(\alpha_2 + 2)}{\pi\theta_0/2}} \right\} \cdot \mathcal{I}(\mathcal{A}_{3n} \cap \mathcal{A}_{4n}) \right]
\end{aligned}
$$

$$= \left\{1 - 2\epsilon_0\sqrt{\frac{2\alpha_2(\alpha_2+2)}{\pi\theta_0}}\right\} \Pi(\mathcal{A}_{3n} \cap \mathcal{A}_{4n} \,| Y_n)$$

$$> (1-\delta)\left\{1 - 2\epsilon_0\sqrt{\frac{2\alpha_2(\alpha_2+2)}{\pi\theta_0}}\right\}, \tag{S.245}$$

where $\mathrm{E}_{\Pi(\mathrm{d}\theta,\mathrm{d}\alpha|Y_n)}$ denotes the posterior expectation with respect to (θ,α); the inequality (i) follows because $\theta > \theta_0 - n^{-1/2}\log^2 n > \theta_0/2$ on the event $\mathcal{A}_{4n}$ for $n > N_{11}'$ and the second expectation in the previous line is nonnegative.

On the right-hand side of (S.245), we can set $\delta = 1/2$ and $\epsilon_0 = \frac{1}{4}\sqrt{\frac{\pi\theta_0}{2\alpha_2(\alpha_2+2)}}$ (ϵ_0 depends on α_2 and hence depends on η), such that (S.245) leads to $\Pr(\Pi(|\beta-\beta_0| > \epsilon_0|Y_n) > 1/4) > 1-\eta$ for all $n > \max(N_{10}', N_{11}')$. The conclusion of Corollary 2.7 follows by taking $\epsilon_0 = \frac{1}{4}\sqrt{\frac{\pi\theta_0}{2\alpha_2(\alpha_2+2)}}$, $\delta_0 = 1/4$, and $N_2 = \max(N_{10}', N_{11}')$. □

S6 Proof of Theorems in Section 3

In this section, we present the proofs of Theorems 3.1, 3.2, 3.3, and 3.4 in Section 3 of the main text.

S6.1 Proof of (30) and Theorem 3.1

Derivation of $\mathrm{v}_n(s^*;\sigma^2,\alpha)$ in (30):

Define $b_\alpha(s^*) = \mathrm{m}(s^*) - M_n^\top R_\alpha^{-1} r_\alpha(s^*)$. First we recall that

$$\beta|Y_n,\sigma^2,\alpha \sim \mathcal{N}\left(\left(M_n^\top R_\alpha^{-1}M_n+\Omega_\beta\right)^{-1}M_n^\top R_\alpha^{-1}Y_n,\ \sigma^2\left(M_n^\top R_\alpha^{-1}M_n+\Omega_\beta\right)^{-1}\right).$$

Given Y_n, the GP predictive distribution for $\widetilde{Y}(s^*)$ is

$$\widetilde{Y}(s^*)|Y_n,\beta,\sigma^2,\alpha \sim \mathcal{N}\left(\widehat{Y}(s^*;\beta,\alpha),\ \sigma^2\left[1 - r_\alpha(s^*)^\top R_\alpha^{-1} r_\alpha(s^*)\right]\right),$$

where

$$\widehat{Y}(s^*;\beta,\alpha) = \mathrm{m}(s^*)^\top\beta + r_\alpha(s^*)^\top R_\alpha^{-1}\left(Y_n - M_n\beta\right) = r_\alpha(s^*)^\top R_\alpha^{-1} Y_n + b_\alpha(s^*)^\top\beta.$$

Therefore, by the law of iterated expectation, we can integrate out β and obtain that $Y(s^*)|Y_n,\sigma^2,\alpha$ still follows a normal distribution, whose mean is

$$\mathrm{E}\left\{\widetilde{Y}(s^*)|Y_n,\sigma^2,\alpha\right\} = \mathrm{E}_{\beta|Y_n,\sigma^2,\alpha}\mathrm{E}\left\{Y(s^*)|Y_n,\beta,\sigma^2,\alpha\right\} = \mathrm{E}_{\beta|Y_n,\sigma^2,\alpha}\left\{\widehat{Y}(s^*;\beta,\alpha)\right\}$$
$$= r_\alpha(s^*)^\top R_\alpha^{-1}Y_n + b_\alpha(s^*)^\top\left(M_n^\top R_\alpha^{-1}M_n+\Omega_\beta\right)^{-1}M_n^\top R_\alpha^{-1}Y_n, \tag{S.246}$$

and by the law of total variance, the variance of $Y(s^*)|Y_n,\sigma^2,\alpha$ is

$$\mathrm{Var}\left\{\widetilde{Y}(s^*)|Y_n,\sigma^2,\alpha\right\}$$
$$= \mathrm{Var}_{\beta|Y_n,\sigma^2,\alpha}\left\{\widehat{Y}(s^*;\beta,\alpha)\right\} + \mathrm{E}_{\beta|Y_n,\sigma^2,\alpha}\left\{\mathrm{Var}[Y(s^*)|Y_n,\beta,\sigma^2,\alpha]\right\}$$
$$= \sigma^2 b_\alpha(s^*)^\top\left(M_n^\top R_\alpha^{-1}M_n+\Omega_\beta\right)^{-1}b_\alpha(s^*) + \sigma^2\left[1 - r_\alpha(s^*)^\top R_\alpha^{-1}r_\alpha(s^*)\right], \tag{S.247}$$

which has proved (30).

Proof of Theorem 3.1. The proof of Part (i) closely follow Theorem 2.1 for a given $\alpha > 0$, and the proof of Part (ii) closely follow Theorem 2.3 for the joint posterior of (θ, α). The two proofs are highly similar and we only show the proof of Part (ii) below, while the proof of Part (i) follows similarly.

We first use the reparameterization $\theta = \sigma^2\alpha^{2\nu}$ to replace σ^2 by θ. By the definition of $\mathrm{v}_n(s^*;\sigma^2,\alpha)$ in (30), we have the following decomposition of ratios:

$$\frac{\mathrm{v}_n(s^*;\sigma^2,\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)} = \frac{\mathrm{v}_n(s^*;\sigma^2,\alpha)}{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)} \cdot \frac{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)}.$$

Using the formula (30) for $\mathrm{v}_n(s^*;\sigma^2,\alpha)$, we can see that for any $s^* \in \mathcal{S}\setminus\mathcal{S}_n$,

$$\begin{aligned} \frac{\mathrm{v}_n(s^*;\sigma^2,\alpha)}{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)} &= \frac{\sigma^2}{\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu}} = \frac{\theta}{\widetilde{\theta}_{\alpha_0}}, \\ \frac{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)} &= \frac{\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu}}{\theta_0/\alpha^{2\nu}} = \frac{\widetilde{\theta}_{\alpha_0}}{\theta_0}, \end{aligned} \tag{S.248}$$

Recall that Lemma S.10 has proved that for the event $\mathcal{E}_4(\epsilon) = \left\{|\widetilde{\theta}_{\alpha_0} - \theta_0| < \epsilon\right\}$,

$$\Pr\left\{\mathcal{E}_4(5\theta_0 n^{-1/2}\log n)^c\right\} \leq 3\exp(-4\log^2 n)$$

for all sufficiently large n. Let $\mathcal{E}_8 = \{|\theta/\widetilde{\theta}_{\alpha_0} - 1| > n^{-1/2}\log n\}$. Then by Theorem 2.3, as $n\to\infty$, almost surely $P_{(\sigma_0^2,\alpha_0)}$,

$$\left|\Pi\left(\mathcal{E}_8\,|Y_n\right) - \int_0^\infty\int_{\mathcal{E}_8} \frac{\sqrt{n}}{2\sqrt{\pi}\theta_0}\mathrm{e}^{-\frac{n(\theta-\widetilde{\theta}_{\alpha_0})^2}{4\theta_0^2}}\cdot\widetilde{\pi}(\alpha|Y_n)\mathrm{d}\theta\mathrm{d}\alpha\right| \to 0. \tag{S.249}$$

(For Part (i), we simply use Theorem 2.1 instead and replace $\Pi\left(\mathcal{E}_8\,|Y_n\right)$ in (S.249) by $\Pi\left(\mathcal{E}_8\,|Y_n,\alpha\right)$ and remove the integral over α, similarly for the rest of the proof.)

On the event $\mathcal{E}_4(5\theta_0 n^{-1/2}\log n)\cap\mathcal{E}_8$, for all sufficiently large n,

$$\left|\theta - \widetilde{\theta}_{\alpha_0}\right| > \widetilde{\theta}_{\alpha_0} n^{-1/2}\log n > (\theta_0 - 5\theta_0 n^{-1/2}\log n)n^{-1/2}\log n > (\theta_0/2)n^{-1/2}\log n.$$

Using the normal tail inequality (S.151), the integral in (S.249) can be bounded by

$$\begin{aligned} &\int_0^\infty\int_{\mathcal{E}_8} \frac{\sqrt{n}}{2\sqrt{\pi}\theta_0}\mathrm{e}^{-\frac{n(\theta-\widetilde{\theta}_{\alpha_0})^2}{4\theta_0^2}}\cdot\widetilde{\pi}(\alpha|Y_n)\mathrm{d}\theta\mathrm{d}\alpha \\ &\leq \int_{|\theta-\widetilde{\theta}_{\alpha_0}|>\frac{\theta_0}{2}n^{-1/2}\log n} \frac{\sqrt{n}}{2\sqrt{\pi}\theta_0}\mathrm{e}^{-\frac{n(\theta-\widetilde{\theta}_{\alpha_0})^2}{4\theta_0^2}}\mathrm{d}\theta\cdot\int_0^\infty\widetilde{\pi}(\alpha|Y_n)\mathrm{d}\alpha \\ &\leq \exp\left(-\log^2 n/16\right) \to 0, \text{ as } n\to\infty. \end{aligned} \tag{S.250}$$

Therefore, by combining (S.248), (S.249) and (S.250) and noticing that $\mathcal{E}_4(5\theta_0 n^{-1/2}\log n,\alpha)$ happens almost surely $P_{(\sigma_0^2,\alpha_0)}$ as $n\to\infty$ by the Borel-Cantelli lemma, we have that

$$\begin{aligned} &\Pi\left(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{v}_n(s^*;\sigma^2,\alpha)}{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)} - 1\right| > n^{-1/2}\log n\middle|Y_n\right) \\ &= \Pi\left(\left|\frac{\theta}{\widetilde{\theta}_{\alpha_0}} - 1\right| > n^{-1/2}\log n\middle|Y_n\right) = \Pi\left(\mathcal{E}_8\,|Y_n\right) \to 0, \text{ a.s. } P_{(\sigma_0^2,\alpha_0)}. \end{aligned} \tag{S.251}$$

The relation of (S.248) and the almost sure convergence property of $\mathcal{E}_4(5\theta_0 n^{-1/2}\log n)$ also implies that

$$\Pi\left(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)}-1\right|>5n^{-1/2}\log n\middle|Y_n\right)$$
$$=\Pi\left(\left|\frac{\widetilde{\theta}_{\alpha_0}}{\theta_0}-1\right|>5n^{-1/2}\log n\middle|Y_n\right)=0,\ \text{a.s. } P_{(\sigma_0^2,\alpha_0)}. \tag{S.252}$$

