1. Supplemental Proofs of Theorems.

1.1. Proofs of Section 4. Throughout this section, we will be using the notation $\Delta^L$ and $\Delta^U$ introduced in (3.15). We also denote by

$$\delta_c = \frac{1}{\lambda_0 - \lambda_1} \log \left[ \frac{1 - \theta \lambda_0}{\theta \lambda_1} \frac{c}{1 - c} \right]$$

and

$$c_+ = \frac{1}{2} \left( 1 + \sqrt{1 - \frac{4}{(\lambda_0 - \lambda_1)^2}} \right) \quad \text{and} \quad c_- = \frac{1}{2} \left( 1 - \sqrt{1 - \frac{4}{(\lambda_0 - \lambda_1)^2}} \right).$$

With this notation, $\delta_{c_-} < \delta_{c_+}$ are the two roots of the equation $pen''(\beta) = 1$.

1.1.1. Proof of Theorem 3.1. First, we assume that $L(\beta, y)$ has two modes. Denote by $\hat{\beta}$ the global mode and by $\tilde{\beta}$ the local mode. First, assume $\hat{\beta} = 0$ and $(|y| - \lambda_1)^2 > 2 \log [1/p^*(0)]$. By (3.10) we have

$$y = \tilde{\beta} + \text{sign}(y) \lambda^*(\tilde{\beta}).$$

Because $\tilde{\beta}$ is a local maximum, the second derivative test $pen''(\tilde{\beta}) < 1$ yields

$$p^*(\tilde{\beta})[1 - p^*(\tilde{\beta})] < 1/(\lambda_0 - \lambda_1)^2.$$

Because $\tilde{\beta} \neq 0$, (1.4) implies that $|\tilde{\beta}| > \delta_{c_+} > 0$, where $\delta_{c_+}$ is defined in (3.14). Since $\hat{\beta} = 0$, we can write

$$y^2 < (y - \tilde{\beta})^2 + 2 \log \left( \frac{\pi(0)}{\pi(\beta)} \right).$$

*Veronika.Rockova@Chicagobooth.edu.
we obtain (1.7), which is equivalent to
\[
\hat{\lambda}^2 \log(1) \leq \log(1)
\]
and invoking (1.3), (1.5) is equivalent to
\[
y^2 < \lambda^\star(\hat{\beta})^2 + 2\lambda_1|y| - 2\lambda_1\lambda^\star(\hat{\beta}) + 2\log\left(\frac{p^\star(\hat{\beta})}{p^\star(0)}\right).
\]
Rearranging the terms above, we obtain
\[
(1.7) \quad (|y| - \lambda_1)^2 < g(\hat{\beta}) + 2\log[1/p^\star(0)],
\]
where \(g(\cdot)\) is defined in (3.13). We will need the following lemma.

**Lemma 1.1.** Assume \((\lambda_0 - \lambda_1) > 2\). The function \(g(x)\) is symmetric and \(g(|x|)\) is negative and monotone-increasing on \([\delta_{c+}; \infty)\). Moreover, \(\lim_{|x| \to \infty} g(x) = 0\) and \(g(|x|)\) has a minimum at \(\delta_{c+}\) if \(g(0) > 0\).

**Proof.** Writing \(g(x) = [1 - p^\star(x)]^2(\lambda_0 - \lambda_1)^2 + 2\log p^\star(x)\), the condition \(g(x) < 0\) will be satisfied when
\[
(1.8) \quad \log p^\star(x) < -\frac{[1 - p^\star(x)]^2}{2} (\lambda_0 - \lambda_1)^2.
\]
Using \(\log(x) < x - 1\) for \(x < 1\), (1.8) will be satisfied when \(p^\star(x) > 1 - \frac{2}{(\lambda_0 - \lambda_1)^2}\). Equating the derivative of \(g(|x|)\) to zero yields
\[
2[1 - p^\star(|x|)](\lambda_0 - \lambda_1) \{1 - p^\star(|x|)[1 - p^\star(|x|)](\lambda_0 - \lambda_1)^2\} = 0
\]
We obtain the two solutions \(\delta_{c-} < \delta_{c+}\) defined by (1.1) and (1.2). Because \(\hat{\beta} > \delta_{c+} > 0\) and \(p|\hat{\beta}| < 1\), we have \(g(|x|) > 0\) on \([\delta_{c+}, \infty)\). Since \(p^\star(x) \geq c_+ > 1 - \frac{2}{(\lambda_0 - \lambda_1)^2}\) on \([\delta_{c+}, \infty)\), we have \(g(x) < 0\) on \([\delta_{c+}, \infty)\). Because \(g(|x|)\) is monotone increasing on \([\delta_{c+}, \infty)\), \(g(|x|)\) has a local minimum at \(\delta_{c+}\). The value \(g(\delta_{c+})\) will be a global minimum if \(g(0) > 0\). Because \(\lim_{|x| \to \infty} p^\star(x) = 1\), we obtain \(\lim_{|x| \to \infty} g(x) = 0\).

