

# Supplementary Materials for ‘Large-Scale Multiple Hypothesis Testing with the Normal-Beta Prime Prior’

Ray Bai <sup>a</sup> and Malay Ghosh <sup>b</sup>

<sup>a</sup> Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, Philadelphia, PA 19104 USA; <sup>b</sup> Department of Statistics, University of Florida, Gainesville, FL 32611 USA

## ARTICLE HISTORY

Compiled July 20, 2019

## ABSTRACT

In Section A, we provide all the proofs of the main results. In Section B, we provide the technical details for implementing the empirical Bayes and fully Bayes variants of the NBP model. In Section C, we compare the posterior shrinkage weights to the theoretical posterior inclusion probabilities.

## Appendix A. Proofs of Main Results

### A.1. Proofs for Section 2.1

**Proof of Proposition 2.1.** As noted by Proposition 2.1 in [1], the beta prime density (7) can be rewritten as a product of independent gamma and inverse gamma densities. We thus reparametrize model (6) for a single observation  $\theta$  as follows:

$$\begin{aligned}\theta|\lambda_i\xi &\sim \mathcal{N}(0, \lambda\xi), \\ \lambda &\sim \mathcal{G}(a, 1), \\ \xi &\sim \mathcal{IG}(b, 1).\end{aligned}\tag{A1}$$

From (A1), we see that the joint distribution of the prior is proportional to

$$\begin{aligned}\pi(\theta, \lambda, \xi) &\propto (\lambda\xi)^{-1/2} \exp\left(-\frac{\theta^2}{2\lambda\xi}\right) \lambda^{a-1} \exp(-\lambda) \exp\left(-\frac{1}{\xi}\right) \xi^{-b-1} \\ &= \lambda^{a-3/2} \exp(-\lambda) \xi^{-b-3/2} \exp\left(-\left(\frac{\theta^2}{2\lambda} + 1\right) \frac{1}{\xi}\right).\end{aligned}$$

Thus,

$$\begin{aligned}\pi(\theta, \lambda) &\propto \lambda^{a-3/2} \exp(-\lambda) \int_{\xi=0}^{\infty} \xi^{-b-3/2} \exp\left(-\left(\frac{\theta^2}{2\lambda} + 1\right) \frac{1}{\xi}\right) d\xi \\ &\propto \left(\frac{\theta^2}{2\lambda} + 1\right)^{-(b+1/2)} \lambda^{a-3/2} e^{-\lambda},\end{aligned}$$

and thus, the marginal density of  $\theta$  is proportional to

$$\pi(\theta) \propto \int_0^{\infty} \left(\frac{\theta^2}{2\lambda} + 1\right)^{-(b+1/2)} \lambda^{a-3/2} e^{-\lambda} d\lambda. \quad (\text{A2})$$

As  $|\theta| \rightarrow 0$ , the expression in (A2) is bounded below by

$$C \int_0^{\infty} \lambda^{a-3/2} e^{-\lambda} d\lambda, \quad (\text{A3})$$

where  $C$  is a constant that depends on  $a$  and  $b$ . The integral expression in (A3) clearly diverges to  $\infty$  for any  $0 < a \leq 1/2$ . Therefore, (A2) diverges to infinity as  $|\theta| \rightarrow 0$ , by the monotone convergence theorem.  $\square$

**Proof of Theorem 2.1.** From (9), the posterior distribution of  $\kappa_i$  under  $\text{NBP}_n$  is proportional to

$$\pi(\kappa_i | X_i) \propto \exp\left(-\frac{\kappa_i X_i^2}{2}\right) \kappa_i^{b-1/2} (1 - \kappa_i)^{a_n-1}, \quad \kappa_i \in (0, 1). \quad (\text{A4})$$

Hence,

$$\begin{aligned}\mathbb{E}(1 - \kappa_i | X_i) &= \frac{\int_0^1 \kappa_i^{b-1/2} (1 - \kappa_i)^{a_n} \exp\left(-\frac{\kappa_i X_i^2}{2}\right) d\kappa_i}{\int_0^1 \kappa_i^{b-1/2} (1 - \kappa_i)^{a_n-1} \exp\left(-\frac{\kappa_i X_i^2}{2}\right) d\kappa_i} \\ &\leq \frac{e^{X_i^2/2} \int_0^1 \kappa_i^{b-1/2} (1 - \kappa_i)^{a_n} d\kappa_i}{\int_0^1 \kappa_i^{b-1/2} (1 - \kappa_i)^{a_n-1} d\kappa_i} \\ &= e^{X_i^2/2} \frac{\Gamma(a_n + 1) \Gamma(b + 1/2)}{\Gamma(a_n + b + 3/2)} \times \frac{\Gamma(a_n + b + 1/2)}{\Gamma(a_n) \Gamma(b + 1/2)} \\ &= e^{X_i^2/2} \left(\frac{a_n}{a_n + b + 1/2}\right).\end{aligned}$$

$\square$

**Proof of Theorem 2.2.** Note that since  $b \in (\frac{1}{2}, \infty)$ ,  $\kappa_i^{b-1/2}$  is increasing in  $\kappa_i$  on  $(0, 1)$ . Additionally, since  $a_n \in (0, 1)$ ,  $(1 - \kappa_i)^{a_n-1}$  is increasing in  $\kappa_i$  on  $(0, 1)$ . Using these facts, we have

$$\begin{aligned}
\Pr(\kappa_i < \epsilon | X_i) &\leq \frac{\int_0^\epsilon \exp\left(-\frac{\kappa_i X_i^2}{2}\right) \kappa_i^{b-1/2} (1-\kappa_i)^{a_n-1} d\kappa_i}{\int_\epsilon^1 \exp\left(-\frac{\kappa_i X_i^2}{2}\right) \kappa_i^{b-1/2} (1-\kappa_i)^{a_n-1} d\kappa_i} \\
&\leq \frac{e^{X_i^2/2} \int_0^\epsilon \kappa_i^{b-1/2} (1-\kappa_i)^{a_n-1} d\kappa_i}{\int_\epsilon^1 \kappa_i^{b-1/2} (1-\kappa_i)^{a_n-1} d\kappa_i} \\
&\leq \frac{e^{X_i^2/2} (1-\epsilon)^{a_n-1} \int_0^\epsilon \kappa_i^{b-1/2} d\kappa_i}{\epsilon^{b-1/2} \int_\epsilon^1 (1-\kappa_i)^{a_n-1} d\kappa_i} \\
&= \frac{e^{X_i^2/2} (1-\epsilon)^{a_n-1} \left(b + \frac{1}{2}\right)^{-1} \epsilon^{b+1/2}}{a_n^{-1} \epsilon^{b-1/2} (1-\epsilon)^{a_n}} \\
&= e^{X_i^2/2} \frac{a_n \epsilon}{(b + 1/2)(1-\epsilon)}.
\end{aligned}$$

□

**Proof of Theorem 2.3.** Letting  $C$  denote the normalizing constant, we have

$$\begin{aligned}
\int_0^\eta \pi(\kappa_i | X_i) d\kappa_i &= C \int_0^\eta \exp\left(-\frac{\kappa_i X_i^2}{2}\right) \kappa_i^{b-1/2} (1-\kappa_i)^{a_n-1} d\kappa_i \\
&\geq C \int_0^{\eta\delta} \exp\left(-\frac{\kappa_i X_i^2}{2}\right) \kappa_i^{b-1/2} (1-\kappa_i)^{a_n-1} d\kappa_i \\
&\geq C \exp\left(-\frac{\eta\delta}{2} X_i^2\right) \int_0^{\eta\delta} \kappa_i^{b-1/2} d\kappa_i \\
&= C \exp\left(-\frac{\eta\delta}{2} X_i^2\right) \left(b + \frac{1}{2}\right)^{-1} (\eta\delta)^{b+\frac{1}{2}}. \tag{A5}
\end{aligned}$$

Also, since  $b \in (\frac{1}{2}, \infty)$ ,  $\kappa_i^{b-1/2}$  is increasing in  $\kappa_i$  on  $(0, 1)$ .

