Corrigendum to "High-dimensional multivariate posterior consistency under global-local shrinkage priors" [J. Multivariate Anal. 167 (2018) 157-170]

Shao-Hsuan Wang[a], Ray Bai[b], Hsin-Hsiung Huang[c,*]

[a] *Graduate Institute of Statistics, National Central University, Taoyuan City, Taiwan 320*
[b] *Department of Statistics, University of South Carolina, Columbia, SC 29208, United States*
[c] *Department of Statistics and Data Science, University of Central Florida, Orlando, FL 32816, United States*

Theorems 3 and 4 of Bai and Ghosh [2] present posterior consistency results for their proposed multivariate Bayesian model with shrinkage priors (MBSP). However, we found an error in the proofs of these theorems (in the Supplementary Material [1]). Specifically, in the proof of Theorem 3, there is a mistake in (2.10) of the Supplement [1]. The center of b_{jk}^n is 0 but the center of X is b_{jk}^0. The area in the first equation is *not* the area under the normal density function around the center; however, the area in the second equation *is* the area under the normal density around the center. Therefore, these areas are generally unequal except for when $b_{jk}^0 = 0$, so the second "=" does not hold. Since the step cannot go through, the subsequent inequalities cannot be achieved, and the given proof of posterior consistency is incorrect. Similarly, the proof of Theorem 4, which follows along much of the same lines, is not correct.

In this note, we provide a corrected proof of posterior consistency for the MBSP model. To recall some of the notation from Bai and Ghosh [2], $\mathcal{A}_n \subset \{1, \ldots, p_n\}$ is the set of indices of the true nonzero rows in the regression coefficients matrix $\mathbf{B}_0$, with cardinality $s_n = |\mathcal{A}_n|$. We similarly define the complement of $\mathcal{A}_n$ as the set of indices of the true zero rows in $\mathbf{B}_0$, i.e. $\mathcal{A}_n^c \equiv \{1, \ldots, p_n\} \setminus \mathcal{A}_n$, which has cardinality $p_n - s_n$.

For simplicity, we only provide the corrected proof for Theorem 4 of Bai and Ghosh [2] in the ultra high-dimensional case where $p_n \gg n$, $\ln(p_n) = o(n)$, and $s_n = o(n/\ln(p_n))$. We note that a correct proof for the low-dimensional case where $p_n \leq n$ and $p_n = o(n)$ (Theorem 3 of Bai and Ghosh [2]) is nearly identical to our proof for Theorem 4, except with the constant $\widetilde{\Delta}$ replaced with Δ from Theorem 3 of Bai and Ghosh [2]. Therefore, this note establishes the validity of *both* Theorems 3 and 4 of Bai and Ghosh [2] and shows that the MBSP model achieves strong posterior consistency, provided that p_n grows at a sub-exponential rate with n.

We first state and prove two lemmas. Lemma 1 shows that a slowly varying function $L(x)$ has a polynomial tail asymptotically, while Lemma 2 establishes a lower bound for an upper-tail probability for the hyperprior π in (3) of Bai and Ghosh [2].

Lemma 1. *Assume that a positive measurable function $L(x)$ is slowly varying, i.e. for each fixed*

*Corresponding author.

Email addresses: picowang@gmail.com (Shao-Hsuan Wang), rbai@mailbox.sc.edu (Ray Bai), hsin-hsiung.huang@ucf.edu (Hsin-Hsiung Huang)

$a > 0$, $L(ax)/L(x) \to 1$ *as* $x \to \infty$. *Then for each* $\varepsilon \in (0,1)$ *and* $a > 1$, *there exists* x_0 *such that*

$$c_1 x^{\ln(1-\varepsilon)/\ln(a)} \leq L(x) \leq c_2 x^{\ln(1+\varepsilon)/\ln(a)} \text{ whenever } x > x_0,$$

for some constants $c_1, c_2 > 0$.