For n sufficiently large, we have $5n^{-1/2}\log n<1/5$. Hence, $|\widetilde{\theta}_{\alpha_0}/\theta_0-1|<1/5$ and $\widetilde{\theta}_{\alpha_0}/\theta_0<6/5$ as $n\to\infty$ almost surely $P_{(\sigma_0^2,\alpha_0)}$. We combine (S.251) and (S.252) to obtain that

$$\begin{aligned}
&\Pi\left(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{v}_n(s^*;\sigma^2,\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)}-1\right|>7n^{-1/2}\log n\middle|Y_n\right)\\
&=\Pi\left(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{v}_n(s^*;\sigma^2,\alpha)}{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)}\cdot\frac{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)}-1\right|>7n^{-1/2}\log n\middle|Y_n\right)\\
&\le\Pi\Bigg(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\Bigg\{\left|\frac{\mathrm{v}_n(s^*;\sigma^2,\alpha)}{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)}-1\right|\cdot\left|\frac{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)}\right|\\
&\qquad+\left|\frac{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)}-1\right|\Bigg\}>7n^{-1/2}\log n\Bigg|Y_n\Bigg)\\
&\le\Pi\Bigg(\frac{6}{5}\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{v}_n(s^*;\sigma^2,\alpha)}{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)}-1\right|\\
&\qquad+\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)}-1\right|>7n^{-1/2}\log n\Bigg|Y_n\Bigg)\\
&\le\Pi\left(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{v}_n(s^*;\sigma^2,\alpha)}{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)}-1\right|>n^{-1/2}\log n\middle|Y_n\right)\\
&\qquad+\Pi\left(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{v}_n(s^*;\widetilde{\theta}_{\alpha_0}/\alpha^{2\nu},\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)}-1\right|>5n^{-1/2}\log n\middle|Y_n\right)\\
&\to 0,\ \text{a.s. } P_{(\sigma_0^2,\alpha_0)}.
\end{aligned}$$

Since $\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{v}_n(s^*;\sigma^2,\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)}-1\right|=|\theta/\theta_0-1|$, this has also proved that

$$\Pi\left(\left|\frac{\theta}{\theta_0}-1\right|>7n^{-1/2}\log n\middle|Y_n\right)\to 0,\ \text{a.s. } P_{(\sigma_0^2,\alpha_0)}. \tag{S.253}$$

This completes the proof. □

S6.2 Proof of Theorems 3.2 and 3.3

Proof of Theorem 3.2. Proof of Part (i):

First, we show the existence of the sequence $\varsigma_n(\alpha)$. Since the two Gaussian measures $\mathrm{GP}(0,(\theta_0/\alpha^{2\nu})K_{\alpha,\nu})$ and $\mathrm{GP}(0,\sigma_0^2K_{\alpha_0,\nu})$ are equivalent, by Assumption (A.5), Equation (3.4) in [Stein, 1990a] implies that there exists a positive sequence $\varsigma_{1n}(\alpha)\to 0$ as $n\to\infty$, such that

$$\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left\{e_n(s^*;\alpha)^2\right\}}{\mathrm{E}_{(\theta_0/\alpha^{2\nu},\alpha)}\left\{e_n(s^*;\alpha)^2\right\}}-1\right|<\frac{1}{2}\varsigma_{1n}(\alpha).$$

Notice that for a small $\epsilon \in (0, 1/2)$, $|a/b - 1| < \epsilon$ implies that $a/b \geq 1 - \epsilon$ and hence $|b/a - 1| \leq |a/b - 1|/|a/b| \leq \epsilon/(1-\epsilon) < 2\epsilon$. Therefore, for sufficiently large n, $\varsigma_{1n}(\alpha) < 1/7$ and

$$\sup_{s^* \in \mathcal{S} \setminus \mathcal{S}_n} \left| \frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu}, \alpha)} \left\{ e_n(s^*; \alpha)^2 \right\}}{\mathrm{E}_{(\sigma_0^2, \alpha_0)} \left\{ e_n(s^*; \alpha)^2 \right\}} - 1 \right| \leq \varsigma_{1n}(\alpha). \tag{S.254}$$

Theorem 1 and Lemma 2 of [Stein, 1990b] further imply that there exists a positive sequence $\varsigma_{2n}(\alpha) \to 0$ as $n \to \infty$, such that

$$\sup_{s^* \in \mathcal{S} \setminus \mathcal{S}_n} \left| \frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu}, \alpha)} \left\{ e_n(s^*; \alpha)^2 \right\}}{\mathrm{E}_{(\sigma_0^2, \alpha_0)} \left\{ e_n(s^*; \alpha_0)^2 \right\}} - 1 \right| < \varsigma_{2n}(\alpha). \tag{S.255}$$

See our Lemma S.36 below for more details. Therefore, we can set $\varsigma_n(\alpha) = \max\{\varsigma_{1n}(\alpha), \varsigma_{2n}(\alpha)\}$ and $\varsigma_n(\alpha) \to 0$ as $n \to \infty$.

For abbreviation, let $\epsilon_{2n}(\alpha) = \max\left\{8n^{-1/2} \log n, \varsigma_n(\alpha)\right\}$. Then based on (S.254) and Theorem 3.1, we have that

$$\begin{aligned}
&\Pi\left(\sup_{s^* \in \mathcal{S} \setminus \mathcal{S}_n} \left| \frac{\mathrm{E}_{(\sigma^2, \alpha)} \left\{ e_n(s^*; \alpha)^2 \right\}}{\mathrm{E}_{(\sigma_0^2, \alpha_0)} \left\{ e_n(s^*; \alpha)^2 \right\}} - 1 \right| > 2\epsilon_{2n}(\alpha) \middle| Y_n, \alpha \right) \\
&= \Pi\left(\sup_{s^* \in \mathcal{S} \setminus \mathcal{S}_n} \left| \frac{\mathrm{v}_n(s^*; \theta/\alpha^{2\nu}, \alpha)}{\mathrm{v}_n(s^*; \theta_0/\alpha^{2\nu}, \alpha)} \cdot \frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu}, \alpha)} \left\{ e_n(s^*; \alpha)^2 \right\}}{\mathrm{E}_{(\sigma_0^2, \alpha_0)} \left\{ e_n(s^*; \alpha)^2 \right\}} - 1 \right| > 2\epsilon_{2n}(\alpha) \middle| Y_n, \alpha \right) \\
&\leq \Pi\left(\sup_{s^* \in \mathcal{S} \setminus \mathcal{S}_n} \left| \frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu}, \alpha)} \left\{ e_n(s^*; \alpha)^2 \right\}}{\mathrm{E}_{(\sigma_0^2, \alpha_0)} \left\{ e_n(s^*; \alpha)^2 \right\}} \right| \cdot \left| \frac{\mathrm{v}_n(s^*; \theta/\alpha^{2\nu}, \alpha)}{\mathrm{v}_n(s^*; \theta_0/\alpha^{2\nu}, \alpha)} - 1 \right| > \epsilon_{2n}(\alpha) \middle| Y_n, \alpha \right) \\
&\quad + \Pi\left(\sup_{s^* \in \mathcal{S} \setminus \mathcal{S}_n} \left| \frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu}, \alpha)} \left\{ e_n(s^*; \alpha)^2 \right\}}{\mathrm{E}_{(\sigma_0^2, \alpha_0)} \left\{ e_n(s^*; \alpha)^2 \right\}} - 1 \right| \geq \epsilon_{2n}(\alpha) \middle| Y_n, \alpha \right).
\end{aligned} \tag{S.256}$$

The second term on the right-hand side of (S.256) is zero, due to (S.254) and $\epsilon_{2n}(\alpha) \geq \varsigma_n(\alpha) \geq \varsigma_{1n}(\alpha)$. In the first term on the right-hand side of (S.256), using (S.254) and the fact that $\varsigma_{1n}(\alpha) < 1/7$ for sufficiently large n, we have from (S.256) that

$$\begin{aligned}
&\Pi\left(\sup_{s^* \in \mathcal{S} \setminus \mathcal{S}_n} \left| \frac{\mathrm{E}_{(\sigma^2, \alpha)} \left\{ e_n(s^*; \alpha)^2 \right\}}{\mathrm{E}_{(\sigma_0^2, \alpha_0)} \left\{ e_n(s^*; \alpha)^2 \right\}} - 1 \right| > 2\epsilon_{2n}(\alpha) \middle| Y_n, \alpha \right) \\
&\leq \Pi\left(\sup_{s^* \in \mathcal{S} \setminus \mathcal{S}_n} \left| \frac{\mathrm{v}_n(s^*; \theta/\alpha^{2\nu}, \alpha)}{\mathrm{v}_n(s^*; \theta_0/\alpha^{2\nu}, \alpha)} - 1 \right| > \frac{7}{8}\epsilon_{2n}(\alpha) \middle| Y_n, \alpha \right) \\
&\leq \Pi\left(\sup_{s^* \in \mathcal{S} \setminus \mathcal{S}_n} \left| \frac{\mathrm{v}_n(s^*; \theta/\alpha^{2\nu}, \alpha)}{\mathrm{v}_n(s^*; \theta_0/\alpha^{2\nu}, \alpha)} - 1 \right| > 7n^{-1/2} \log n \middle| Y_n, \alpha \right) \to 0, \text{ a.s. } P_{(\sigma_0^2, \alpha_0)},
\end{aligned}$$

following the result of Theorem 3.1 Part (i). This has proved the first convergence in Theorem 3.2 Part (i). The proof of the second convergence in Theorem 3.2 Part (i) is similar, by instead using (S.255) and replacing all $\mathrm{E}_{(\sigma_0^2, \alpha_0)} \left\{ e_n(s^*; \alpha)^2 \right\}$ in the display above by $\mathrm{E}_{(\sigma_0^2, \alpha_0)} \left\{ e_n(s^*; \alpha_0)^2 \right\}$.

Proof of Part (ii):

Let $\epsilon_{3n} = \max(8n^{-1/2} \log n, \varsigma_n)$. Let $\mathcal{E}_9 = \{\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]\}$. By Assumption (A.6), for all sufficiently large n, on the event $\mathcal{E}_9$,

$$\sup_{s^* \in \mathcal{S} \setminus \mathcal{S}_n} \left| \frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu}, \alpha)} \left\{ e_n(s^*; \alpha)^2 \right\}}{\mathrm{E}_{(\sigma_0^2, \alpha_0)} \left\{ e_n(s^*; \alpha)^2 \right\}} - 1 \right| \leq \varsigma_n < 1/7,$$

$$\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu},\alpha)}\left\{e_n(s^*;\alpha)^2\right\}}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left\{e_n(s^*;\alpha_0)^2\right\}}-1\right|\le\varsigma_n<1/7.$$

Therefore, we have that

$$\begin{aligned}
&\Pi\left(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{E}_{(\sigma^2,\alpha)}\left\{e_n(s^*;\alpha)^2\right\}}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left\{e_n(s^*;\alpha)^2\right\}}-1\right|>2\epsilon_{3n},\mathcal{E}_9\,\middle|Y_n\right)\\
&=\Pi\left(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{v}_n(s^*;\theta/\alpha^{2\nu},\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)}\cdot\frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu},\alpha)}\left\{e_n(s^*;\alpha)^2\right\}}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left\{e_n(s^*;\alpha)^2\right\}}-1\right|>2\epsilon_{3n},\mathcal{E}_9\,\middle|Y_n\right)\\
&\le\Pi\left(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu},\alpha)}\left\{e_n(s^*;\alpha)^2\right\}}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left\{e_n(s^*;\alpha)^2\right\}}\right|\cdot\left|\frac{\mathrm{v}_n(s^*;\theta/\alpha^{2\nu},\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)}-1\right|>\epsilon_{3n},\mathcal{E}_9\,\middle|Y_n\right)\\
&\quad+\Pi\left(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu},\alpha)}\left\{e_n(s^*;\alpha)^2\right\}}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left\{e_n(s^*;\alpha)^2\right\}}-1\right|>\epsilon_{3n},\mathcal{E}_9\,\middle|Y_n\right)\\
&\le\Pi\left(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{v}_n(s^*;\theta/\alpha^{2\nu},\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)}-1\right|>7n^{-1/2}\log n,\mathcal{E}_9\,\middle|Y_n\right)\\
&\quad+\Pi\left(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu},\alpha)}\left\{e_n(s^*;\alpha)^2\right\}}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left\{e_n(s^*;\alpha)^2\right\}}-1\right|>\varsigma_n,\mathcal{E}_9\,\middle|Y_n\right)\\
&\le\Pi\left(\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{v}_n(s^*;\theta/\alpha^{2\nu},\alpha)}{\mathrm{v}_n(s^*;\theta_0/\alpha^{2\nu},\alpha)}-1\right|>7n^{-1/2}\log n\middle|Y_n\right)\to0,\ \text{a.s. }P_{(\sigma_0^2,\alpha_0)},
\end{aligned}\tag{S.257}$$

where the last convergence follows from Theorem 3.1 Part (ii).

