




Global-Local Shrinkage Priors for Asymptotic Point and Interval Estimation of Normal Means under Sparsity

Zikun Qin^{} and Malay Ghosh
University of Florida, Gainesville, USA

Abstract

The paper addresses asymptotic estimation of normal means under sparsity. The primary focus is estimation of multivariate normal means where we obtain exact asymptotic minimax error under global-local shrinkage prior. This extends the corresponding univariate work of Ghosh and Chakrabarti (2017). In addition, we obtain similar results for the Dirichlet-Laplace prior as considered in Bhattacharya et al. (2015). Also, following van der Pas et al. (2017), we have been able to derive credible sets for multivariate normal means under global-local priors.

AMS (2000) subject classification. 62F15.

Keywords and phrases. Exponential-inverse-gamma, Beta prime priors, Concentration inequalities, Exact rate, Minimax.

1 Introduction

Estimation of normal means under sparsity started a while ago. Its importance is felt for the analysis of high dimensional data. For example, in microarray experiments, there is a multitude of genes, but only a few have impact on a certain disease. A foundational article appears in Donoho et al. (1992), who provided an asymptotic minimax estimation rate for estimation of normal means with a large majority of zeros, but with also a few significant departures from zeros. The idea was pursued in a Bayesian framework by Castillo and van der Vaart (2012) who provided the same asymptotic minimax estimation rate under a class of priors Π_n with “exponential decay” (see (2.2) of their paper for its definition).

The present work addresses the same problem, but stems from several recent excellent articles of Bhattacharya et al. (2015), van der Pas et al. (2014), van der Pas et al. (2016) and Ghosh and Chakrabarti (2017). In particular, our paper has more direct structural connection with the work of Ghosh and Chakrabarti (2017), but extends their work in certain directions.

It may be pointed out that the priors of Bhattacharya et al. (2015) or Ghosh and Chakrabarti (2017) can be brought under the general framework of van der Pas et al. (2014), but each has its own salient features which enable one to provide a more concrete set of results. In particular, these priors, now commonly referred to as “global-local” priors following Carvalho et al. (2009, 2010), are scale mixtures of normal priors with the scale parameters involving both global and local components. The global components try to shrink the normal means towards zero, while the local parameters try to balance the same with the end of identifying and distinguishing the true signals from the noises. While the work of Ghosh and Chakrabarti (2017) considers a single global parameter and utilizes the same as a tuning parameter, Bhattacharya et al. (2015) considered essentially multiple global parameters and assigned certain priors on them. These ideas will be made more specific in the following sections.

We first find the asymptotic minimax error for estimation of multivariate normal means under sparsity in the nearly-black sense (Castillo and van der Vaart, 2012). It is the same as the asymptotic minimax error in the univariate case, which was proved by Donoho et al. (1992).

We then consider estimation of multivariate normal means under global-local priors. Like Ghosh and Chakrabarti (2017), we obtain exact asymptotic minimaxity results as well in this situation. Further, in the framework of Bhattacharya et al. (2015), we obtain asymptotic minimaxity results in the multivariate case. This is the case where we put priors to the global parameters.

The final contribution of this paper is finding credible sets for multivariate normal means following the framework of van der Pas et al. (2017), who considered the univariate case. We have considered a general class of global-local priors which includes the now famous horseshoe prior, as well as a more specific class of priors which is in the framework of Bhattacharya et al. (2015). Like van der Pas et al. (2017), we have been able to identify parameter vectors for which the posteriors give good coverage, and others for which they do not.

The outline of the remaining sections is as follows. In Section 2.1, we find the asymptotic minimax error in the multivariate setting. In Section 2.2, we consider estimation of multivariate normal means and obtain exact asymptotic minimax error. We also find out the corresponding posterior contraction rates around both the estimator and the true means. Section 3 addresses results related to credible sets of multivariate normal means. Some final remarks are made in Section 4. The proofs of some technical lemmas

are given in Appendix A. The proofs of the main theorems are given in Appendix B.

2 Point Estimation of Multivariate Normal Means

2.1 Asymptotic Minimax Error under Nearly-Black Sparsity Suppose $X_i \stackrel{\text{ind}}{\sim} N(\theta_i, 1), i = 1, \dots, n$. To estimate multiple normal means, Ghosh and Chakrabarti (2017) used a general global-local prior in which the global parameter is treated as a tuning parameter. The exact asymptotic minimaxity was established under the prior. There, the true means θ_{0i} ($1 \leq i \leq n$) are assumed to be sparse in the nearly-black sense (Castillo and van der Vaart, 2012, Donoho et al., 1992), meaning that the cardinality of the non-zero θ_{0i} 's, say q_n , is $o(n)$, as $n \rightarrow \infty$. The set of nearly-black mean vectors is denoted by $l_0[q_n] = \{\boldsymbol{\theta} \in \mathbb{R}^n : \sum_{i=1}^n \mathbb{1}(\theta_i \neq 0) \leq q_n\}$ with $q_n = o(n)$. Donoho et al. (1992) provides the asymptotic minimax error,

$$\inf_{\hat{\boldsymbol{\theta}}} \sup_{\boldsymbol{\theta}_0 \in l_0[q_n]} \sum_{i=1}^n E_{\theta_{0i}} \left(\hat{\theta}_i - \theta_{0i} \right)^2 = 2q_n \log \left(\frac{n}{q_n} \right) (1 + o(1)), \text{ as } n \rightarrow \infty. \quad (1)$$

In the multivariate situation, the true means $\boldsymbol{\theta}_{0i}$ ($1 \leq i \leq n$) being assumed to be sparse in the nearly-black sense also means that $\sum_{i=1}^n \mathbb{1}(\boldsymbol{\theta}_{0i} \neq \mathbf{0}) \leq q_n$ with $q_n = o(n)$. We denote the set of nearly-black multivariate means by $L_0[q_n] = \{\{\boldsymbol{\theta}_{0i}\}_{i=1}^n : \boldsymbol{\theta}_{0i} \in \mathbb{R}^k, i = 1, \dots, n, \sum_{i=1}^n \mathbb{1}(\boldsymbol{\theta}_{0i} \neq \mathbf{0}) \leq q_n\}$. One can prove that, in the multivariate setting, the asymptotic minimax error using the Mahalanobis distance loss is the same as the asymptotic minimax error using the squared error loss in the univariate setting. We use $\|\cdot\|_{\boldsymbol{\Sigma}}$ to denote the Mahalanobis norm, e.g., $\|\mathbf{X}_i\|_{\boldsymbol{\Sigma}}^2 = \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i$, where $\boldsymbol{\Sigma}$ is the positive definite population covariance matrix.

Theorem 1 *Suppose that $\mathbf{X}_i \sim N_k(\boldsymbol{\theta}_i, \boldsymbol{\Sigma})$, independently, for $i = 1, \dots, n$, with a positive definite covariance matrix $\boldsymbol{\Sigma}$, and that the true mean vectors $\{\boldsymbol{\theta}_i\}_{i=1}^n$ are sparse in the nearly-black sense. If we measure the error of an estimator using the Mahalanobis distance loss, then, as $n \rightarrow \infty$,*

$$\inf_{\{\hat{\boldsymbol{\theta}}_i\}} \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]} \sum_{i=1}^n E_{\boldsymbol{\theta}_{0i}} \|\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_{0i}\|_{\boldsymbol{\Sigma}}^2 = 2q_n \log \left(\frac{n}{q_n} \right) (1 + o(1)). \quad (2)$$

Remark 1 When $\boldsymbol{\Sigma}$ is positive definite, the Mahalanobis norm $\|\cdot\|_{\boldsymbol{\Sigma}}$ is equivalent to the l_2 -norm $\|\cdot\|_2$ in the sense of equivalent norms, which means that there exist two positive real constants c and C such that $0 < c \leq C$, for

each $\mathbf{X} \in \mathbb{R}^k$, $c\|\mathbf{X}\|_2 \leq \|\mathbf{X}\|_{\Sigma} \leq C\|\mathbf{X}\|_2$. Specifically, $c = \lambda_{\max}^{-1}(\Sigma)$, inverse of the largest eigenvalue of Σ , and $C = \lambda_{\min}^{-1}(\Sigma)$, inverse of the smallest eigenvalue of Σ . So, Theorem 1 will not give us an exact asymptotic minimax error under the l_2 -norm unless Σ satisfies certain eigenvalue conditions. Instead, we can get both lower and upper bounds of the minimax error under the l_2 -norm. Since both bounds are of the same rate, $2q_n \log(n/q_n)(1+o(1))$, the minimax error under the l_2 -norm must be of the same rate as well, and will only differ from it by up to a constant factor.

2.2 Minimax Estimation of Multivariate Normal Means Now, we first extend the results of Ghosh and Chakrabarti (2017) to the multivariate case. We begin with a general global-local prior model

- (i) $\mathbf{X}_i | \theta_i \stackrel{ind}{\sim} N_k(\theta_i, \Sigma)$, $i = 1, \dots, n$, Σ is known positive definite;
- (ii) $\theta_i | \lambda_i^2 \stackrel{ind}{\sim} N_k(\mathbf{0}, \lambda_i^2 \tau_n \Sigma)$, $i = 1, \dots, n$, where $\tau_n \in (0, 1)$ is a sequence of positive constants to be chosen later, $\tau_n \rightarrow 0$ as $n \rightarrow \infty$;
- (iii) $\pi(\lambda_i^2) = K(\lambda_i^2)^{-a-1} L(\lambda_i^2)$, $i = 1, \dots, n$, where $a > 0$ and L is a slowly varying function.

In this model, the global parameter τ_n is assumed to be a tuning parameter. Note that the horseshoe prior (Carvalho et al., 2009, 2010) is a special case of this prior in the univariate setup with $a = 1/2$.

The following regularity assumptions are made:

- (I) L is non-decreasing in its argument with $0 < m \leq L(u) \leq M < \infty$;
- (II) $0 < \lambda_{\min}(\Sigma) \leq \lambda_{\max}(\Sigma) < \infty$, where $\lambda_{\min}(\Sigma)$ and $\lambda_{\max}(\Sigma)$ denote the minimum and maximum eigenvalues of Σ .

We estimate θ_i using the posterior means under the global-local prior, i.e.,

$$\widehat{\theta}_i = E(\theta_i | \mathbf{X}_i) = E(1 - \kappa_i | \mathbf{X}_i) \mathbf{X}_i, \text{ where } \kappa_i = (1 + \lambda_i^2 \tau_n)^{-1}, \quad (3)$$

and κ_i is the shrinkage factor. The estimators using a prior (iii) are denoted by $\widehat{\theta}_i^R$ specifically.

We prove the following theorem under this model, in which (4) and (5) concern the error contributed by the zero and non-zero true means, respectively, and (6) is then immediate following the previous two results. In particular, when $0 < a \leq 1$, we already have an upper bound for the error.

Theorem 1 provides the minimax lower bound, which matches the upper bound here. This fact actually finishes both the proofs of Theorem 1 and (7). As shown in Theorem 2, this general class of global-local priors attains the asymptotic minimax rate in the multivariate setting, and when $0 < a \leq 1$, it attains the exact asymptotic minimax error.

Theorem 2 *Assume that the true means are sparse in the nearly-black sense. Under the regularity assumptions (I) and (II), using the global-local prior with a tuning parameter, i.e., a model satisfying (i), (ii) and (iii), if $\tau_n = (q_n/n)^{\frac{1+\epsilon}{n}}$, where $\epsilon > 0$ and $0 < \eta < \min(1, a)$, then, for any valid choice of ϵ and η ,*

$$\lim_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n], i: \boldsymbol{\theta}_{0i} = \mathbf{0}} E_0 \|\widehat{\boldsymbol{\theta}}_i^R\|_{\Sigma}^2 / \left(q_n \log \left(\frac{n}{q_n} \right) \right) = 0. \quad (4)$$

and

$$\limsup_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]} \frac{\sum_{i: \boldsymbol{\theta}_{0i} \neq \mathbf{0}} E_{\boldsymbol{\theta}_{0i}} \|\widehat{\boldsymbol{\theta}}_i^R - \boldsymbol{\theta}_{0i}\|_{\Sigma}^2}{2q_n \log(n/q_n)} \leq a / \min(1, a). \quad (5)$$

Consequently,

$$\limsup_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]} \frac{\sum_{i=1}^n E_{\boldsymbol{\theta}_{0i}} \|\widehat{\boldsymbol{\theta}}_i^R - \boldsymbol{\theta}_{0i}\|_{\Sigma}^2}{2q_n \log(n/q_n)} \leq a / \min(1, a). \quad (6)$$

In particular, since the minimax error (2) provides a lower bound, when $0 < a \leq 1$, one gets the result

$$\lim_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]} \frac{\sum_{i=1}^n E_{\boldsymbol{\theta}_{0i}} \|\widehat{\boldsymbol{\theta}}_i^R - \boldsymbol{\theta}_{0i}\|_{\Sigma}^2}{2q_n \log(n/q_n)} = 1. \quad (7)$$

The following theorem provides results on the rates of posterior contraction for this prior around both the Bayes estimators and the true means. By (8), the posterior distributions contracts around the Bayes estimator at least as fast as at the minimax rate. However, by (9), the rate of posterior contraction around the true means would be slower than the minimax rate.

Theorem 3 *Under the assumptions of Theorem 2, we have*

$$\lim_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]} E_{\{\boldsymbol{\theta}_{0i}\}} \Pi \left(\sum_{i=1}^n \|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i^R\|^2 > q_n \log \left(\frac{n}{q_n} \right) \mid \{\mathbf{X}_i\} \right) = 0, \quad (8)$$

and

$$\limsup_{n \rightarrow \infty} \sup_{\{\theta_{0i}\} \in L_0[q_n]} E_{\{\theta_{0i}\}} \Pi \left(\sum_{i=1}^n \|\theta_i - \theta_{0i}\|^2 > M_n q_n \log\left(\frac{n}{q_n}\right) \mid \{\mathbf{X}_i\} \right) = 0, \quad (9)$$

for any $\{M_n\}$ such that $\lim_{n \rightarrow \infty} M_n = \infty$.

Next we extend the work of Bhattacharya et al. (2015) in the present multivariate framework. We consider the following prior in which, while (i) remains the same in our earlier formulation, we replace (ii) and (iii) respectively by

$$(ii') \quad \theta_i \mid \lambda_i^2, \tau_i \stackrel{ind}{\sim} N_k(\mathbf{0}, \lambda_i^2 \tau_i \boldsymbol{\Sigma}), \quad i = 1, \dots, n;$$

$$(iii') \quad \lambda_i^2 \text{ and } \tau_i \text{ are mutually independent. Also, } \lambda_i^2 \text{'s are independent with } \pi(\lambda_i^2) \propto \exp(-\lambda_i^2/2), \quad i = 1, \dots, n, \text{ while } \tau_i \text{'s are also independent with } \pi(\tau_i) \propto \exp(-c_n/(2\tau_i))\tau_i^{-d-1}, \text{ where } c_n \rightarrow 0 \text{ and will be chosen later and } 0 < d < 1.$$

As noted by Bhattacharya et al. (2015) as well, the Dirichlet-Laplace priors can be rewritten in the above formulation, except for that the authors put a Gamma prior on τ_i while we put an Inverse-Gamma prior on it. Due to this discrepancy, we refer the prior defined by (ii') and (iii') as the Exponential-Inverse-Gamma prior.

It is worth mentioning that writing $u_i = \lambda_i^2 \tau_i$, one gets $\pi(u) \propto (u_i + c_n)^{-d-1}$, and one can directly use the u_i for inferential purposes. Further, this particular formulation is a special case of van der Pas et al. (2016), who has a very general result concerning asymptotic minimaxity of univariate normal means. However, it seems more convenient to work with separate priors for λ_i^2 and τ_i , and the explicit nature of these priors makes the calculation smooth. As an aside, u_i/c_n has a beta prime prior with $a = 1$ and $b = d$, and this is the prior considered in Armagan et al. (2011) and Griffin and Brown (2017).

With the above formulation, the estimators of the θ_i are denoted as $\widehat{\theta}_i^{EIG}$. We will prove Theorems 4 in this setup. It shows that the prior attains the exact asymptotic minimax error as well.

Theorem 4 *Assume that the true means are sparse in the nearly-black sense. Under the regularity assumption (II), using the Exponential-Inverse-*

Gamma prior, i.e., a model satisfying (i), (ii') and (iii') above, if $c_n = (q_n/n)^{\frac{1+\epsilon}{d}}$, where $\epsilon > 0$, then, for any valid choice of ϵ ,

$$\lim_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]_{i:\boldsymbol{\theta}_{0i}=\mathbf{0}}} \sum E_0 \|\widehat{\boldsymbol{\theta}}_i^{EIG}\|^2 / \left(q_n \log \left(\frac{n}{q_n} \right) \right) = 0. \quad (10)$$

and

$$\limsup_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]} \frac{\sum_{i:\boldsymbol{\theta}_{0i} \neq \mathbf{0}} E_{\boldsymbol{\theta}_{0i}} \|\widehat{\boldsymbol{\theta}}_i^{EIG} - \boldsymbol{\theta}_{0i}\|^2}{2q_n \log(n/q_n)} \leq 1. \quad (11)$$

Consequently,

$$\lim_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]} \frac{\sum_{i=1}^n E_{\boldsymbol{\theta}_{0i}} \|\widehat{\boldsymbol{\theta}}_i^{EIG} - \boldsymbol{\theta}_{0i}\|^2}{2q_n \log(n/q_n)} = 1. \quad (12)$$

We also have the following results regarding the posterior contraction rate around the Bayes estimator and the true means. The same contraction rates are observed as using the tuning parameter model.