Proof of Lemma 1. By the definition of a slowly varying function, for each $\varepsilon \in (0,1)$ and $a > 1$, there exists u_0 so that $|L(au)/L(u) - 1| < \varepsilon$ whenever $u > u_0$. Thus, we have

$$L(u_0)(1-\varepsilon) \leq L(au_0) \leq L(u_0)(1+\varepsilon).$$

By induction, for all $k \in \mathbb{N}$,

$$L(u_0)(1-\varepsilon)^k \leq L(a^k u_0) \leq L(u_0)(1+\varepsilon)^k. \tag{1}$$

Note that $a^k u > u_0$ for all $k \in \mathbb{N}$ since $a > 1$. Take $x = a^k u_0$ so that $k = \ln(x/u_0)/\ln(a)$. We can then rewrite (1) as

$$L(u_0)(1-\varepsilon)^{\ln(x/u_0)/\ln(a)} \leq L(x) \leq L(u_0)(1+\varepsilon)^{\ln(x/u_0)/\ln(a)},$$

and thus,

$$L(u_0)\left(\frac{x}{u_0}\right)^{\frac{\ln(1-\varepsilon)}{\ln(a)}} \leq L(x) \leq L(u_0)\left(\frac{x}{u_0}\right)^{\frac{\ln(1+\varepsilon)}{\ln(a)}}.$$

□

Lemma 2. *Suppose that* ξ *follows the hyperprior distribution* $\pi(\xi)$ *in (3) of Bai and Ghosh [2], i.e.* $\pi(\xi) = K\xi^{-a-1}L(\xi)$, *where* $L(\cdot)$ *is slowly varying and* $s_n \ln(p_n) = o(n)$. *Then*

$$\pi(\xi > s_n p_n n^{\rho-1}) > \exp(-A_1 n/s_n),$$

for some finite $A_1 > 0$.

Proof of Lemma 2. By Lemma 1, there exists finite $\delta > 0$ and $K_1 > 0$ so that

$$\begin{aligned}
\pi(\xi > s_n p_n n^{\rho-1}) &= \int_{s_n p_n n^{\rho-1}}^{\infty} K u^{-a-1} L(u) du \\
&\geq \int_{s_n p_n n^{\rho-1}}^{\infty} K u^{-a-1} K_1 u^{-\delta} du \\
&= \frac{KK_1}{a+\delta}(s_n p_n n^{\rho-1})^{-a-\delta} \\
&= \exp\left\{-(a+\delta)\ln(p_n) - (a+\delta)\ln(s_n n^{\rho-1}) + \ln(KK_1/(a+\delta))\right\} \\
&> \exp(-A_1 n/s_n),
\end{aligned}$$

for a sufficiently large $A_1 > 0$. In the last line of the display, we used the fact that $s_n \ln(p_n) = o(n)$. □

We now establish strong posterior consistency for the MBSP model of Bai and Ghosh [2] when $p \gg n$ and $s_n \ln(p_n) = o(n)$. In our proof, we make an additional mild assumption that under the hyperprior $\pi(\xi)$ in Lemma 2 of this note, $\mathbb{E}(\xi) < \infty$. Note that this assumption holds for all practical examples for $\pi(\xi)$ in scale-mixture shrinkage priors (see Table 1 of Bai and Ghosh [2]).

Proof of Theorem 4 of Bai and Ghosh [2]. In light of Theorem 2 of Bai and Ghosh [2], it is sufficient to show that the marginal prior $\pi_n(\mathbf{B}_n)$ satisfies

$$\Pi_n\left(\mathbf{B}_n : \|\mathbf{B}_n - \mathbf{B}_0\|_F < \frac{\widetilde{\Delta}}{n^{\rho/2}}\right) > \exp(-kn), \tag{2}$$

for $k > 0$.

Using notation from Bai and Ghosh [2], we let $\mathbf{b}_j^n \in \mathbb{R}^q$ denote the jth row of $\mathbf{B}_n$; analogously, $\mathbf{b}_j^0$ is the jth row of the true parameter $\mathbf{B}_0$. We let b_{jk}^n denote the kth entry of $\mathbf{b}_j^n$ and b_{jk}^0 denote the kth entry of $\mathbf{b}_j^0$. Since the rows of $\mathbf{B}_n$ are independent,

$$\pi_n(\mathbf{B}_n) = \prod_{j=1}^{p} \pi_n(\mathbf{b}_j^n).$$

Under the MBSP model, $\mathbf{b}_j^n \mid \xi_j \overset{ind}{\sim} \mathcal{N}_q(\mathbf{0}_q, \tau_n \xi_j \mathbf{\Sigma})$ for $j \in \{1, \ldots, p_n\}$. Therefore, the marginal prior for $\pi_n(\mathbf{b}_j^n)$ is