According to Lemma 1.1 we have \(g(x) < 0\) for \(|x| \geq \delta_c\). Because \(\hat{\beta} > \delta_c\), we have \(g(\hat{\beta}) < 0\) and (1.7) contradicts the assumption \((|y| - \lambda_1)^2 > 2\log[1/p^\star(0)]\). Therefore \(\hat{\beta}\) has to be nonzero. Next, assume \((|y| - \lambda_1)^2 < 2\log[1/p^\star(0)] - d\). Since \(-d = g(\delta_{c+})\) is the minimum of the function \(g(x)\), we obtain (1.7), which is equivalent to \(\beta = 0\). The selection threshold can be
written as $\Delta = \sqrt{2 \log[1/p^*(0)]} + g(\beta) + \lambda_1$. Note that $|\beta| > \delta_\epsilon$. By Lemma 1.1, $g(x)$ is negative and monotone increasing on $[\delta_\epsilon; \infty)$ with a minimum at $\delta_\epsilon$, we have $\Delta^L < \Delta \leq \Delta^U$.

Now, assume that the posterior is unimodal. First, let $\beta \neq 0$ and $|y| < \Delta$. Then $|y| > \lambda^*(0)$ and because $g(0) > 0$ we have $\lambda^*(0) > \Delta$, which yields a contradiction. Next, assume $\beta = 0$ and $|y| > \Delta$. Denote by $x_0 \in (0, \delta_\epsilon)$ the root of $g(x) = 0$. Then, it is necessary that $|y| < \lambda^*(x_0) + x_0$. We show that $\lambda^*(x_0) + x_0 < \Delta$, yielding a contradiction. Because $g(x_0) = 0$, we have $\lambda^*(x_0) + x_0 = \lambda_1 + \sqrt{2 \log[1/p^*(x_0)]} + x_0$. Define $h(x) = x + \sqrt{2 \log[1/p^*(x)]}$, then $h'(x_0) = 0$ and $h'(x) < 0$ for $x \in (0, x_0)$. Therefore $h(x_0) < h(0)$ and $\lambda^*(x_0) + x_0 < \Delta$.

**Lemma 1.2.** Let $d = -g(\delta_\epsilon)$. Then $0 < d < 2 - \left(\frac{1}{\lambda_0 - \lambda_1} - \sqrt{2}\right)^2$.

**Proof.** Because $p^*(\delta_\epsilon) = c_\epsilon = (1 - c_-)$. We can invoke the inequality $1 - \sqrt{1 - x^2} > x^2/2$ for $0 < x < 1$ to obtain $(\lambda_0 - \lambda_1)^2(1 - c_\epsilon)^2 > 1/(\lambda_0 - \lambda_1)^2$. Because $c_\epsilon > 1 - \frac{2}{(\lambda_0 - \lambda_1)^2} > 0.5$ and $\log(1 - x^2) > -x$ for $0 < x < 0.5$, we have $\log c_\epsilon > -\frac{\sqrt{2}}{\lambda_0 - \lambda_1}$. Altogether, we have

$$g(\delta_\epsilon) > \frac{1}{(\lambda_0 - \lambda_1)^2} - \frac{\sqrt{2}}{\lambda_0 - \lambda_1} = \left(\frac{1}{\lambda_0 - \lambda_1} - \sqrt{2}\right)^2 - 2.$$ 

Lemma 1.1 yields $f(\delta_\epsilon) < 0$, which completes the proof. \qed

1.2. Proofs of Section 4.

1.2.1. **Proof of Theorem 4.1.** The risk of the Spike-and-Slab LASSO estimator $\hat{\beta}$ (3.7) can be written as

\begin{equation}
R(\hat{\beta}) = \sum_{i=1}^{p_0} E_{\beta_0 \mid i} [\hat{\beta}_i - \beta_0_i]^2 + \sum_{i=p_0+1}^{n} E_{\beta_0 \mid i} [\hat{\beta}_i^2],
\end{equation}

where $p_n = ||\beta_0||_0$. We begin by finding an upper bound to the first summand in (1.9). We have

\begin{equation}
E_{\beta_0 \mid i} [\hat{\beta}_i - \beta_0_i]^2 \leq 2E_{\beta_0 \mid i} [\hat{\beta}_i - Y_i]^2.
\end{equation}