On the other hand, for the event $\mathcal{E}_9^c$, Theorem 2.3 implies that for the event

$$\mathcal{E}_{10}=\left\{\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{E}_{(\sigma^2,\alpha)}\left\{e_n(s^*;\alpha)^2\right\}}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left\{e_n(s^*;\alpha)^2\right\}}-1\right|>\max\left(16n^{-1/2}\log n,2\varsigma_n\right)\right\}\cap\mathcal{E}_9^c,$$

as $n\to\infty$, almost surely $P_{(\sigma_0^2,\alpha_0)}$,

$$\left|\Pi\left(\mathcal{E}_{10}\left|Y_n\right.\right)-\int_{\mathcal{E}_{10}}\frac{\sqrt{n}}{2\sqrt{\pi}\theta_0}\mathrm{e}^{-\frac{n(\theta-\widetilde{\theta}_{\alpha_0})^2}{4\theta_0^2}}\cdot\widetilde{\pi}(\alpha|Y_n)\mathrm{d}\theta\mathrm{d}\alpha\right|\to0.\tag{S.258}$$

But from (S.194) and (S.195) in the proof of Theorem 2.3, it follows that as $n\to\infty$, almost surely $P_{(\sigma_0^2,\alpha_0)}$,

$$\int_{\mathcal{E}_{10}}\frac{\sqrt{n}}{2\sqrt{\pi}\theta_0}\mathrm{e}^{-\frac{n(\theta-\widetilde{\theta}_{\alpha_0})^2}{4\theta_0^2}}\cdot\widetilde{\pi}(\alpha|Y_n)\mathrm{d}\theta\mathrm{d}\alpha\le\int_{\mathcal{E}_9^c}\widetilde{\pi}(\alpha|Y_n)\mathrm{d}\theta\mathrm{d}\alpha\to0.\tag{S.259}$$

Therefore, (S.258) and (S.259) imply that $\Pi\left(\mathcal{E}_{10}\left|Y_n\right.\right)\to0$ almost surely $P_{(\sigma_0^2,\alpha_0)}$ as $n\to\infty$. The first convergence in Theorem 3.2 Part (ii) follows by combining this with (S.257). The second convergence in Theorem 3.2 Part (ii) follows from the similar argument as above by replacing all $\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left\{e_n(s^*;\alpha)^2\right\}$ by $\mathrm{E}_{(\sigma_0^2,\alpha_0)}\left\{e_n(s^*;\alpha_0)^2\right\}$. □

Define $\mathrm{KL}(P_1,P_2)=\int\log(\mathrm{d}P_1/\mathrm{d}P_2)\mathrm{d}P_1$ to be the Kullback-Leibler divergence between two measures P_1 and P_2, where $\mathrm{d}P_1/\mathrm{d}P_2$ is the Radon-Nickdym derivative of P_1 with respect to P_2. For two mean zero Gaussian processes with Matérn covariance functions $\sigma_i^2K_{\alpha_i,\nu}$ $(i=1,2)$, let

$P^{(n)}_{(\sigma_i^2,\alpha_i)}$ be the joint Gaussian distribution of the observations $X(s_1),\dots,X(s_n)$. Then one can show that

$$\mathrm{KL}\left(P^{(n)}_{(\sigma_1^2,\alpha_1)}, P^{(n)}_{(\sigma_2^2,\alpha_2)}\right) = \frac{1}{2}\left\{\log\frac{|\sigma_2^2 R_{\alpha_2}|}{|\sigma_1^2 R_{\alpha_1}|} - n + \frac{\sigma_1^2}{\sigma_2^2}\operatorname{tr}\left(R_{\alpha_2}^{-1}R_{\alpha_1}\right)\right\}.$$

For $d \in \{1,2,3\}$, let us consider two equivalent Gaussian measures with Matérn covariance functions $\sigma_0^2 K_{\alpha_0,\nu}$ and $\sigma^2 K_{\alpha,\nu}$, such that $\sigma_0^2\alpha_0^{2\nu} = \theta_0 = \sigma^2\alpha^{2\nu}$. Let

$$\begin{aligned}\mathrm{r}_n(\alpha) &= \mathrm{KL}\left(P^{(n)}_{(\sigma_0^2,\alpha_0)}, P^{(n)}_{(\sigma^2,\alpha)}\right) + \mathrm{KL}\left(P^{(n)}_{(\sigma^2,\alpha)}, P^{(n)}_{(\sigma_0^2,\alpha_0)}\right) \\ &= -n + \frac{\alpha^{2\nu}}{2\alpha_0^{2\nu}}\operatorname{tr}\left(R_\alpha^{-1}R_{\alpha_0}\right) + \frac{\alpha_0^{2\nu}}{2\alpha^{2\nu}}\operatorname{tr}\left(R_{\alpha_0}^{-1}R_\alpha\right). \end{aligned} \tag{S.260}$$

Then due to the equivalence, for any given $\alpha > 0$, under Assumption (A.5) that $\mathcal{S}_n$ is dense in $\mathcal{S} = [0,T]^d$ as $n \to \infty$, the sequence $\{\mathrm{r}_n(\alpha)\}_{n=1}^\infty$ is increasing with n to a finite limit $\mathrm{r}(\alpha) = \lim_{n\to\infty}\mathrm{r}_n(\alpha)$ ([Ibragimov and Rozanov, 1978]), which satisfies $\mathrm{r}(\alpha) = \mathrm{KL}(P_{(\sigma_0^2,\alpha_0)}, P_{(\sigma^2,\alpha)}) + \mathrm{KL}(P_{(\sigma^2,\alpha)}, P_{(\sigma_0^2,\alpha_0)})$, where $\mathrm{KL}(P_{(\sigma_0^2,\alpha_0)}, P_{(\sigma^2,\alpha)})$ and $\mathrm{KL}(P_{(\sigma^2,\alpha)}, P_{(\sigma_0^2,\alpha_0)})$ are the limits of $\mathrm{KL}(P^{(n)}_{(\sigma_0^2,\alpha_0)}, P^{(n)}_{(\sigma^2,\alpha)})$ and $\mathrm{KL}(P^{(n)}_{(\sigma^2,\alpha)}, P^{(n)}_{(\sigma_0^2,\alpha_0)})$ as $n \to \infty$ ([Kullback et al., 1987]); see Section 3 of [Stein, 1990b].

The following lemma is a result from [Stein, 1990b].

Lemma S.36. *Suppose that $d \in \{1,2,3\}$, $\nu \in \mathbb{R}^+$, and Assumption (A.5) holds. Consider two mean zero Gaussian processes with Matérn covariance functions $\sigma_0^2 K_{\alpha_0,\nu}$ and $\sigma^2 K_{\alpha,\nu}$, where $\sigma_0^2\alpha_0^{2\nu} = \theta_0 = \sigma^2\alpha^{2\nu}$ and $\alpha > 0$ is given. If $\mathrm{r}_n(\cdot)$ is defined as in* (S.260) *and* $\mathrm{r}(\alpha) = \lim_{n\to\infty}\mathrm{r}_n(\alpha)$*, then as $n \to \infty$,*

$$\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha)^2\}}{\mathrm{E}_{(\sigma^2,\alpha)}\{e_n(s^*;\alpha)^2\}} - 1\right| \le 2\sqrt{\mathrm{r}(\alpha)-\mathrm{r}_n(\alpha)} \to 0, \tag{S.261}$$

$$\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\frac{\mathrm{E}_{(\sigma^2,\alpha)}\{e_n(s^*;\alpha)^2\}}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha_0)^2\}} - 1\right| \le 4\sqrt{\mathrm{r}(\alpha)-\mathrm{r}_n(\alpha)} \to 0. \tag{S.262}$$

Proof of Lemma S.36. Using similar notation to [Stein, 1990b], we let

$$\begin{aligned} a_n(s^*;\alpha) &= \frac{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha)^2\}}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha_0)^2\}} - 1, \quad & \widetilde{a}_n(s^*;\alpha) &= \frac{\mathrm{E}_{(\sigma^2,\alpha)}\{e_n(s^*;\alpha_0)^2\}}{\mathrm{E}_{(\sigma^2,\alpha)}\{e_n(s^*;\alpha)^2\}} - 1, \\ b_n(s^*;\alpha) &= \frac{\mathrm{E}_{(\sigma^2,\alpha)}\{e_n(s^*;\alpha_0)^2\}}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha_0)^2\}} - 1, \quad & \widetilde{b}_n(s^*;\alpha) &= \frac{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha)^2\}}{\mathrm{E}_{(\sigma^2,\alpha)}\{e_n(s^*;\alpha)^2\}} - 1. \end{aligned}$$

In [Stein, 1990b], their Theorem 1, Lemma 2 and the analysis in Section 3 imply that for every given $\alpha > 0$, as $n \to \infty$,

$$0 \le \sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left[b_n(s^*;\alpha) + \widetilde{b}_n(s^*;\alpha)\right] \le 2[\mathrm{r}(\alpha)-\mathrm{r}_n(\alpha)], \tag{S.263}$$

$$\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}|b_n(s^*;\alpha)| \le \sqrt{4[\mathrm{r}(\alpha)-\mathrm{r}_n(\alpha)]\max\{1, 4[\mathrm{r}(\alpha)-\mathrm{r}_n(\alpha)]\}} \overset{(i)}{\le} 2\sqrt{\mathrm{r}(\alpha)-\mathrm{r}_n(\alpha)}, \tag{S.264}$$

$$\text{and similarly } \sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\widetilde{b}_n(s^*;\alpha)\right| \le 2\sqrt{\mathrm{r}(\alpha)-\mathrm{r}_n(\alpha)}, \tag{S.265}$$

where (i) follows because $\mathrm{r}_n(\alpha)$ increases to $\mathrm{r}(\alpha)$ as $n \to \infty$. Therefore, (S.261) follows from (S.265) and the definition of $\widetilde{b}_n(s^*;\alpha)$.

Using the relation

$$[1 + a_n(s^*;\alpha)]\,[1 + \widetilde{a}_n(s^*;\alpha)] = [1 + b_n(s^*;\alpha)]\left[1 + \widetilde{b}_n(s^*;\alpha)\right],$$

and the fact that $\widetilde{a}_n(s^*;\alpha) \geq 0$, we can obtain that

$$
\begin{aligned}
\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} a_n(s^*;\alpha) &\leq \sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} \left[b_n(s^*;\alpha) + \widetilde{b}_n(s^*;\alpha) + b_n(s^*;\alpha)\widetilde{b}_n(s^*;\alpha)\right] \\
&\leq \sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} \left|b_n(s^*;\alpha) + \widetilde{b}_n(s^*;\alpha)\right| + \sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} \left|b_n(s^*;\alpha)\widetilde{b}_n(s^*;\alpha)\right| \\
&\overset{(i)}{\leq} 2\left[\mathrm{r}(\alpha) - \mathrm{r}_n(\alpha)\right] + 4\left[\mathrm{r}(\alpha) - \mathrm{r}_n(\alpha)\right] \\
&\leq 6\left[\mathrm{r}(\alpha) - \mathrm{r}_n(\alpha)\right], \qquad \text{(S.266)}
\end{aligned}
$$

where (i) follows from (S.263), (S.264) and (S.265).