$$\pi_n(\mathbf{b}_j^n) = \int_0^\infty (2\pi\tau_n\xi_j)^{-q/2} |\mathbf{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2\xi_j\tau_n}\|\mathbf{\Sigma}^{-1/2}\mathbf{b}_j^n\|_2^2\right)\pi(\xi_j)d\xi_j, \tag{3}$$

where $\pi(\xi_j)$ is the hyperprior on ξ_j as in Lemma 2 of this note. We have

$$\begin{aligned}
\Pi_n\left(\mathbf{B}_n : \|\mathbf{B}_n - \mathbf{B}_0\|_F^2 < \frac{\widetilde{\Delta}^2}{n^\rho}\right) &= \Pi_n\left(\mathbf{B}_n : \sum_{j\in\mathcal{A}_n}\|\mathbf{b}_j^n - \mathbf{b}_j^0\|_2^2 + \sum_{j\in\mathcal{A}_n^c}\|\mathbf{b}_j^n\|_2^2 < \frac{\widetilde{\Delta}^2}{n^\rho}\right) \\
&\geq \Pi_n\left(\mathbf{B}_n : \sum_{j\in\mathcal{A}_n}\|\mathbf{b}_j^n - \mathbf{b}_j^0\|_2^2 < \frac{\widetilde{\Delta}^2}{2n^\rho}, \sum_{j\in\mathcal{A}_n^c}\|\mathbf{b}_j^n\|_2^2 < \frac{\widetilde{\Delta}^2}{2n^\rho}\right) \\
&\geq \left\{\prod_{j\in\mathcal{A}_n}\Pi_n\left(\mathbf{b}_j^n : \|\mathbf{b}_j^n - \mathbf{b}_j^0\|_2^2 < \frac{\widetilde{\Delta}^2}{2s_n n^\rho}\right)\right\}\left\{\Pi_n\left(\mathbf{B}_n : \sum_{j\in\mathcal{A}_n^c}\|\mathbf{b}_j^n\|_2^2 < \frac{\widetilde{\Delta}^2}{2n^\rho}\right)\right\} \\
&\triangleq \mathcal{I}_1 \times \mathcal{I}_2.
\end{aligned} \tag{4}$$

We first bound $\mathcal{I}_1$ from below. In what follows, we let C_1 and C_2 be appropriate constants that do not depend on n, while A_1 is the constant from Lemma 2 of this note. From (3), we have

$$\begin{aligned}
\mathcal{I}_1 &= \prod_{j\in\mathcal{A}_n}\Pi_n\left(\mathbf{b}_j^n : \|\mathbf{b}_j^n - \mathbf{b}_j^0\|_2^2 < \frac{\widetilde{\Delta}^2}{2s_n n^\rho}\right) \\
&= \prod_{j\in\mathcal{A}_n}\left[\int_0^\infty \int_{\|\mathbf{b}_j^n - \mathbf{b}_j^0\|_2^2 < \frac{\widetilde{\Delta}^2}{2s_n n^\rho}} (2\pi\tau_n\xi_j)^{-q/2}|\mathbf{\Sigma}|^{-1/2}\exp\left(-\frac{1}{2\xi_j\tau_n}\|\mathbf{\Sigma}^{-1/2}\mathbf{b}_j^n\|_2^2\right)\pi(\xi_j)d\mathbf{b}_j^n d\xi_j\right]
\end{aligned}$$