By Theorem 3.1 and because $|\hat{\beta}_i| < |Y_i|$, we can write

$$E_{\beta_0} [\hat{\beta}_i - Y_i]^2 \leq E_{\beta_0} [Y_i^2 \mathbb{I}(|Y_i| \leq \Delta)] + E_{\beta_0} [\lambda^* (\hat{\beta}_i^2 \mathbb{I}(|Y_i| > \Delta])]$$.
Because $|\hat{\beta}_i| > \delta_{c+}$, we have $p^*(\hat{\beta}_i) > c_+$ when $|Y_i| > \Delta$, where $c_+$ is defined in the proof of Lemma 3.1. Then, we have $\lambda^*(\hat{\beta}_i) < c_+(\lambda_1 - \lambda_0) + \lambda_0$ when $|Y_i| > \Delta$. Next, we obtain

$$E_{\beta_0}[\hat{\beta}_i - Y_i]^2 \leq \Delta^2 + [\lambda_1 + (1 - c_+)(\lambda_0 - \lambda_1)]^2.$$ 

Because $c_+(1 - c_+) = 1/(\lambda_0 - \lambda_1)^2$ and $c_+ > 0.5$, we have $(1 - c_+)(\lambda_0 - \lambda_1) < 2/(\lambda_0 - \lambda_1)$. Since $\Delta^2 < 4\log[1/p^*(0)] + 2\lambda_1^2$ and because $\lambda_1 < e^{-2} < 1$ and $\lambda_0 - \lambda_1 > 2$, then

$$E_{\beta_0}[\hat{\beta}_i - Y_i]^2 \leq 4\log[1/p^*(0)] + 2.$$ 

Together with (1.10), this yields

$$E_{\beta_0}[\hat{\beta}_i - \beta_0]^2 \leq 8\log[1/p^*(0)] + 1.$$ 

Now we continue with the second summand in (1.9). Denote by $\phi(x)$ the standard normal density. Because $|\hat{\beta}_i| < |Y_i|$ and $\Delta^L < \Delta \leq \Delta^U$, we can write

$$(1.11) \quad E_0[\hat{\beta}^2_i] < E_0[Y_i^2 I(|Y_i| > \Delta)] < E_0[Y_i^2 I(|Y_i| > \Delta^L)] \leq 4\Delta^L \phi(\Delta^L),$$

where the last inequality uses Mills’s ratio (as in [? ]). Because $\Delta^L > \sqrt{2\log[1/p^*(0)] - d}$, we have $\phi(\Delta^L) < 1/\sqrt{2\pi e^{d/2} p^*(0)}$. Because $\Delta^L < \sqrt{2\log[1/p^*(0)]} + \lambda_1$ and $\lambda_1 < e^{-2} < \sqrt{2}$, (1.11) can be further simplified as

$$E_0[\hat{\beta}^2_i] < \frac{4e^{d/2}}{\sqrt{\pi}} p^*(0)\sqrt{\log[1/p^*(0)] + 1}.$$ 

Altogether, we obtain

$$(1.12) \quad E_{\beta_0}[|\beta - \beta_0|^2] < 8p_n (1 + \log[1/p^*(0)]) + 4/\sqrt{\pi}(n - p_n)e^d p^*(0)(1 + \sqrt{\log[1/p^*(0)]}).$$

1.2.2. **Proof of Theorem 4.2.**

**Proof.** Without loss of generality, we will assume $\frac{1 - \theta}{\sigma_{\lambda_1}} = n^{\alpha + \nu}/\lambda_1$. Denote by $Q(\beta) = -\frac{1}{2}\|y^{(n)} - \beta\|^2 + \log \pi(\beta | \theta)$, where $\pi(\beta | \theta)$ is the SSL prior arising from (2.1) and (2.2). Using the global optimality $0 \geq Q(\beta_0) - Q(\hat{\beta})$, we can write

$$(1.13) \quad 0 \geq ||\hat{\beta} - \beta_0||^2 - 2e(\hat{\beta} - \beta_0) + 2 \log \left( \frac{\pi(\beta_0 | \theta)}{\pi(\hat{\beta} | \theta)} \right).$$
Now, we focus on a set \( \tau_0 = \{ y^{(n)} : \|y^{(n)} - \beta_0\|_{\infty} \leq \tilde{\lambda} \} \), where \( \tilde{\lambda} = 2\sqrt{\log n} \). This set has a large probability under the generative model, i.e. \( P(\tau_0) \geq 1 - \frac{2}{n} \). Conditioning on the set \( \tau_0 \), we use the Hölder inequality \( |\alpha'\beta| \leq |\alpha|_{\infty}||\beta||_1 \) to find that

\[
(1.14) \quad 0 \geq ||\hat{\beta} - \beta_0||^2 - 2\tilde{\lambda}||\hat{\beta} - \beta_0||_1 + 2 \log \left( \frac{\pi(\beta_0 | \theta)}{\pi(\beta | \theta)} \right).
\]