$$\begin{aligned}
\int_\eta^1 \pi(\kappa_i | X_i) d\kappa_i &= C \int_\eta^1 \exp\left(-\frac{\kappa_i X_i^2}{2}\right) \kappa_i^{b-1/2} (1-\kappa_i)^{a_n-1} d\kappa_i \\
&\leq C \exp\left(-\frac{\eta X_i^2}{2}\right) \int_\eta^1 \kappa_i^{b-1/2} (1-\kappa_i)^{a_n-1} d\kappa_i \\
&\leq C \exp\left(-\frac{\eta X_i^2}{2}\right) \int_\eta^1 (1-\kappa_i)^{a_n-1} d\kappa_i \\
&= C \exp\left(-\frac{\eta X_i^2}{2}\right) a_n^{-1} (1-\eta)^{a_n}. \tag{A6}
\end{aligned}$$

Combining (A5) and (A6), we have

$$\Pr(\kappa_i > \eta | X_i) \leq \frac{\int_{\eta}^1 \pi(\kappa_i | X_i) d\kappa_i}{\int_0^{\eta} \pi(\kappa_i | X_i) d\kappa_i} \leq \frac{(b + \frac{1}{2})(1 - \eta)^{a_n}}{a_n(\eta\delta)^{b + \frac{1}{2}}} \exp\left(-\frac{\eta(1 - \delta)}{2} X_i^2\right).$$

□

### A.2. Proofs for Section 3.3

Our proof methods follow those of [2–4], except our arguments rely on control of the sequence of hyperparameters  $a_n$ , rather than on specifying a rate or an estimate for a rescaling parameter  $\tau$ , as in the class of priors (5). Moreover, we make explicit use of Theorems 2.1–2.3 in the present manuscript in our proofs.

**Proof of Theorem 3.1.** By Theorem 2.1, the event  $\{\mathbb{E}(1 - \kappa_i | X_i) > \frac{1}{2}\}$  implies the event

$$\left\{ e^{X_i^2/2} \left( \frac{a_n}{a_n + b + 1/2} \right) > \frac{1}{2} \right\} \Leftrightarrow \left\{ X_i^2 > 2 \log \left( \frac{a_n + b + 1/2}{2a_n} \right) \right\}.$$

Therefore, noting that under  $H_{0i}$ ,  $X_i \sim \mathcal{N}(0, 1)$  and using Mill's ratio, i.e.  $P(|Z| > x) \leq \frac{2\phi(x)}{x}$ , we have

$$\begin{aligned} t_{1i} &\leq \Pr \left( X_i^2 > 2 \log \left( \frac{a_n + b + 1/2}{2a_n} \right) \middle| H_{0i} \text{ is true} \right) \\ &= \Pr \left( |Z| > \sqrt{2 \log \left( \frac{a_n + b + 1/2}{2a_n} \right)} \right) \\ &\leq \frac{2\phi \left( \sqrt{2 \log \left( \frac{a_n + b + 1/2}{2a_n} \right)} \right)}{\sqrt{2 \log \left( \frac{a_n + b + 1/2}{2a_n} \right)}} \\ &= \frac{2\sqrt{2}a_n}{\sqrt{\pi}(a_n + b + 1/2)} \left[ \log \left( \frac{a_n + b + 1/2}{2a_n} \right) \right]^{-1/2}. \end{aligned} \tag{A7}$$

□

**Proof of Theorem 3.2.** By definition, the probability of a Type I error for the  $i$ th decision is given by

$$t_{1i} = \Pr \left[ \mathbb{E}(1 - \kappa_i | X_i) > \frac{1}{2} \middle| H_{0i} \text{ is true} \right].$$

Fix  $\xi \in (0, 1/2)$ . By Theorem 2.3,

$$\mathbb{E}(\kappa_i | X_i) \leq \xi + \frac{(b + \frac{1}{2})(1 - \xi)^{a_n}}{a_n(\xi\delta)^{b + \frac{1}{2}}} \exp\left(-\frac{\xi(1 - \delta)}{2} X_i^2\right).$$

Hence,

$$\left\{ \mathbb{E}(1 - \kappa_i | X_i) > \frac{1}{2} \right\} \supseteq \left\{ \frac{(b + \frac{1}{2})(1 - \xi)^{a_n}}{a_n(\xi\delta)^{b + \frac{1}{2}}} \exp\left(-\frac{\xi(1 - \delta)}{2} X_i^2\right) < \frac{1}{2} - \xi \right\}.$$

Thus, using the definition of  $t_{1i}$  and noting that under  $H_{0i}$ ,  $X_i \sim \mathcal{N}(0, 1)$ , as  $n \rightarrow \infty$ ,

$$\begin{aligned} t_{1i} &\geq \Pr\left(\frac{(b + \frac{1}{2})(1 - \xi)^{a_n}}{a_n(\xi\delta)^{b + \frac{1}{2}}} \exp\left(-\frac{\xi(1 - \delta)}{2} X_i^2\right) < \frac{1}{2} - \xi \mid H_{0i} \text{ is true}\right) \\ &= \Pr\left(X_i^2 > \frac{2}{\xi(1 - \delta)} \left[\log\left(\frac{(b + \frac{1}{2})(1 - \xi)^{a_n}}{a_n(\xi\delta)^{b + \frac{1}{2}}(\frac{1}{2} - \xi)}\right)\right]\right) \\ &= 2 \Pr\left(Z > \sqrt{\frac{2}{\xi(1 - \delta)} \left[\log\left(\frac{(b + \frac{1}{2})(1 - \xi)^{a_n}}{a_n(\xi\delta)^{b + \frac{1}{2}}(\frac{1}{2} - \xi)}\right)\right]}\right) \\ &= 2 \left(1 - \Phi\left(\sqrt{\frac{2}{\xi(1 - \delta)} \left[\log\left(\frac{(b + \frac{1}{2})(1 - \xi)^{a_n}}{a_n(\xi\delta)^{b + \frac{1}{2}}(\frac{1}{2} - \xi)}\right)\right]}\right)\right), \end{aligned}$$

where for the second to last inequality, we used the fact that  $a_n \rightarrow 0$  as  $n \rightarrow \infty$ , and the fact that both  $\xi$  and  $\xi\delta \in (0, \frac{1}{2})$ , so that the  $\log(\cdot)$  term in final equality is greater than zero for sufficiently large  $n$ .  $\square$

Before proving the asymptotic upper bound on Type II error in Theorem 3.3, we first prove a lemma that bounds the quantity  $\mathbb{E}(\kappa_i | X_i)$  from above for a single  $X_i$ .

**Lemma A.1.** *Suppose we observe  $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\theta}, \mathbf{I}_n)$  and we place an NBP $_n$  prior (10) on  $\boldsymbol{\theta}$ , with and  $a_n \in (0, 1)$  where  $a_n \rightarrow 0$  as  $n \rightarrow \infty$ , and fixed  $b \in (1/2, \infty)$ . Fix constants  $\eta \in (0, 1)$ ,  $\delta \in (0, 1)$ , and  $d > 2$ . Then for a single observation  $x$  and any  $n$ , the posterior shrinkage coefficient  $\mathbb{E}(\kappa | x)$  can be bounded above by a measurable, non-negative real-valued function  $h_n(x)$ , given by*

$$h_n(x) = \begin{cases} C_{n,\eta} \left[ x^2 \int_0^{x^2} t^{b-1/2} e^{-t/2} dt \right]^{-1} + \frac{(b + \frac{3}{2})^{-1} (1 - \eta)^{a_n}}{a_n (\eta \delta)^{b + 3/2}} \exp\left(-\frac{\eta(1 - \delta)}{2} x^2\right), & \text{if } |x| > 0, \\ 1, & \text{if } x = 0, \end{cases} \quad (\text{A8})$$

where  $C_{n,\eta} = (1 - \eta)^{a_n - 1} \Gamma(b + \frac{3}{2}) 2^{b + 3/2}$ . For any  $\rho > \frac{2}{\eta(1 - \delta)}$ ,  $h_n(x)$  also satisfies

$$\lim_{n \rightarrow \infty} \sup_{|x| > \sqrt{\rho \log\left(\frac{1}{a_n}\right)}} h_n(x) = 0. \quad (\text{A9})$$