Theorem 5 *Under the assumptions of Theorem 4, we have*

$$\lim_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in l_0[q_n]} E_{\{\boldsymbol{\theta}_{0i}\}} \Pi \left(\sum_{i=1}^n \|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i^{EIG}\|^2 > q_n \log \left(\frac{n}{q_n} \right) \mid \{\mathbf{X}_i\} \right) = 0, \quad (13)$$

and

$$\limsup_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in l_0[q_n]} E_{\{\boldsymbol{\theta}_{0i}\}} \Pi \left(\sum_{i=1}^n \|\boldsymbol{\theta}_i - \boldsymbol{\theta}_{0i}\|^2 > M_n q_n \log \left(\frac{n}{q_n} \right) \mid \{\mathbf{X}_i\} \right) = 0, \quad (14)$$

for any $\{M_n\}$ such that $\lim_{n \rightarrow \infty} M_n = \infty$.

3 Credible Sets of Multivariate Normal Means

In this section, we first study coverage probabilities of credible sets constructed under global-local priors defined by (ii) and (iii). The global parameter is treated as a tuning parameter. We consider credible sets of the form:

$$\widehat{C}_{n,i}^R = \{\boldsymbol{\theta}_i : \|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i^R\|_{\Sigma}^2 \leq L \widehat{r}_{n,i}^{\alpha/(1+\rho)}(\alpha, \tau_n)\}, \quad (15)$$

for some $\rho(> 0)$ to be chosen later and $\widehat{r}_{n,i}(\alpha, \tau_n)$ is determined from

$$\Pi(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i^R\|_{\Sigma}^2 \leq \widehat{r}_{n,i}(\alpha, \tau_n) \mid \mathbf{X}_i) = 1 - \alpha.$$

In the following, we will omit the subscript n in $\widehat{C}_{n,i}^R$ and $\widehat{r}_{n,i}$ for notational simplicity.

Following van der Pas et al. (2017), we view the true mean vectors as in three categories:

$$\begin{aligned} \mathcal{S} &:= \{\boldsymbol{\theta}_{0i} : \|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \leq K_S \tau_n\} \\ \mathcal{M} &:= \{\boldsymbol{\theta}_{0i} : f_{\tau_n} \tau_n \leq \|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \leq K_M \log \frac{1}{\tau_n}\} \\ \mathcal{L} &:= \{\boldsymbol{\theta}_{0i} : \|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \geq K_L \log \frac{1}{\tau_n}\} \end{aligned}$$

for some positive constants K_S , K_M and K_L , and some f_{τ_n} that goes to infinity as τ_n goes to zero. We will show that, the proposed credible sets will cover the true means in either \mathcal{S} or \mathcal{L} with a desired probability, while the true means in \mathcal{M} will not be covered with probability tending to one. The results are summarized in the following theorem.

Theorem 6 *Consider the global-local prior with a tuning parameter τ_n , i.e., a model satisfying (i), (ii) and (iii), with $a < 1$, under the regularity assumptions (I) and (II). Suppose that $K_S > 0$, $K_M < 2a$ and $K_L > 2a$, and that $f_{\tau_n} \rightarrow \infty$ and $f_{\tau_n} \tau_n \rightarrow 0$ as $\tau_n \rightarrow 0$. Then, given α , for the credible sets of form (15) with $L > \chi_{k,\alpha}^2 (\chi_{k,\beta}^2)^{-a/(1+\rho)}$ for some fixed $\beta > \alpha$ and $\rho > 0$,*

$$P_{\boldsymbol{\theta}_{0i}}(\boldsymbol{\theta}_{0i} \in \widehat{C}_i^R) \geq 1 - \alpha, \text{ if } \boldsymbol{\theta}_{0i} \in \mathcal{S}, \quad (16)$$

$$P_{\boldsymbol{\theta}_{0i}}(\boldsymbol{\theta}_{0i} \notin \widehat{C}_i^R) \rightarrow 1, \text{ if } \boldsymbol{\theta}_{0i} \in \mathcal{M}, \quad (17)$$

$$P_{\boldsymbol{\theta}_{0i}}(\boldsymbol{\theta}_{0i} \in \widehat{C}_i^R) \geq 1 - \alpha, \text{ if } \boldsymbol{\theta}_{0i} \in \mathcal{L}. \quad (18)$$

as $\tau_n \rightarrow 0$.

Remark 2 From the proof of the theorem, the conclusions for $\boldsymbol{\theta}_{0i}$ in either \mathcal{S} or \mathcal{M} do not rely on any specific choice of L , while only the conclusion for \mathcal{L} has the requirement on L . To make the credible sets as narrow as possible, noticing that $L > \chi_{k,\alpha}^2 (\chi_{k,\beta}^2)^{-a/(1+\rho)}$, we should choose β to be as close to α as possible. As for ρ , noticing that

$$L \widehat{r}_i^{a/(1+\rho)}(\alpha, \tau_n) > \chi_{k,\alpha}^2 (\widehat{r}_i(\alpha, \tau_n) / \chi_{k,\beta}^2)^{a/(1+\rho)},$$

the choice should depend on $\widehat{r}_i(\alpha, \tau_n)/\chi_{k,\beta}^2$. For instance, when $\widehat{r}_i(\alpha, \tau_n)/\chi_{k,\beta}^2 > 1$, we can choose as large ρ as possible, so that $L\widehat{r}_i^{\alpha/(1+\rho)}(\alpha, \tau_n)$ would essentially become $\chi_{k,\alpha}^2$. On the other hand, when $\widehat{r}_i(\alpha, \tau_n)/\chi_{k,\beta}^2 < 1$, it would be more preferable to choose ρ closer to 0. This observation also motivates an individualized choice of L_i instead of a common L among all the subjects, so that each credible set can be narrowed as much as possible while maintaining the theoretical coverage probability.

Assuming sparsity in the nearly-black sense, most true means would be in the set \mathcal{S} . This fact immediately leads to a high overall coverage probability, i.e., the following corollary.

Corollary 1 *Under the setup of Theorem 6, further assume that the true means $\boldsymbol{\theta}_{0i}$ are sparse in the nearly-black sense. Then, for almost all $i = 1, \dots, n$, as $\tau_n \rightarrow 0$,*

$$P_{\boldsymbol{\theta}_{0i}}(\boldsymbol{\theta}_{0i} \in \widehat{C}_i^R) \geq 1 - \alpha.$$

Next, we study coverage probabilities of credible sets constructed under the Exponential-Inverse-Gamma priors defined by (ii') and (iii'). We consider credible sets of the same form as in the previous setup:

$$\widehat{C}_i^{EIG} = \{\boldsymbol{\theta}_i : \|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i^{EIG}\|_{\Sigma}^2 \leq L\widehat{r}_i^{d/(1+\rho)}(\alpha, c_n)\}, \quad (19)$$

for some $\rho(> 0)$ to be chosen later and $\widehat{r}_i(\alpha, c_n)$ is determined from

$$\Pi(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i^{EIG}\|_{\Sigma}^2 \leq \widehat{r}_i(\alpha, c_n) \mid \mathbf{X}_i) = 1 - \alpha.$$

Here, we divide the true mean vectors as in the following three categories:

$$\begin{aligned} \mathcal{S}' &:= \{\boldsymbol{\theta}_{0i} : \|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \leq K'_S c_n\} \\ \mathcal{M}' &:= \{\boldsymbol{\theta}_{0i} : f'_{c_n} c_n \leq \|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \leq K'_M \log \frac{1}{c_n}\} \\ \mathcal{L}' &:= \{\boldsymbol{\theta}_{0i} : \|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \geq K'_L \log \frac{1}{c_n}\} \end{aligned}$$

for some positive constants K'_S , K'_M and K'_L , and some f'_{c_n} that goes to infinity as c_n goes to zero. And similar results regarding the coverage probabilities are observed under this prior.

Theorem 7 *Consider the Exponential-Inverse-Gamma prior, i.e., a model satisfying (i), (ii') and (iii'), under the regularity assumption (II). Suppose that $K'_S > 0$, $K'_M < 2d$ and $K'_L > 2d$, and that $f'_{c_n} \rightarrow \infty$ and $f'_{c_n} c_n \rightarrow 0$*

as $c_n \rightarrow 0$. Then, given α , for the credible sets of form (19) with $L > \chi_{k,\alpha}^2 (\chi_{k,\beta}^2)^{-d/(1+\rho)}$ for some fixed $\beta > \alpha$ and $\rho > 0$, for $\alpha < 1/2$

$$P_{\theta_{0i}}(\theta_{0i} \in \widehat{C}_i^{EIG}) \geq 1 - \alpha, \text{ if } \theta_{0i} \in \mathcal{S}', \quad (20)$$

$$P_{\theta_{0i}}(\theta_{0i} \notin \widehat{C}_i^{EIG}) \rightarrow 1, \text{ if } \theta_{0i} \in \mathcal{M}', \quad (21)$$

$$P_{\theta_{0i}}(\theta_{0i} \in \widehat{C}_i^{EIG}) \geq 1 - \alpha, \text{ if } \theta_{0i} \in \mathcal{L}'. \quad (22)$$

as $\tau_n \rightarrow 0$.

The following corollary is also immediate due to the nearly-black sparsity.

Corollary 2 *Under the setup of Theorem 7, further assume that the true means θ_{0i} are sparse in the nearly-black sense. Then, for almost all $i = 1, \dots, n$, as $c_n \rightarrow 0$,*

$$P_{\theta_{0i}}(\theta_{0i} \in \widehat{C}_i^{EIG}) \geq 1 - \alpha.$$

4 Final Remarks

The paper addresses asymptotic estimation of multivariate normal means under global-local priors. We first find the asymptotic minimax error in the multivariate setup. Then, the asymptotic minimax error is obtained by treating the global parameter as a tuning parameter. The same result is obtained under Dirichlet-Laplace priors. Also, credible sets are obtained under global-local priors extending the idea of van der Pas et al. (2017) in the multivariate case.

Acknowledgements. We thank the anonymous reviewers for their careful reading of our manuscript and their many insightful comments and suggestions.

Funding The authors did not receive support from any organization for the submitted work.

Declarations

Competing Interests The authors have no competing interests to declare that are relevant to the content of this article.

References

- Armagan, A., Dunson, D.B., and Clyde, M. (2011). Generalized beta mixtures of gaussians. *Proceedings of the 24th international conference on neural information processing systems* (p. 523–531). Red Hook, NY, USA: Curran Associates Inc.
- Bhattacharya, A., Pati, D., Pillai, N.S., Dunson, D.B. (2015). Dirichlet-Laplace priors for optimal shrinkage. *J. Amer. Statist. Assoc.*, *110*(512), 1479–1490. Retrieved from <https://doi.org/10.1080/01621459.2014.960967>
- Carvalho, C.M., Polson, N.G., Scott, J.G. (2009). Handling sparsity via the horseshoe. D. van Dyk & M. Welling (Eds.), *Proceedings of the twelfth international conference on artificial intelligence and statistics* (Vol. 5, pp. 73–80). Hilton Clearwater Beach Resort, Clearwater Beach, Florida USA: PMLR. Retrieved from <https://proceedings.mlr.press/v5/carvalho09a.html>
- Carvalho, C.M., Polson, N.G., Scott, J.G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, *97*(2), 465–480. Retrieved from <https://doi.org/10.1093/biomet/asq017>
- Castillo, I. van der Vaart, A. (2012). Needles and straw in a haystack: posterior concentration for possibly sparse sequences. *Ann. Statist.*, *40*(4), 2069–2101. Retrieved from <https://doi.org/10.1214/12-AOS1029>
- Donoho, D.L., Johnstone, I.M., Hoch, J.C., Stern, A.S. (1992). Maximum entropy and the nearly black object. *J. Roy. Statist. Soc. Ser. B*, *54*(1), 41–81. Retrieved from [http://links.jstor.org/sici?sici=0035-9246\(1992\)54:1<41:MEATNB>2.0.CO;2-I&origin=MSN](http://links.jstor.org/sici?sici=0035-9246(1992)54:1<41:MEATNB>2.0.CO;2-I&origin=MSN) (With discussion and a reply by the authors)
- Ghosh, P., & Chakrabarti, A. (2017). Asymptotic optimality of one-group shrinkage priors in sparse high-dimensional problems. *Bayesian Anal.*, *12*(4), 1133–1161. Retrieved from <https://doi.org/10.1214/16-BA1029>
- Griffin, J., & Brown, P. (2017). Hierarchical shrinkage priors for regression models. *Bayesian Analysis*, *12*(1), 135–159.
- van der Pas, S.L., Kleijn, B.J.K., van der Vaart, A.W. (2014). The horseshoe estimator: posterior concentration around nearly black vectors. *Electron. J. Stat.*, *8*(2), 2585–2618. Retrieved from <https://doi.org/10.1214/14-EJS962>
- van der Pas, S.L., Salomond, J.-B., Schmidt-Hieber, J. (2016). Conditions for posterior contraction in the sparse normal means problem. *Electron. J. Stat.*, *10*(1), 976–1000. Retrieved from <https://doi.org/10.1214/16-EJS1130>
- van der Pas, S.L., Szabó, B., van der Vaart, A. (2017). Uncertainty quantification for the horseshoe (with discussion). *Bayesian Anal.*, *12*(4), 1221–1274. Retrieved from <https://doi.org/10.1214/17-BA1065> (With a rejoinder by the authors)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Appendix A: Lemmas

A.1 Lemmas for the Multivariate Tuning Parameter Model Regarding the multivariate tuning parameter model, we can establish the following lemmas. Under the model (i) - (iii), the posterior density of κ_i is given by

$$\Pi(\kappa_i | \mathbf{X}_i) \propto \kappa_i^{k/2+a-1} (1 - \kappa_i)^{-a-1} L\left(\frac{1 - \kappa_i}{\kappa_i \tau_n}\right) \exp\left(-\frac{\kappa_i}{2} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right).$$

In the following, we will use $K(> 0)$ to denote a generic constant.