On the other hand, by the definition of $a_n(s^*;\alpha)$ and $\widetilde{b}_n(s^*;\alpha)$, we have

$$
\begin{aligned}
&\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} \left|\frac{\mathrm{E}_{(\sigma^2,\alpha)}\{e_n(s^*;\alpha)^2\}}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha_0)^2\}} - 1\right| = \sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} \left|\frac{a_n(s^*;\alpha)+1}{1+\widetilde{b}_n(s^*;\alpha)} - 1\right| \\
&= \sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} \left|\frac{a_n(s^*;\alpha) - \widetilde{b}_n(s^*;\alpha)}{1+\widetilde{b}_n(s^*;\alpha)}\right| \overset{(i)}{\leq} \frac{3}{2}\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} a_n(s^*;\alpha) + \frac{3}{2}\sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n}\left|\widetilde{b}_n(s^*;\alpha)\right| \\
&\overset{(ii)}{\leq} 9\left[\mathrm{r}(\alpha) - \mathrm{r}_n(\alpha)\right] + 3\sqrt{\mathrm{r}(\alpha) - \mathrm{r}_n(\alpha)} \\
&\overset{(iii)}{\leq} 4\sqrt{\mathrm{r}(\alpha) - \mathrm{r}_n(\alpha)},
\end{aligned}
$$

where (i) follows from that $a_n(s^*;\alpha) \geq 0$ and for all sufficiently large n, $\mathrm{r}(\alpha) - \mathrm{r}_n(\alpha) < 1/81$ so $\left|\widetilde{b}_n(s^*;\alpha)\right| \leq 1/3$ by (S.265); (ii) follows from (S.265) and (S.266); and (iii) follows from $\sqrt{\mathrm{r}(\alpha) - \mathrm{r}_n(\alpha)} < 1/9$ as $n \to \infty$. This has proved (S.262). □

Proof of Theorem 3.3. We verify Assumption (A.6) for this special case. We can calculate that

$$
\begin{aligned}
\mathrm{r}_n(\alpha) &= \frac{\alpha}{2\alpha_0}\operatorname{tr}\left(R_\alpha^{-1}R_{\alpha_0}\right) + \frac{\alpha_0}{2\alpha}\operatorname{tr}\left(R_{\alpha_0}^{-1}R_\alpha\right) - n \\
&= \frac{\alpha}{2\alpha_0}\left[n + \frac{(n-1)\alpha}{\alpha_0}\frac{\mathrm{e}^{-\alpha/n}(\mathrm{e}^{-\alpha/n} - \mathrm{e}^{-\alpha_0/n})}{1-\mathrm{e}^{-2\alpha/n}}\right] \\
&\quad + \frac{\alpha_0}{2\alpha}\left[n + \frac{(n-1)\alpha_0}{\alpha}\frac{\mathrm{e}^{-\alpha_0/n}(\mathrm{e}^{-\alpha_0/n} - \mathrm{e}^{-\alpha/n})}{1-\mathrm{e}^{-2\alpha_0/n}}\right] - n. \qquad \text{(S.267)}
\end{aligned}
$$

The Taylor series expansion of the first term in (S.267) over all $\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]$ gives

$$
\begin{aligned}
&\frac{\alpha}{2\alpha_0}\left[n + \frac{(n-1)\alpha}{\alpha_0}\frac{\mathrm{e}^{-\alpha/n}(\mathrm{e}^{-\alpha/n} - \mathrm{e}^{-\alpha_0/n})}{1-\mathrm{e}^{-2\alpha/n}}\right] \\
&= \frac{n}{2} + \frac{(\alpha-\alpha_0)(\alpha+\alpha_0+2)}{4\alpha_0} + \frac{(\alpha_0^2-\alpha^2)(\alpha_0+3)}{12\alpha_0 n} \\
&\quad + \frac{(\alpha^2-\alpha_0^2)(\alpha_0^2+4\alpha_0-\alpha^2)}{48\alpha_0 n^2} + O\left(\frac{1}{n^{5/2}}\right). \qquad \text{(S.268)}
\end{aligned}
$$

The order of the remainder is at most $O(n^{-5/2})$ since $\overline{\alpha}_n \preceq n^{0.02}$ and $\underline{\alpha}_n \succeq n^{-0.05}$.

By symmetry, for the second term in (S.267), we have

$$
\frac{\alpha_0}{2\alpha}\left[n + \frac{(n-1)\alpha_0}{\alpha}\frac{\mathrm{e}^{-\alpha_0/n}(\mathrm{e}^{-\alpha_0/n} - \mathrm{e}^{-\alpha/n})}{1-\mathrm{e}^{-2\alpha_0/n}}\right]
$$

$$
= \frac{n}{2} + \frac{(\alpha_0 - \alpha)(\alpha + \alpha_0 + 2)}{4\alpha} + \frac{(\alpha^2 - \alpha_0^2)(\alpha + 3)}{12\alpha n}
+ \frac{(\alpha_0^2 - \alpha^2)(\alpha^2 + 4\alpha - \alpha_0^2)}{48\alpha n^2} + O\left(\frac{1}{n^{5/2}}\right). \tag{S.269}
$$

Therefore, (S.267), (S.268), and (S.269) together imply that

$$
\mathrm{r}_n(\alpha) = \frac{(\alpha - \alpha_0)^2(\alpha + \alpha_0 + 2)}{4\alpha\alpha_0} - \frac{(\alpha - \alpha_0)^2(\alpha + \alpha_0)}{4\alpha\alpha_0 n}
- \frac{(\alpha - \alpha_0)^2(\alpha + \alpha_0)^3}{48\alpha\alpha_0 n^2} + O\left(\frac{1}{n^{5/2}}\right),
$$

and

$$
\mathrm{r}(\alpha) = \lim_{n\to\infty} \mathrm{r}_n(\alpha) = \frac{(\alpha - \alpha_0)^2(\alpha + \alpha_0 + 2)}{4\alpha\alpha_0}.
$$

Therefore, uniformly over all $\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]$,

$$
\mathrm{r}(\alpha) - \mathrm{r}_n(\alpha) = \frac{(\alpha - \alpha_0)^2(\alpha + \alpha_0)}{4\alpha\alpha_0 n} + \frac{(\alpha - \alpha_0)^2(\alpha + \alpha_0)^3}{48\alpha\alpha_0 n^2} + O\left(\frac{1}{n^{5/2}}\right). \tag{S.270}
$$

By (S.261) in Lemma S.36 and the uniformity over all $\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]$, we obtain that for sufficiently large n,

$$
\begin{aligned}
&\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} \left[\frac{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha)^2\}}{\mathrm{E}_{(\theta_0/\alpha^{2\nu},\alpha)}\{e_n(s^*;\alpha)^2\}} - 1\right]^2 \\
&\leq \sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} 4[\mathrm{r}(\alpha) - \mathrm{r}_n(\alpha)] \leq \sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \frac{2(\alpha - \alpha_0)^2(\alpha + \alpha_0)}{n\alpha\alpha_0} \\
&\leq \sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \frac{2(\alpha + \alpha_0)}{n} \cdot \left(\frac{\alpha}{\alpha_0} + \frac{\alpha_0}{\alpha} - 2\right) \leq \frac{2(\overline{\alpha}_n + \alpha_0)}{n} \cdot \max\left\{\frac{(\overline{\alpha}_n - \alpha_0)^2}{\overline{\alpha}_n\alpha_0}, \frac{(\underline{\alpha}_n - \alpha_0)^2}{\underline{\alpha}_n\alpha_0}\right\} \\
&\leq \frac{3\overline{\alpha}_n \max\left(\frac{\overline{\alpha}_n}{\alpha_0}, \frac{\alpha_0}{\underline{\alpha}_n}\right)}{n} \leq 3n^{-1}\overline{\alpha}_n\left(\frac{\overline{\alpha}_n}{\alpha_0} + \frac{\alpha_0}{\underline{\alpha}_n}\right) \leq 4n^{2\overline{\kappa}+\underline{\kappa}-1}.
\end{aligned} \tag{S.271}
$$

Since $\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha)^2\} \geq \mathrm{E}_{(\theta_0/\alpha^{2\nu},\alpha)}\{e_n(s^*;\alpha)^2\}$, it follows from (S.271) that

$$
\begin{aligned}
&\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} \left|\frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu},\alpha)}\{e_n(s^*;\alpha)^2\}}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha)^2\}} - 1\right| \\
&= \sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} \frac{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha)^2\}\Big/\mathrm{E}_{(\theta_0/\alpha^{2\nu},\alpha)}\{e_n(s^*;\alpha)^2\} - 1}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha)^2\}\Big/\mathrm{E}_{(\theta_0/\alpha^{2\nu},\alpha)}\{e_n(s^*;\alpha)^2\}} \\
&\leq \sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} \left[\frac{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha)^2\}}{\mathrm{E}_{(\theta_0/\alpha^{2\nu},\alpha)}\{e_n(s^*;\alpha)^2\}} - 1\right] \\
&\leq 2n^{(\overline{\kappa}+\underline{\kappa}/2)-1/2}.
\end{aligned} \tag{S.272}
$$

From (S.262) in Lemma S.36, we obtain that for sufficiently large n,

$$
\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} \left|\frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu},\alpha)}\{e_n(s^*;\alpha)^2\}}{\mathrm{E}_{(\sigma_0^2,\alpha_0)}\{e_n(s^*;\alpha_0)^2\}} - 1\right|^2
$$

$$\le \sup_{\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]} \frac{5(\alpha-\alpha_0)^2(\alpha+\alpha_0)}{n\alpha\alpha_0} \le \sup_{\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]} \frac{5(\alpha+\alpha_0)}{n} \cdot \left(\frac{\alpha}{\alpha_0} + \frac{\alpha_0}{\alpha} - 2 \right)$$

$$\le \frac{5(\overline{\alpha}_n+\alpha_0)}{n} \cdot \max\left\{ \frac{(\overline{\alpha}_n-\alpha_0)^2}{\overline{\alpha}_n\alpha_0}, \frac{(\underline{\alpha}_n-\alpha_0)^2}{\underline{\alpha}_n\alpha_0} \right\} \le \frac{6\overline{\alpha}_n \max\left(\frac{\overline{\alpha}_n}{\alpha_0}, \frac{\alpha_0}{\underline{\alpha}_n} \right)}{n} \le 7n^{2\overline{\kappa}+\underline{\kappa}-1}.$$

Therefore, for sufficiently large n,

$$\sup_{\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]} \sup_{s^* \in \mathcal{S} \setminus \mathcal{S}_n} \left| \frac{\mathrm{E}_{(\theta_0/\alpha^{2\nu}, \alpha)}\{e_n(s^*;\alpha)^2\}}{\mathrm{E}_{(\sigma_0^2, \alpha_0)}\{e_n(s^*;\alpha_0)^2\}} - 1 \right| \le 3n^{(\overline{\kappa}+\underline{\kappa}/2)-1/2}. \tag{S.273}$$

Based on (S.272) and (S.273), we conclude that Assumption (A.6) is satisfied with $\varsigma_n = 3n^{-1/2+(\overline{\kappa}+\underline{\kappa}/2)}$. Therefore, the posterior convergence rates of asymptotic efficiency in Theorem 3.2 become $\max\left(16n^{-1/2}\log n, 2\varsigma_n\right) = 6n^{-1/2+(\overline{\kappa}+\underline{\kappa}/2)}$ as $n \to \infty$. This completes the proof of Theorem 3.3. □