**Proof of Lemma A.1.** We first focus on the case where  $|x| > 0$ . Fix  $\eta \in (0, 1)$ ,  $\delta \in (0, 1)$ , and observe that

$$\mathbb{E}(\kappa | x) = \mathbb{E}(\kappa 1\{\kappa < \eta\} | x) + \mathbb{E}(\kappa 1\{\kappa \geq \eta\} | x). \quad (\text{A10})$$

We consider the two terms in (A10) separately. To bound the first term, we have from

(9) and the fact that  $(1 - \kappa)^{a_n - 1}$  is increasing in  $\kappa \in (0, 1)$  for  $a_n \in (0, 1)$  that

$$\begin{aligned}
\mathbb{E}(\kappa 1\{\kappa < \eta\}) &= \frac{\int_0^\eta \kappa \cdot \kappa^{b-1/2} (1 - \kappa)^{a_n - 1} e^{-\kappa x^2/2} d\kappa}{\int_0^1 \kappa^{b-1/2} (1 - \kappa)^{a_n - 1} e^{-\kappa x^2/2} d\kappa} \\
&\leq (1 - \eta)^{a_n - 1} \frac{\int_0^\eta \kappa^{b+1/2} e^{-\kappa x^2/2} d\kappa}{\int_0^1 \kappa^{b-1/2} e^{-\kappa x^2/2} d\kappa} \\
&= (1 - \eta)^{a_n - 1} \frac{1}{x^2} \frac{\int_0^{\eta x^2} t^{b+1/2} e^{-t/2} dt}{\int_0^{x^2} t^{b-1/2} e^{-t/2} dt} \\
&\leq (1 - \eta)^{a_n - 1} \frac{1}{x^2} \frac{\int_0^\infty t^{b+1/2} e^{-t/2} dt}{\int_0^{x^2} t^{b-1/2} e^{-t/2} dt} \\
&= C(n) \left[ x^2 \int_0^{x^2} t^{b-1/2} e^{-t/2} dt \right]^{-1} \\
&:= h_1(x) \quad (\text{say}), \tag{A11}
\end{aligned}$$

where we use a change of variables  $t = \kappa x^2$  in the second equality, and  $C(n) = (1 - \eta)^{a_n - 1} \Gamma(b + \frac{3}{2}) 2^{b+3/2}$ .

To bound the second term in (A10) from above, we follow the same steps as the proof of Theorem 2.3, except we replace  $\kappa_i^{b-1/2}$  in the numerators of the integrands with  $\kappa_i^{b+1/2}$  to obtain an upper bound,

$$\frac{(b + \frac{3}{2}) (1 - \eta)^{a_n}}{a_n (\eta \delta)^{b+3/2}} \exp\left(-\frac{\eta(1 - \delta)}{2} x^2\right) := h_2(x) \quad (\text{say}). \tag{A12}$$

Combining (A10)-(A12), we set  $h_n(x) = h_1(x) + h_2(x)$  for any  $|x| > 0$ , and we easily see that for any  $x \neq 0$  and fixed  $n$ ,  $\mathbb{E}(\kappa|x) \leq h_n(x)$ . On the other hand, if  $x = 0$ , then

$$\mathbb{E}(\kappa|x) = \frac{\int_0^1 \kappa^{b+1/2} (1 - \kappa)^{a_n - 1} d\kappa}{\int_0^1 \kappa^{b-1/2} (1 - \kappa)^{a_n - 1} d\kappa} = \frac{b + 1/2}{a_n + b + 1/2} \leq 1,$$

so we can set  $h_n(x) = 1$  when  $x = 0$ . Therefore,  $\mathbb{E}(\kappa|x)$  is bounded above by the function  $h_n(x)$  in (A8) for all  $x \in \mathbb{R}$ .

Now, observe from (A11) that for fixed  $n$ ,  $h_1(x)$  is strictly decreasing in  $|x|$ . Therefore,

$$\sup_{|x| > \sqrt{\rho \log\left(\frac{1}{a_n}\right)}} h_1(x) \leq C_{n,\eta} \left[ \rho \log\left(\frac{1}{a_n}\right) \int_0^{\rho \log\left(\frac{1}{a_n}\right)} t^{b-1/2} e^{-t/2} dt \right]^{-1},$$

for any fixed  $n$  and  $\rho > 0$ . Since  $a_n \rightarrow 0$  as  $n \rightarrow \infty$ , this implies that

$$\lim_{n \rightarrow \infty} \sup_{|x| > \sqrt{\rho \log\left(\frac{1}{a_n}\right)}} h_1(x) = 0. \tag{A13}$$

Letting  $K \equiv K(b, \eta, \delta) = (b + \frac{3}{2}) / (\eta\delta)^{b+3/2}$ , we have from (A12) and the fact that  $0 < a_n < 1$  for all  $n$  and  $a_n \rightarrow 0$  as  $n \rightarrow \infty$  that

$$\begin{aligned} \lim_{n \rightarrow \infty} h_2 \left( \sqrt{\rho \log \left( \frac{1}{a_n} \right)} \right) &= K \lim_{n \rightarrow \infty} \frac{(1-\eta)^{a_n}}{a_n} \sqrt{\rho \log \left( \frac{1}{a_n} \right)} e^{-\frac{\eta(1-\delta)}{2} \rho \log \left( \frac{1}{a_n} \right)} \\ &\leq K \sqrt{\rho} \lim_{n \rightarrow \infty} \frac{1}{a_n} \sqrt{\log \left( \frac{1}{a_n} \right)} e^{-\frac{\eta(1-\delta)}{2} \log(a_n^{-\rho})} \\ &= K \sqrt{\rho} \lim_{n \rightarrow \infty} \sqrt{\log \left( \frac{1}{a_n} \right)} (a_n)^{\frac{\eta(1-\delta)}{2} (\rho - \frac{2}{\eta(1-\delta)})} \\ &= \begin{cases} 0 & \text{if } \rho > \frac{2}{\eta(1-\delta)}, \\ \infty & \text{otherwise,} \end{cases} \end{aligned}$$

from which it follows that

$$\lim_{n \rightarrow \infty} \sup_{|x| > \sqrt{\rho \log \left( \frac{1}{a_n} \right)}} h_2(x) = \begin{cases} 0 & \text{if } \rho > \frac{2}{\eta(1-\delta)}, \\ \infty & \text{otherwise.} \end{cases} \quad (\text{A14})$$

Combining (A13) and (A14), it is clear that

$$\lim_{n \rightarrow \infty} \sup_{|x| > \sqrt{\rho \log \left( \frac{1}{a_n} \right)}} h_n(x) = \begin{cases} 0 & \text{if } \rho > \frac{2}{\eta(1-\delta)}, \\ \infty & \text{otherwise,} \end{cases}$$

that is,  $h_n(x)$  satisfies (A9).  $\square$

**Proof of Theorem 3.3.** Fix  $\eta \in (0, 1)$  and  $\delta \in (0, 1)$ , and choose any  $\rho > \frac{2}{\eta(1-\delta)}$ . By Lemma A.1, we have that the event  $\{\mathbb{E}(\kappa_i | X_i) \geq 0.5\}$  implies  $\{h_n(X_i) \geq 0.5\}$ , where  $h_n(x)$  is as defined in (A8). Therefore,

$$\begin{aligned} t_{2i} &= \Pr[\mathbb{E}(\kappa_i | X_i) \geq 0.5 | H_{1i} \text{ is true}] \\ &\leq \Pr(h_n(X_i) \geq 0.5 | H_{1i} \text{ is true}) \\ &= \Pr \left( h_n(X_i) \geq 0.5, |X_i| > \sqrt{\rho \log \left( \frac{1}{a_n} \right)} \mid H_{1i} \text{ is true} \right) + \\ &\quad \Pr \left( h_n(X_i) \geq 0.5, |X_i| \leq \sqrt{\rho \log \left( \frac{1}{a_n} \right)} \mid H_{1i} \text{ is true} \right) \\ &\leq \Pr \left( h_n(X_i) \geq 0.5 \mid |X_i| > \sqrt{\rho \log \left( \frac{1}{a_n} \right)}, H_{1i} \text{ is true} \right) + \\ &\quad \Pr \left( |X_i| \leq \sqrt{\rho \log \left( \frac{1}{a_n} \right)} \mid H_{1i} \text{ is true} \right) \end{aligned} \quad (\text{A15})$$