Lemma 1 *Under the multivariate global-local prior model with treating the global parameter as a tuning parameter and under the regularity assumption (I), assuming $\tau_n \rightarrow 0$ as $n \rightarrow \infty$, for arbitrary η such that $0 < \eta < \min(1, a)$, when n is sufficiently large,*

$$E(1 - \kappa_i | \mathbf{X}_i) \leq K \tau_n^\eta \exp\left(\frac{\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right). \quad (\text{A1})$$

Proof For an arbitrary constant $\xi > 0$,

$$\begin{aligned} & E(1 - \kappa_i | \mathbf{X}_i) \\ &= \tau_n E\left[\lambda_i^2 (1 + \lambda_i^2 \tau_n)^{-1} | \mathbf{X}_i\right] \\ &= \tau_n \frac{\int_0^\infty (1 + \lambda_i^2 \tau_n)^{-\left(\frac{k}{2}+1\right)} (\lambda_i^2)^{-a} L(\lambda_i^2) \exp\left(-\frac{\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2(1+\lambda_i^2 \tau_n)}\right) d\lambda_i^2}{\int_0^\infty (1 + \lambda_i^2 \tau_n)^{-\frac{k}{2}} (\lambda_i^2)^{-a-1} L(\lambda_i^2) \exp\left(-\frac{\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2(1+\lambda_i^2 \tau_n)}\right) d\lambda_i^2} \\ &\leq \tau_n \exp\left(\frac{\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) I_1 / (I_2 L(\xi)) \end{aligned} \quad (\text{A2})$$

where

$$I_1 = \int_0^\infty (1 + \lambda_i^2 \tau_n)^{-\left(\frac{k}{2}+1\right)} (\lambda_i^2)^{-a} L(\lambda_i^2) d\lambda_i^2 \quad (\text{A3})$$

and

$$I_2 = \int_\xi^\infty (1 + \lambda_i^2 \tau_n)^{-\frac{k}{2}} (\lambda_i^2)^{-a-1} d\lambda_i^2 \quad (\text{A4})$$

But

$$I_2 \geq \int_\xi^\infty (1 + \lambda_i^2)^{-\frac{k}{2}} (\lambda_i^2)^{-a-1} d\lambda_i^2 = K. \quad (\text{A5})$$

Next, by choosing $0 < \eta < \min(1, a)$,

$$\begin{aligned}
 I_1 &\leq \int_0^1 (\lambda_i^2)(\lambda_i^2)^{-a-1} L(\lambda_i^2) d\lambda_i^2 + M \int_1^\infty (\lambda_i^2 \tau_n)^{-1+\eta} (1 + \lambda_i^2 \tau_n)^{-\frac{k}{2}-\eta} (\lambda_i^2)^{-a} d\lambda_i^2 \\
 &\leq K^{-1} + M \tau_n^{\eta-1} (a - \eta)^{-1}
 \end{aligned} \tag{A6}$$

Combining (A2)-(A6), $E(1 - \kappa_i \mid \mathbf{X}_i) \leq K \tau_n^\eta \exp\left(\frac{\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right)$, when n is sufficiently large. \square

Lemma 2 *Under the multivariate global-local prior model with treating the global parameter as a tuning parameter and under the regularity assumption (I), for arbitrary constants $0 < \xi < 1$ and $0 < \delta < 1$,*

$$E(\kappa_i \mathbb{1}_{[\kappa_i > \xi]} \mid \mathbf{X}_i) \leq K \tau_n^{-a} \exp\left[-\frac{\xi(1-\delta)}{2} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right]. \tag{A7}$$

Proof For an arbitrary constant $0 < \xi < 1$,

$$\begin{aligned}
 &E(\kappa_i \mathbb{1}_{[\kappa_i > \xi]} \mid \mathbf{X}_i) \\
 &= \frac{\int_\xi^1 \kappa_i^{k/2+a} (1 - \kappa_i)^{-a-1} L\left(\frac{1-\kappa_i}{\kappa_i \tau_n}\right) \exp\left(-\frac{\kappa_i}{2} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right) d\kappa_i}{\int_0^1 \kappa_i^{k/2+a-1} (1 - \kappa_i)^{-a-1} L\left(\frac{1-\kappa_i}{\kappa_i \tau_n}\right) \exp\left(-\frac{\kappa_i}{2} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right) d\kappa_i} \\
 &\leq \frac{\int_\xi^1 \kappa_i^{k/2+a} (1 - \kappa_i)^{-a-1} L\left(\frac{1-\kappa_i}{\kappa_i \tau_n}\right) \exp\left(-\frac{\kappa_i}{2} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right) d\kappa_i}{\int_0^{\xi\delta} \kappa_i^{k/2+a-1} (1 - \kappa_i)^{-a-1} L\left(\frac{1-\kappa_i}{\kappa_i \tau_n}\right) \exp\left(-\frac{\kappa_i}{2} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right) d\kappa_i} \\
 &\leq \exp\left[-\frac{\xi(1-\delta)}{2} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right] \frac{\int_\xi^1 \kappa_i^{k/2+a} (1 - \kappa_i)^{-a-1} L\left(\frac{1-\kappa_i}{\kappa_i \tau_n}\right) d\kappa_i}{\int_0^{\xi\delta} \kappa_i^{k/2+a-1} (1 - \kappa_i)^{-a-1} L\left(\frac{1-\kappa_i}{\kappa_i \tau_n}\right) d\kappa_i}.
 \end{aligned} \tag{A8}$$

Now observe that

$$\begin{aligned}
& \int_{\xi}^1 \kappa_i^{k/2+a} (1 - \kappa_i)^{-a-1} L\left(\frac{1 - \kappa_i}{\kappa_i \tau_n}\right) d\kappa_i \\
&= \int_0^{(1-\xi)/(\xi\tau_n)} (1 + \lambda_i^2 \tau_n)^{-(k/2+a)} (\lambda_i^2 \tau_n (1 + \lambda_i^2 \tau_n)^{-1})^{-a-1} L(\lambda_i^2) \frac{\tau_n d\lambda_i^2}{(1 + \lambda_i^2 \tau_n)^2} \\
&= \tau_n^{-a} \int_0^{(1-\xi)/(\xi\tau_n)} (1 + \lambda_i^2 \tau_n)^{-(k/2+1)} (\lambda_i^2)^{-a-1} L(\lambda_i^2) d\lambda_i^2 \\
&\leq \tau_n^{-a} \int_0^{\infty} (\lambda_i^2)^{-a-1} L(\lambda_i^2) d\lambda_i^2 = \tau_n^{-a} K^{-1}.
\end{aligned} \tag{A9}$$

By assumption (I) and $\tau_n < 1$,

$$\begin{aligned}
& \int_0^{\xi\delta} \kappa_i^{k/2+a-1} (1 - \kappa_i)^{-a-1} L\left(\frac{1 - \kappa_i}{\kappa_i \tau_n}\right) d\kappa_i \\
&\geq \int_0^{\xi\delta} \kappa_i^{k/2+a-1} L\left(\frac{1 - \kappa_i}{\kappa_i \tau_n}\right) d\kappa_i \\
&\geq L\left(\frac{1 - \xi\delta}{\xi\delta}\right) \int_0^{\xi\delta} \kappa_i^{k/2+a-1} d\kappa_i \\
&= L\left(\frac{1 - \xi\delta}{\xi\delta}\right) \frac{(\xi\delta)^{k/2+a}}{k/2 + a}.
\end{aligned} \tag{A10}$$

Combining (A8) - (A10),

$$E(\kappa_i \mathbb{1}_{[\kappa_i > \xi]} | \mathbf{X}_i) \leq K \tau_n^{-a} \exp\left[-\frac{\xi(1 - \delta)}{2} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right]. \tag{A11}$$

□

Lemma 3 *Under the multivariate global-local prior model with treating the global parameter as a tuning parameter and under the regularity assumption (I), for an arbitrary constant $0 < \xi < 1$,*

$$E(\kappa_i \mathbb{1}_{[\kappa_i \leq \xi]} | \mathbf{X}_i) \leq K / \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i. \tag{A12}$$

Proof For an arbitrary constant $0 < \xi < 1$,

$$\begin{aligned}
 & E \left(\kappa_i \mathbb{1}_{[\kappa_i \leq \xi]} \mid \mathbf{X}_i \right) \\
 &= \frac{\int_0^\xi \kappa_i^{k/2+a} (1 - \kappa_i)^{-a-1} L \left(\frac{1-\kappa_i}{\kappa_i \tau_n} \right) \exp \left(-\frac{\kappa_i}{2} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right) d\kappa_i}{\int_0^1 \kappa_i^{k/2+a-1} (1 - \kappa_i)^{-a-1} L \left(\frac{1-\kappa_i}{\kappa_i \tau_n} \right) \exp \left(-\frac{\kappa_i}{2} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right) d\kappa_i} \\
 &\leq \frac{(1 - \xi)^{-a-1} M}{L \left(\frac{1-\xi}{\xi \tau_n} \right)} \frac{\int_0^\xi \kappa_i^{k/2+a} \exp \left(-\frac{\kappa_i}{2} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right) d\kappa_i}{\int_0^\xi \kappa_i^{k/2+a-1} \exp \left(-\frac{\kappa_i}{2} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \right) d\kappa_i} \\
 &= \frac{(1 - \xi)^{-a-1} M}{L \left(\frac{1-\xi}{\xi \tau_n} \right)} \frac{\int_0^{\xi \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2} \left(\frac{2t}{\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i} \right)^{k/2+a} \exp(-t) dt}{\int_0^{\xi \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2} \left(\frac{2t}{\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i} \right)^{k/2+a-1} \exp(-t) dt} \\
 &= \frac{(1 - \xi)^{-a-1} M}{L \left(\frac{1-\xi}{\xi \tau_n} \right) \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i} \frac{\int_0^{\xi \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2} t^{k/2+a} \exp(-t) dt}{\int_0^{\xi \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2} t^{k/2+a-1} \exp(-t) dt}.
 \end{aligned} \tag{A13}$$

Integrating the numerator by parts,

$$\begin{aligned}
 & \int_0^{\xi \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2} t^{k/2+a} \exp(-t) dt = (k/2 + a) \times \\
 & \int_0^{\xi \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2} t^{k/2+a-1} \exp(-t) dt - \left(\frac{\xi \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2} \right)^{k/2+a} \exp \left(-\frac{\xi \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2} \right).
 \end{aligned} \tag{A14}$$

Combining (A13) and (A14),

$$E \left(\kappa_i \mathbb{1}_{[\kappa_i \leq \xi]} \mid \mathbf{X}_i \right) \leq K / \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i. \tag{A15}$$

□

A.2 Lemmas for the Multivariate Exponential-Inverse-Gamma Model
 Under an Exponential-Inverse-Gamma model (i), (ii') and (iii') the posterior density of κ_i is given by

$$\Pi(\kappa_i \mid \mathbf{X}_i) \propto \kappa_i^{d+\frac{k}{2}-1} (1 - \kappa_i + \kappa_i c_n)^{-d-1} \exp \left(-\frac{\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2} \right).$$

Now, we give three basic inequalities involving the Exponential-Inverse-Gamma prior.

Lemma 4 *Under the multivariate Exponential-Inverse-Gamma model, assuming $c_n \rightarrow 0$ as $n \rightarrow \infty$, for large n ,*

$$E(1 - \kappa_i | \mathbf{X}_i) \leq K c_n^d \exp\left(\frac{\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right). \quad (\text{A16})$$

Proof Firstly,

$$\begin{aligned} E(1 - \kappa_i | \mathbf{X}_i) &= \frac{\int_0^1 (1 - \kappa_i) \kappa_i^{d+\frac{k}{2}-1} \exp\left(-\frac{\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) (1 - \kappa_i + \kappa_i c_n)^{-d-1} d\kappa_i}{\int_0^1 \kappa_i^{d+\frac{k}{2}-1} \exp\left(-\frac{\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) (1 - \kappa_i + \kappa_i c_n)^{-d-1} d\kappa_i} \\ &\leq \exp(\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2) \frac{\int_0^1 (1 - \kappa_i) \kappa_i^{d+\frac{k}{2}-1} (1 - \kappa_i + \kappa_i c_n)^{-d-1} d\kappa_i}{\int_0^1 \kappa_i^{d+\frac{k}{2}-1} (1 - \kappa_i + \kappa_i c_n)^{-d-1} d\kappa_i} \\ &:= \exp(\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2) \frac{N}{D}. \end{aligned}$$

Now,

$$\begin{aligned} D &\geq \int_{1-c_n}^1 \kappa_i^{d+\frac{k}{2}-1} (1 - \kappa_i + \kappa_i c_n)^{-d-1} d\kappa_i \\ &\geq \int_{1-c_n}^1 \kappa_i^{d+\frac{k}{2}-1} (c_n + c_n)^{-d-1} d\kappa_i \\ &= (2c_n)^{-(d+1)} \int_{1-c_n}^1 \kappa_i^{d+\frac{k}{2}-1} d\kappa_i \\ &= (2c_n)^{-(d+1)} \frac{1 - (1 - c_n)^{d+\frac{k}{2}}}{d + \frac{k}{2}} \\ &\geq K c_n^{-(d+1)} \left[c_n - \frac{d + \frac{k}{2} - 1}{2} c_n^2 \right] \\ &= K c_n^{-d} \left[1 - \frac{d + \frac{k}{2} - 1}{2} c_n \right] \\ &\geq K c_n^{-d}. \end{aligned}$$

Next,

$$\begin{aligned}
 N &= \int_0^1 (1 - \kappa_i) \kappa_i^{d+\frac{k}{2}-1} (1 - \kappa_i + \kappa_i c_n)^{-d-1} d\kappa_i \\
 &\leq \int_0^1 (1 - \kappa_i) \kappa_i^{d+\frac{k}{2}-1} (1 - \kappa_i)^{-d-1} d\kappa_i \\
 &= \int_0^1 \kappa_i^{d+\frac{k}{2}-1} (1 - \kappa_i)^{-d} d\kappa_i \\
 &= \text{Beta}(1 - d, \frac{k}{2} + d).
 \end{aligned}$$

Therefore, when n is sufficiently large,

$$E(1 - \kappa_i | \mathbf{X}_i) \leq K c_n^d \exp(\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2).$$

□

Lemma 5 *Under the multivariate Exponential-Inverse-Gamma model, for arbitrary constants $0 < \xi < 1$ and $0 < \delta < 1$,*

$$E(\kappa_i \mathbb{1}_{[\kappa_i > \xi]} | \mathbf{X}_i) \leq K c_n^{-d} \exp\left(-\frac{1}{2}\xi(1 - \delta)\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right). \quad (\text{A17})$$

Proof For arbitrary constants $0 < \xi < 1$ and $0 < \delta < 1$,

$$\begin{aligned}
 &E(\kappa_i \mathbb{1}_{[\kappa_i > \xi]} | \mathbf{X}_i) \\
 &= \frac{\int_{\xi}^1 \kappa_i \kappa_i^{d+\frac{k}{2}-1} \exp\left(-\frac{\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) (1 - \kappa_i + \kappa_i c_n)^{-d-1} d\kappa_i}{\int_0^1 \kappa_i^{d+\frac{k}{2}-1} \exp\left(-\frac{\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) (1 - \kappa_i + \kappa_i c_n)^{-d-1} d\kappa_i} \\
 &\leq \frac{\int_{\xi}^1 \kappa_i^{d+\frac{k}{2}} \exp\left(-\frac{\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) (1 - \kappa_i + \kappa_i c_n)^{-d-1} d\kappa_i}{\int_0^{\xi\delta} \kappa_i^{d+\frac{k}{2}-1} \exp\left(-\frac{\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) (1 - \kappa_i + \kappa_i c_n)^{-d-1} d\kappa_i} \\
 &\leq \exp\left(-\frac{1}{2}\xi(1 - \delta)\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right) \frac{\int_{\xi}^1 \kappa_i^{d+\frac{k}{2}} (1 - \kappa_i + \kappa_i c_n)^{-d-1} d\kappa_i}{\int_0^{\xi\delta} \kappa_i^{d+\frac{k}{2}-1} (1 - \kappa_i + \kappa_i c_n)^{-d-1} d\kappa_i} \\
 &\leq \exp\left(-\frac{1}{2}\xi(1 - \delta)\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i\right) \frac{\int_{\xi}^1 \kappa_i^{d+\frac{k}{2}} (1 - \kappa_i + \kappa_i c_n)^{-d-1} d\kappa_i}{\int_0^{\xi\delta} \kappa_i^{d+\frac{k}{2}-1} (1 + c_n)^{-d-1} d\kappa_i}
 \end{aligned}$$

$$\begin{aligned}
&\leq \exp\left(-\frac{1}{2}\xi(1-\delta)\mathbf{X}_i^T\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right)(1+c_n)^{d+1}\frac{d+k/2}{(\xi\delta)^{d+k/2}}\int_{\xi}^1(1-\kappa_i+\kappa_ic_n)^{-d-1}d\kappa_i \\
&\leq K\exp\left(-\frac{1}{2}\xi(1-\delta)\mathbf{X}_i^T\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right)\left[\frac{(1-\kappa_i+\xi c_n)^{-d}}{d}\right]_{\xi}^1 \\
&\leq K\exp\left(-\frac{1}{2}\xi(1-\delta)\mathbf{X}_i^T\boldsymbol{\Sigma}^{-1}\mathbf{X}_i\right)c_n^{-d}.
\end{aligned}$$

□

Lemma 6 Under the multivariate Exponential-Inverse-Gamma model, for an arbitrary constant $0 < \xi < 1$,

$$E(\kappa_i \mathbb{1}_{[\kappa_i \leq \xi]} | \mathbf{X}_i) \leq K / \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i. \quad (\text{A18})$$

Proof For an arbitrary constant $0 < \xi < 1$,

$$\begin{aligned}
&E(\kappa_i \mathbb{1}_{[\kappa_i \leq \xi]} | \mathbf{X}_i) \\
&= \frac{\int_0^{\xi} \kappa_i^{d+\frac{k}{2}} \exp\left(-\frac{\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) (1-\kappa_i+\kappa_ic_n)^{-d-1} d\kappa_i}{\int_0^1 \kappa_i^{d+\frac{k}{2}-1} \exp\left(-\frac{\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) (1-\kappa_i+\kappa_ic_n)^{-d-1} d\kappa_i} \\
&\leq \frac{\int_0^{\xi} \kappa_i^{d+\frac{k}{2}} \exp\left(-\frac{\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) (1-\kappa_i+\kappa_ic_n)^{-d-1} d\kappa_i}{\int_0^{\xi} \kappa_i^{d+\frac{k}{2}-1} \exp\left(-\frac{\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) (1-\kappa_i+\kappa_ic_n)^{-d-1} d\kappa_i} \\
&\leq \frac{(1-\xi)^{-d-1} \int_0^{\xi} \kappa_i^{d+\frac{k}{2}} \exp\left(-\frac{\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) d\kappa_i}{(1+\xi c_n)^{-d-1} \int_0^{\xi} \kappa_i^{d+\frac{k}{2}-1} \exp\left(-\frac{\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) d\kappa_i} \\
&\leq \frac{(1-\xi)^{-d-1} \int_0^{\xi} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2 \exp(-t) t^{d+\frac{k}{2}} dt}{(1+\xi c_n)^{-d-1} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i \int_0^{\xi} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2 \exp(-t) t^{d+\frac{k}{2}-1} dt}.
\end{aligned}$$

Integrating by parts,

$$\begin{aligned}
&\int_0^{\xi} \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2 \exp(-t) t^{d+\frac{k}{2}} dt \\
&= \left[-\exp(-t) t^{d+\frac{k}{2}}\right]_{t=0}^{t=\xi \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2} + \left(d+\frac{k}{2}\right) \int_0^{\xi \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2} \exp(-t) t^{d+\frac{k}{2}-1} dt.
\end{aligned}$$

Therefore,

$$\frac{\int_0^{\xi \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2} \exp(-t) t^{d+\frac{k}{2}} dt}{\int_0^{\xi \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2} \exp(-t) t^{d+\frac{k}{2}-1} dt} \leq d + \frac{k}{2}.$$

Hence,

$$E(\kappa_i \mathbb{1}_{[\kappa_i \leq \xi]} | \mathbf{X}_i) \leq K / \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i.$$

□

Appendix B: Proofs of the Main Theorems

B.1 Proofs of Theorems in Section 2

Proof of Theorem 1 For each $\mathbf{X}_i \stackrel{\text{ind}}{\sim} N_k(\boldsymbol{\theta}_i, \boldsymbol{\Sigma})$, define $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{ik})^\top := \boldsymbol{\Sigma}^{-1/2} \mathbf{X}_i$ and $\boldsymbol{\lambda}_i = (\lambda_{i1}, \dots, \lambda_{ik})^\top := \boldsymbol{\Sigma}^{-1/2} \boldsymbol{\theta}_i$. So, all the components of \mathbf{Y}_i 's are independently normally distributed, i.e., $Y_{ij} \stackrel{\text{ind}}{\sim} N(\lambda_{ij}, 1), i = 1, \dots, n, j = 1, \dots, k$.