Denote by \( \Delta \equiv \hat{\beta} - \beta_0 \). Using the fact \( ||\Delta||_1 \leq ||\Delta|| \|\Delta\|_0^{1/2} \), we have

\[
(1.15) \quad 0 \geq ||\Delta||^2 - 2\tilde{\lambda}||\Delta|| ||\Delta||_0^{1/2} + 2 \log \left( \frac{\pi(\beta_0 | \theta)}{\pi(\beta | \theta)} \right).
\]

Because \( p^*(\hat{\beta}_j) > c_+ \) whenever \( \hat{\beta}_j \neq 0 \), we can write

\[
\log \left[ \frac{\pi(\beta_0 | \theta)}{\pi(\beta | \theta)} \right] \geq -\lambda_1 ||\beta_0 - \hat{\beta}||_1 + \sum_{j=1}^{n} \log \left[ \frac{p^*(\hat{\beta}_j)}{p^*(\hat{\beta}_0)} \right] \geq -\lambda_1 ||\beta_0 - \hat{\beta}||_1 + \hat{p}_n b + (\hat{p}_n - p_n) \log[1/p^*(0)],
\]

where \( 0 > b = \log c_+ > \log 0.5 \) is a constant very close to 0 and where \( \lim_{\lambda_0 \to \infty} b = 0 \). To continue with (1.15), we can write

\[
(1.16) \quad \left[ ||\Delta|| - (\lambda + \lambda_1)||\Delta||_0^{1/2} \right]^2 + 2 (\hat{p}_n - p_n) \log[1/p^*(0)] \leq (\lambda + \lambda_1)^2 ||\Delta||_0 - 2\hat{p}_n b.
\]

With (1.16) and using the fact \( ||\Delta||_0 \leq \hat{p}_n + p_n \), we obtain

\[
2 (\hat{p}_n - p_n) \log[1/p^*(0)] \leq (\lambda + \lambda_1)^2 (\hat{p}_n + p_n) - 2\hat{p}_n b
\]

which is equivalent to writing

\[
\hat{p}_n \leq p_n (1 + C),
\]

where

\[
C = \frac{2(\lambda + \lambda_1)^2}{2\log[1/p^*(0)] - (\lambda + \lambda_1)^2 + 2b}
\]

and where \( C = \mathcal{O}(1) \) under the given choice of hyper-parameters. Moreover, with \( \lambda_1 < e^{-1} \) and \( 1 - \frac{\theta}{\lambda_1} = \lambda_1^{\alpha + \nu} / \lambda_1 \), we have \( 2 \log[1/p^*(0)] - 2(\lambda + \lambda_1)^2 + 2b > 2(\alpha + \nu - 4 - 4\lambda_1) \log n + 2 \log(1/\lambda_1) - 2\lambda_1^2 + 2b > 0 \) for some suitable \( c > 0 \). This inequality yields \( 0 < C < 2 \).
What remains to be shown is that $p_n \leq \hat{p}$ under the suitable beta-min condition. We prove this statement by contradiction. Assume $\hat{p}_n < p_n$ and denote by $q_n = p_n - \hat{p}_n > 0$. We continue with (1.15) and write

$$0 \geq \|\Delta\| \left[ \|\Delta\| - 2 (\bar{\lambda} + \lambda_1) \|\Delta\|^{1/2}_0 \right] - 2 q_n \log[1/p^*(0)] + 2 \hat{p}_n b.$$ 

Because $\|\Delta\|_0 < (1 + C) p_n$, we can write

(1.17)

$$0 \geq \|\Delta\| \left[ \|\Delta\| - 2 (\bar{\lambda} + \lambda_1) \sqrt{(1 + C) p_n} \right] - 2 q_n \log[1/p^*(0)] + 2 \hat{p}_n \log 0.5.$$ 

Assuming the minimal-strength condition $|\beta_{0i}| > b_0 \geq D \sqrt{p_n \log n}$ when $\beta_{0i} \neq 0$, we can write

$$\frac{1}{2} \|\Delta\| > \frac{\sqrt{q_n}}{2} b_0 \geq D \sqrt{q_n p_n \log n} > 2 (\bar{\lambda} + \lambda_1) \sqrt{p_n (1 + C)}$$

for a suitably large $D > 0$. Assuming $\frac{1 - \theta}{\theta} \lambda_0 = n^{\alpha+\nu}/\lambda_1$, (1.17) yields a contradiction

$$0 \geq \frac{1}{2} \|\Delta\|^2 - 2 q_n \log \left( 1 + n^{\alpha+\nu}/\lambda_1 \right) + 2 \hat{p}_n \log 0.5$$

$$> \frac{D^2}{2} q_n p_n \log n - 2 q_n \log \left( 1 + n^{\alpha+\nu}/\lambda_1 \right) + 2 \hat{p}_n \log 0.5 > 0$$

for $D$ sufficiently large. \hfill \Box

Veronika Ročková
Booth School of Business
University of Chicago
369 Carles M. Harper Center
5807 S. Woodlawn Avenue
Chicago, IL 60637