We will consider the two terms in (A15) separately. Recall that  $h_n(x)$  from (A8) is a measurable and nonnegative. We also see that (A8) is decreasing in  $|x|$ , and thus,

$$\mathbb{E} \left( h_n(X_i) \middle| |X_i| > \sqrt{\rho \log \left( \frac{1}{a_n} \right)}, H_{1i} \text{ is true} \right)$$

is well-defined and bounded for sufficiently large  $n$ . By Markov's inequality, we have for sufficiently large  $n$ ,

$$\begin{aligned} & \Pr \left( h_n(X_i) \geq 0.5 \middle| |X_i| > \sqrt{\rho \log \left( \frac{1}{a_n} \right)}, H_{1i} \text{ is true} \right) \\ & \leq 2 \mathbb{E} \left( h_n(X_i) \middle| |X_i| > \sqrt{\rho \log \left( \frac{1}{a_n} \right)}, H_{1i} \text{ is true} \right) \\ & \leq 2 \left( \sup_{|X_i| > \sqrt{\rho \log \left( \frac{1}{a_n} \right)}} h_n(X_i) \right), \end{aligned}$$

from which it follows, by Lemma A.1, that

$$\lim_{n \rightarrow \infty} \Pr \left( h_n(X_i) \geq 0.5 \middle| |X_i| > \sqrt{\rho \log \left( \frac{1}{a_n} \right)}, H_{1i} \text{ is true} \right) = 0. \quad (\text{A16})$$

By assumption,  $\lim_{n \rightarrow \infty} \frac{a_n}{p_n} \in (0, \infty)$ . Thus, by the third and fourth conditions of Assumption 1, we have  $\lim_{n \rightarrow \infty} \log \left( \frac{1}{a_n} \right) / \psi_n^2 = C/2$ . To see this, note that  $\frac{1-p_n}{p_n} \sim \frac{1}{p_n}$ . Combining this with the third and fourth conditions implies that  $\frac{2 \log(1/p_n)}{\psi_n^2} \rightarrow C$ , and then we use our assumption that  $a_n/p_n \rightarrow d, d > 0$ . Thus, for all sufficiently large  $n$ ,

$$\begin{aligned} & \Pr \left( |X_i| \leq \sqrt{\rho \log \left( \frac{1}{a_n} \right)} \middle| H_{1i} \text{ is true} \right) = \Pr \left( |Z| \leq \sqrt{\rho} \sqrt{\frac{\log \left( \frac{1}{a_n} \right)}{1 + \psi_n^2}} \right) \\ & = \Pr \left( |Z| \leq \sqrt{\rho} \sqrt{\frac{\log \left( \frac{1}{a_n} \right)}{\psi_n^2}} (1 + o(1)) \right) \text{ as } n \rightarrow \infty \\ & = \Pr \left( |Z| \leq \sqrt{\frac{\rho C}{2}} (1 + o(1)) \right) \text{ as } n \rightarrow \infty \\ & = \left[ 2\Phi \left( \sqrt{\frac{\rho C}{2}} \right) - 1 \right] (1 + o(1)) \text{ as } n \rightarrow \infty. \end{aligned} \quad (\text{A17})$$



Combining (A15)-(A17), we thus have

$$t_{2i} \leq \left[ 2\Phi \left( \sqrt{\frac{\rho C}{2}} \right) - 1 \right] (1 + o(1)),$$

as  $n \rightarrow \infty$ . □

**Proof of Theorem 3.4.** By definition, the probability of a Type II error for the  $i$ th decision is given by

$$t_{2i} = P \left( \mathbb{E}(1 - \kappa_i) \leq \frac{1}{2} \middle| H_{1i} \text{ is true} \right).$$

For any  $n$ , we have by Theorem 2.1 that

$$\left\{ e^{X_i^2/2} \left( \frac{a_n}{a_n + b + 1/2} \right) \leq \frac{1}{2} \right\} \subseteq \left\{ \mathbb{E}(1 - \kappa_i | X_i) \leq \frac{1}{2} \right\}.$$

Therefore,

$$\begin{aligned} t_{2i} &= \Pr \left( \mathbb{E}(1 - \kappa_i | X_i) \leq \frac{1}{2} \middle| H_{1i} \text{ is true} \right) \\ &\geq \Pr \left( e^{X_i^2/2} \left( \frac{a_n}{a_n + b + 1/2} \right) \leq \frac{1}{2} \middle| H_{1i} \text{ is true} \right) \\ &= \Pr \left( X_i^2 \leq 2 \log \left( \frac{a_n + b + 1/2}{2a_n} \right) \middle| H_{1i} \text{ is true} \right). \end{aligned} \quad (\text{A18})$$

Since  $X_i \sim N(0, 1 + \psi^2)$  under  $H_{1i}$ , we have by the second condition in Assumption 1 that  $\lim_{n \rightarrow \infty} \frac{\psi_n^2}{1 + \psi_n^2} \rightarrow 1$ . From (A18), we have for sufficiently large  $n$ ,

$$\begin{aligned} t_{2i} &\geq \Pr \left( |Z| \leq \sqrt{\frac{2 \log \left( \frac{a_n + b + 1/2}{2a_n} \right)}{\psi^2}} (1 + o(1)) \right) \text{ as } n \rightarrow \infty \\ &\geq \Pr \left( |Z| \leq \sqrt{\frac{\log \left( \frac{1}{2a_n} \right)}{\psi^2}} (1 + o(1)) \right) \text{ as } n \rightarrow \infty \\ &= \Pr(|Z| \leq \sqrt{C})(1 + o(1)) \text{ as } n \rightarrow \infty \\ &= 2[\Phi(\sqrt{C}) - 1](1 + o(1)) \text{ as } n \rightarrow \infty, \end{aligned}$$

where we used the assumption that  $\lim_{n \rightarrow \infty} \frac{a_n}{p_n} \in (0, \infty)$  and Assumption 1. □

**Proof of Theorem 3.5.** Fix  $\eta \in (0, 1)$ ,  $\delta \in (0, 1)$ , and  $\xi \in (0, 1/2)$ , and choose  $\rho > \frac{2}{\eta(1-\delta)}$ . Since the  $\kappa_i$ 's,  $i = 1, \dots, n$  are *a posteriori* independent, the Type I and Type II error probabilities  $t_{1i}$  and  $t_{2i}$  are the same for every test  $i$ ,  $i = 1, \dots, n$ . By Theorems 3.1 and 3.2, we have for large enough  $n$ ,

$$\begin{aligned}
& 2 \left( 1 - \Phi \left( \sqrt{\frac{2}{\xi(1-\delta)} \left[ \log \left( \frac{(b+\frac{1}{2})(1-\xi)^{a_n}}{a_n(\xi\delta)^{b+\frac{1}{2}}(\frac{1}{2}-\xi)} \right) \right]} \right) \right) \leq t_{1i} \\
& \leq \frac{2\sqrt{2}a_n}{\sqrt{\pi}(a_n+b+1/2)} \left[ \log \left( \frac{a_n+b+1/2}{2a_n} \right) \right]^{-1/2}.
\end{aligned}$$

Taking the limit as  $n \rightarrow \infty$  of all the terms above, we have

$$\lim_{n \rightarrow \infty} t_{1i} = 0 \quad (\text{A19})$$

for the  $i$ th test, under the assumptions on the hyperparameters  $a_n$  and  $b$ .