Given estimators $\widehat{\boldsymbol{\lambda}}_i$, if we use $\widehat{\boldsymbol{\theta}}_i = \boldsymbol{\Sigma}^{1/2} \widehat{\boldsymbol{\lambda}}_i$ to estimate $\boldsymbol{\theta}_i$'s, then

$$E \|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i\|_{\boldsymbol{\Sigma}}^2 = E \|\widehat{\boldsymbol{\lambda}}_i - \boldsymbol{\lambda}_i\|^2 = \sum_{j=1}^k E(\widehat{\lambda}_{ij} - \lambda_{ij})^2.$$

Also, since $\{\lambda_{ij}\} \in l_0[q'_{nk}]$ would imply that $\{\boldsymbol{\theta}_i\} \in L_0[q_n]$, when we let $q'_{nk} = q_n$, we have

$$\sup_{\{\lambda_{ij}\} \in l_0[q'_{nk}]} \sum_{i=1}^n \sum_{j=1}^k E(\widehat{\lambda}_{ij} - \lambda_{ij})^2 \leq \sup_{\{\boldsymbol{\theta}_i\} \in L_0[q_n]} \sum_{i=1}^n E \|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i\|_{\boldsymbol{\Sigma}}^2.$$

Finally, since $\boldsymbol{\Sigma}$ is positive definite, there is a one-to-one correspondence between $\{\{\widehat{\boldsymbol{\theta}}_i\}_{i=1}^n\}$ and $\{\{\widehat{\lambda}_{ij}\}_{i=1, j=1}^n, k\}$. So the above inequality still hold, if we further take the infimum over all possible estimators:

$$\inf_{\{\widehat{\lambda}_{ij}\}} \sup_{\{\lambda_{ij}\} \in l_0[q'_{nk}]} \sum_{i=1}^n \sum_{j=1}^k E(\widehat{\lambda}_{ij} - \lambda_{ij})^2 \leq \inf_{\{\widehat{\boldsymbol{\theta}}_i\}} \sup_{\{\boldsymbol{\theta}_i\} \in L_0[q_n]} \sum_{i=1}^n E \|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i\|_{\boldsymbol{\Sigma}}^2.$$

Donoho et al. (1992) provides the result for the left hand side, which is, as $n \rightarrow \infty$,

$$\begin{aligned} \inf_{\{\widehat{\lambda}_{ij}\}} \sup_{\{\lambda_{ij}\} \in l_0[q'_{nk}]} \sum_{i=1}^n \sum_{j=1}^k E(\widehat{\lambda}_{ij} - \lambda_{ij})^2 &= 2q'_{nk} \log(nk/q'_{nk})(1 + o(1)) \\ &= 2q_n \log(n/q_n)(1 + o(1)). \end{aligned}$$

The last equality holds because we let $q'_{nk} = q_n$.

So we have a minimax lower bound now,

$$\inf_{\{\widehat{\boldsymbol{\theta}}_i\}} \sup_{\{\boldsymbol{\theta}_i\} \in L_0[q_n]} \sum_{i=1}^n E \|\widehat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i\|_{\Sigma}^2 \geq 2q_n \log(n/q_n)(1 + o(1)).$$

As we will see in (6) and (12), when using some particular priors, the error of the Bayes estimate of $\boldsymbol{\theta}_i$'s will be at most $2q_n \log(n/q_n)(1 + o(1))$. This fact provides an upper bound for the minimax error, which coincides with the lower bound, and finishes this proof. \square

Proof of Theorem 2 We first prove (4). Observe that

$$\|\widehat{\boldsymbol{\theta}}_i^R\|_{\Sigma}^2 = E(1 - \kappa_i | \mathbf{X}_i)^2 \|\mathbf{X}_i\|_{\Sigma}^2 \leq E(1 - \kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2.$$

Denote the l_2 -norm by $\|\cdot\|_2$. Then, making use of Lemma 1, for a sequence of positive constants $\{a_n, n \geq 1\}$ to be specified later with $a_n \rightarrow \infty$ as $n \rightarrow \infty$,

$$\begin{aligned} & E_0 \left[E(1 - \kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_2^2 \leq a_n]} \right] \\ & \leq K\tau_n^\eta E_0 \left[\exp(\mathbf{X}_i^T \Sigma^{-1} \mathbf{X}_i / 2) \|\mathbf{X}_i\|_2^2 \mathbb{1}_{[\|\mathbf{X}_i\|_2^2 \leq a_n]} \right] \\ & = K\tau_n^\eta \int_{\|\mathbf{X}_i\|_2^2 \leq a_n} \|\mathbf{X}_i\|_2^2 \exp(\mathbf{X}_i^T \Sigma^{-1} \mathbf{X}_i / 2 - \mathbf{X}_i^T \Sigma^{-1} \mathbf{X}_i / 2) d\mathbf{X}_i \\ & = K\tau_n^\eta \int_{\|\mathbf{X}_i\|_2^2 \leq a_n} \|\mathbf{X}_i\|_2^2 d\mathbf{X}_i \\ & \leq K\tau_n^\eta a_n V_k(a_n) \\ & = K\tau_n^\eta a_n^{k+1}, \end{aligned} \tag{B19}$$

where $V_k(r) = \frac{\pi^{k/2}}{\Gamma(k/2+1)} r^k$ is the volume of a Euclidean ball of radius r in k -dimensional Euclidean space.

Moreover, under assumption (II),

$$\begin{aligned} & E_0 \left[E(1 - \kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_2^2 > a_n]} \right] \\ & \leq E_0 \left(\|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 > \lambda_{\min}(\Sigma^{-1})a_n]} \right) \end{aligned}$$

Noting that under $\boldsymbol{\theta}_{0i} = \mathbf{0}$, $\|\mathbf{X}_i\|_{\Sigma}^2 \sim \chi_k^2$, one gets

$$\begin{aligned}
 & E_0 \left(\|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 > \lambda_{\min}(\Sigma^{-1})a_n]} \right) \\
 &= \int_{\lambda_{\min}(\Sigma^{-1})a_n}^{\infty} x \exp\left(-\frac{x}{2}\right) \frac{x^{k/2-1}}{\Gamma(k/2)2^{k/2}} dx \\
 &= 2 \int_{(\lambda_{\min}(\Sigma^{-1})a_n)/2}^{\infty} \exp(-x) \frac{x^{k/2}}{\Gamma(k/2)} dx \\
 &\leq 2 \exp\left(-\lambda_{\min}(\Sigma^{-1})a_n/4\right) \int_{(\lambda_{\min}(\Sigma^{-1})a_n)/2}^{\infty} \exp\left(-\frac{x}{2}\right) \frac{x^{k/2}}{\Gamma(k/2)} dx \\
 &\leq \exp\left(-\lambda_{\min}(\Sigma^{-1})a_n/4\right) \left(\frac{k}{2}\right) 2^{\frac{k}{2}+2}.
 \end{aligned}$$

Now choosing $a_n = 4\lambda_{\min}^{-1}(\Sigma^{-1})(1+\epsilon)\log(n/q_n)$, with $\epsilon > 0$, one gets

$$E_0 \left[E(1 - \kappa_i \mid \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 > a_n]} \right] \leq K(q_n/n)^{1+\epsilon}. \quad (\text{B20})$$

Finally, choosing $\tau_n = (q_n/n)^{\frac{1+\epsilon}{n}}$, the theorem follows from (B19) and (B20),

$$\begin{aligned}
 E_0 \left[E(1 - \kappa_i \mid \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \right] &\leq K \left[(q_n/n)^{1+\epsilon} (\log(n/q_n))^{k+1} + (q_n/n)^{1+\epsilon} \right] \\
 &\lesssim (q_n/n)^{1+\epsilon} (\log(n/q_n))^{k+1}, \text{ as } n \rightarrow \infty.
 \end{aligned}$$

Summing over all i 's for which $\boldsymbol{\theta}_{0i} = \mathbf{0}$, one gets

$$\sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]_{i:\boldsymbol{\theta}_{0i}=\mathbf{0}}} \sum E_0 \|\widehat{\boldsymbol{\theta}}_i^R\|_{\Sigma}^2 / \left(q_n \log\left(\frac{n}{q_n}\right) \right) = o(1), \text{ as } n \rightarrow \infty.$$

Now we prove (5). Use the inequality

$$\begin{aligned}
 & E_{\boldsymbol{\theta}_{0i}} \|\widehat{\boldsymbol{\theta}}_i^R - \boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \\
 &= E_{\boldsymbol{\theta}_{0i}} \left\| (1 - E(\kappa_i \mid \mathbf{X}_i)) \mathbf{X}_i - \mathbf{X}_i + \mathbf{X}_i - \boldsymbol{\theta}_{0i} \right\|_{\Sigma}^2 \\
 &= E_{\boldsymbol{\theta}_{0i}} \left[E(\kappa_i \mid \mathbf{X}_i)^2 \|\mathbf{X}_i\|_{\Sigma}^2 + \|\mathbf{X}_i - \boldsymbol{\theta}_{0i}\|_{\Sigma}^2 - 2\langle \Sigma^{-1/2}(\mathbf{X}_i - \boldsymbol{\theta}_{0i}), E(\kappa_i \mid \mathbf{X}_i) \Sigma^{-1/2} \mathbf{X}_i \rangle \right] \\
 &\leq E_{\boldsymbol{\theta}_{0i}} \left[E(\kappa_i \mid \mathbf{X}_i)^2 \|\mathbf{X}_i\|_{\Sigma}^2 \right] + E_{\boldsymbol{\theta}_{0i}} \|\mathbf{X}_i - \boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \\
 &+ 2 \left(E_{\boldsymbol{\theta}_{0i}} \|\mathbf{X}_i - \boldsymbol{\theta}_{0i}\|_{\Sigma}^2 E_{\boldsymbol{\theta}_{0i}} \left[E(\kappa_i \mid \mathbf{X}_i)^2 \|\mathbf{X}_i\|_{\Sigma}^2 \right] \right)^{\frac{1}{2}}
 \end{aligned}$$

But $E_{\theta_{0i}} \|\mathbf{X}_i - \theta_{0i}\|_{\Sigma}^2 = k$, the above becomes

$$\begin{aligned} E_{\theta_{0i}} \|\widehat{\theta}_i^R - \theta_{0i}\|_{\Sigma}^2 &\leq E_{\theta_{0i}} \left[E(\kappa_i | \mathbf{X}_i)^2 \|\mathbf{X}_i\|_{\Sigma}^2 \right] \\ &\quad + 2k^{\frac{1}{2}} E_{\theta_{0i}}^{\frac{1}{2}} \left[E(\kappa_i | \mathbf{X}_i)^2 \|\mathbf{X}_i\|_{\Sigma}^2 \right] + k \\ &\leq E_{\theta_{0i}} \left[E(\kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \right] \\ &\quad + 2k^{\frac{1}{2}} E_{\theta_{0i}}^{\frac{1}{2}} \left[E(\kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \right] + k \end{aligned}$$

Since $q_n k / q_n \log(n/q_n) \rightarrow 0$, as $n \rightarrow \infty$, it suffices to show that

$$\limsup_{n \rightarrow \infty} \sup_{\{\theta_{0i}\} \in L_0[q_n]} \frac{\sum_{i: \theta_{0i} \neq \mathbf{0}} E_{\theta_{0i}} [E(\kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2]}{4a q_n \log(n/q_n)} \leq 1/(2 \min(1, a)). \quad (\text{B21})$$

and

$$\limsup_{n \rightarrow \infty} \sup_{\{\theta_{0i}\} \in L_0[q_n]} \frac{\sum_{i: \theta_{0i} \neq \mathbf{0}} E_{\theta_{0i}}^{\frac{1}{2}} [E(\kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2]}{q_n \log(n/q_n)} = 0. \quad (\text{B22})$$

In view of Lemma 2, for sufficiently large $b_n > 0$, uniformly in $\theta_{0i} \neq \mathbf{0}$,

$$\begin{aligned} &E(\kappa_i \mathbb{1}_{[\kappa_i > \xi]} | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 > b_n]} \\ &\leq K \tau_n^{-a} \|\mathbf{X}_i\|_{\Sigma}^2 \exp \left[-\frac{\xi(1-\delta)}{2} \|\mathbf{X}_i\|_{\Sigma}^2 \right] \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 > b_n]} \\ &\leq K \tau_n^{-a} b_n \exp \left[-\frac{\xi(1-\delta)}{2} b_n \right] \\ &\leq K (q_n/n)^{a(\rho-\epsilon)/n} \log \left(\frac{n}{q_n} \right), \end{aligned} \quad (\text{B23})$$

by choosing $b_n = \frac{1+\rho}{\eta} \frac{2a}{\xi(1-\delta)} \log \left(\frac{n}{q_n} \right)$, with $\rho > \epsilon$, and recalling that $\tau_n = (q_n/n)^{\frac{1+\epsilon}{n}}$.

Then summing over all i 's for which $\theta_{0i} \neq \mathbf{0}$, one gets

$$\limsup_{n \rightarrow \infty} \sup_{\{\theta_{0i}\} \in L_0[q_n]} \frac{\sum_{i: \theta_{0i} \neq \mathbf{0}} E_{\theta_{0i}} [E(\kappa_i \mathbb{1}_{[\kappa_i > \xi]} | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 > b_n]}]}{q_n \log(n/q_n)} = 0. \quad (\text{B24})$$

Finally,

$$\begin{aligned}
 & E_{\theta_{0i}} [E(\kappa_i \mathbb{1}_{[\kappa_i > \xi]} | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 \leq b_n]}] \\
 & \leq E_{\theta_{0i}} [\|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 \leq b_n]}] \\
 & \leq b_n \\
 & \leq \frac{1 + \rho}{\eta} \frac{2a}{\xi(1 - \delta)} \log\left(\frac{n}{q_n}\right).
 \end{aligned} \tag{B25}$$

Since the result above is independent of any specific choice of the parameters, by making $\epsilon \rightarrow 0$, $\rho \rightarrow 0$, $\eta \rightarrow \min(1, a)$, $\xi \rightarrow 1$ and $\delta \rightarrow 0$, and summing over all i 's for which $\theta_{0i} \neq \mathbf{0}$, one gets

$$\begin{aligned}
 \limsup_{n \rightarrow \infty} \sup_{\{\theta_{0i}\} \in L_0[q_n]} & \frac{\sum_{i: \theta_{0i} \neq \mathbf{0}} E_{\theta_{0i}} [E(\kappa_i \mathbb{1}_{[\kappa_i > \xi]} | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 \leq b_n]}]}{4aq_n \log(n/q_n)} \\
 & \leq 1/(2 \min(1, a)).
 \end{aligned} \tag{B26}$$

Together with Lemma 3 and (B24), this leads to (B21).

Altogether Lemma 2, (B23) and (B25) also imply that

$$E_{\theta_{0i}} [E(\kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2] \leq K \log\left(\frac{n}{q_n}\right) (1 + o(1)), \text{ as } n \rightarrow \infty. \tag{B27}$$

Again summing over all i 's for which $\theta_{0i} \neq \mathbf{0}$, one gets (B22). This completes the proof of (5).

