Supplemental Materials for:
"The Likelihood Ratio Test in High-Dimensional Logistic Regression Is Asymptotically a *Rescaled* Chi-Square"

Pragya Sur* Yuxin Chen† Emmanuel J. Candès*‡

June 2017

Abstract

This document presents the proof of Lemma 6(ii) given in the paper [1]: "The Likelihood Ratio Test in High-Dimensional Logistic Regression Is Asymptotically a *Rescaled* Chi-Square".

1 Proof of Lemma 6(ii)

We shall prove that $\mathcal{V}(\tau^2) < \tau^2$ whenever τ^2 is sufficiently large. Before proceeding, we recall from the main text and [2, Proposition 6.4] that

$$\mathcal{V}(\tau^2) := \frac{1}{\kappa}\mathbb{E}\left[\Psi^2(\tau Z; b(\tau))\right] = \frac{1}{\kappa}\mathbb{E}\left[\left(b(\tau)\rho'\left(\mathsf{prox}_{b(\tau)\rho}\left(\tau Z\right)\right)\right)^2\right], \tag{1}$$

where $b(\tau)$ obeys

$$\kappa \quad = \quad \mathbb{E}\left[\Psi'\left(\tau Z;\ b(\tau)\right)\right] = 1 - \mathbb{E}\left[\frac{1}{1 + b(\tau)\rho''\left(\mathsf{prox}_{b(\tau)\rho}\left(\tau Z\right)\right)}\right]. \tag{2}$$

In what follows, we study the logistic and probit models separately.

1.1 The logistic case

Consider the bivariate functions

$$h\left(b,\tau\right) := \mathbb{E}\left[\frac{1}{1 + b\rho''\left(\mathsf{prox}_{b\rho}\left(\tau Z\right)\right)}\right],$$
$$w\left(b,\tau\right) = \mathbb{E}\left[\left(\rho'\left(\mathsf{prox}_{b\rho}\left(\tau Z\right)\right)\right)^2\right],$$

which plays a central role in (1) and (2). In the sequel, we will first analyze these two functions for any b obeying

$$b = c_0\tau \tag{3}$$

for some constant $c_0 > 0$. The result is this:

*Department of Statistics, Stanford University, Stanford, CA 94305, U.S.A.
†Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, U.S.A.
‡Department of Mathematics, Stanford University, Stanford, CA 94305, U.S.A.

Lemma 1. *For any constant $c_0 > 0$, one has*

$$\lim_{\tau\to\infty} h\left(c_0\tau,\tau\right) = \mathbb{P}\left\{Z<0 \text{ or } Z>c_0\right\}; \tag{4}$$

$$\lim_{\tau\to\infty} w\left(c_0\tau,\tau\right) = \mathbb{P}\left\{Z>c_0\right\} + \frac{1}{c_0^2}\mathbb{E}\left[Z^2\mathbf{1}_{\{0<Z<c_0\}}\right].$$

Recall that $0<\kappa<1/2$. One can easily find two constants $c_0 > \tilde{c}_0 > 0$ such that

$$\mathbb{P}\left\{Z<0 \text{ or } Z>c_0\right\} < 1-\kappa < \mathbb{P}\left\{Z<0 \text{ or } Z>\tilde{c}_0\right\}.$$

In view of Lemma 1, for any sufficiently large $\tau>0$ one has

$$h(c_0\tau,\tau) < 1-\kappa = h(b(\tau),\tau) < h(\tilde{c}_0\tau,\tau).$$

According to [1, Lemma 5], $h(b,\tau)$ is a monotonic function in b for any given $\tau>0$, thus indicating that

$$b(\tau)\in[\tilde{c}_0\tau, c_0\tau];$$

that said, $b(\tau)$ scales linearly in τ as $\tau\to\infty$. Furthermore, since $b(\tau)$ is the solution to $h(b_\tau,\tau)=1-\kappa$, one has

$$\lim_{\tau\to\infty}\mathbb{P}\left\{Z<0 \text{ or } Z>\frac{b(\tau)}{\tau}\right\} = 1-\kappa,$$

which leads to the closed-form expression

$$\lim_{\tau\to\infty}\frac{b(\tau)}{\tau} = \Phi^{-1}\left(\kappa+0.5\right). \tag{5}$$