By Theorems 3.1 and 3.2, we also have

$$\left[ 2\Phi(\sqrt{C}) - 1 \right] (1 + o(1)) \leq t_{2i} \leq \left[ 2\Phi \left( \sqrt{\frac{\rho C}{2}} \right) - 1 \right] (1 + o(1)). \quad (\text{A20})$$

Therefore, we have by (A19) and (A20) that as  $n \rightarrow \infty$ , the asymptotic risk (14) of the classification rule (19),  $R_{NBP}$ , can be bounded as follows:

$$np(2\Phi(\sqrt{C}) - 1)(1 + o(1)) \leq R_{NBP} \leq np \left[ 2\Phi \left( \sqrt{\frac{\rho C}{2}} \right) - 1 \right] (1 + o(1)). \quad (\text{A21})$$

Therefore, from (16) and (A21), we have as  $n \rightarrow \infty$ ,

$$1 \leq \liminf_{n \rightarrow \infty} \frac{R_{NBP}}{R_{Opt}^{BO}} \leq \limsup_{n \rightarrow \infty} \frac{R_{NBP}}{R_{Opt}^{BO}} \leq \frac{2\Phi \left( \sqrt{\frac{\rho C}{2}} \right) - 1}{2\Phi(\sqrt{C}) - 1}. \quad (\text{A22})$$

The testing rule (18) does not depend on how  $\eta \in (0, 1)$ ,  $\delta \in (0, 1)$  and  $\rho > 2/(\eta(1-\delta))$  are chosen, and thus, the ratio  $R_{NBP}/R_{Opt}^{BO}$  is also free of these constants. By continuity of  $\Phi$ , we can take the infimum over all  $\rho$ 's in the rightmost term in (A22), and the inequalities remain valid. The infimum of  $\rho$  is obviously 2, and so from (A22), we have

$$1 \leq \liminf_{n \rightarrow \infty} \frac{R_{NBP}}{R_{Opt}^{BO}} \leq \limsup_{n \rightarrow \infty} \frac{R_{NBP}}{R_{Opt}^{BO}} \leq \frac{2\Phi(\sqrt{C}) - 1}{2\Phi(\sqrt{C}) - 1}. \quad (\text{A23})$$

We clearly see from (A23) that classification rule (19) under the  $NBP_n$  prior (10) is ABOS, i.e.

$$\frac{R_{NBP}}{R_{Opt}^{BO}} \rightarrow 1 \text{ as } n \rightarrow \infty.$$

□

### A.3. Proofs for Section 3.4

Our proofs in this section follow from the proof of Theorem 10 of [4], as well as Theorems 2.1 through Theorem 2.3 established in this paper.

**Proof of Theorem 3.6.** Under thresholding rule (23), the probability of a Type I error for the  $i$ th decision is given by

$$\begin{aligned}\tilde{t}_{1i} &= \Pr\left(\mathbb{E}(1 - \kappa_i | X_i, \hat{a}_n^{ES}) > \frac{1}{2} \middle| H_{0i} \text{ is true}\right) \\ &= \Pr\left(\mathbb{E}(1 - \kappa_i | X_i, \hat{a}_n^{ES}) > \frac{1}{2}, \hat{a}_n^{ES} \leq 2\alpha_n \middle| H_{0i} \text{ is true}\right) \\ &\quad + \Pr\left(\mathbb{E}(1 - \kappa_i | X_i, \hat{a}_n^{ES}) > \frac{1}{2}, \hat{a}_n^{ES} > 2\alpha_n \middle| H_{0i} \text{ is true}\right),\end{aligned}\tag{A24}$$

where  $\alpha_n$  is defined in (24). To obtain an upper bound on  $\tilde{t}_{1i}$ , we consider the two terms in (A24) separately. By Theorem 2.1, we see that  $\mathbb{E}(1 - \kappa_i | X_i)$  is nondecreasing in  $a_n$ . Thus,  $\mathbb{E}(1 - \kappa_i | X_i, \hat{a}_n^{ES}) \leq \mathbb{E}(1 - \kappa_i | X_i, 2\alpha_n)$  whenever  $\hat{a}_n^{ES} \leq 2\alpha_n$ . We have

$$\begin{aligned}\Pr\left(\mathbb{E}(1 - \kappa_i | X_i, \hat{a}_n^{ES}) > \frac{1}{2}, \hat{a}_n^{ES} \leq 2\alpha_n \middle| H_{0i} \text{ is true}\right) \\ \leq \Pr\left(\mathbb{E}(1 - \kappa_i | X_i, 2\alpha_n) > \frac{1}{2} \middle| H_{0i} \text{ is true}\right) \\ \leq \frac{4\alpha_n}{\sqrt{\pi}(2\alpha_n + b + 1/2)} \left[\log\left(\frac{2\alpha_n + b + 1/2}{4\alpha_n}\right)\right]^{-1/2} (1 + o(1)).\end{aligned}\tag{A25}$$

For the second term in (A24), we have

$$\begin{aligned}\Pr\left(\mathbb{E}(1 - \kappa_i | X_i, \hat{a}_n^{ES}) > \frac{1}{2}, \hat{a}_n^{ES} > 2\alpha_n \middle| H_{0i} \text{ is true}\right) \\ \leq \Pr(\hat{a}_n^{ES} > 2\alpha_n | H_{0i} \text{ is true}) \\ \leq \frac{1/\sqrt{\pi}}{n^{c_1/2}\sqrt{\log n}} + e^{-(2\log 2 - 1)\alpha_n(1+o(1))},\end{aligned}\tag{A26}$$

where the last inequality follows from the proof of Theorem 10 in [4]. Thus, since  $\alpha_n \sim 2\beta p_n$  by (25), we combine (A25) and (A26) to obtain an upper bound on  $\tilde{t}_{1i}$ ,

$$\begin{aligned}\tilde{t}_{1i} &\leq \frac{4\alpha_n}{\sqrt{\pi}(2\alpha_n + b + 1/2)} \left[\log\left(\frac{2\alpha_n + b + 1/2}{4\alpha_n}\right)\right]^{-1/2} (1 + o(1)) \\ &\quad + \frac{1/\sqrt{\pi}}{n^{c_1/2}\sqrt{\log n}} + e^{-2(2\log 2 - 1)\beta n p_n(1+o(1))}.\end{aligned}$$

To obtain the lower bound, note that by (A24), we immediately have

$$\tilde{t}_{1i} \geq \Pr\left(\mathbb{E}(1 - \kappa_i | X_i, \hat{a}_n^{ES}) > \frac{1}{2}, \hat{a}_n^{ES} \leq 2\alpha_n \middle| H_{0i} \text{ is true}\right).\tag{A27}$$

By the proof for Theorem 3.2, we have that for fixed  $\xi \in (0, 1/2)$  and  $\delta \in (0, 1)$ ,

$$\mathbb{E}(\kappa_i | X_i) \leq \xi + \frac{(b + \frac{1}{2})(1 - \xi)^{a_n}}{a_n(\xi\delta)^{b+\frac{1}{2}}} \exp\left(-\frac{\xi(1 - \delta)}{2} X_i^2\right). \quad (\text{A28})$$

The right-hand side of (A28) is a nonincreasing function in  $a_n$ . Thus, whenever  $\widehat{a}_n^{ES} \leq 2\alpha_n$ , we have

$$\left\{ \mathbb{E}(1 - \kappa_i | X_i, \widehat{a}_n) > \frac{1}{2}, \widehat{a}_n^{ES} \leq 2\alpha_n \right\} \supseteq \left\{ \frac{(b + \frac{1}{2})(1 - \xi)^{2\alpha_n}}{2\alpha_n(\xi\delta)^{b+1/2}} \exp\left(-\frac{\xi(1 - \delta)}{2} X_i^2\right) < \frac{1}{2} - \xi \right\},$$

from which, by Theorem 3.2 and (A27), we automatically attain the lower bound,

$$\tilde{t}_{1i} \geq 1 - \Phi\left(\sqrt{\frac{2}{\xi(1 - \delta)} \left[ \log\left(\frac{(b + \frac{1}{2})(1 - \xi)^{2\alpha_n}}{2\alpha_n(\xi\delta)^{b+1/2}}\right) \right]}\right) (1 + o(1)) \text{ as } n \rightarrow \infty.$$