(6) follows (4) and (5), immediately. In particular, when $0 < a \leq 1$, Theorem 1 provides a lower bound for the estimation error and (7) follows. \square

Proof of Theorem 3 By Markov's inequality and the independence of samples,

$$\begin{aligned}
 & E_{\{\theta_{0i}\}} \Pi\left(\sum_{i=1}^n \|\theta_i - \hat{\theta}_i^R\|_{\Sigma}^2 > q_n \log\left(\frac{n}{q_n}\right) \mid \{\mathbf{X}_i\}\right) \\
 & \leq E_{\{\theta_{0i}\}} E\left(\sum_{i=1}^n \|\theta_i - \hat{\theta}_i^R\|_{\Sigma}^2 \mid \{\mathbf{X}_i\}\right) / q_n \log\left(\frac{n}{q_n}\right) \\
 & = \sum_{i=1}^n E_{\theta_{0i}} E(\|\theta_i - \hat{\theta}_i^R\|_{\Sigma}^2 \mid \mathbf{X}_i) / q_n \log\left(\frac{n}{q_n}\right).
 \end{aligned}$$

Since

$$\boldsymbol{\theta}_i \mid \mathbf{X}_i, \kappa_i \sim N_k(\widehat{\boldsymbol{\theta}}_i^R, (1 - \kappa_i)\boldsymbol{\Sigma}),$$

we have

$$\begin{aligned} & E(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i^R\|_{\boldsymbol{\Sigma}}^2 \mid \mathbf{X}_i) \\ &= E\{E(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i^R\|_{\boldsymbol{\Sigma}}^2 \mid \mathbf{X}_i, \kappa_i) \mid \mathbf{X}_i\} \\ &= E\{(1 - \kappa_i)E\left(\frac{\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i^R\|_{\boldsymbol{\Sigma}}^2}{1 - \kappa_i} \mid \mathbf{X}_i, \kappa_i\right) \mid \mathbf{X}_i\} \\ &= kE(1 - \kappa_i \mid \mathbf{X}_i). \end{aligned}$$

Now we only need to find a bound for $E_{\boldsymbol{\theta}_{0i}}E(1 - \kappa_i \mid \mathbf{X}_i)$. When $\boldsymbol{\theta}_{0i} \neq 0$,

$$E_{\boldsymbol{\theta}_{0i}}E(1 - \kappa_i \mid \mathbf{X}_i) \leq 1.$$

When $\boldsymbol{\theta}_{0i} = 0$, letting $\{a_n\}$ be a sequence of positive numbers that will be chosen later, using Lemma 1, for $0 < \eta < \min(0, a)$,

$$\begin{aligned} & E_{\boldsymbol{\theta}_{0i}}E(1 - \kappa_i \mid \mathbf{X}_i) \\ &= E_{\mathbf{0}}\{E(1 - \kappa_i \mid \mathbf{X}_i)\mathbb{1}_{\{\|\mathbf{X}_i\|_2^2 \leq a_n\}}\} + E_{\mathbf{0}}\{E(1 - \kappa_i \mid \mathbf{X}_i)\mathbb{1}_{\{\|\mathbf{X}_i\|_2^2 > a_n\}}\} \\ &\leq E_{\mathbf{0}}\{K\tau_n^\eta \exp(\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i/2)\mathbb{1}_{\{\|\mathbf{X}_i\|_2^2 \leq a_n\}}\} + E_{\mathbf{0}}\{\mathbb{1}_{\{\|\mathbf{X}_i\|_2^2 > a_n\}}\} \\ &= K\tau_n^\eta V_k(a_n) + KPr(\chi_k^2 > a_n) \\ &\leq K\tau_n^\eta a_n^k + K a_n^{k/2} \exp(-a_n/2), \end{aligned}$$

where $V_k(r)$ is the volume of a Euclidean ball of radius r in k -dimensional Euclidean space, and the probability is bounded using the Chernoff bound for the χ^2 random variables.

For some $\epsilon > 0$, choose $a_n = 2(1 + \epsilon) \log(n/q_n)$ and $\tau_n = (q_n/n)^{(1+\epsilon)/\eta}$. Then,

$$\begin{aligned} & \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]} \sum_{i=1}^n E_{\boldsymbol{\theta}_{0i}}E(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i^R\|_{\boldsymbol{\Sigma}}^2 \mid \mathbf{X}_i)/q_n \log\left(\frac{n}{q_n}\right) \\ & \leq \frac{q_n + (n - q_n)(K\tau_n^\eta a_n^k + K a_n^{k/2} \exp(-a_n/2))}{q_n \log\left(\frac{n}{q_n}\right)} \rightarrow 0, \end{aligned}$$

as $n \rightarrow \infty$. This proves (8).

Next, since

$$E(\|\boldsymbol{\theta}_i - \boldsymbol{\theta}_{0i}\|_{\boldsymbol{\Sigma}}^2 \mid \mathbf{X}_i) \leq 2E(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i^R\|_{\boldsymbol{\Sigma}}^2 \mid \mathbf{X}_i) + 2E(\|\widehat{\boldsymbol{\theta}}_i^R - \boldsymbol{\theta}_{0i}\|_{\boldsymbol{\Sigma}}^2 \mid \mathbf{X}_i),$$

by Markov's inequality, (6) and (8) immediately lead to (9). \square

Proof of Theorem 4 The proof of (10) is similar to the proof of (4), but now using Lemma 4. We start from

$$\|\widehat{\boldsymbol{\theta}}_i^{EIG}\|_{\Sigma}^2 = E(1 - \kappa_i | \mathbf{X}_i)^2 \|\mathbf{X}_i\|_{\Sigma}^2 \leq E(1 - \kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2.$$

Then, for a sequence of positive constants $\{a_n, n \geq 1\}$ to be specified later with $a_n \rightarrow \infty$ as $n \rightarrow \infty$,

$$\begin{aligned} & E_0 \left[E(1 - \kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 \leq a_n]} \right] \\ & \leq K c_n^d E_0 \left[\exp(\mathbf{X}_i^T \Sigma^{-1} \mathbf{X}_i / 2) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 \leq a_n]} \right] \\ & \leq K c_n^d a_n^{k+1}, \end{aligned}$$

and,

$$\begin{aligned} & E_0 \left[E(1 - \kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 > a_n]} \right] \\ & \leq E_0 \left(\|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 > \lambda_{\min}(\Sigma^{-1}) a_n]} \right) \\ & \leq \exp(-\lambda_{\min}(\Sigma^{-1}) a_n / 4). \end{aligned}$$

By choosing $a_n = 4\lambda_{\min}^{-1}(\Sigma^{-1})(1 + \epsilon) \log(n/q_n)$ and $\tau_n = (q_n/n)^{\frac{1+\epsilon}{d}}$, with $\epsilon > 0$, one gets

$$\begin{aligned} E_0 \left[E(1 - \kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \right] & \leq K \left[(q_n/n)^{1+\epsilon} (\log(n/q_n))^{k+1} + (q_n/n)^{1+\epsilon} \right] \\ & \lesssim (q_n/n)^{1+\epsilon} (\log(n/q_n))^{k+1}, \text{ as } n \rightarrow \infty. \end{aligned}$$

and hence

$$\sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]_{i:\boldsymbol{\theta}_{0i}=\mathbf{0}}} \sum E_0 \|\widehat{\boldsymbol{\theta}}_i^{EIG}\|_{\Sigma}^2 / \left(q_n \log \left(\frac{n}{q_n} \right) \right) = o(1), \text{ as } n \rightarrow \infty.$$

The proof of (11) is similar to the proof of (5), but now using Lemmas 5 and 6. Again, it suffices to show that

$$\limsup_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]} \frac{\sum_{i:\boldsymbol{\theta}_{0i} \neq \mathbf{0}} E_{\boldsymbol{\theta}_{0i}} [E(\kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2]}{2\lambda_{\max}(\Sigma) q_n \log(n/q_n)} \leq 1. \quad (\text{B28})$$

and

$$\limsup_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]} \frac{\sum_{i:\boldsymbol{\theta}_{0i} \neq \mathbf{0}} E_{\boldsymbol{\theta}_{0i}}^{\frac{1}{2}} [E(\kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2]}{q_n \log(n/q_n)} = 0. \quad (\text{B29})$$

By Lemma 5, for sufficiently large $b_n > 0$, uniformly in $\boldsymbol{\theta}_{0i} \neq \mathbf{0}$,

$$\begin{aligned} & E(\kappa_i \mathbb{1}_{[\kappa_i > \xi]} | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 > b_n]} \\ & \leq K c_n^{-d} \|\mathbf{X}_i\|_{\Sigma}^2 \exp\left[-\frac{\xi(1-\delta)}{2} \|\mathbf{X}_i\|_{\Sigma}^2\right] \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 > b_n]} \\ & \leq K c_n^{-d} b_n \exp\left[-\frac{\xi(1-\delta)}{2} b_n\right] \\ & \leq K (q_n/n)^{\rho-\epsilon} \log\left(\frac{n}{q_n}\right), \end{aligned}$$

by choosing $b_n = \frac{2(1+\rho)}{\xi(1-\delta)} \log\left(\frac{n}{q_n}\right)$, with $\rho > \epsilon$, and recalling that $c_n = (q_n/n)^{\frac{1+\epsilon}{d}}$. So,

$$\limsup_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]} \frac{\sum_{i:\boldsymbol{\theta}_{0i} \neq \mathbf{0}} E_{\boldsymbol{\theta}_{0i}} [E(\kappa_i \mathbb{1}_{[\kappa_i > \xi]} | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 > b_n]}]}{q_n \log(n/q_n)} = 0. \quad (\text{B30})$$

Also,

$$\begin{aligned} & E_{\boldsymbol{\theta}_{0i}} [E(\kappa_i \mathbb{1}_{[\kappa_i > \xi]} | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 \leq b_n]}] \\ & \leq E_{\boldsymbol{\theta}_{0i}} [\|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 \leq b_n]}] \\ & \leq b_n \\ & \leq \frac{2(1+\rho)}{\xi(1-\delta)} \log\left(\frac{n}{q_n}\right). \end{aligned} \quad (\text{B31})$$

By making $\epsilon \rightarrow 0$, $\rho \rightarrow 0$, $\eta \rightarrow \min(1, a)$, $\xi \rightarrow 1$ and $\delta \rightarrow 0$, one gets

$$\limsup_{n \rightarrow \infty} \sup_{\{\boldsymbol{\theta}_{0i}\} \in L_0[q_n]} \frac{\sum_{i:\boldsymbol{\theta}_{0i} \neq \mathbf{0}} E_{\boldsymbol{\theta}_{0i}} [E(\kappa_i \mathbb{1}_{[\kappa_i > \xi]} | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \mathbb{1}_{[\|\mathbf{X}_i\|_{\Sigma}^2 \leq b_n]}]}{2q_n \log(n/q_n)} \leq 1. \quad (\text{B32})$$

Together with Lemma 6 and (B30), this leads to (B28).

Altogether Lemma 5, (B30) and (B31) also imply that

$$E_{\theta_{0i}}[E(\kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2] \leq K \log\left(\frac{n}{q_n}\right) (1 + o(1)), \text{ as } n \rightarrow \infty. \quad (\text{B33})$$

Consequently, one gets (B29). This completes the proof of (11).

Finally, (12) follows the previous two results and Theorem 1. □

Proof of Theorem 5 Similar to the Proof of Theorem 3, we only need to find a bound for $E_{\theta_{0i}}E(1 - \kappa_i | \mathbf{X}_i)$. When $\theta_{0i} \neq 0$,

$$E_{\theta_{0i}}E(1 - \kappa_i | \mathbf{X}_i) \leq 1.$$

When $\theta_{0i} = 0$, letting $\{a_n\}$ be a sequence of positive numbers that will be chosen later, using Lemma 4,

$$\begin{aligned} & E_{\theta_{0i}}E(1 - \kappa_i | \mathbf{X}_i) \\ &= E_{\mathbf{0}}\{E(1 - \kappa_i | \mathbf{X}_i) \mathbb{1}_{\{\|\mathbf{X}_i\|_2^2 \leq a_n\}}\} + E_{\mathbf{0}}\{E(1 - \kappa_i | \mathbf{X}_i) \mathbb{1}_{\{\|\mathbf{X}_i\|_2^2 > a_n\}}\} \\ &\leq K c_n^d a_n^k + K a_n^{k/2} \exp(-a_n/2). \end{aligned}$$

For some $\epsilon > 0$, choose $a_n = 2(1 + \epsilon) \log(n/q_n)$ and $c_n = (q_n/n)^{(1+\epsilon)/d}$. Then,

$$\begin{aligned} & \sup_{\{\theta_{0i}\} \in L_0[q_n]} \sum_{i=1}^n E_{\theta_{0i}}E(\|\theta_i - \hat{\theta}_i^{EIG}\|_{\Sigma}^2 | \mathbf{X}_i) / q_n \log\left(\frac{n}{q_n}\right) \\ &\leq \frac{q_n + (n - q_n)(K c_n^d a_n^k + K a_n^{k/2} \exp(-a_n/2))}{q_n \log\left(\frac{n}{q_n}\right)} \rightarrow 0, \end{aligned}$$

as $n \rightarrow \infty$. This proves (13).

Next, since

$$E(\|\theta_i - \theta_{0i}\|_{\Sigma}^2 | \mathbf{X}_i) \leq 2E(\|\theta_i - \hat{\theta}_i^{EIG}\|_{\Sigma}^2 | \mathbf{X}_i) + 2E(\|\hat{\theta}_i^{EIG} - \theta_{0i}\|_{\Sigma}^2 | \mathbf{X}_i),$$

by Markov's inequality, (12) and (13) immediately lead to (14). □

B.2 Proofs of Theorems in Section 3

Proof of Theorem 6 We use $\hat{\theta}_i$ for $\hat{\theta}_i^R$ in this proof.

Proof of (16). Look at the case where $\|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \leq K_S \tau_n$. Note that we could write $\mathbf{X}_i = \boldsymbol{\theta}_{0i} + \boldsymbol{\epsilon}_i$, where $\boldsymbol{\epsilon}_i \sim \mathbf{N}_k(\mathbf{0}, \boldsymbol{\Sigma})$ and hence $\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \sim \chi_k^2$. So,

$$\begin{aligned} \|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 &= \|\boldsymbol{\theta}_{0i} - E(1 - \kappa_i | \mathbf{X}_i) \mathbf{X}_i\|_{\Sigma}^2 \\ &= \|E(\kappa_i | \mathbf{X}_i) \boldsymbol{\theta}_{0i} - E(1 - \kappa_i | \mathbf{X}_i) \boldsymbol{\epsilon}_i\|_{\Sigma}^2 \\ &\leq 2E(\kappa_i | \mathbf{X}_i)^2 \|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 + 2E(1 - \kappa_i | \mathbf{X}_i)^2 \|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \\ &\leq 2K_s \tau_n + 2E(1 - \kappa_i | \mathbf{X}_i) \|\boldsymbol{\epsilon}_i\|_{\Sigma}^2. \end{aligned}$$

By Lemma 1, for any fixed $\eta < a$,

$$\begin{aligned} E(1 - \kappa_i | \mathbf{X}_i) &\leq K \tau_n^{\eta} \exp\left(\frac{\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2}\right) \\ &\leq \tau_n^{\eta} \exp(\|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2) \exp(\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2). \end{aligned}$$

We will show that, for small enough τ_n ,

$$\widehat{r}_i(\alpha, \tau_n) \geq \chi_{k, A\alpha}^2 c \tau_n (1 + o(1)), \quad (\text{B34})$$

for some $c > 0$ which we will fix later and any fixed $A > Km/(ac^a)$ with, specifically, $K = \int_0^{\infty} u^{-a-1} L(u) du$ here. Given this, for fixed $a/(1 + \rho) < \eta < a$ and for any fixed L ,

$$\begin{aligned} P_{\boldsymbol{\theta}_{0i}}(\boldsymbol{\theta}_{0i} \in \widehat{C}_i^R) &= P_{\boldsymbol{\theta}_{0i}}(\|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 \leq L \widehat{r}_i^{\alpha/(1+\rho)}(\alpha, \tau_n)) \\ &\geq P_{\boldsymbol{\theta}_{0i}}(K \tau_n + K \tau_n^{\eta} e^{\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2} \|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq K \tau_n^{\alpha/(1+\rho)} (1 + o(1))) \\ &\geq P_{\boldsymbol{\theta}_{0i}}(K \tau_n + K \tau_n^{\eta} e^{\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2} \|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq K \tau_n^{\alpha/(1+\rho)} (1 + o(1)) \mid \|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k, \alpha}^2) \\ &\quad \times P(\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k, \alpha}^2) \\ &\rightarrow 1 \times (1 - \alpha) = 1 - \alpha, \end{aligned}$$

as $\tau_n \rightarrow 0$, since the left hand side of the inequality in the conditional probability is of a higher order of infinitesimal.