S6.3 Proof of Theorem 3.4

We introduce some concepts from scattered data approximation. For technical details, we refer the readers to the book [Wendland, 2005]. For a generic kernel function $K(\cdot,\cdot)$ on $\mathcal{S}$, we define the *power function* (Chapter 11 of [Wendland, 2005]) as

$$\mathsf{P}(s; K, \mathcal{S}_n) = \left\{ K(s,s) - K(\mathcal{S}_n, s)^\top K(\mathcal{S}_n, \mathcal{S}_n)^{-1} K(\mathcal{S}_n, s) \right\}^{1/2}, \quad \text{for any } s \in \mathcal{S}, \tag{S.274}$$

where $\mathcal{S}_n = \{s_1, \ldots, s_n\}$, $K(\mathcal{S}_n, s) = (K(s_1,s), \ldots, K(s_n,s))^\top$, and $K(\mathcal{S}_n, \mathcal{S}_n)$ is the $n \times n$ covariance matrix with entries $\{K(\mathcal{S}_n, \mathcal{S}_n)\}_{ij} = K(s_i, s_j)$ for $i,j = 1, \ldots, n$. The power function $\mathsf{P}_{K,\mathcal{S}_n}(s)$ plays an important role in error estimates of kriging interpolation. We cite the following results from [Wendland, 2005]:

Lemma S.37. *([Wendland, 2005] Theorem 11.4) For any $f \in \mathcal{H}_K$, let $f_n = (f(s_1), \ldots, f(s_n))^\top$. Then*

$$\left| f(s) - f_n^\top K(\mathcal{S}_n, \mathcal{S}_n)^{-1} K(\mathcal{S}_n, s) \right| \le \|f\|_{\mathcal{H}_K} \mathsf{P}(s; K, \mathcal{S}_n), \text{ for any } s \in \mathcal{S}\,.$$

Proof of Theorem 3.4. By Assumption (A.1) and Lemma S.11, $\|\mathrm{m}_j\|_{\mathcal{H}_{\sigma^2 K_{\alpha,\nu}}}$ is finite for any $(\sigma^2, \alpha) \in \mathbb{R}^+ \times \mathbb{R}^+$ and for all $j = 1, \ldots, p$.

Define $\mathrm{m}_{j,n} = (\mathrm{m}_j(s_1), \ldots, \mathrm{m}_j(s_n))^\top$, for $j = 1, \ldots, p$. Then Lemma S.37 shows that for any $\alpha > 0$, any $s \in \mathcal{S}$ and each $j = 1, \ldots, p$,

$$\left| \mathrm{m}_j(s) - r_\alpha(s)^\top R_\alpha^{-1} \mathrm{m}_{j,n} \right| \le \|\mathrm{m}_j\|_{\mathcal{H}_{(\theta_0/\alpha^{2\nu})K_{\alpha,\nu}}} \mathsf{P}(s; (\theta_0/\alpha^{2\nu})K_{\alpha,\nu}, \mathcal{S}_n). \tag{S.275}$$

Therefore, using the definition of $b_\alpha(s)$ in (30), for any $s^* \in \mathcal{S} \setminus \mathcal{S}_n$,

$$b_\alpha(s^*)^\top b_\alpha(s^*) = \sum_{j=1}^p \left| \mathrm{m}_j(s^*) - r_\alpha(s^*)^\top R_\alpha^{-1} \mathrm{m}_{j,n} \right|^2$$

$$\le \mathsf{P}(s^*; (\theta_0/\alpha^{2\nu})K_{\alpha,\nu}, \mathcal{S}_n)^2 \sum_{j=1}^p \|\mathrm{m}_j\|^2_{\mathcal{H}_{(\theta_0/\alpha^{2\nu})K_{\alpha,\nu}}}. \tag{S.276}$$

By the inequality (B.4) and the subsequent argument in the proof of Theorem 2 in [Wang et al., 2019], for any $\alpha > 0$,

$$\lambda_{\min}\left(M_n^\top R_\alpha^{-1} M_n\right) \geq \max_{\mathcal{I}\subseteq\{1,\ldots,n\},|\mathcal{I}|=p} \lambda_{\min}\left(M_{\mathcal{I}}^\top M_{\mathcal{I}}\right)/p = \underline{\lambda}(M_n, p). \tag{S.277}$$

Therefore, using the reparameterization $\theta = \sigma^2\alpha^{2\nu}$ and the definition of $\mathrm{v}_n(s^*; \sigma^2, \alpha)$ in (30), we can combine (S.276) and (S.277) and obtain that for any $(\sigma^2, \alpha) \in \mathbb{R}^+ \times \mathbb{R}^+$,

$$\begin{aligned}
&\mathrm{v}_n(s^*; \sigma^2, \alpha) \\
&= \sigma^2\left\{1 - r_\alpha(s^*)^\top R_\alpha^{-1} r_\alpha(s^*)\right\} + \sigma^2 b_\alpha(s^*)^\top \left(M_n^\top R_\alpha^{-1} M_n + \Omega_\beta\right)^{-1} b_\alpha(s^*) \\
&\leq \frac{\theta}{\theta_0}\mathsf{P}(s^*; (\theta_0/\alpha^{2\nu})K_{\alpha,\nu}, \mathcal{S}_n)^2 + \frac{\theta}{\theta_0}\cdot\frac{\theta_0}{\alpha^{2\nu}}\cdot\lambda_{\min}\left(M_n^\top R_\alpha^{-1} M_n\right)^{-1} b_\alpha(s^*)^\top b_\alpha(s^*) \\
&\leq \frac{\theta}{\theta_0}\mathsf{P}(s^*; (\theta_0/\alpha^{2\nu})K_{\alpha,\nu}, \mathcal{S}_n)^2\left\{\underline{\lambda}(M_n, p)^{-1}\frac{\theta_0}{\alpha^{2\nu}}\sum_{j=1}^p \|\mathrm{m}_j\|^2_{\mathcal{H}_{(\theta_0/\alpha^{2\nu})K_{\alpha,\nu}}} + 1\right\}.
\end{aligned} \tag{S.278}$$

Because $\mathcal{S} = [0, T]^d$ is compact and convex with positive Lebesgue measure, and $\mathcal{S}_n$ is dense in $\mathcal{S}$ by Assumption (A.5), Theorem 5.14 of [Wu and Schaback, 1993] has shown that for a constant $C_{\mathrm{v},1} > 0$ that depends on $\sigma_0^2, \alpha_0, \nu, d, T$ and for all sufficiently large n,

$$\sup_{s^*\in\mathcal{S}} \mathsf{P}(s^*; \sigma_0^2 K_{\alpha_0,\nu}, \mathcal{S}_n) \leq C_{\mathrm{v},1} h^\nu_{\mathcal{S}_n}. \tag{S.279}$$

Therefore, the upper bound for $\mathrm{v}_n(s^*; \sigma_0^2, \alpha_0)$ in the first convergence of Theorem 3.4 in the main text immediately follows from the upper bounds in (S.278) and (S.279), by setting $\theta = \theta_0$ and $\alpha = \alpha_0$:

$$\sup_{s^*\in\mathcal{S}} \mathrm{v}_n(s^*; \sigma_0^2, \alpha_0) \leq C_{\mathrm{v},1}\left[C_{\mathrm{m}}\sigma_0^2\underline{\lambda}(M_n, p)^{-1} + 1\right] h^{2\nu}_{\mathcal{S}_n}.$$

Now we turn to the second convergence in Theorem 3.4, where (σ^2, α) is randomly drawn from the posterior distribution $\Pi(\cdot|Y_n)$. We notice that Assumption (A.6') can be equivalently written as

$$\sup_{\alpha\in[\underline{\alpha}_n,\overline{\alpha}_n]} \sup_{s^*\in\mathcal{S}\setminus\mathcal{S}_n} \left|\frac{\mathsf{P}(s^*; (\theta_0/\alpha^{2\nu})K_{\alpha,\nu}, \mathcal{S}_n)^2}{\mathsf{P}(s^*; \sigma_0^2 K_{\alpha_0,\nu}, \mathcal{S}_n)^2} - 1\right| \leq \tilde{\varsigma}_n, \tag{S.280}$$

for a deterministic sequence $\tilde{\varsigma}_n \to 0$ as $n \to \infty$. It is trivial to see that if $s^* \in \mathcal{S}_n$, then $\mathsf{P}(s^*; (\theta_0/\alpha^{2\nu})K_{\alpha,\nu}, \mathcal{S}_n) = \mathsf{P}(s^*; \sigma_0^2 K_{\alpha_0,\nu}, \mathcal{S}_n) = 0$.

We recall from the proof of Theorem 3.2 Part (ii) that $\mathcal{E}_9 = \{\alpha \in [\underline{\alpha}_n, \overline{\alpha}_n]\}$, and that Theorem 2.3 and its proof implies that $\Pi(\mathcal{E}_9^c|Y_n) \to 0$ as $n \to \infty$ almost surely $P_{(\sigma_0^2,\alpha_0)}$. This implies that given any $\eta \in (0, 1)$, any $\delta \in (0, 1)$, there exist two numbers $0 < \alpha_1 < \alpha_2 < \infty$ and a sufficiently large integer N'_{12} ($\alpha_1, \alpha_2, N'_{12}$ are dependent on η, δ), such that for all $n > N'_{12}$, $\Pr(\Pi(\alpha \in [\alpha_1, \alpha_2] \mid Y_n) \leq 1 - \delta/2) < \eta/2$.

On the other hand, in the proof of Theorem 3.1, (S.253) has shown that $\Pi(|\theta/\theta_0 - 1| > 7n^{-1/2}\log n \mid Y_n) \to 0$ as $n \to \infty$ almost surely in $P_{(\sigma_0^2,\alpha_0)}$. This implies that for a sufficiently large N'_{13}, such that for all $n > N'_{13}$, $\Pr(\Pi(|\theta/\theta_0 - 1| \leq 7n^{-1/2}\log n) \leq 1 - \delta/2) < \eta/2$. Define the event

$$\mathcal{E}_{11} = \left\{|\theta/\theta_0 - 1| \leq 7n^{-1/2}\log n, \text{ and } \alpha \in [\alpha_1, \alpha_2]\right\}. \tag{S.281}$$

Then for all $n > N'_{14} = \max(N'_{12}, N'_{13})$,

$$\Pr\left(\Pi(\mathcal{E}_{11}^c|Y_n) > \delta\right)$$

$$\leq \Pr\Big(\Pi(|\theta/\theta_0 - 1| > 7n^{-1/2}\log n, \text{ or } \alpha \notin [\alpha_1,\alpha_2] \mid Y_n) > \delta\Big)$$
$$\leq \Pr\Big(\Pi(|\theta/\theta_0 - 1| > 7n^{-1/2}\log n \mid Y_n) > \delta/2, \text{ or } \Pi(\alpha \notin [\alpha_1,\alpha_2] \mid Y_n) > \delta/2\Big)$$
$$\leq \Pr\Big(\Pi(|\theta/\theta_0 - 1| > 7n^{-1/2}\log n \mid Y_n) > \delta/2\Big) + \Pr\left(\Pi(\alpha \notin [\alpha_1,\alpha_2] \mid Y_n) > \delta/2\right)$$
$$< \frac{\eta}{2} + \frac{\eta}{2} = \eta, \tag{S.282}$$

which can be equivalently written as $\Pr\left(\Pi(\mathcal{E}_{11} \,|Y_n) > 1-\delta\right) > 1-\eta$.