□

**Proof of Theorem 3.7.** Fix  $\gamma \in (0, 1/c_2)$ . Decompose the probability of a Type II error under (23) as

$$\begin{aligned} \tilde{t}_{2i} &= \Pr\left(\mathbb{E}(\kappa_i | X_i, \widehat{a}_n^{ES}) \geq \frac{1}{2} \mid H_{1i} \text{ is true}\right) \\ &= \Pr\left(\mathbb{E}(\kappa_i | X_i, \widehat{a}_n^{ES}) \geq \frac{1}{2}, \widehat{a}_n^{ES} \leq \gamma\alpha_n \mid H_{1i} \text{ is true}\right) + \\ &\quad \Pr\left(\mathbb{E}(\kappa_i | X_i, \widehat{a}_n^{ES}) \geq \frac{1}{2}, \widehat{a}_n^{ES} > \gamma\alpha_n \mid H_{1i} \text{ is true}\right). \end{aligned} \quad (\text{A29})$$

To obtain an upper bound on  $\tilde{t}_{2i}$ , we consider the two terms in (A29) separately. For the first term in (A29), we have

$$\begin{aligned} &\Pr\left(\mathbb{E}(\kappa_i | X_i, \widehat{a}_n^{ES}) \geq \frac{1}{2}, \widehat{a}_n^{ES} \leq \gamma\alpha_n \mid H_{1i} \text{ is true}\right) \\ &\leq \Pr(\widehat{a}_n^{ES} \leq \gamma\alpha_n \mid H_{1i} \text{ is true}) \\ &\leq \frac{(1 - c_2\gamma)^{-2}(1 - \alpha_n)}{n\alpha_n} (1 + o(1)) \\ &\rightarrow 0 \text{ as } n \rightarrow \infty, \end{aligned} \quad (\text{A30})$$

where the last two steps follow from the proof of Theorem 11 in [4].

We now focus on bounding the second term in (A29). By Theorem 2.1,  $\mathbb{E}(1 - \kappa_i | X_i)$  is nondecreasing in  $a_n$ , and so  $\mathbb{E}(\kappa_i | X_i)$  is nonincreasing in  $a_n$ . Thus, for sufficiently large  $n$ , we have  $\mathbb{E}(\kappa_i | X_i, \widehat{a}_n^{ES}) \leq \mathbb{E}(\kappa_i | X_i, \gamma\alpha_n)$  for  $\widehat{a}_n^{ES} > \gamma\alpha_n$  and that

$$\{\mathbb{E}(\kappa_i | X_i, \gamma\alpha_n) \geq 0.5 \mid H_{1i} \text{ is true}\} \subseteq \{h_n(X_i, \gamma\alpha_n) \geq 0.5 \mid H_{1i} \text{ is true}\},$$

where  $h_n(X_i, \gamma\alpha_n)$  denotes that we substitute  $a_n$  with  $\gamma\alpha_n$  in (A8). Using the same arguments as in the proof of Theorem 3.3, along with the fact that  $\alpha_n \sim 2\beta p_n$  (by (24)), we obtain as an upper bound for the second term in (A29),

$$\begin{aligned} & \Pr\left(\mathbb{E}(\kappa_i|X_i, \widehat{a}_n^{ES}) \geq \frac{1}{2}, \widehat{a}_n^{ES} > \gamma\alpha_n \mid H_{1i} \text{ is true}\right) \\ & \leq \Pr\left(\mathbb{E}(\kappa_i|X_i, \gamma\alpha_n) \geq \frac{1}{2} \mid H_{1i} \text{ is true}\right) \\ & \leq \left[2\Phi\left(\sqrt{\frac{\rho C}{2}}\right) - 1\right] (1 + o(1)) \text{ as } n \rightarrow \infty. \end{aligned} \quad (\text{A31})$$

From (A29)-(A31), an upper bound on the probability of Type II error under (23) is

$$\tilde{t}_{2i} \leq \left[2\Phi\left(\sqrt{\frac{\rho C}{2}}\right) - 1\right] (1 + o(1)) \text{ as } n \rightarrow \infty.$$

To obtain a lower bound on  $\tilde{t}_{2i}$ , we note that by (A29),

$$\begin{aligned} \tilde{t}_{2i} & \geq \Pr\left(\mathbb{E}(\kappa_i|X_i, \widehat{a}_n^{ES}) \geq \frac{1}{2}, \widehat{a}_n^{ES} > \gamma\alpha_n \mid H_{1i} \text{ is true}\right) \\ & \geq \Pr\left(\mathbb{E}(\kappa_i|X_i, \widehat{a}_n^{ES}) \geq \frac{1}{2}\right) - \Pr(\widehat{a}_n^{ES} \leq \gamma\alpha_n) \\ & \rightarrow 2\left[\Phi(\sqrt{C}) - 1\right] (1 + o(1)) - o(1), \end{aligned}$$

where we use the result in Theorem 3.4, the fact that  $\mathbb{E}(\kappa_i|X_i)$  is nondecreasing in  $a_n$ , and the fact that  $\Pr(\widehat{a}_n^{ES} \leq \gamma\alpha_n)$  is asymptotically vanishing (by (A30)) to arrive at the final inequality.  $\square$

## Appendix B. Sampling from the NBP Model

### B.1. No Prior on the Hyperparameter $a$

Suppose that there is no prior placed on the hyperparameter  $a$ . By the reparametrization of  $\sigma_i^2 = \lambda_i \xi_i$ ,  $i = 1, \dots, n$ , given in (A1) and letting  $\kappa_i = 1/(1 + \lambda_i \xi_i)$ , the full conditional distributions for (6) are

$$\begin{aligned} \theta_i \mid \text{rest} & \sim \mathcal{N}((1 - \kappa_i)X_i, 1 - \kappa_i), i = 1, \dots, n, \\ \lambda_i \mid \text{rest} & \sim \mathcal{GIG}\left(\frac{\theta_i^2}{\xi_i}, 2, a - \frac{1}{2}\right), i = 1, \dots, n, \\ \xi_i \mid \text{rest} & \sim \mathcal{IG}\left(b + \frac{1}{2}, \frac{\theta_i^2}{2\lambda_i} + 1\right), i = 1, \dots, n, \end{aligned} \quad (\text{B1})$$

where  $\mathcal{GIG}(c, d, p)$  denotes a generalized inverse Gaussian (giG) density with  $f(x; c, d, p) \propto x^{(p-1)} e^{-(c/x + dx)/2}$ . Therefore, the NBP model (6) – and consequently, thresholding rules (19) and (23) – can be implemented straightforwardly with Gibbs sampling utilizing the full conditionals in (B1). Moreover, since the full conditionals are independent, we can update the  $\theta_i$ 's,  $\lambda_i$ 's, and  $\xi_i$ 's efficiently using block updates.

### B.2. Uniform Prior on the Hyperparameter $a$

In the case that a prior is placed on  $a$ , the steps for sampling from the full conditionals for  $(\theta_i, \lambda_i, \xi_i), i = 1, \dots, n$  from (B1) remain the same. However, we now also need to sample from the full conditional of  $a$ . When  $a \sim \mathcal{U}(1/n, 1)$ , the full conditional for  $a$  is proportional to

$$\pi(a|\text{rest}) \propto \left( \frac{\Gamma(a+b)}{\Gamma(a)} \right)^n \left( \prod_{i=1}^n (\sigma_i^2)^{a-1} (1 + \sigma_i^2)^{-a-b} \right) \mathbb{I}\{1/n \leq a \leq 1\}, \quad (\text{B2})$$

where  $\sigma_i^2 = \lambda_i \xi_i$ . Using (B2), we update  $a$  using a Metropolis-Hastings random walk. For our proposal distribution, we use a truncated normal density on the interval  $[1/n, 1]$ . If  $a$  is the current value of the chain, a new value  $a^*$  will be generated from the proposal distribution,

$$q(a^*|a) = \frac{\phi\left(\frac{a^*-a}{\omega}\right)}{\omega \left( \Phi\left(\frac{1-a}{\omega}\right) - \Phi\left(\frac{1/n-a}{\omega}\right) \right)} \mathbb{I}\{1/n \leq a^* \leq 1\}, \quad (\text{B3})$$