Now, it remains to prove (B34). Due to the normality of the posterior

$$\boldsymbol{\theta}_i \mid \mathbf{X}_i, \lambda_i^2 \sim \mathbf{N}_k((1 - \kappa_i) \mathbf{X}_i, (1 - \kappa_i) \boldsymbol{\Sigma}),$$

applying Anderson's lemma, we have

$$\Pi(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 > \widehat{r}_i(\alpha, \tau_n) \mid \mathbf{X}_i, \lambda_i^2) \geq \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i) \mathbf{X}_i\|_{\Sigma}^2 > \widehat{r}_i(\alpha, \tau_n) \mid \mathbf{X}_i, \lambda_i^2).$$

Thus,

$$\begin{aligned} \alpha &= \int_0^\infty \Pi(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 > \widehat{r}_i(\alpha, \tau_n) \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\ &\geq \int_0^\infty \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 > \widehat{r}_i(\alpha, \tau_n) \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2. \end{aligned}$$

Recall that $\pi(\lambda_i^2 \mid \mathbf{X}_i) \propto (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) \exp(-\frac{\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i}{2(1 + \lambda_i^2 \tau_n)})$. Let $\widetilde{\pi}(\lambda_i^2 \mid \mathbf{X}_i) \propto (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2)$ be another density. Then, since $\pi(\lambda_i^2 \mid \mathbf{X}_i) / \widetilde{\pi}(\lambda_i^2 \mid \mathbf{X}_i)$ and $\Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 > \widehat{r}_i(\alpha, \tau_n) \mid \mathbf{X}_i, \lambda_i^2)$ are both increasing in λ_i^2

$$\begin{aligned} &\int_0^\infty \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 > \widehat{r}_i(\alpha, \tau_n) \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\ &\geq \int_0^\infty \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 > \widehat{r}_i(\alpha, \tau_n) \mid \mathbf{X}_i, \lambda_i^2) \widetilde{\pi}(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2. \end{aligned}$$

On the other hand, since $\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 / (1 - \kappa_i) \mid \mathbf{X}_i, \lambda_i^2 \sim \chi_k^2$, and $1 - \kappa_i \geq c\tau_n(1 + o(1))$ when $\lambda_i^2 \geq c$ for any $c > 0$. Thus, for some fixed $A(> 0)$ to be determined later,

$$\begin{aligned} &\int_0^\infty \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 > \chi_{k, A\alpha}^2 c\tau_n(1 + o(1)) \mid \mathbf{X}_i, \lambda_i^2) \widetilde{\pi}(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\ &\geq \int_c^\infty \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 > \chi_{k, A\alpha}^2 c\tau_n(1 + o(1)) \mid \mathbf{X}_i, \lambda_i^2) \widetilde{\pi}(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\ &\geq \int_c^\infty \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 > \chi_{k, A\alpha}^2 (1 - \kappa_i) \mid \mathbf{X}_i, \lambda_i^2) \widetilde{\pi}(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\ &= A\alpha \widetilde{\Pi}(\lambda_i^2 \geq c \mid \mathbf{X}_i). \end{aligned}$$

Since, by the dominated convergence theorem,

$$\begin{aligned} \widetilde{\Pi}(\lambda_i^2 \geq c \mid \mathbf{X}_i) &= \frac{\int_c^\infty (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) d\lambda_i^2}{\int_0^\infty (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) d\lambda_i^2} \\ &\rightarrow \frac{\int_c^\infty (\lambda_i^2)^{-a-1} L(\lambda_i^2) d\lambda_i^2}{\int_0^\infty (\lambda_i^2)^{-a-1} L(\lambda_i^2) d\lambda_i^2} \\ &\geq Km \int_c^\infty (\lambda_i^2)^{-a-1} d\lambda_i^2 = Km / (ac^a). \end{aligned}$$

First fixing c to be such that $Km/(ac^a) > \alpha$, if we further fix $A > ac^a/(Km)$ with $K = \int_0^\infty u^{-a-1}L(u)du$, then

$$\begin{aligned} & \int_0^\infty \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 > \chi_{k, A\alpha}^2 \tau_n / 2(1 + o(1)) \mid \mathbf{X}_i, \lambda_i^2) \tilde{\pi}(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\ & \geq \alpha AKm/(ac^a) \\ & > \alpha \\ & \geq \int_0^\infty \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 > \hat{r}_i(\alpha, \tau_n) \mid \mathbf{X}_i, \lambda_i^2) \tilde{\pi}(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2. \end{aligned}$$

This implies that $\hat{r}_i(\alpha, \tau_n) \geq \chi_{k, A\alpha}^2 c \tau_n (1 + o(1))$ which completes the proof of the first part.

Proof of (17) For the case where $f_{\tau_n} \tau_n \leq \|\boldsymbol{\theta}_{0i}\|_\Sigma^2 \leq K_M \log \frac{1}{\tau_n}$, we start with the inequality

$$\|\mathbf{X}_i\|_\Sigma \leq \|\boldsymbol{\theta}_{0i}\|_\Sigma + \|\boldsymbol{\epsilon}_{0i}\|_\Sigma.$$

For $K_0 > K_M$,

$$\|\mathbf{X}_i\|_\Sigma - \left(K_0 \log \frac{1}{\tau_n}\right)^{1/2} \leq \|\boldsymbol{\epsilon}_{0i}\|_\Sigma + \left(\sqrt{K_M} - \sqrt{K_0}\right) \left(\log \frac{1}{\tau_n}\right)^{1/2} \leq 0,$$

if

$$\|\boldsymbol{\epsilon}_{0i}\|_\Sigma \leq \left(\sqrt{K_0} - \sqrt{K_M}\right) \left(\log \frac{1}{\tau_n}\right)^{1/2},$$

the probability of which converges to 1 as $\tau_n \rightarrow 0$.

We now study $\hat{r}_i(\alpha, \tau_n)$ in this case. We will find an upper bound for $\hat{r}_i(\alpha, \tau_n)$. As the first ingredient, for $B > 0$,

$$\begin{aligned} & \Pi(\lambda_i^2 \geq B \mid \mathbf{X}_i) \\ &= \frac{\int_B^\infty (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) \exp\left(-\frac{\|\mathbf{X}_i\|_\Sigma^2}{2(1 + \lambda_i^2 \tau_n)}\right) d\lambda_i^2}{\int_0^\infty (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) \exp\left(-\frac{\|\mathbf{X}_i\|_\Sigma^2}{2(1 + \lambda_i^2 \tau_n)}\right) d\lambda_i^2}. \end{aligned}$$

The denominator

$$\begin{aligned}
 & \int_0^\infty (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) \exp\left(-\frac{\|\mathbf{X}_i\|_\Sigma^2}{2(1 + \lambda_i^2 \tau_n)}\right) d\lambda_i^2 \\
 & \geq \exp\left(-\frac{\|\mathbf{X}_i\|_\Sigma^2}{2}\right) \int_1^2 (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) d\lambda_i^2 \\
 & \geq \exp\left(-\frac{\|\mathbf{X}_i\|_\Sigma^2}{2}\right) (3)^{-k/2} L(1) \int_1^2 (\lambda_i^2)^{-a-1} d\lambda_i^2 \\
 & = K \exp\left(-\frac{\|\mathbf{X}_i\|_\Sigma^2}{2}\right).
 \end{aligned}$$

Hence,

$$\begin{aligned}
 \Pi(\lambda_i^2 \geq B \mid \mathbf{X}_i) & \leq K \exp\left(\frac{\|\mathbf{X}_i\|_\Sigma^2}{2}\right) \int_B^\infty (\lambda_i^2)^{-a-1} d\lambda_i^2 \\
 & = K \exp\left(\frac{\|\mathbf{X}_i\|_\Sigma^2}{2}\right) B^{-a}.
 \end{aligned} \tag{B35}$$

Next, by the inequality

$$\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 \leq 2\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 + 2\|\widehat{\boldsymbol{\theta}}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2,$$

for $r > 0$,

$$\frac{1}{2}\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 - \|\widehat{\boldsymbol{\theta}}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 \geq r$$

would imply that

$$\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 \geq r.$$

Now,

$$\begin{aligned}
 & \Pi\left(\frac{1}{2}\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 - \|\widehat{\boldsymbol{\theta}}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 \geq r \mid \mathbf{X}_i, \lambda_i^2\right) \\
 & \leq \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 \geq r \mid \mathbf{X}_i, \lambda_i^2).
 \end{aligned} \tag{B36}$$

Using (B36),

$$\begin{aligned}
& \int_0^\infty \Pi(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 \geq 2r+2 \sup_{\lambda_i^2 \leq B} \|\widehat{\boldsymbol{\theta}}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\
& \leq \int_0^B \Pi(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 \geq 2r+2 \sup_{\lambda_i^2 \leq B} \|\widehat{\boldsymbol{\theta}}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\
& \quad + \Pi(\lambda_i^2 \geq B \mid \mathbf{X}_i) \\
& \leq \int_0^B \Pi(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 \geq 2r + 2\|\widehat{\boldsymbol{\theta}}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\
& \quad + \Pi(\lambda_i^2 \geq B \mid \mathbf{X}_i) \\
& \leq \int_0^B \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 \geq r \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 + \Pi(\lambda_i^2 \geq B \mid \mathbf{X}_i).
\end{aligned}$$

Since $1 - \kappa_i \leq \frac{B\tau_n}{1+B\tau_n}$ when $\lambda_i^2 \leq B$, we can bound the first term above by $\alpha/2$ via

$$\begin{aligned}
& \int_0^B \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 \geq \chi_{k,\alpha/2}^2 \frac{B\tau_n}{1+B\tau_n} \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\
& \leq \int_0^B \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 \geq \chi_{k,\alpha/2}^2 (1 - \kappa_i) \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\
& \leq \alpha/2.
\end{aligned}$$

As for the second term, by (B35), when \mathbf{X}_i is fixed, for large enough B ,

$$\Pi(\lambda_i^2 \geq B \mid \mathbf{X}_i) \leq \alpha/2. \quad (\text{B37})$$

This leads to

$$\begin{aligned}
& \int_0^\infty \Pi(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 > \widehat{r}_i(\alpha, \tau_n) \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\
& = \alpha \\
& \geq \int_0^\infty \Pi(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 \geq 2r+2 \sup_{\lambda_i^2 \leq B} \|\widehat{\boldsymbol{\theta}}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2,
\end{aligned}$$

if $r = \chi_{k,\alpha/2}^2 B\tau_n / (1 + B\tau_n)$ and B is large enough. Thus, when $\|\mathbf{X}_i\|_\Sigma^2 \leq K_0 \log \frac{1}{\tau_n}$, for small enough τ_n ,

$$\begin{aligned} \widehat{r}_i(\alpha, \tau_n) &\leq 2\chi_{k,\alpha/2}^2 \frac{B\tau_n}{1 + B\tau_n} + 2 \sup_{\lambda_i^2 \leq B} \|\widehat{\boldsymbol{\theta}}_i - (1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 \\ &\leq 2\chi_{k,\alpha/2}^2 \frac{B\tau_n}{1 + B\tau_n} + 4 \sup_{\lambda_i^2 \leq B} \|E(1 - \kappa_i | \mathbf{X}_i)\mathbf{X}_i\|_\Sigma^2 + 4 \sup_{\lambda_i^2 \leq B} \|(1 - \kappa_i)\mathbf{X}_i\|_\Sigma^2 \\ &\leq 2\chi_{k,\alpha/2}^2 B\tau_n + 8B\tau_n \|\mathbf{X}_i\|_\Sigma^2 \\ &\leq KB\tau_n \log \frac{1}{\tau_n}. \end{aligned}$$

Again, applying Lemma 1, when $\|\mathbf{X}_i\|_\Sigma^2 \leq K_0 \log \frac{1}{\tau_n}$, for fixed $\eta < a$,

$$\begin{aligned} \|\widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 &\leq E(1 - \kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_\Sigma^2 \\ &\leq K\tau_n^\eta \exp\left(\|\mathbf{X}_i\|_\Sigma^2/2\right) \|\mathbf{X}_i\|_\Sigma^2 \\ &\leq K\tau_n^\eta \exp\left(K_0 \log \frac{1}{\tau_n}/2\right) K_0 \log \frac{1}{\tau_n} \\ &= K\tau_n^{\eta - K_0/2} \log \frac{1}{\tau_n}. \end{aligned}$$

If we choose B to be such that $B^a = K\tau_n^{-K_0/2}$, in which the factor K makes (B37) hold, we then have, as $\tau_n \rightarrow 0$,

$$2\|\widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 + 2\widehat{r}_i^{\alpha/(1+\rho)}(\alpha, \tau_n) \leq K\tau_n^{\eta - K_0/2} \log \frac{1}{\tau_n} + K \left(\tau_n^{1 - K_0/2a} \log \frac{1}{\tau_n}\right)^{\alpha/(1+\rho)} = o(1),$$

if we require $K_M < 2a$ and fix η and K_0 to be such that $K_M < K_0 < 2\eta < 2a$. This leads to that, when $\|\mathbf{X}_i\|_\Sigma^2 \leq K_0 \log \frac{1}{\tau_n}$, for some f_{τ_n} such that $f_{\tau_n} \rightarrow \infty$ and $f_{\tau_n}\tau_n \rightarrow 0$ as $\tau_n \rightarrow 0$,

$$\|\boldsymbol{\theta}_{0i}\|_\Sigma^2 \geq f_{\tau_n}\tau_n \geq 2\|\widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 + 2\widehat{r}_i^{\alpha/(1+\rho)}(\alpha, \tau_n).$$

Finally, by the inequality

$$\|\boldsymbol{\theta}_{0i}\|_\Sigma^2 \leq 2\|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 + 2\|\widehat{\boldsymbol{\theta}}_i\|_\Sigma^2,$$

the fact that

$$\|\boldsymbol{\theta}_{0i}\|_\Sigma^2 \geq 2\|\widehat{\boldsymbol{\theta}}_i\|_\Sigma^2 + 2\widehat{r}_i^{\alpha/(1+\rho)}(\alpha, \tau_n),$$

would imply that

$$\|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 \geq \widehat{r}_i^{a/(1+\rho)}(\alpha, \tau_n).$$

So,

$$\begin{aligned} & P_{\boldsymbol{\theta}_{0i}}(\boldsymbol{\theta}_{0i} \notin \widehat{C}_i^R) \\ &= P_{\boldsymbol{\theta}_{0i}}(\|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 > L\widehat{r}_i^{a/(1+\rho)}(\alpha, \tau_n)) \\ &\geq P_{\boldsymbol{\theta}_{0i}}(\|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 > L\widehat{r}_i^{a/(1+\rho)}(\alpha, \tau_n) \mid \|\boldsymbol{\epsilon}_{0i}\|_{\Sigma}^2 \leq (\sqrt{K_0} - \sqrt{K_M})^2 \log 1/\tau_n) \\ &\times P(\|\boldsymbol{\epsilon}_{0i}\|_{\Sigma}^2 \leq (\sqrt{K_0} - \sqrt{K_M})^2 \log 1/\tau_n) \\ &= P(\|\boldsymbol{\epsilon}_{0i}\|_{\Sigma}^2 \leq (\sqrt{K_0} - \sqrt{K_M})^2 \log 1/\tau_n) \\ &\rightarrow 1, \end{aligned}$$

as $\tau_n \rightarrow 0$, for any fixed $L > 0$.