Therefore, we combine (S.278), (S.279), (S.280), and the posterior convergence of θ to θ_0 above together, and obtain that on the event $\mathcal{E}_{11}$, for all $n > N_{14}'$,

$$\begin{aligned}
&\sup_{s^*\in\mathcal{S}} \mathrm{v}_n(s^*;\sigma^2,\alpha) \\
&\leq \frac{\theta}{\theta_0}\cdot\left[\sup_{s^*\in\mathcal{S}} \mathsf{P}(s^*;(\theta_0/\alpha^{2\nu})K_{\alpha,\nu},\mathcal{S}_n)^2\right]\cdot\left\{\underline{\lambda}(M_n,p)^{-1}\frac{\theta_0}{\alpha^{2\nu}}\sum_{j=1}^p \|\mathrm{m}_j\|^2_{\mathcal{H}_{(\theta_0/\alpha^{2\nu})K_{\alpha,\nu}}}+1\right\} \\
&\overset{(i)}{\leq} \left(1+7n^{-1/2}\log n\right)\cdot\left[(1+\tilde{\varsigma}_n)\sup_{s^*\in\mathcal{S}} \mathsf{P}(s^*;\sigma_0^2K_{\alpha_0,\nu},\mathcal{S}_n)^2\right] \\
&\quad\times\left[\underline{\lambda}(M_n,p)^{-1}\frac{\theta_0}{\alpha_1^{2\nu}}\max\left\{\left(\frac{\alpha_2}{\alpha_0}\right)^{2\nu+d},1\right\}\sum_{j=1}^p \|\mathrm{m}_j\|^2_{\mathcal{H}_{\sigma_0^2K_{\alpha_0,\nu}}}+1\right] \\
&\overset{(ii)}{\leq} \left(1+7n^{-1/2}\log n\right)(1+\tilde{\varsigma}_n)\,C_{\mathrm{v},1}^2 h_{\mathcal{S}_n}^{2\nu} \\
&\quad\times\left[\underline{\lambda}(M_n,p)^{-1}\frac{\theta_0}{\alpha_1^{2\nu}}\max\left\{\left(\frac{\alpha_2}{\alpha_0}\right)^{2\nu+d},1\right\}c_2(\sigma_0,\alpha_0)^2\sum_{j=1}^p \|\mathrm{m}_j\|^2_{\mathcal{W}_2^{\nu+d/2}}+1\right], \\
&\overset{(iii)}{\leq} C_{\mathrm{v},2}\left[C_{\mathrm{v},3}C_{\mathrm{m}}\underline{\lambda}(M_n,p)^{-1}+1\right]h_{\mathcal{S}_n}^{2\nu},
\end{aligned} \tag{S.283}$$

where the inequality (i) follows from $\alpha_1 \leq \alpha \leq \alpha_2$ on $\mathcal{E}_{11}$, the inequality (S.280), and the relation between the RKHS norms of $\mathcal{H}_{(\theta_0/\alpha^{2\nu})K_{\alpha,\nu}}$ and $\mathcal{H}_{\sigma_0^2K_{\alpha_0,\nu}}$ in Lemma S.13; (ii) follows from (S.279) and the equivalence between the Matérn RKHS norm and the Sobolev norm in Lemma S.11; in (iii), the constant $C_{\mathrm{v},2} \geq \sup_{n\geq 1}\left(1+7n^{-1/2}\log n\right)(1+\tilde{\varsigma}_n)\,C_{\mathrm{v},1}^2$, which depends on $\sigma_0^2,\alpha_0,\nu,d,T$, and $C_{\mathrm{v},3} = (\theta_0/\alpha_1^{2\nu})\max\left\{(\alpha_2/\alpha_0)^{2\nu+d},1\right\}c_2(\sigma_0,\alpha_0)^2$, which depends on $\eta,\delta,\sigma_0^2,\alpha_0,\nu,d,T$.

Finally, we combine (S.282) and (S.283) to conclude that for all $n > N_{14}'$,

$$\Pr\left(\Pi\left[\sup_{s^*\in\mathcal{S}} \mathrm{v}_n(s^*;\sigma^2,\alpha) \leq C_{\mathrm{v},2}\left[C_{\mathrm{v},3}C_{\mathrm{m}}\underline{\lambda}(M_n,p)^{-1}+1\right]h_{\mathcal{S}_n}^{2\nu} \;\middle|\; Y_n\right] > 1-\delta\right) > 1-\eta.$$

Setting $N_3 = N_{14}'$ completes the proof. □

S7 Additional Simulation Results for Universal Kriging Model with Regression Terms

We present additional simulation results for the universal kriging model (1) with regression terms $Y(s) = \mathrm{m}(s)^\top\beta_0 + X(s)$ for $s \in \mathcal{S}$, and $X(\cdot) \sim \mathrm{GP}(0,\sigma_0^2K_{\alpha_0,\nu})$. We consider three values of the smoothness parameter $\nu = 1/2$, $\nu = 1/4$ and $\nu = 3/2$. We still set $\mathcal{S} = [0,1]^d$ and $\mathcal{S}_n$ to be the regular grid as in Section 4 for $d = 1, 2$. For the $d = 1$ case, we let $\mathrm{m}(s) = \left(1, s, s^2, s^3\right)^\top$

for $s \in [0,1]$ and $\beta_0 = (1, 0.66, -1.5, 1)^\top$. For $d = 2$, we let $\mathrm{m}(s) = \left(1, s_1, s_2, s_1^2, s_1 s_2, s_2^2\right)^\top$ and $\beta_0 = (1, -1.5, -1.5, 2, 1, 2)^\top$ for $s = (s_1, s_2) \in [0,1]^2$. The true covariance parameters are $\alpha_0 = 1, \sigma_0^2 = 2, \theta_0 = \sigma_0^2 \alpha_0^{2\nu} = 2$ for $\nu = 1/2, 1/4, 3/2$. We impose the noninformative improper prior $\pi(\beta|\sigma^2, \alpha) \propto 1$ on β, corresponding to $\Omega_\beta = 0_{p\times p}$. The prior specification for (θ, α) and the posterior sampling and estimation procedures are all the same as in Section 4.

For $\nu = 1/2$, we report the posterior means and variances of (θ, α) from both the true posterior distribution and the limiting posterior from Theorem 2.3, as well as the W_2 distance between these two distributions in Tables S.1 and S.2. We have similar observation to Tables 1 and 2 for the model without regression terms in Section 4 of the main text. The marginal posterior of θ is close to the normal limiting distribution whose center is increasingly close to $\theta_0 = 2$ with a shrinking variance as n increases. The marginal posterior of α maintains a large posterior variance. The approximation errors from the limiting marginal posterior distributions of θ and α decrease as n increases. Figure S.1 illustrates the convergence of posterior densities for the $d = 1$ case, which shows similar convergence to that in Figure 2 in the main text.

Table S.1: Parameter estimation and Wasserstein-2 distances between the true posterior and the limiting posteriors in Theorem 2.3 for the model with $\nu = 1/2$, $d = 1$ and with regression terms $\mathrm{m}(\cdot)^\top \beta$. $\mathrm{E}(\cdot|Y_n)$, $\mathrm{Var}(\cdot|Y_n)$, $\widetilde{\mathrm{E}}(\cdot|Y_n)$, and $\widetilde{\mathrm{Var}}(\cdot|Y_n)$ are the posterior means and variances under the true posterior, the limiting posterior in Theorem 2.3, and the limiting posterior in Theorem 2.6. The true parameter values are $\theta_0 = 2$ and $\alpha_0 = 1$. All numbers are averaged over 100 macro replications. The standard errors are in the parentheses.

$d = 1$	$n = 25$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
$\mathrm{E}(\theta\|Y_n)$	2.7152 (0.0826)	2.3743 (0.0505)	2.2113 (0.0333)	2.0659 (0.0193)	2.0334 (0.0130)
$\mathrm{Var}(\theta\|Y_n)$	1.3074 (0.0800)	0.3269 (0.0143)	0.1162 (0.0036)	0.0465 (0.0009)	0.0214 (0.0003)
$\widetilde{\mathrm{E}}(\theta\|Y_n)$	1.9597 (0.0630)	2.0529 (0.0436)	2.0664 (0.0311)	1.9983 (0.0188)	2.0004 (0.0127)
$\widetilde{\mathrm{Var}}(\theta\|Y_n)$	0.3204 (0.0007)	0.1604 (0.0003)	0.0802 (0.0002)	0.0399 (0.0001)	0.0200 (0.0000)
$\mathrm{E}(\alpha\|Y_n)$	9.4697 (0.3697)	8.5853 (0.4230)	8.2324 (0.4578)	8.1489 (0.4035)	7.5458 (0.3547)
$\mathrm{Var}(\alpha\|Y_n)$	67.3011 (4.6625)	46.8428 (4.0312)	39.2354 (3.5382)	36.8218 (2.9984)	31.6578 (2.4480)
$\widetilde{\mathrm{E}}(\alpha\|Y_n)$	8.5783 (0.3278)	8.1409 (0.3872)	8.0208 (0.4350)	8.0678 (0.4043)	7.5148 (0.3569)
$\widetilde{\mathrm{Var}}(\alpha\|Y_n)$	56.5017 (4.0339)	42.9290 (3.6248)	37.2086 (3.3029)	35.9076 (2.9028)	31.5305 (2.4689)

$d = 1$	$n = 25$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
$W_2\left(\Pi(\mathrm{d}\theta\|Y_n), \mathcal{N}\left(\mathrm{d}\theta\middle\|\widetilde{\theta}_{\alpha_0}, \frac{2\theta_0^2}{n}\right)\right)$	0.9812 (0.0360)	0.3834 (0.0132)	0.1672 (0.0050)	0.0753 (0.0018)	0.0370 (0.0008)
$W_2(\Pi(\mathrm{d}\alpha\|Y_n), \widetilde{\Pi}(\mathrm{d}\alpha\|Y_n))$	1.3446 (0.0679)	0.7517 (0.0521)	0.5161 (0.0380)	0.4720 (0.0269)	0.3998 (0.0203)

Table S.2: Parameter estimation and Wasserstein-2 distances between the true posterior and the limiting posteriors in Theorem 2.3 for the model with $\nu = 1/2$, $d = 2$ and with regression terms $\mathrm{m}(\cdot)^\top \beta$. $\mathrm{E}(\cdot|Y_n)$, $\mathrm{Var}(\cdot|Y_n)$, $\widetilde{\mathrm{E}}(\cdot|Y_n)$, and $\widetilde{\mathrm{Var}}(\cdot|Y_n)$ are the posterior means and variances under the true posterior and the limiting posterior in Theorem 2.3. The true parameter values are $\theta_0 = 2$ and $\alpha_0 = 1$. All numbers are averaged over 100 macro replications. The standard errors are in the parentheses.