We are now ready to characterize the variance map. Note that when τ is sufficiently large,

$$\begin{aligned}
\frac{\mathcal{V}\left(\tau^2\right)}{\tau^2} &= \frac{b^2(\tau)}{\tau^2}\cdot\frac{\mathbb{E}\left[\left(\rho'(\mathsf{prox}_{b(\tau)\rho}\left(\tau Z\right))\right)^2\right]}{1-\mathbb{E}\left[\frac{1}{1+b(\tau)\rho''\left(\mathsf{prox}_{b(\tau)\rho}(\tau Z)\right)}\right]} && (6)\\
&= (1+o(1))\,\frac{b^2(\tau)}{\tau^2}\,\frac{\left\{\mathbb{P}\left\{Z>\frac{b(\tau)}{\tau}\right\}+\frac{\tau^2}{b^2(\tau)}\mathbb{E}\left[Z^2\mathbf{1}_{\left\{0<Z<\frac{b(\tau)}{\tau}\right\}}\right]\right\}}{\mathbb{P}\left\{0<Z<\frac{b(\tau)}{\tau}\right\}} && (7)\\
&= (1+o(1))\left.\frac{x^2\mathbb{P}\left\{Z>x\right\}+\mathbb{E}\left[Z^2\mathbf{1}_{\{0<Z<x\}}\right]}{\mathbb{P}\left\{0<Z<x\right\}}\right|_{x=\frac{b(\tau)}{\tau}}. && (8)
\end{aligned}$$

This together with the expression of $\frac{b(\tau)}{\tau}$ in (5) gives

$$\lim_{\tau\to\infty}\frac{\mathcal{V}\left(\tau^2\right)}{\tau^2} = \left.\frac{x^2\mathbb{P}\left\{Z>x\right\}+\mathbb{E}\left[Z^2\mathbf{1}_{\{0<Z<x\}}\right]}{\mathbb{P}\left\{0<Z<x\right\}}\right|_{x=\Phi^{-1}(\kappa+0.5)}. \tag{9}$$

In order to prove that $\mathcal{V}(\tau^2)\le\tau^2$ for large τ, it suffices to show that the function

$$g\left(x\right) := x^2\mathbb{P}\left\{Z>x\right\}+\mathbb{E}\left[Z^2\mathbf{1}_{\{0<Z<x\}}\right]-\mathbb{P}\left\{0<Z<x\right\}$$

obeys $g(x)<0$ for all $x>0$. To this end, some algebra gives

$$\begin{aligned}
g\left(x\right) &= x^2\int_x^\infty\phi\left(z\right)\mathrm{d}z+\int_0^x z^2\phi\left(z\right)\mathrm{d}z-\int_0^x\phi\left(z\right)\mathrm{d}z\\
&= x^2\int_x^\infty\phi\left(z\right)\mathrm{d}z - z\phi\left(z\right)\Big|_0^x+\int_0^x\phi\left(z\right)\mathrm{d}z-\int_0^x\phi\left(z\right)\mathrm{d}z && (10)\\
&= x\left(x\int_x^\infty\phi\left(z\right)\mathrm{d}z-\phi\left(x\right)\right)<0,
\end{aligned}$$

where (10) comes from integration by parts, and the last inequality follows from $\int_x^\infty \phi(z)\,\mathrm{d}z < \frac{1}{x}\phi(x)$. This establishes that $\mathcal{V}(\tau^2) \le \tau^2$ for any sufficiently large $\tau > 0$.

Finally, we prove Lemma 1.

Proof of Lemma 1. Take $\varepsilon > 0$ to be an arbitrarily small constant. We study $\frac{1}{1+b\rho''\left(\mathsf{prox}_{b\rho}(\tau Z)\right)}$ and $\left(\rho'\left(\mathsf{prox}_{b\rho}(\tau Z)\right)\right)^2$ in three separate cases.