Proof of (18). Consider the case where $\|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \geq K_L \log \frac{1}{\tau_n}$. We write

$$\begin{aligned} \|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma} &= \|\boldsymbol{\theta}_{0i} - \mathbf{X}_i - E(\kappa_i \mid \mathbf{X}_i)\mathbf{X}_i\|_{\Sigma} \\ &\leq \|\boldsymbol{\epsilon}_i\|_{\Sigma} + \|E(\kappa_i \mid \mathbf{X}_i)\mathbf{X}_i\|_{\Sigma}. \end{aligned}$$

Applying both Lemmas 2 and 3, for any fixed $\xi, \delta \in (0, 1)$,

$$\begin{aligned} \|E(\kappa_i \mid \mathbf{X}_i)\mathbf{X}_i\|_{\Sigma}^2 &\leq 2 \left(\frac{K}{\|\mathbf{X}_i\|_{\Sigma}^2} \right)^2 \|\mathbf{X}_i\|_{\Sigma}^2 \\ &\quad + 2 \left(K\tau_n^{-a} \exp\left(-\frac{\xi(1-\delta)}{2}\|\mathbf{X}_i\|_{\Sigma}^2\right) \right)^2 \|\mathbf{X}_i\|_{\Sigma}^2, \end{aligned}$$

Further,

$$\|\mathbf{X}_i\|_{\Sigma} = \|\boldsymbol{\theta}_{0i} + \boldsymbol{\epsilon}_i\|_{\Sigma} \geq \left| \|\boldsymbol{\theta}_{0i}\|_{\Sigma} - \|\boldsymbol{\epsilon}_i\|_{\Sigma} \right|,$$

in which $\|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \geq K_L \log \frac{1}{\tau_n}$, by assumption, and $\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k,\alpha}^2$ with probability $1 - \alpha$. Since $\tau_n \rightarrow 0$, when $\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k,\alpha}^2$, we will have

$$\|\mathbf{X}_i\|_{\Sigma} \geq \left(K_L \log \frac{1}{\tau_n} \right)^{1/2} - (\chi_{k,\alpha}^2)^{1/2}$$

and, consequently,

$$\|\mathbf{X}_i\|_{\Sigma}^2 \geq K_L \log \frac{1}{\tau_n} (1 + o(1)).$$

By choosing K_L , ξ , and δ to be such that $K_L(1+o(1)) > \frac{2a}{\xi(1-\delta)}$, e.g., $K_L = 3a$ and $\xi(1-\delta) = 3/4$, when $\|\epsilon_i\|_{\Sigma}^2 \leq \chi_{k,\alpha}^2$, we will have, as $\tau_n \rightarrow 0$,

$$\|E(\kappa_i | \mathbf{X}_i)\mathbf{X}_i\|_{\Sigma}^2 \leq o(1),$$

hence,

$$\|\theta_{0i} - \hat{\theta}_i\|_{\Sigma}^2 \leq \left(\sqrt{\chi_{k,\alpha}^2} + o(1)\right)^2 = \chi_{k,\alpha}^2(1 + o(1)). \tag{B38}$$

Then, we find a lower bound for $\hat{r}_i(\alpha, \tau_n)$ in this case. Making use of the posterior normality and Anderson’s lemma again,

$$\begin{aligned} \alpha &= \int_0^\infty \Pi(\|\theta_i - \hat{\theta}_i\|_{\Sigma}^2 > \hat{r}_i(\alpha, \tau_n) | \mathbf{X}_i, \lambda_i^2)\pi(\lambda_i^2 | \mathbf{X}_i)d\lambda_i^2 \\ &\geq \int_0^\infty \Pi(\|\theta_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \hat{r}_i(\alpha, \tau_n) | \mathbf{X}_i, \lambda_i^2)\pi(\lambda_i^2 | \mathbf{X}_i)d\lambda_i^2. \end{aligned}$$

On the other hand, since $1 - \kappa_i \geq g_{\tau_n}/(1 + g_{\tau_n}) = 1 + o(1)$, if $\lambda_i^2 \geq g_{\tau_n}/\tau_n$ with $g_{\tau_n} = (\log 1/\tau_n)^{1/3}$, for some fixed A ,

$$\begin{aligned} &\int_0^\infty \Pi(\|\theta_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \chi_{k,A\alpha}^2 g_{\tau_n}/(1 + g_{\tau_n}) | \mathbf{X}_i, \lambda_i^2)\pi(\lambda_i^2 | \mathbf{X}_i)d\lambda_i^2 \\ &\geq \int_{g_{\tau_n}/\tau_n}^\infty \Pi(\|\theta_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \chi_{k,A\alpha}^2 g_{\tau_n}/(1 + g_{\tau_n}) | \mathbf{X}_i, \lambda_i^2)\pi(\lambda_i^2 | \mathbf{X}_i)d\lambda_i^2 \\ &\geq \int_{g_{\tau_n}/\tau_n}^\infty \Pi(\|\theta_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \chi_{k,A\alpha}^2(1 - \kappa_i) | \mathbf{X}_i, \lambda_i^2)\pi(\lambda_i^2 | \mathbf{X}_i)d\lambda_i^2 \\ &= A\alpha \Pi(\lambda_i^2 \geq g_{\tau_n}/\tau_n | \mathbf{X}_i). \end{aligned}$$

We now study the posterior probability in a situation where $\|\mathbf{X}_i\|_{\Sigma}^2 \geq K_L \log \frac{1}{\tau_n}(1 + o(1))$:

$$\begin{aligned} &\Pi(\lambda_i^2 < g_{\tau_n}/\tau_n | \mathbf{X}_i) \\ &= \frac{\int_0^{g_{\tau_n}/\tau_n} (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + \lambda_i^2 \tau_n)}\right) d\lambda_i^2}{\int_0^\infty (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + \lambda_i^2 \tau_n)}\right) d\lambda_i^2} \\ &:= N/D. \end{aligned}$$

Next,

$$\begin{aligned}
D &\geq \int_{(g_{\tau_n}+1)/\tau_n}^{2(g_{\tau_n}+1)/\tau_n} (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + \lambda_i^2 \tau_n)}\right) d\lambda_i^2 \\
&\geq m(2g_{\tau_n} + 3)^{-k/2} \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(2 + g_{\tau_n})}\right) \int_{(g_{\tau_n}+1)/\tau_n}^{2(g_{\tau_n}+1)/\tau_n} (\lambda_i^2)^{-a-1} d\lambda_i^2 \\
&= K(2g_{\tau_n} + 3)^{-k/2} \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(2 + g_{\tau_n})}\right) \left(\frac{\tau_n}{g_{\tau_n} + 1}\right)^a.
\end{aligned}$$

Next, fixing a constant $c > 0$, when g_{τ_n} is large enough,

$$\begin{aligned}
N &= \int_0^{g_{\tau_n}/\tau_n} (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + \lambda_i^2 \tau_n)}\right) d\lambda_i^2 \\
&= \left(\int_0^{c/\tau_n} + \int_{c/\tau_n}^{g_{\tau_n}/\tau_n}\right) (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + \lambda_i^2 \tau_n)}\right) d\lambda_i^2 \\
&=: N_1 + N_2.
\end{aligned}$$

The first term

$$\begin{aligned}
N_1 &= \int_0^{c/\tau_n} (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + \lambda_i^2 \tau_n)}\right) d\lambda_i^2 \\
&\leq \int_0^{c/\tau_n} (\lambda_i^2)^{-a-1} L(\lambda_i^2) \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + c)}\right) d\lambda_i^2 \\
&\leq \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + c)}\right) \int_0^{\infty} (\lambda_i^2)^{-a-1} L(\lambda_i^2) d\lambda_i^2 \\
&= K \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + c)}\right).
\end{aligned}$$

When $\|\mathbf{X}_i\|_{\Sigma}^2 \geq K_L \log \frac{1}{\tau_n} (1 + o(1))$,

$$\begin{aligned}
N_1/D &\leq K(2g_{\tau_n} + 3)^{k/2} \left(\frac{g_{\tau_n} + 1}{\tau_n}\right)^a \exp\left(\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(2 + g_{\tau_n})} - \frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + c)}\right) \\
&\leq K(2g_{\tau_n} + 3)^{k/2} \left(\frac{g_{\tau_n} + 1}{\tau_n}\right)^a \exp\left(-\frac{K_L}{2} \log \frac{1}{\tau_n} \frac{1}{(1 + c)} (1 + o(1))\right) \\
&= K(2g_{\tau_n} + 3)^{k/2} (g_{\tau_n} + 1)^a \tau_n^{\frac{K_L}{2(1+c)}(1+o(1))-a} = o(1),
\end{aligned} \tag{B39}$$

as $\tau_n \rightarrow 0$, if c is chosen properly, e.g., $c = 1/3$. Our choice of K_L makes sure that the above is $o(1)$ as $\tau_n \rightarrow 0$, given that $g_{\tau_n} = (\log 1/\tau_n)^{1/3}$. Moreover,

$$\begin{aligned} N_2 &= \int_{c/\tau_n}^{g_{\tau_n}/\tau_n} (1 + \lambda_i^2 \tau_n)^{-k/2} (\lambda_i^2)^{-a-1} L(\lambda_i^2) \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + \lambda_i^2 \tau_n)}\right) d\lambda_i^2 \\ &\leq M(1 + c)^{-k/2} \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + g_{\tau_n})}\right) \int_{c/\tau_n}^{g_{\tau_n}/\tau_n} (\lambda_i^2)^{-a-1} d\lambda_i^2 \\ &\leq M(1 + c)^{-k/2} \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + g_{\tau_n})}\right) (\tau_n/c)^{a+1} \frac{g_{\tau_n} - c}{\tau_n}. \end{aligned}$$

It follows that

$$N_2/D \leq K(2g_{\tau_n} + 3)^{k/2} (g_{\tau_n} + 1)^a (g_{\tau_n} - c) \exp\left(-\frac{\|\mathbf{X}_i\|_{\Sigma}^2}{2(1 + g_{\tau_n})(2 + g_{\tau_n})}\right).$$

Since we have chosen $g_{\tau_n} = (\log 1/\tau_n)^{1/3}$, when $\|\mathbf{X}_i\|_{\Sigma}^2 \geq K_L \log \frac{1}{\tau_n} (1 + o(1))$, we will have, as $\tau_n \rightarrow 0$,

$$N_2/D = o(1). \tag{B40}$$

Combining (B39) and (B40), when $\|\mathbf{X}_i\|_{\Sigma}^2 \geq K_L \log \frac{1}{\tau_n} (1 + o(1))$, as $\tau_n \rightarrow 0$,

$$\Pi(\lambda_i^2 < g_{\tau_n}/\tau_n \mid \mathbf{X}_i) = o(1),$$

and hence

$$\begin{aligned} &\int_0^{\infty} \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \chi_{k, A\alpha}^2 g_{\tau_n}/(1 + g_{\tau_n}) \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\ &\geq A\alpha \Pi(\lambda_i^2 \geq g_{\tau_n}/\tau_n \mid \mathbf{X}_i) \\ &= A\alpha(1 - o(1)). \end{aligned}$$

Here, let $A = \beta/\alpha$ for some fixed $\beta > \alpha$. For small enough τ_n , we will have

$$\begin{aligned} &\int_0^{\infty} \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \chi_{k, \beta}^2 g_{\tau_n}/(1 + g_{\tau_n}) \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2 \\ &\geq \beta(1 - o(1)) \\ &> \alpha \\ &\geq \int_0^{\infty} \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \hat{r}_i(\alpha, \tau_n) \mid \mathbf{X}_i, \lambda_i^2) \pi(\lambda_i^2 \mid \mathbf{X}_i) d\lambda_i^2. \end{aligned}$$

It follows that

$$\widehat{r}_i(\alpha, \tau_n) \geq \chi_{k,\beta}^2 g_{\tau_n} / (1 + g_{\tau_n}) = \chi_{k,\beta}^2 (1 + o(1)), \quad (\text{B41})$$

when $\|\mathbf{X}_i\|_{\Sigma}^2 \geq K_L \log \frac{1}{\tau_n} (1 + o(1))$.

As a reminder, when τ_n is small enough, $\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k,\alpha}^2$ implies that $\|\mathbf{X}_i\|_{\Sigma}^2 \geq K_L \log \frac{1}{\tau_n} (1 + o(1))$.

So, by (B38) along with (B41), in the case where $\|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \geq K_L \log \frac{1}{\tau_n}$, for any fixed $\beta > \alpha$, as $\tau_n \rightarrow 0$,

$$\begin{aligned} P_{\boldsymbol{\theta}_{0i}}(\boldsymbol{\theta}_{0i} \in \widehat{C}_i^R) &= P_{\boldsymbol{\theta}_{0i}}(\|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 \leq L \widehat{r}_i^{a/(1+\rho)}(\alpha, \tau_n)) \\ &\geq P_{\boldsymbol{\theta}_{0i}}(\|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 \leq L \widehat{r}_i^{a/(1+\rho)}(\alpha, \tau_n) \mid \|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k,\alpha}^2) \\ &\quad \times P(\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k,\alpha}^2) \\ &\geq P_{\boldsymbol{\theta}_{0i}}(\chi_{k,\alpha}^2 (1 + o(1)) \leq L (\chi_{k,\beta}^2)^{a/(1+\rho)} (1 + o(1)) \mid \|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k,\alpha}^2) \\ &\quad \times P(\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k,\alpha}^2) \\ &\rightarrow 1 \times (1 - \alpha) = 1 - \alpha, \end{aligned}$$

if we choose $L > \chi_{k,\alpha}^2 (\chi_{k,\beta}^2)^{-a/(1+\rho)}$, e.g., $L = 2\chi_{k,\alpha}^2 (\chi_{k,\beta}^2)^{-a/(1+\rho)}$. \square

Proof of Theorem 7 We use $\widehat{\boldsymbol{\theta}}_i$ for $\widehat{\boldsymbol{\theta}}_i^{EIG}$ in this proof.

Proof of (20). We first find a lower bound for $\widehat{r}_i(\alpha, c_n)$. Recall that $\pi(\kappa_i \mid \mathbf{X}_i) \propto \kappa_i^{d+k/2-1} (1 - \kappa_i + c_n \kappa_i)^{-d-1} \exp(-\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i / 2)$. Similar to the proof of (16), let $\widetilde{\pi}(\kappa_i \mid \mathbf{X}_i) \propto \kappa_i^{d+k/2-1} (1 - \kappa_i + c_n \kappa_i)^{-d-1}$ be another density. Also, under the setup of this theorem,

$$\boldsymbol{\theta}_i \mid \mathbf{X}_i, \kappa_i \sim N_k((1 - \kappa_i) \mathbf{X}_i, (1 - \kappa_i) \boldsymbol{\Sigma}),$$

Now we can proceed similarly as the proof of (B34). For some $A(> 0)$ and some $v(> 1)$ to be determined later,

$$\begin{aligned} &\int_0^1 \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i) \mathbf{X}_i\|_{\Sigma}^2 > \chi_{k,A\alpha}^2 c_n^v \mid \mathbf{X}_i, \kappa_i) \widetilde{\pi}(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\ &\geq \int_0^{1-c_n^v} \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i) \mathbf{X}_i\|_{\Sigma}^2 > \chi_{k,A\alpha}^2 c_n^v \mid \mathbf{X}_i, \kappa_i) \widetilde{\pi}(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\ &\geq \int_0^{1-c_n^v} \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i) \mathbf{X}_i\|_{\Sigma}^2 > \chi_{k,A\alpha}^2 (1 - \kappa_i) \mid \mathbf{X}_i, \kappa_i) \widetilde{\pi}(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\ &= A\alpha \widetilde{\Pi}(\kappa_i \leq 1 - c_n^v \mid \mathbf{X}_i). \end{aligned}$$

Since

$$\begin{aligned}
 \tilde{\Pi}(\kappa_i > 1 - c_n^v \mid \mathbf{X}_i) &= \frac{\int_{1-c_n^v}^1 \kappa_i^{d+k/2-1} (1 - \kappa_i + c_n \kappa_i)^{-d-1} d\kappa_i}{\int_0^1 \kappa_i^{d+k/2-1} (1 - \kappa_i + c_n \kappa_i)^{-d-1} d\kappa_i} \\
 &\leq \frac{\int_{1-c_n^v}^1 c_n^{-d-1} d\kappa_i}{\int_{1-c_n}^1 (1 - c_n)^{d+k/2-1} (c_n + c_n(1 - c_n))^{-d-1} d\kappa_i} \\
 &= c_n^{v-1} \frac{(2 - c_n)^{d+1}}{(1 - c_n)^{d+k/2-1}},
 \end{aligned}$$

suppose $\alpha < 1/2$, when c_n is small enough, if we fix $A = 2$, for example,

$$A\alpha \tilde{\Pi}(\kappa_i \leq 1 - c_n^v \mid \mathbf{X}_i) > \alpha.$$

On the other hand, by Anderson's lemma,

$$\begin{aligned}
 \alpha &= \int_0^1 \Pi(\|\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 > \hat{r}_i(\alpha, c_n) \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\
 &\geq \int_0^1 \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \hat{r}_i(\alpha, c_n) \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i.
 \end{aligned}$$

So,

$$\begin{aligned}
 &\int_0^1 \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \chi_{k, A\alpha}^2 c_n^v \mid \mathbf{X}_i, \kappa_i) \tilde{\pi}(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\
 &> \alpha \geq \int_0^1 \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \hat{r}_i(\alpha, c_n) \mid \mathbf{X}_i, \kappa_i) \tilde{\pi}(\kappa_i \mid \mathbf{X}_i) d\kappa_i.
 \end{aligned}$$

This implies that

$$\hat{r}_i(\alpha, c_n) \geq \chi_{k, A\alpha}^2 c_n^v. \quad (\text{B42})$$

For the case where $\|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \leq K'_S c_n$. Again,

$$\begin{aligned}
 \|\boldsymbol{\theta}_{0i} - \hat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 &= \|\boldsymbol{\theta}_{0i} - E(1 - \kappa_i \mid \mathbf{X}_i)\mathbf{X}_i\|_{\Sigma}^2 \\
 &= \|E(\kappa_i \mid \mathbf{X}_i)\boldsymbol{\theta}_{0i} - E(1 - \kappa_i \mid \mathbf{X}_i)\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \\
 &\leq 2E(\kappa_i \mid \mathbf{X}_i)^2 \|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 + 2E(1 - \kappa_i \mid \mathbf{X}_i)^2 \|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \\
 &\leq 2K_s \tau_n + 2E(1 - \kappa_i \mid \mathbf{X}_i) \|\boldsymbol{\epsilon}_i\|_{\Sigma}^2.
 \end{aligned}$$

Using Lemma 4 and (B42), when we choose v to be such that $vd/(1+\rho) < d < 1$, e.g., $v = 1 + \rho/2$,

$$\begin{aligned}
P_{\theta_{0i}}(\boldsymbol{\theta}_{0i} \in \widehat{C}_i^{EIG}) &= P_{\theta_{0i}}(\|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 \leq L\widehat{r}_i^{d/(1+\rho)}(\alpha, c_n)) \\
&\geq P_{\theta_{0i}}(Kc_n + Kc_n^d e^{\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2} \|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq (\chi_{k, A\alpha}^2 c_n^v)^{d/(1+\rho)}) \\
&\geq P_{\theta_{0i}}(Kc_n + Kc_n^d e^{\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2} \|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq (\chi_{k, A\alpha}^2 c_n^v)^{d/(1+\rho)} \mid \|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k, \alpha}^2) \\
&\quad \times P(\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k, \alpha}^2) \\
&\rightarrow 1 \times (1 - \alpha) = 1 - \alpha,
\end{aligned}$$

as $c_n \rightarrow 0$, since the left hand side of the inequality in the conditional probability is of a higher order of infinitesimal.