$d = 2$	$n = 10^2$	$n = 20^2$	$n = 30^2$
$\mathrm{E}(\theta\|Y_n)$	2.0309 (0.0307)	2.0139 (0.0146)	1.9947 (0.0096)
$\mathrm{Var}(\theta\|Y_n)$	0.0884 (0.0026)	0.0210 (0.0003)	0.0090 (0.0001)
$\widetilde{\mathrm{E}}(\theta\|Y_n)$	2.0223 (0.0320)	2.0099 (0.0146)	1.9927 (0.0097)
$\widetilde{\mathrm{Var}}(\theta\|Y_n)$	0.0800 (0.0001)	0.0200 (0.0000)	0.0089 (0.0000)
$\mathrm{E}(\alpha\|Y_n)$	1.1007 (0.0179)	1.0905 (0.0197)	1.0981 (0.0252)
$\mathrm{Var}(\alpha\|Y_n)$	1.0745 (0.0441)	1.0086 (0.0352)	1.0276 (0.0578)
$\widetilde{\mathrm{E}}(\alpha\|Y_n)$	1.1028 (0.0179)	1.0767 (0.0427)	1.0871 (0.0240)
$\widetilde{\mathrm{Var}}(\alpha\|Y_n)$	1.0952 (0.0462)	1.0019 (0.0444)	1.0192 (0.0588)
$W_2\left(\Pi(\mathrm{d}\theta\|Y_n), \mathcal{N}\left(\mathrm{d}\theta\middle\|\widetilde{\theta}_{\alpha_0}, \frac{2\theta_0^2}{n}\right)\right)$	0.0595 (0.0215)	0.0167 (0.0055)	0.0086 (0.0025)
$W_2(\Pi(\mathrm{d}\alpha\|Y_n), \widetilde{\Pi}(\mathrm{d}\alpha\|Y_n))$	0.1011 (0.0365)	0.1020 (0.0459)	0.0963 (0.0458)

Similar to Table 3 in Section 4 of the main text, we further compute the asymptotic efficiency measure for the model with regression terms, using the relative error of GP predictive variance

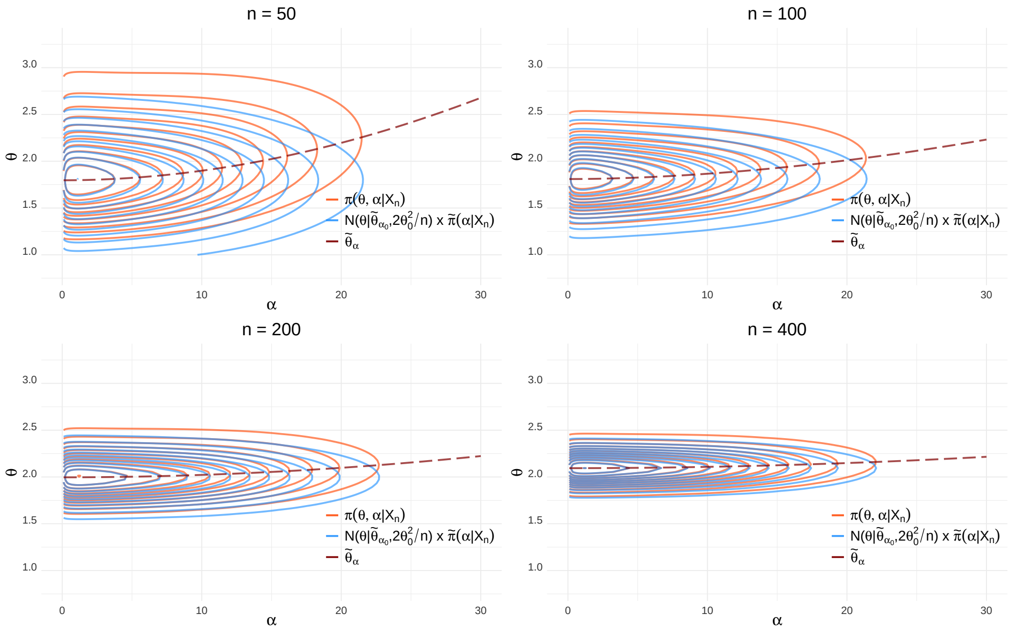


Figure S.1: Contour plots of the true joint posterior density $\pi(\theta, \alpha|Y_n)$ (in red) and the limiting posterior density $\mathcal{N}(\theta|\widetilde{\theta}_{\alpha_0}, 2\theta_0^2/n) \times \widetilde{\pi}(\alpha|Y_n)$ (in blue) in Theorem 2.3, for the 1-d Ornstein-Uhlenbeck process with sample size $n = 50, 100, 200, 400$ in the model with regression terms $\mathrm{m}(\cdot)^\top \beta$. The dashed line is the "ridge" REML $\widetilde{\theta}_\alpha$ (given in (7)), the value of θ that maximizes the joint likelihood for each given α. The true parameter values are $\theta_0 = 2$ and $\alpha_0 = 1$.

to the oracle predictive variance defined as

$$\mathsf{r}_n(s^*) = \left| \frac{\mathrm{v}_n(s^*; \sigma^2, \alpha)}{\mathrm{v}_n(s^*; \sigma_0^2, \alpha_0)} - 1 \right|, \tag{S.284}$$

over a large number of testing points s^* from the Latin hypercube design, where $\mathrm{v}_n(s^*; \sigma^2, \alpha)$ is given in (30). We again use 1000 testing points in $\mathcal{S} = [0, 1]$ for the $d = 1$ case, and 2500 testing points in $\mathcal{S} = [0, 1]^2$ for the $d = 2$ case. The posterior expectations of $\mathsf{r}_n(s^*)$ are reported in Table S.3. We can see that for both $d = 1$ and $d = 2$ cases, the GP predictive variance based on a randomly drawn (σ^2, α) from the posterior has a decreasing relative error to the oracle predictive variance as n increases.

Table S.3: The posterior means of the ratio of predictive variance defined in (37) maximized over 2500 testing points s^* for the model with $\nu = 1/2$ and with regression terms $\mathrm{m}(\cdot)^\top \beta$, averaged over 100 macro replications. The standard errors are in the parentheses.

$d = 1$	$n = 25$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
$\mathrm{E}\left[\max_{s^* \in \mathcal{S}^*} \mathsf{r}_n(s^*) \mid Y_n\right]$	0.5452 (0.0520)	0.3197 (0.0245)	0.2055 (0.0142)	0.1201 (0.0082)	0.0795 (0.0055)
$d = 2$	$n = 10^2$	$n = 20^2$	$n = 30^2$		
$\mathrm{E}\left[\max_{s^* \in \mathcal{S}^*} \mathsf{r}_n(s^*) \mid Y_n\right]$	0.1458 (0.0105)	0.0861 (0.0054)	0.0696 (0.0041)		

For $\nu = 1/4$, we summarize the estimation and prediction results of in Tables S.4, S.5, and S.6. For $\nu = 3/2$, we summarize the results in Tables S.7, S.8, and S.9. All results are averaged over 100 macro simulations. In particular, Tables S.4, S.5, S.7, and S.8 provide the estimation results for marginal posterior means, variances, and the W_2 distance to the limiting distribution for the parameters θ and α, in $d = 1$ and $d = 2$ cases. Tables S.6 and S.9 provide the prediction results for the asymptotic efficiency measure $\mathsf{r}_n(s^*)$ defined in (S.284).

Overall, the tables for $\nu = 1/4$ and $\nu = 3/2$ show similar trends as the tables for $\nu = 1/2$. The marginal posterior distribution of θ becomes concentrated around the true value $\theta_0 = 2$

as n increases in all cases, and the normal limiting distribution is accurate in approximation. The marginal posterior of α does not converge to the true value $\alpha_0 = 1$ with a non-shrinking variance. The asymptotic efficiency measure $\mathsf{r}_n(s^*)$ decreases quickly to zero as n increases for all cases except the case of $\nu = 3/2, d = 1$, where $\mathsf{r}_n(s^*)$ seems to decrease slower with n.

Table S.4: Parameter estimation and Wasserstein-2 distances between the true posterior and the limiting posteriors in Theorem 2.3 for the model with $\nu = 1/4$, $d = 1$ and with regression terms $\mathrm{m}(\cdot)^\top\beta$. $\mathrm{E}(\cdot|Y_n)$, $\mathrm{Var}(\cdot|Y_n)$, $\widetilde{\mathrm{E}}(\cdot|Y_n)$, and $\widetilde{\mathrm{Var}}(\cdot|Y_n)$ are the posterior means and variances under the true posterior and the limiting posterior in Theorem 2.3. The true parameter values are $\theta_0 = 2$ and $\alpha_0 = 1$. All numbers are averaged over 100 macro replications. The standard errors are in the parentheses.

$d = 1$	$n = 25$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
$\mathrm{E}(\theta\|Y_n)$	2.5894 (0.0795)	2.3458 (0.0493)	2.2013 (0.0332)	2.0612 (0.0193)	2.0331 (0.0129)
$\mathrm{Var}(\theta\|Y_n)$	1.0144 (0.0620)	0.3020 (0.0124)	0.1143 (0.0035)	0.0457 (0.0009)	0.0215 (0.0003)
$\widetilde{\mathrm{E}}(\theta\|Y_n)$	1.9593 (0.0630)	2.0576 (0.0437)	2.0680 (0.0312)	1.9980 (0.0187)	2.0004 (0.0127)
$\widetilde{\mathrm{Var}}(\theta\|Y_n)$	0.3204 (0.0007)	0.1604 (0.0003)	0.0802 (0.0002)	0.0399 (0.0001)	0.0200 (0.0000)
$\mathrm{E}(\alpha\|Y_n)$	10.6603 (0.2895)	10.0808 (0.4566)	9.2329 (0.4538)	9.3252 (0.4351)	8.7252 (0.3941)
$\mathrm{Var}(\alpha\|Y_n)$	96.3202 (4.4944)	78.058 (6.4254)	59.2596 (5.3221)	56.6530 (4.9028)	49.7329 (4.0853)
$\widetilde{\mathrm{E}}(\alpha\|Y_n)$	9.8786 (0.2713)	9.6640 (0.4221)	8.9950 (0.4364)	9.2310 (0.4351)	8.7241 (0.3947)
$\widetilde{\mathrm{Var}}(\alpha\|Y_n)$	83.3761 (3.9361)	71.4120 (5.7540)	56.1845 (5.0395)	55.3350 (4.7184)	49.4396 (4.0132)

$d = 1$	$n = 25$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
$W_2\left(\Pi(\mathrm{d}\theta\|Y_n), \mathcal{N}\left(\mathrm{d}\theta\middle\|\widetilde{\theta}_{\alpha_0}, \frac{2\theta_0^2}{n}\right)\right)$	0.8028 (0.0291)	0.3450 (0.0100)	0.1562 (0.0041)	0.0709 (0.0016)	0.0366 (0.0008)
$W_2(\Pi(\mathrm{d}\alpha\|Y_n), \widetilde{\Pi}(\mathrm{d}\alpha\|Y_n))$	1.4547 (0.0580)	0.9177 (0.0667)	0.6772 (0.0403)	0.5813 (0.0336)	0.5626 (0.0287)

Table S.5: Parameter estimation and Wasserstein-2 distances between the true posterior and the limiting posteriors in Theorem 2.3 for the model with $\nu = 1/4$, $d = 2$ and with regression terms $\mathrm{m}(\cdot)^\top\beta$. $\mathrm{E}(\cdot|Y_n)$, $\mathrm{Var}(\cdot|Y_n)$, $\widetilde{\mathrm{E}}(\cdot|Y_n)$, and $\widetilde{\mathrm{Var}}(\cdot|Y_n)$ are the posterior means and variances under the true posterior and the limiting posterior in Theorem 2.3. The true parameter values are $\theta_0 = 2$ and $\alpha_0 = 1$. All numbers are averaged over 100 macro replications. The standard errors are in the parentheses.

$d = 2$	$n = 10^2$	$n = 20^2$	$n = 30^2$
$\mathrm{E}(\theta\|Y_n)$	2.0277 (0.0303)	2.0138 (0.0146)	1.9951 (0.0096)
$\mathrm{Var}(\theta\|Y_n)$	0.0880 (0.0026)	0.0209 (0.0003)	0.0089 (0.0001)
$\widetilde{\mathrm{E}}(\theta\|Y_n)$	2.0228 (0.0316)	2.0104 (0.0146)	1.9928 (0.0097)
$\widetilde{\mathrm{Var}}(\theta\|Y_n)$	0.0800 (0.0001)	0.0200 (0.0000)	0.0089 (0.0000)
$\mathrm{E}(\alpha\|Y_n)$	1.1063 (0.0134)	1.1009 (0.0154)	1.1035 (0.0196)
$\mathrm{Var}(\alpha\|Y_n)$	1.1027 (0.0328)	1.0606 (0.0366)	1.0937 (0.0519)
$\widetilde{\mathrm{E}}(\alpha\|Y_n)$	1.0986 (0.0125)	1.0844 (0.0157)	1.0903 (0.0186)
$\widetilde{\mathrm{Var}}(\alpha\|Y_n)$	1.0958 (0.0319)	1.0411 (0.0392)	1.0632 (0.0511)
$W_2\left(\Pi(\mathrm{d}\theta\|Y_n), \mathcal{N}\left(\mathrm{d}\theta\middle\|\widetilde{\theta}_{\alpha_0}, \frac{2\theta_0^2}{n}\right)\right)$	0.0584 (0.0239)	0.0169 (0.0053)	0.0086 (0.0026)
$W_2(\Pi(\mathrm{d}\alpha\|Y_n), \widetilde{\Pi}(\mathrm{d}\alpha\|Y_n))$	0.1099 (0.0422)	0.1075 (0.0475)	0.1037 (0.0433)

Table S.6: The posterior means of the ratio of predictive variance defined in (37) maximized over 2500 testing points s^* for the model with $\nu = 1/4$ and with regression terms $\mathrm{m}(\cdot)^\top\beta$, averaged over 100 macro replications. The standard errors are in the parentheses.