- **Case 1:** $Z \le -\varepsilon$**.** Recall that $\mathsf{prox}_{b\rho}(\tau Z)$ is the solution to

$$b\frac{e^t}{1+e^t} + t = \tau Z, \tag{11}$$

  which implies that

$$\mathsf{prox}_{b\rho}(\tau Z) = \tau Z - b\frac{e^t}{1+e^t}\bigg|_{t=\mathsf{prox}_{b\rho}(\tau Z)} < \tau Z \le -\varepsilon\tau. \tag{12}$$

  When $\tau \to \infty$, this yields

$$\begin{aligned} 0 \;&\le\; b\rho''(\mathsf{prox}_{b\rho}(\tau Z)) = b\frac{e^t}{(1+e^t)^2}\bigg|_{t=\mathsf{prox}_{b\rho}(\tau Z)} \le be^t\big|_{t=\mathsf{prox}_{b\rho}(\tau Z)} \\ &\le\; c_0\tau e^{-\varepsilon\tau} \to 0, \end{aligned}$$

  or equivalently,

$$1 - \frac{1}{1+b\rho''(\mathsf{prox}_{b\rho}(\tau Z))} \to 0 \qquad \text{as} \quad \tau \to \infty.$$

  Similarly, one can derive

$$\left(\rho'\left(\mathsf{prox}_{b\rho}(\tau Z)\right)\right)^2 = \frac{e^{2t}}{(1+e^t)^2}\bigg|_{t=\mathsf{prox}_{b\rho}(\tau Z)} \le e^{2\mathsf{prox}_{b\rho}(\tau Z)} \overset{\text{(a)}}{\le} e^{-2\varepsilon\tau} \to 0,$$

  where (a) follows from (12).

- **Case 2:** $Z \ge \frac{b}{\tau} + \varepsilon$**.** In this case, it holds that

$$\mathsf{prox}_{b\rho}(\tau Z) = \tau Z - b\frac{e^t}{1+e^t}\bigg|_{t=\mathsf{prox}_{\rho}(\tau Z; b)} > \tau\left(\frac{b}{\tau} + \varepsilon\right) - b = \varepsilon\tau.$$

  Applying a similar argument as in the previous case, we see that as $\tau \to \infty$,

$$1 - \frac{1}{1+b\rho''(\mathsf{prox}_{b\rho}(\tau Z))} \to 0 \qquad \text{and} \qquad \left(\rho'\left(\mathsf{prox}_{b\rho}(\tau Z)\right)\right)^2 \to 1.$$

- **Case 3:** $\varepsilon < Z < \frac{b}{\tau} - \varepsilon$**.** We can first rule out the possibility of $\left|\mathsf{prox}_{b\rho}(\tau Z)\right| \gtrsim \tau$. In fact, if $\left|\mathsf{prox}_{b\rho}(\tau Z)\right| \gtrsim \tau$ and $\mathsf{prox}_{b\rho}(\tau Z) \ge 0$, then

$$\begin{aligned} b\frac{e^t}{1+e^t}\bigg|_{t=\mathsf{prox}_{b\rho}(\tau Z)} + \mathsf{prox}_{b\rho}(\tau Z) \ge b\frac{e^t}{1+e^t}\bigg|_{t=\mathsf{prox}_{b\rho}(\tau Z)} &= b - \frac{b}{1+e^{\mathsf{prox}_{b\rho}(\tau Z)}} \\ &\overset{\text{(b)}}{=} b - \frac{c_0\tau}{e^{\Theta(\tau)}} \overset{\text{(c)}}{>} b - \varepsilon\tau > \tau Z, \end{aligned}$$

  where (b) follows from the assumptions $b_0 = c\tau$ and $\left|\mathsf{prox}_{b\rho}(\tau Z)\right| \gtrsim \tau$, and (c) holds when τ is sufficiently large. This violates the identity (11). Similarly, if $\left|\mathsf{prox}_{b\rho}(\tau Z)\right| \gtrsim \tau$ and $\mathsf{prox}_{b\rho}(\tau Z) < 0$, then

$$\begin{aligned} b\frac{e^t}{1+e^t}\bigg|_{t=\mathsf{prox}_{b\rho}(\tau Z)} + \mathsf{prox}_{b\rho}(\tau Z) < b\frac{e^{\mathsf{prox}_{b\rho}(\tau Z)}}{1+e^{\mathsf{prox}_{b\rho}(\tau Z)}} &= c_0\tau\frac{e^{-\left|\mathsf{prox}_{b\rho}(\tau Z)\right|}}{1+e^{-\left|\mathsf{prox}_{b\rho}(\tau Z)\right|}} \\ &\overset{\text{(d)}}{<} \varepsilon\tau \le \tau Z, \end{aligned}$$

where (d) follows when τ is sufficiently large. This inequality contradicts (11) as well. As a result, we reach $\left|\mathsf{prox}_{b\rho}\left(\tau Z\right)\right| = o(\tau)$ in this case, which combined with (11) gives

$$b\frac{e^t}{1+e^t}\bigg|_{t=\mathsf{prox}_{b\rho}(\tau Z)} = \left(1+o\left(1\right)\right)\tau Z. \tag{13}$$