Proof of (21) For the case where $f'_{c_n} c_n \leq \|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \leq K'_M \log \frac{1}{c_n}$, similar to the proof of (17), for $K'_0 > K'_M$,

$$\|\mathbf{X}_i\|_{\Sigma} - \left(K'_0 \log \frac{1}{c_n}\right)^{1/2} \leq \|\boldsymbol{\epsilon}_{0i}\|_{\Sigma} + \left(\sqrt{K'_M} - \sqrt{K'_0}\right) \left(\log \frac{1}{c_n}\right)^{1/2} \leq 0,$$

if

$$\|\boldsymbol{\epsilon}_{0i}\|_{\Sigma} \leq \left(\sqrt{K'_0} - \sqrt{K'_M}\right) \left(\log \frac{1}{c_n}\right)^{1/2},$$

the probability of which converges to 1 as $c_n \rightarrow 0$.

Using (B36) again,

$$\begin{aligned}
&\int_0^1 \Pi(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 \geq 2r + 2 \sup_{\kappa_i \geq B_n} \|\widehat{\boldsymbol{\theta}}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\
&\leq \int_{B_n}^1 \Pi(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 \geq 2r + 2 \sup_{\kappa_i \geq B_n} \|\widehat{\boldsymbol{\theta}}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\
&\quad + \Pi(\kappa_i < B_n \mid \mathbf{X}_i) \\
&\leq \int_{B_n}^1 \Pi(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 \geq 2r + 2\|\widehat{\boldsymbol{\theta}}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\
&\quad + \Pi(\kappa_i < B_n \mid \mathbf{X}_i) \\
&\leq \int_{B_n}^1 \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 \geq r \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i + \Pi(\kappa_i < B_n \mid \mathbf{X}_i).
\end{aligned}$$

For the first term, let $r = \chi_{k,\alpha/2}^2(1 - B_n)$,

$$\begin{aligned} & \int_{B_n}^1 \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 \geq \chi_{k,\alpha/2}^2(1 - B_n) \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\ & \leq \int_{B_n}^1 \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 \geq \chi_{k,\alpha/2}^2(1 - \kappa_i) \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\ & \leq \alpha/2. \end{aligned}$$

For the other term, when \mathbf{X}_i is fixed, if $B_n = 1 - c_n^{d/(2d+2)}$, as $c_n \rightarrow 0$,

$$\begin{aligned} & \Pi(\kappa_i < B_n \mid \mathbf{X}_i) \\ & = \frac{\int_0^{1-c_n^{d/(2d+2)}} \kappa_i^{d+k/2-1} (1 - \kappa_i + c_n \kappa_i)^{-d-1} \exp(-\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i/2) d\kappa_i}{\int_0^1 \kappa_i^{d+k/2-1} (1 - \kappa_i + c_n \kappa_i)^{-d-1} \exp(-\kappa_i \mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i/2) d\kappa_i} \\ & \leq \frac{\int_0^{1-c_n^{d/(2d+2)}} (c_n^{d/(2d+2)} + c_n(1 - c_n^{d/(2d+2)}))^{-d-1} d\kappa_i}{\exp(-\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i/2) \int_{1-c_n}^1 (1 - c_n)^{d+k/2-1} (c_n + c_n(1 - c_n))^{-d-1} d\kappa_i} \\ & = \frac{c_n^{-d/2} (1 + c_n^{(d+2)/(2d+2)} - c_n)^{-d-1}}{\exp(-\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i/2) c_n^{-d} (1 - c_n)^{d+k/2-1} (2 - c_n)^{-d-1}} \\ & \leq K \exp(\mathbf{X}_i^T \boldsymbol{\Sigma}^{-1} \mathbf{X}_i/2) c_n^{d/2} \leq \alpha/2. \end{aligned}$$

So,

$$\begin{aligned} & \int_0^1 \Pi(\|\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 > \hat{r}_i(\alpha, c_n) \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\ & = \alpha \\ & \geq \int_0^1 \Pi(\|\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 \geq 2r + 2 \sup_{\kappa_i \geq B_n} \|\hat{\boldsymbol{\theta}}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i, \end{aligned}$$

if $r = \chi_{k,\alpha/2}^2(1 - B_n)$ and $B_n = 1 - c_n^{d/(2d+2)}$. Thus, when $\|\mathbf{X}_i\|_{\Sigma}^2 \leq K'_0 \log \frac{1}{c_n}$, for small enough c_n ,

$$\begin{aligned} \hat{r}_i(\alpha, c_n) & \leq 2\chi_{k,\alpha/2}^2(1 - B_n) + 2 \sup_{\kappa_i \geq B_n} \|\hat{\boldsymbol{\theta}}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 \\ & \leq 2\chi_{k,\alpha/2}^2 c_n^{d/(2d+2)} + 4 \sup_{\kappa_i \geq B_n} E(1 - \kappa_i \mid \mathbf{X}_i) \mathbf{X}_i^T \boldsymbol{\Sigma} \mathbf{X}_i + 4 \sup_{\kappa_i \geq B_n} \|(1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 \\ & \leq 2\chi_{k,\alpha/2}^2 c_n^{d/(2d+2)} + 8c_n^{d/(2d+2)} \|\mathbf{X}_i\|_{\Sigma}^2 \\ & \leq K c_n^{d/(2d+2)} \log \frac{1}{c_n}. \end{aligned}$$

Finally, by the inequality

$$\|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \leq 2\|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 + 2\|\widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2,$$

the fact that

$$\|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \geq 2\|\widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 + 2\widehat{r}_i^{d/(1+\rho)}(\alpha, \tau_n),$$

would imply that

$$\|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 \geq \widehat{r}_i^{d/(1+\rho)}(\alpha, \tau_n).$$

Applying Lemma 4, when $\|\mathbf{X}_i\|_{\Sigma}^2 \leq K'_0 \log \frac{1}{c_n}$,

$$\begin{aligned} \|\widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 &\leq E(1 - \kappa_i | \mathbf{X}_i) \|\mathbf{X}_i\|_{\Sigma}^2 \\ &\leq K c_n^d \exp(\|\mathbf{X}_i\|_{\Sigma}^2/2) \|\mathbf{X}_i\|_{\Sigma}^2 \\ &\leq K c_n^d \exp\left(K'_0 \log \frac{1}{c_n}/2\right) K'_0 \log \frac{1}{c_n} \\ &= K c_n^{d-K'_0/2} \log \frac{1}{c_n}. \end{aligned}$$

We then have, as $c_n \rightarrow 0$,

$$2\|\widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 + 2\widehat{r}_i^{d/(1+\rho)}(\alpha, c_n) \leq K c_n^{d-K'_0/2} \log \frac{1}{c_n} + K \left(c_n^{d/(2d+2)} \log \frac{1}{c_n}\right)^{d/(1+\rho)} = o(1),$$

if we require $K_M < 2d$ and fix K'_0 to be such that $K_M < K'_0 < 2d$. This implies that when $\|\mathbf{X}_i\|_{\Sigma}^2 \leq K'_0 \log \frac{1}{c_n}$, for some f_{τ_n} such that $f'_{c_n} \rightarrow \infty$ and $f'_{c_n} c_n \rightarrow 0$ as $c_n \rightarrow 0$,

$$\|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \geq f'_{c_n} c_n \geq 2\|\widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 + 2\widehat{r}_i^{d/(1+\rho)}(\alpha, c_n).$$

Finally,

$$\begin{aligned} &P_{\boldsymbol{\theta}_{0i}}(\boldsymbol{\theta}_{0i} \notin \widehat{C}_i^{EIG}) \\ &= P_{\boldsymbol{\theta}_{0i}}(\|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 > L\widehat{r}_i^{d/(1+\rho)}(\alpha, \tau_n)) \\ &\geq P_{\boldsymbol{\theta}_{0i}}(\|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 > L\widehat{r}_i^{d/(1+\rho)}(\alpha, \tau_n) \mid \|\boldsymbol{\epsilon}_{0i}\|_{\Sigma}^2 \leq (\sqrt{K_0} - \sqrt{K_M})^2 \log 1/\tau_n) \\ &\times P(\|\boldsymbol{\epsilon}_{0i}\|_{\Sigma}^2 \leq (\sqrt{K_0} - \sqrt{K_M})^2 \log 1/\tau_n) \\ &= P(\|\boldsymbol{\epsilon}_{0i}\|_{\Sigma}^2 \leq (\sqrt{K_0} - \sqrt{K_M})^2 \log 1/\tau_n) \\ &\rightarrow 1, \end{aligned}$$

as $c_n \rightarrow 0$, for any fixed $L > 0$.

Proof of (22). Similar to the proof of (18), using Lemmas 5 and 6, when $\|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \geq K'_L \log \frac{1}{c_n}$ and $\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k,\alpha}^2$, we have, as $c_n \rightarrow 0$,

$$\|\mathbf{X}_i\|_{\Sigma}^2 \geq K'_L \log \frac{1}{c_n} (1 + o(1)),$$

and, for $K'_L(1 + o(1)) > \frac{2d}{\xi(1-\delta)}$,

$$\|\boldsymbol{\theta}_{0i} - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 \leq \chi_{k,\alpha}^2 (1 + o(1)). \tag{B43}$$

Then, we find a lower bound for $\widehat{r}_i(\alpha, c_n)$ in this case. Making use of the posterior normality and Anderson's lemma again,

$$\begin{aligned} \alpha &= \int_0^1 \Pi(\|\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 > \widehat{r}_i(\alpha, c_n) \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\ &\geq \int_0^1 \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \widehat{r}_i(\alpha, c_n) \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i. \end{aligned}$$

On the other hand, since $1 - \kappa_i \geq 1 - c_n$ if $\kappa_i \leq c_n$, for some fixed A ,

$$\begin{aligned} &\int_0^1 \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \chi_{k,A\alpha}^2 (1 - c_n) \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\ &\geq \int_0^{c_n} \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \chi_{k,A\alpha}^2 (1 - c_n) \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\ &\geq \int_0^{c_n} \Pi(\|\boldsymbol{\theta}_i - (1 - \kappa_i)\mathbf{X}_i\|_{\Sigma}^2 > \chi_{k,A\alpha}^2 (1 - \kappa_i) \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\ &= A\alpha \Pi(\kappa_i \leq c_n \mid \mathbf{X}_i). \end{aligned}$$

When $\|\mathbf{X}_i\|_{\Sigma}^2 \geq K'_L \log \frac{1}{c_n} (1 + o(1))$, as $n \rightarrow \infty$,

$$\begin{aligned}
& P(\kappa_i \geq c_n \mid \mathbf{X}_i) \\
& \leq \frac{\int_{c_n}^1 \kappa_i^{d+k/2-1} (1 - \kappa_i + c_n \kappa_i)^{-d-1} \exp(-\kappa_i \mathbf{X}_i^T \Sigma^{-1} \mathbf{X}_i / 2) d\kappa_i}{\int_{1-c_n^{1/d}}^1 \kappa_i^{d+k/2-1} (1 - \kappa_i + c_n \kappa_i)^{-d-1} \exp(-\kappa_i \mathbf{X}_i^T \Sigma^{-1} \mathbf{X}_i / 2) d\kappa_i} \\
& \leq \frac{\int_{c_n}^1 c_n^{-d-1} \exp(-c_n \mathbf{X}_i^T \Sigma^{-1} \mathbf{X}_i / 2) d\kappa_i}{\int_{1-c_n^{1/d}}^1 (1 - c_n^{1/d})^{d+k/2-1} (c_n^{1/d} + c_n(1 - c_n^{1/d}))^{-d-1} \exp(-\mathbf{X}_i^T \Sigma^{-1} \mathbf{X}_i / 2) d\kappa_i} \\
& = c_n^{-d} \exp(-(1 - c_n) \mathbf{X}_i^T \Sigma^{-1} \mathbf{X}_i / 2) (1 + o(1)) \\
& \leq c_n^{-d} \exp\left(- (1 - c_n) K'_L \log \frac{1}{c_n} (1 + o(1)) / 2\right) (1 + o(1)) \\
& = c_n^{(1-c_n)K'_L(1+o(1))/2-d} = o(1).
\end{aligned}$$

since we have already chosen $K'_L(1 + o(1)) > \frac{2d}{\xi(1-\delta)} > 2d$.

Let $A = \beta/\alpha$ for some fixed $\beta > \alpha$. For small enough c_n , we will have

$$\begin{aligned}
& \int_0^1 P(\|\boldsymbol{\theta}_i - (1 - \kappa_i) \mathbf{X}_i\|_{\Sigma}^2 > \chi_{k, A\alpha}^2 (1 - c_n) \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i \\
& \geq \beta(1 - o(1)) \\
& > \alpha \\
& \geq \int_0^1 P(\|\boldsymbol{\theta}_i - (1 - \kappa_i) \mathbf{X}_i\|_{\Sigma}^2 > \hat{r}_i(\alpha, c_n) \mid \mathbf{X}_i, \kappa_i) \pi(\kappa_i \mid \mathbf{X}_i) d\kappa_i.
\end{aligned}$$

It follows that, as $c_n \rightarrow 0$,

$$\hat{r}_i(\alpha, c_n) \geq \chi_{k, \beta}^2 (1 - c_n) = \chi_{k, \beta}^2 (1 + o(1)), \quad (\text{B44})$$

when $\|\mathbf{X}_i\|_{\Sigma}^2 \geq K'_L \log \frac{1}{c_n} (1 + o(1))$.

So, by (B43) along with (B44), in the case where $\|\boldsymbol{\theta}_{0i}\|_{\Sigma}^2 \geq K'_L \log \frac{1}{c_n}$, for any fixed $\beta > \alpha$, as $c_n \rightarrow 0$,

$$\begin{aligned}
P_{\boldsymbol{\theta}_{0i}}(\boldsymbol{\theta}_{0i} \in \hat{\mathcal{C}}_i^{EIG}) & = P_{\boldsymbol{\theta}_{0i}}(\|\boldsymbol{\theta}_{0i} - \hat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 \leq L \hat{r}_i^{d/(1+\rho)}(\alpha, c_n)) \\
& \geq P_{\boldsymbol{\theta}_{0i}}(\|\boldsymbol{\theta}_{0i} - \hat{\boldsymbol{\theta}}_i\|_{\Sigma}^2 \leq L \hat{r}_i^{d/(1+\rho)}(\alpha, c_n) \mid \|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k, \alpha}^2) \\
& \quad \times P(\|\boldsymbol{\epsilon}_i\|_{\Sigma}^2 \leq \chi_{k, \alpha}^2)
\end{aligned}$$

$$\begin{aligned}
&\geq P_{\theta_{0i}}(\chi_{k,\alpha}^2(1+o(1)) \leq L(\chi_{k,\beta}^2)^{d/(1+\rho)}(1+o(1)) \mid \|\epsilon_i\|_{\Sigma}^2 \leq \chi_{k,\alpha}^2) \\
&\times P(\|\epsilon_i\|_{\Sigma}^2 \leq \chi_{k,\alpha}^2) \\
&\rightarrow 1 \times (1 - \alpha) = 1 - \alpha,
\end{aligned}$$

if we choose $L > \chi_{k,\alpha}^2(\chi_{k,\beta}^2)^{-d/(1+\rho)}$, e.g., $L = 2\chi_{k,\alpha}^2(\chi_{k,\beta}^2)^{-d/(1+\rho)}$. □

ZIKUN QIN AND MALAY GHOSH

DEPARTMENT OF STATISTICS, UNIVERSITY OF FLORIDA, 102 GRIFFIN-FLOYD HALL,
GAINESVILLE 32611 FLORIDA, USA

E-mail: qinzikun@ufl.edu

E-mail: ghoshm@ufl.edu

Paper received: 5 June 2022