where  $\phi(\cdot)$  and  $\Phi(\cdot)$  denote the standard normal probability density function (pdf) and cumulative distribution function (cdf) respectively, and  $\omega > 0$  is a scaling parameter that is properly calibrated to control the Metropolis-Hastings acceptance rate. Given a candidate state  $a^*$  drawn from  $q(a^*|a)$ , it then follows from (B2) and (B3) that  $a^*$  is accepted with probability,

$$\min \left\{ 1, \left( \frac{\Gamma(a^*+b)\Gamma(a)}{\Gamma(a+b)\Gamma(a^*)} \right)^n \left( \prod_{i=1}^n \left( \frac{\sigma_i^2}{1 + \sigma_i^2} \right)^{a^*-a} \right) \left[ \frac{\Phi\left(\frac{1-a}{\omega}\right) - \Phi\left(\frac{1/n-a}{\omega}\right)}{\Phi\left(\frac{1-a^*}{\omega}\right) - \Phi\left(\frac{1/n-a^*}{\omega}\right)} \right] \right\},$$

where  $\sigma_i^2, i = 1, \dots, n$ , is taken as the product of the  $\lambda_i$  and  $\xi_i$  from the most recent Gibbs sampling updates for  $(\lambda_i, \xi_i), i = 1, \dots, n$ . We tune  $\omega$  so that the acceptance rate is between 20 and 40 percent.

### B.3. Truncated Cauchy Prior on the Hyperparameter $a$

If we place a truncated Cauchy prior on  $a$  where  $a \in [1/n, 1]$ , i.e.  $\pi(a) = [\arctan(1) - \arctan(1/n)]^{-1} (1+a)^{-1} \mathbb{I}\{1/n < a < 1\}$ , the full conditional for  $a$  is proportional to

$$\pi(a|\text{rest}) \propto \left( \frac{\Gamma(a+b)}{(1+a)\Gamma(a)} \right)^n \left( \prod_{i=1}^n (\sigma_i^2)^{a-1} (1 + \sigma_i^2)^{-a-b} \right) \mathbb{I}\{1/n \leq a \leq 1\}. \quad (\text{B4})$$

As before, we use Metropolis-Hastings to update  $a$ . We use the truncated normal density  $q(a^*|a)$  from (B3) as the proposal distribution. Given a candidate state  $a^*$  drawn from  $q(a^*|a)$  in (B3), it follows from (B4) that  $a^*$  is accepted with probability,

$$\min \left\{ 1, \left( \frac{(1+a)\Gamma(a^*+b)\Gamma(a)}{(1+a^*)\Gamma(a+b)\Gamma(a^*)} \right)^n \left( \prod_{i=1}^n \left( \frac{\sigma_i^2}{1 + \sigma_i^2} \right)^{a^*-a} \right) \left[ \frac{\Phi\left(\frac{1-a}{\omega}\right) - \Phi\left(\frac{1/n-a}{\omega}\right)}{\Phi\left(\frac{1-a^*}{\omega}\right) - \Phi\left(\frac{1/n-a^*}{\omega}\right)} \right] \right\},$$

where  $\sigma_i^2, i = 1, \dots, n$ , is taken as the product of the  $\lambda_i$  and  $\xi_i$  from the most recent Gibbs sampling updates for  $(\lambda_i, \xi_i), i = 1, \dots, n$ . We tune  $\omega$  so that the acceptance rate is between 20 and 40 percent.

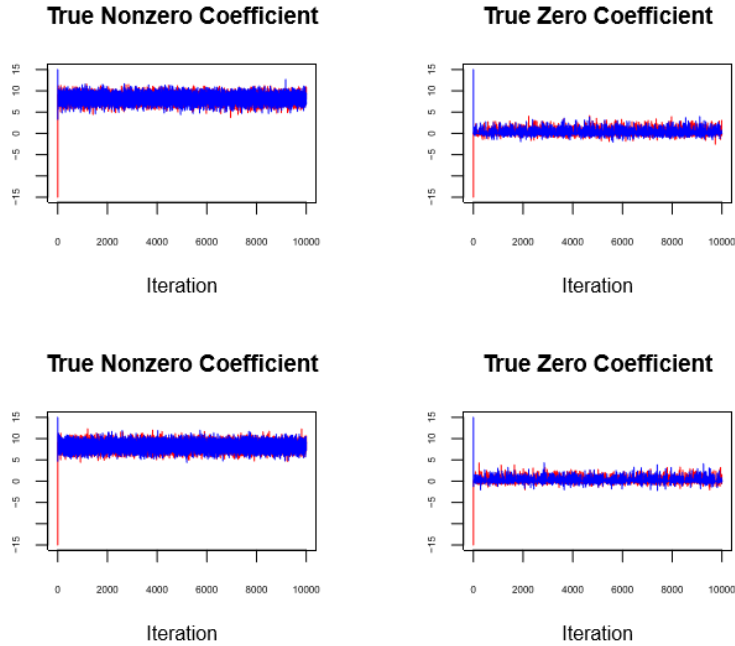
#### B.4. Convergence of the MCMC Algorithm

To assess the convergence and the mixing of the MCMC algorithms for the hierarchical Bayes approaches described in Sections D.2 and D.3, we consider two chains with different starting values: 1)  $\theta_i^{(0)} = -15, i = 1, \dots, n$ , and 2)  $\theta_i^{(0)} = 15, i = 1, \dots, n$ . In our simulation studies, the true  $\theta_0$  was generated from

$$\theta_{0i} \stackrel{iid}{\sim} (1 - p)\delta_0 + p\mathcal{N}(0, \psi^2), i = 1, \dots, n,$$

with  $\psi = \sqrt{2 \log(500)} = 3.53$ . Thus, these initial values for  $\theta_i^{(0)}, i = 1, \dots, n$ , are all far away from a ‘typical’ value of  $\theta_{0i}$ . We found that in both cases, the MCMC algorithms still converged very rapidly (usually within 100 iterations), giving very similar posterior estimates for  $\theta$  after discarding the first 5000 iterations as burnin.

To illustrate this, we plot in Figure B1 the history plots for one nonzero coefficient ( $\theta_{0i} = 7.225$ ) and one null coefficient ( $\theta_{0i} = 0$ ) when the sparsity level is  $p = 0.2$ . For the nonnull coefficient, we see that the chains mix well and rapidly converge to a stationary distribution centered around the true value of  $\theta_{0i}$ . For the null coefficient, the chains rapidly converge to a stationary distribution centered around zero.



**Figure B1.** History plots of the 10,000 draws from the MCMC algorithm for the NBP-UNIF (top panel) and NBP-TC models (bottom panel) for a single  $\theta_{0i}$ . The plots on the left are for a  $\theta_{0i}$  whose true value is equal to 7.225, and the plots on the right are for a  $\theta_{0i}$  whose true value is equal to 0.

## Appendix C. Comparison Between Posterior Inclusion Probabilities and Posterior Shrinkage Weights

If  $(p, \psi)$  were known, then a natural thresholding rule for selection of nonnull values in  $\boldsymbol{\theta}$  would be to threshold the posterior inclusion probabilities,  $\pi(\nu_i = 1|X_i)$ ,  $i = 1, \dots, n$ , for the two-groups model (11). That is, we would reject  $H_{0i}$  if  $\pi(\nu_i = 1|X_i) > 0.5$ . In our paper, we have used the posterior shrinkage weights  $\mathbb{E}(1 - \kappa_i|X_1, \dots, X_n)$  in (19) as a proxy for these posterior inclusion probabilities, and one may wonder how well the shrinkage weights approximate the true inclusion probabilities.