$$
\begin{aligned}
&\geq \prod_{j\in\mathcal{A}_n}\left[\int_0^\infty \int_{\|\mathbf{b}_j^n-\mathbf{b}_j^0\|_2^2<\frac{\tilde{\Delta}^2}{2s_n n^\rho}} (2\pi\tau_n\xi_j d_2)^{-q/2}\exp\left(-\frac{\|\mathbf{b}_j^n\|_2^2}{2\xi_j\tau_n d_1}\right)\pi(\xi_j)d\mathbf{b}_j^n d\xi_j\right] \\
&\geq \prod_{j\in\mathcal{A}_n}\prod_{k=1}^q\left[\int_0^\infty \int_{|b_{jk}^n-b_{jk}^0|^2<\frac{\tilde{\Delta}^2}{2qs_n n^\rho}} \frac{1}{\sqrt{2\pi\tau_n\xi_j d_2}}\exp\left(-\frac{(b_{jk}^n)^2}{2\xi_j\tau_n d_1}\right)\pi(\xi_j)db_{jk}^n d\xi_j\right] \\
&= \prod_{j\in\mathcal{A}_n}\prod_{k=1}^q\left[\int_0^\infty \int_{b_{jk}^0-\frac{\tilde{\Delta}}{\sqrt{2qs_n n^\rho}}}^{b_{jk}^0+\frac{\tilde{\Delta}}{\sqrt{2qs_n n^\rho}}} \frac{1}{\sqrt{2\pi\tau_n\xi_j d_2}}\exp\left(-\frac{(b_{jk}^n)^2}{2\xi_j\tau_n d_1}\right)\pi(\xi_j)db_{jk}^n d\xi_j\right] \\
&= \prod_{j\in\mathcal{A}_n}\prod_{k=1}^q\left[\int_0^\infty \frac{C_1}{\sqrt{q\tau_n\xi_j s_n n^\rho}}\exp\left(-\frac{(b_{jk}^0)^2}{2\xi_j\tau_n d_1}\right)\pi(\xi_j)d\xi_j\right](1+o(1)) \\
&\geq \prod_{j\in\mathcal{A}_n}\prod_{k=1}^q\left[\int_{s_n p_n n^{\rho-1}}^\infty \frac{C_1}{\sqrt{q\tau_n\xi_j s_n n^\rho}}\exp\left(-\frac{(b_{jk}^0)^2}{2\xi_j\tau_n d_1}\right)\pi(\xi_j)d\xi_j\right] \\
&\geq \prod_{j\in\mathcal{A}_n}\prod_{k=1}^q\left[\frac{C_1}{C_2\sqrt{q\tau_n s_n n^\rho}}\exp\left(-\frac{(b_{jk}^0)^2}{2d_1\tau_n s_n p_n n^{\rho-1}}\right)\int_{s_n p_n n^{\rho-1}}^\infty \pi^\star(\xi_j)d\xi_j\right], \\
&\qquad\qquad \text{where } \pi^\star(\xi_j)=C_2\xi_j^{-1/2}\pi(\xi_j) \text{ and } \int_0^\infty \pi^\star(\xi_j)d\xi_j=1, \\
&> \prod_{j\in\mathcal{A}_n}\prod_{k=1}^q\left[\frac{C_1}{C_2\sqrt{q\tau_n s_n n^\rho}}\exp\left(-\frac{M^2}{2d_1\tau_n s_n p_n n^{\rho-1}}\right)\exp\left(-\frac{A_1 n}{s_n}\right)\right] \\
&= \exp\left(qs_n\left[\ln\left(\frac{C_1}{C_2}\right)-\frac{1}{2}\ln(\tau_n)-\frac{1}{2}\ln(qs_n n^\rho)-\frac{M^2}{2d_1\tau_n s_n p_n n^{\rho-1}}-\frac{A_1 n}{s_n}\right]\right) \\
&= \exp\left(-qA_1 n-\frac{qM^2}{2d_1\tau_n p_n n^{\rho-1}}+qs_n\left[\ln\left(\frac{C_1}{C_2}\right)-\frac{1}{2}\ln(\tau_n)-\frac{1}{2}\ln(qs_n n^\rho)\right]\right). \qquad (5)
\end{aligned}
$$

In the third line, we used Assumption (B5) of Bai and Ghosh [2] that $0 < d_1 < \lambda_{\min}(\mathbf{\Sigma}) \leq \lambda_{\max}(\mathbf{\Sigma}) < d_2 < \infty$. In the ninth line, we used Assumption (C2) of Bai and Ghosh [2] that $\max_{j,k}|b_{jk}^0| \leq M < \infty$ and the fact that $\pi^\star(\xi_j) := C_2\xi_j^{-1/2}\pi(\xi_j) = C_2\xi_j^{-(a+1/2)-1}L(\xi)$, where C_2 is the normalizing constant to ensure that $\pi^\star(\xi_j)$ is a valid probability density; therefore, we can apply Lemma 2 to the integral term in the eighth line.