$d = 1$	$n = 25$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
$\mathrm{E}\left[\max_{s^*\in\mathcal{S}^*} \mathsf{r}_n(s^*)\|Y_n\right]$	0.4872 (0.0434)	0.3088 (0.0232)	0.2019 (0.0141)	0.1188 (0.0082)	0.0794 (0.0055)
$d = 2$	$n = 10^2$	$n = 20^2$	$n = 30^2$		
$\mathrm{E}\left[\max_{s^*\in\mathcal{S}^*} \mathsf{r}_n(s^*)\|Y_n\right]$	0.1480 (0.0111)	0.0868 (0.0053)	0.0697 (0.0041)		

Table S.7: Parameter estimation and Wasserstein-2 distances between the true posterior and the limiting posteriors in Theorem 2.3 for the model with $\nu = 3/2$, $d = 1$ and with regression terms $\mathrm{m}(\cdot)^\top \beta$. $\mathrm{E}(\cdot|Y_n)$, $\mathrm{Var}(\cdot|Y_n)$, $\widetilde{\mathrm{E}}(\cdot|Y_n)$, and $\widetilde{\mathrm{Var}}(\cdot|Y_n)$ are the posterior means and variances under the true posterior and the limiting posterior in Theorem 2.3. The true parameter values are $\theta_0 = 2$ and $\alpha_0 = 1$. All numbers are averaged over 100 macro replications. The standard errors are in the parentheses.

$d = 1$	$n = 25$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
$\mathrm{E}(\theta\|Y_n)$	2.8495 (0.0913)	2.3841 (0.0507)	2.2177 (0.0344)	2.0674 (0.0196)	2.0305 (0.0142)
$\mathrm{Var}(\theta\|Y_n)$	1.6724 (0.1238)	0.3364 (0.0162)	0.1167 (0.0038)	0.0466 (0.0009)	0.0215 (0.0003)
$\widetilde{\mathrm{E}}(\theta\|Y_n)$	1.9658 (0.0664)	2.0504 (0.0427)	2.0693 (0.0319)	1.9972 (0.0189)	1.9983 (0.0139)
$\widetilde{\mathrm{Var}}(\theta\|Y_n)$	0.3204 (0.0007)	0.1604 (0.0003)	0.0802 (0.0002)	0.0399 (0.0001)	0.0200 (0.0000)
$\mathrm{E}(\alpha\|Y_n)$	6.0370 (0.3310)	5.0376 (0.2576)	4.8005 (0.3143)	5.0102 (0.2890)	4.2705 (0.2080)
$\mathrm{Var}(\alpha\|Y_n)$	17.9624 (1.5519)	11.4256 (0.8790)	9.3582 (0.7143)	9.6294 (0.6663)	7.9125 (0.5840)
$\widetilde{\mathrm{E}}(\alpha\|Y_n)$	5.5050 (0.2907)	4.8298 (0.2484)	4.7126 (0.3044)	4.9523 (0.2834)	4.2357 (0.2059)
$\widetilde{\mathrm{Var}}(\alpha\|Y_n)$	15.6473 (1.3884)	10.5893 (0.7763)	9.0112 (0.6664)	9.5040 (0.6710)	7.8498 (0.5704)

$d = 1$	$n = 25$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
$W_2\left(\Pi(\mathrm{d}\theta\|Y_n), \mathcal{N}\left(\mathrm{d}\theta\middle\|\widetilde{\theta}_{\alpha_0}, \frac{2\theta_0^2}{n}\right)\right)$	1.1592 (0.0540)	0.3972 (0.0155)	0.1710 (0.0063)	0.0781 (0.0021)	0.0366 (0.0009)
$W_2(\Pi(\mathrm{d}\alpha\|Y_n), \widetilde{\Pi}(\mathrm{d}\alpha\|Y_n))$	0.6743 (0.0486)	0.3400 (0.0185)	0.2608 (0.0193)	0.2630 (0.0165)	0.2205 (0.0124)

Table S.8: Parameter estimation and Wasserstein-2 distances between the true posterior and the limiting posteriors in Theorem 2.3 for the model with $\nu = 3/2$, $d = 2$ and with regression terms $\mathrm{m}(\cdot)^\top \beta$. $\mathrm{E}(\cdot|Y_n)$, $\mathrm{Var}(\cdot|Y_n)$, $\widetilde{\mathrm{E}}(\cdot|Y_n)$, and $\widetilde{\mathrm{Var}}(\cdot|Y_n)$ are the posterior means and variances under the true posterior and the limiting posterior in Theorem 2.3. The true parameter values are $\theta_0 = 2$ and $\alpha_0 = 1$. All numbers are averaged over 100 macro replications. The standard errors are in the parentheses.

$d = 2$	$n = 10^2$	$n = 20^2$	$n = 30^2$
$\mathrm{E}(\theta\|Y_n)$	2.0504 (0.0315)	2.0162 (0.0148)	1.9956 (0.0096)
$\mathrm{Var}(\theta\|Y_n)$	0.0953 (0.0029)	0.0211 (0.0003)	0.0091 (0.0001)
$\widetilde{\mathrm{E}}(\theta\|Y_n)$	2.0293 (0.0328)	2.0118 (0.0149)	1.9936 (0.0097)
$\widetilde{\mathrm{Var}}(\theta\|Y_n)$	0.0800 (0.0001)	0.0200 (0.0000)	0.0089 (0.0000)
$\mathrm{E}(\alpha\|Y_n)$	1.1005 (0.0451)	0.9758 (0.0304)	1.0077 (0.0445)
$\mathrm{Var}(\alpha\|Y_n)$	0.8706 (0.0676)	0.6304 (0.0363)	0.6375 (0.0474)
$\widetilde{\mathrm{E}}(\alpha\|Y_n)$	1.1206 (0.0481)	0.9722 (0.0306)	0.9976 (0.0434)
$\widetilde{\mathrm{Var}}(\alpha\|Y_n)$	0.9128 (0.0737)	0.6283 (0.0356)	0.6353 (0.0479)
$W_2\left(\Pi(\mathrm{d}\theta\|Y_n), \mathcal{N}\left(\mathrm{d}\theta\middle\|\widetilde{\theta}_{\alpha_0}, \frac{2\theta_0^2}{n}\right)\right)$	0.0732 (0.0327)	0.0184 (0.0066)	0.0102 (0.0046)
$W_2(\Pi(\mathrm{d}\alpha\|Y_n), \widetilde{\Pi}(\mathrm{d}\alpha\|Y_n))$	0.0801 (0.0470)	0.0672 (0.0349)	0.0651 (0.0394)

Table S.9: The posterior means of the ratio of predictive variance defined in (37) maximized over 2500 testing points s^* for the model with $\nu = 3/2$ and with regression terms $\mathrm{m}(\cdot)^\top \beta$, averaged over 100 macro replications. The standard errors are in the parentheses.

$d = 1$	$n = 25$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
$\mathrm{E}\left[\max_{s^* \in \mathcal{S}^*} \mathsf{r}_n(s^*)\|Y_n\right]$	0.8196 (0.5916)	0.4218 (0.1615)	0.3957 (0.2152)	0.3152 (0.1874)	0.2998 (0.1255)
$d = 2$	$n = 10^2$	$n = 20^2$	$n = 30^2$		
$\mathrm{E}\left[\max_{s^* \in \mathcal{S}^*} \mathsf{r}_n(s^*)\|Y_n\right]$	0.1773 (0.0122)	0.0935 (0.0052)	0.0806 (0.0083)		

References

E. Anderes. On the consistent separation of scale and variance in Gaussian random fields. *Ann. Stat.*, 38(2):870–893, 2010.

J. O. Berger, V. De Oliveria, and B. Sansó. Objective Bayesian analysis of spatially correlated data. *J. Amer. Statist. Assoc.*, 96(456):1361–1374, 2001.

L. Devroye, A. Mehrabian, and T. Reddad. The total variation distance between high-dimensional Gaussians. *arXiv preprint arXiv:1810.08693*, 2018.

S. Ghosal and A. W. van der Vaart. *Fundamentals of Nonparametric Bayesian Inference.* Cambridge University Press, 2017.

M. Gu, X. Wang, and J. O. Berger. Robust Gaussian stochastic process emulation. *Ann. Stat.*, 46(6A):3038–3066, 2018.

R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambrige University Press, 1985.

D. Hsu, S. M. Kakade, and T. Zhang. A tail inequality for quadratic forms of subgaussian random vectors. *Electron. Commun. Probab.*, 17(52):1–6, 2012.

I. A. Ibragimov and Y. A. Rozanov. *Gaussian Random Processes (translated by A. B. Aries).* Springer, New York, 1978.

S. Kullback, J. C. Keegel, and J. H. Kullback. *Topics in Statistical Information Theory. Lecture Notes in Statist.*, volume 42. Springer, New York, 1987.

B. Laurent and P. Massart. Adaptive estimation of a quadratic functional by model selection. *Ann. Stat.*, 28(5):1302–1338, 2000.

W.-L. Loh. Estimating the smoothness of a Gaussian random field from irregularly spaced data via higher-order quadratic variations. *Ann. Stat.*, 43(6):2766–2794, 2015.

W.-L. Loh, S. Sun, and J. Wen. On fixed-domain asymptotics, parameter estimation and isotropic Gaussian random fields with Matérn covariance functions. *Ann. Stat.*, 49(6):3127–3152, 2021.

C. E. Rasmussen and C. K. Williams. *Gaussian Process for Machine Learning.* MIT press, 2006.

L. Schwartz. On Bayes procedures. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 4:10 − 26, 1965.

M. L. Stein. Uniform asymptotic optimality of linear predictions of a random field using an incorrect second-order structure. *Ann. Stat.*, 18(2):850–872, 1990a.

M. L. Stein. Bounds on the efficiency of linear predictions using an incorrect covariance function. *Ann. Stat.*, 18(3):1116–1138, 1990b.

M. L. Stein. *Interpolation for Spatial Data: Some Theory for Kriging.* Springer, New York, 1999.

D. Wang and W.-L. Loh. On fixed-domain asymptotics and covariance tapering in Gaussian random field models. *Electron. J. Stat.*, 5:238–269, 2011.

W. Wang, R. Tuo, and C. F. J. Wu. On prediction properties of kriging: Uniform error bounds and robustness. *J. Amer. Statist. Assoc.*, 115(530):920–930, 2019.

H. Wendland. *Scattered Data Approximation.* Cambridge University Press, 2005.

Z. Wu and R. Schaback. Local error estimates for radial basis function interpolation of scattered data. *IMA J. Numer. Anal.*, 13(1):13–27, 1993.

H. Zhang. Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *J. Amer. Statist. Assoc.*, 99(465):250–261, 2004.