Taking different choices of  $p \in \{0.05, 0.10, 0.20, 0.30\}$ , we plot in Figures C1 and Figure C2 the theoretical posterior inclusion probabilities  $\omega_i(X_i) = P(\nu_i = 1|X_i)$  for the two-groups model (11) given by

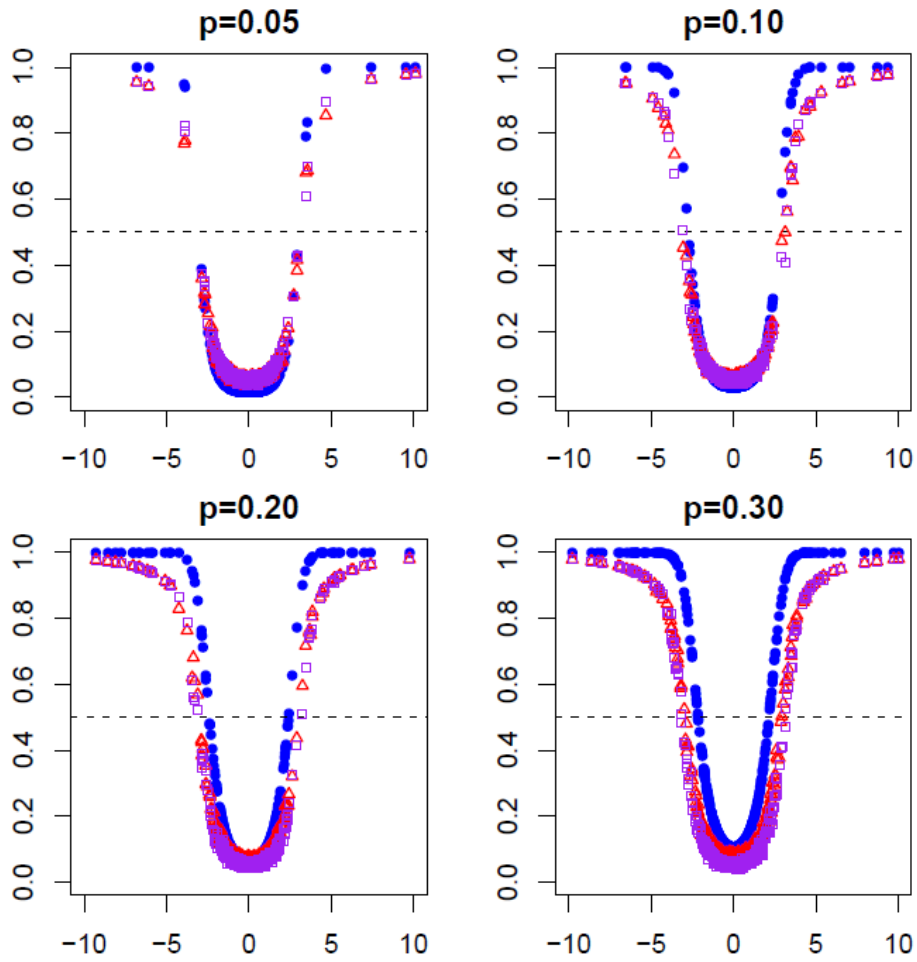
$$\omega_i(X_i) = \pi(\nu_i = 1|X_i) = \left\{ \left( \frac{1-p}{p} \right) \sqrt{1 + \psi^2} e^{-\frac{x_i^2}{2} \frac{\psi^2}{1+\psi^2}} + 1 \right\}^{-1},$$

along with the shrinkage weights  $\mathbb{E}(1 - \kappa_i|\hat{a}^{ES}, X_i)$ ,  $\mathbb{E}(1 - \kappa_i|\hat{a}^{REML}, X_i)$ , and  $\mathbb{E}(1 - \kappa_i|X_1, \dots, X_n)$  for the NBP-ES, NBP-REML, NBP-UNIF, and NBP-TC models. These plots show that for small values of the sparsity level  $p$ , the shrinkage weights are in close proximity to the posterior inclusion probabilities. This offers empirical support for the use of these posterior shrinkage weights as an approximation to the corresponding posterior inclusion probabilities  $\omega_i(X_i)$  in sparse situations.

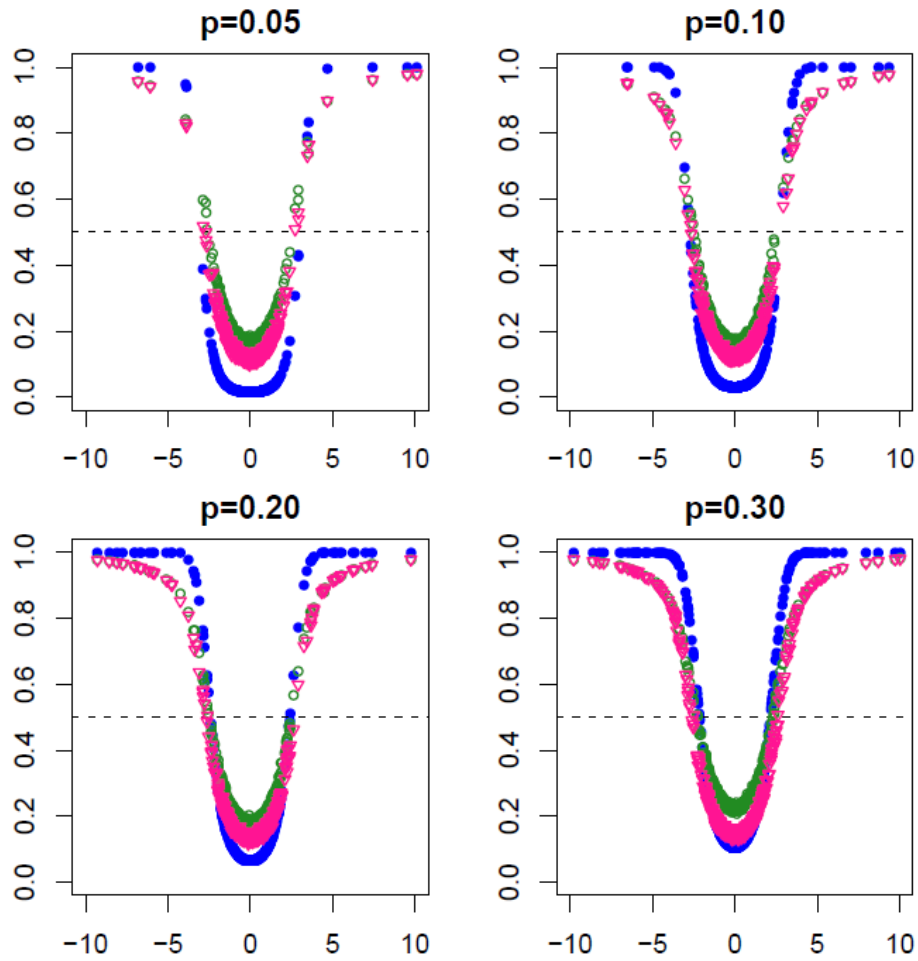
## References

- [1] Armagan A, Clyde M, Dunson DB (2011). Generalized beta mixtures of gaussians. *NeurIPS 2011*; 24:523-531.
- [2] Datta J, Ghosh JK. Asymptotic properties of bayes risk for the horseshoe prior. *Bayesian Anal.* 2013;8:111-132.
- [3] Ghosh P, Chakrabarti A. Asymptotic optimality of one-group shrinkage priors in sparse high-dimensional problems. *Bayesian Anal.* 2017;12:1133-1161.
- [4] Ghosh P, Tang X, Ghosh M, Chakrabarti A. Asymptotic properties of bayes risk of a general class of shrinkage priors in multiple hypothesis testing under sparsity. *Bayesian Anal.* 2016;11:753-796.





**Figure C1.** Comparison between the posterior inclusion probabilities  $\omega_i(X_i) = \pi(\nu_i = 1|X_i)$  and the posterior shrinkage weights  $\mathbb{E}(1 - \kappa_i|\hat{a}^{ES}, X_i)$ ,  $\mathbb{E}(1 - \kappa_i|\hat{a}^{REML}, X_i)$ . The solid circles are the posterior inclusion probabilities, while the empty triangles correspond to NBP-ES and the empty squares correspond to NBP-REML.



**Figure C2.** Comparison between the posterior inclusion probabilities  $\omega_i(X_i) = \pi(\nu_i = 1|X_i)$  and the posterior shrinkage weights  $\mathbb{E}(1 - \kappa_i|X_1, \dots, X_n)$  under the hierarchical Bayes approaches. The solid circles are the posterior inclusion probabilities, while the empty circles correspond to NBP-UNIF and the empty upside-down triangles correspond to NBP-TC.