By Assumption (C3) of Bai and Ghosh [2], $\tau_n = o(p_n^{-1}n^{-\rho})$, and by Assumption (B6) of Bai and Ghosh [2], $s_n\ln(p_n) = o(n)$. Therefore, besides $-qA_1n$, all the other terms in the exponent of (5) are of order $o(n)$. Noting that q is fixed (i.e. it does not depend on n) and the strict inequality in the ninth line of the above display, we see from (5) that $\mathcal{I}_1$ in (4) can be bounded from below as

$$
\mathcal{I}_1 > \exp(-kn), \qquad (6)
$$

for sufficiently large n. Next, we bound $\mathcal{I}_2$ in (4) from below. By Markov's inequality and the fact that $\mathbf{b}_j^n \mid \xi_j \sim \mathcal{N}_q(\mathbf{0}_q, \tau_n\xi_j\mathbf{\Sigma})$ for $j \in \{1,\ldots,p_n\}$, we have

$$
\mathcal{I}_2 = \Pi_n\left(\mathbf{B}_n \ : \ \sum_{j\in\mathcal{A}_n^c}\|\mathbf{b}_j^n\|_2^2 < \frac{\tilde{\Delta}^2}{2n^\rho}\right)
$$

$$
\begin{aligned}
&= 1 - \Pi_n \left(\mathbf{B}_n \ : \ \sum_{j \in \mathcal{A}_n^c} \|\mathbf{b}_j^n\|_2^2 > \frac{\tilde{\Delta}^2}{2n^\rho} \right) \\
&= 1 - \mathbb{E} \left[\Pi_n \left(\mathbf{B}_n \ : \ \sum_{j \in \mathcal{A}_n^c} \|\mathbf{b}_j^n\|_2^2 > \frac{\tilde{\Delta}^2}{2n^\rho} \,\middle|\, \xi_1, \ldots, \xi_{p_n} \right) \right] \\
&\geq 1 - \mathbb{E} \left[\frac{2n^\rho \mathbb{E} \left(\sum_{j \in \mathcal{A}_n^c} \|\mathbf{b}_j^n\|_2^2 \mid \xi_1, \ldots, \xi_{p_n} \right)}{\tilde{\Delta}^2} \right] \\
&= 1 - \frac{2n^\rho \tau_n \text{tr}(\mathbf{\Sigma}) \mathbb{E} \left(\sum_{j \in \mathcal{A}_n^c} \xi_j \right)}{\tilde{\Delta}^2} \\
&\geq 1 - \frac{2qd_2 \tau_n n^\rho (p_n - s_n) \mathbb{E}(\xi_1)}{\tilde{\Delta}^2} \to 1 \text{ as } n \to \infty,
\end{aligned} \tag{7}
$$

where the last line of the display follows from Assumption (B5) of Bai and Ghosh [2] that $0 < d_1 < \lambda_{\min}(\mathbf{\Sigma}) \leq \lambda_{\max}(\mathbf{\Sigma}) < d_2 < \infty$, Assumption (C3) of Bai and Ghosh [2] that $\tau_n = o(p_n^{-1} n^{-\rho})$, and our assumption that $\mathbb{E}(\xi) < \infty$ when $\xi \sim \pi(\xi)$. Thus, $2qd_2\tau_n n^\rho (p_n - s_n)\mathbb{E}(\xi_1)/\widetilde{\Delta}^2 \to 0$ as $n \to \infty$.

Combining (4)-(7), it is clear that for sufficiently large n,

$$
\Pi_n \left(\mathbf{B}_n : \|\mathbf{B}_n - \mathbf{B}_0\|_F < \frac{\widetilde{\Delta}}{n^{\rho/2}} \right) > \exp(-kn),
$$

i.e. we have established the prior mass condition (2). Strong posterior consistency for the MBSP model now follows immediately from Theorem 2 of Bai and Ghosh [2]. □

References

[1] Ray Bai and Malay Ghosh. Supplementary materials for "High-dimensional multivariate posterior consistency under global-local shrinkage priors". `https://ars.els-cdn.com/content/image/1-s2.0-S0047259X17306905-mmc1.pdf`. Accessed: 2021-12-30.

[2] Ray Bai and Malay Ghosh. High-dimensional multivariate posterior consistency under global–local shrinkage priors. *Journal of Multivariate Analysis*, 167:157–170, 2018. ISSN 0047-259X.