Additionally, (13) leads to

$$\frac{1}{1+e^t}\bigg|_{t=\mathsf{prox}_{b\rho}(\tau Z)} = \left(1+o\left(1\right)\right)\left(1-\frac{\tau Z}{b}\right), \tag{14}$$

which is bounded away from 0 in this case. Taken together, (13) and (14) yield

$$\frac{1}{1+b\rho''(\mathsf{prox}_{b\rho}\left(\tau Z\right))} = \frac{1}{1+b\frac{e^t}{(1+e^t)^2}\Big|_{t=\mathsf{prox}_{b\rho}(\tau Z)}} = \frac{1}{1+\left(1+o\left(1\right)\right)\tau Z\left(1-\frac{\tau Z}{b}\right)} \to 0$$

and

$$\left(\rho'\left(\mathsf{prox}_{b\rho}\left(\tau Z\right)\right)\right)^2 = \left(\frac{e^t}{1+e^t}\right)^2\bigg|_{t=\mathsf{prox}_{b\rho}(\tau Z)} = (1+o(1))\frac{\tau^2 Z^2}{b^2}.$$

Putting the above cases together and applying dominated convergence gives

$$\begin{aligned}\lim_{\tau\to\infty}&\left\{\mathbb{E}\left[\frac{1}{1+b\rho''(\mathsf{prox}_{b\rho}\left(\tau Z\right))}\right]-\mathbb{E}\left[\frac{1}{1+b\rho''(\mathsf{prox}_{b\rho}\left(\tau Z\right))}\mathbf{1}_{\{|Z|\leq\varepsilon\text{ or }|Z-b/\tau|\leq\varepsilon\}}\right]\right\}\\ &= \lim_{\tau\to\infty}\left\{\mathbb{E}\left[\mathbf{1}_{\{Z<-\varepsilon\}}\right]+\mathbb{E}\left[\mathbf{1}_{\left\{Z>\frac{b}{\tau}-\varepsilon\right\}}\right]\right\} = \lim_{\tau\to\infty}\mathbb{P}\left\{Z<-\varepsilon\text{ or }Z>\frac{b}{\tau}+\varepsilon\right\}\end{aligned}$$

when $b = c_0\tau$ for some constant $c_0 > 0$. Recognizing that

$$\mathbb{E}\left[\frac{1}{1+b\rho''(\mathsf{prox}_{b\rho}\left(\tau Z\right))}\mathbf{1}_{\{|Z|\leq\varepsilon\text{ or }|Z-b/\tau|\leq\varepsilon\}}\right] \leq \mathbb{E}\left[\mathbf{1}_{\{|Z|\leq\varepsilon\text{ or }|Z-b/\tau|\leq\varepsilon\}}\right] \leq 4\varepsilon$$

$$\text{and}\qquad \mathbb{P}\left\{-\varepsilon\leq Z\leq 0\text{ or }\frac{b}{\tau}\leq Z\leq\frac{b}{\tau}+\varepsilon\right\}\leq 2\varepsilon,$$

we arrive at

$$\left|\lim_{\tau\to\infty}\mathbb{E}\left[\frac{1}{1+b\rho''(\mathsf{prox}_{b\rho}\left(\tau Z\right))}\right]-\lim_{\tau\to\infty}\mathbb{P}\left\{Z<0\text{ or }Z>\frac{b}{\tau}\right\}\right|\leq 6\varepsilon.$$

Since $\varepsilon > 0$ can be arbitrarily small, we have

$$\lim_{\tau\to\infty}\mathbb{E}\left[\frac{1}{1+b\rho''(\mathsf{prox}_{b\rho}\left(\tau Z\right))}\right] = \lim_{\tau\to\infty}\mathbb{P}\left\{Z<0\text{ or }Z>\frac{b}{\tau}\right\} \tag{15}$$

when $b = c_0\tau$. Similarly,

$$\lim_{\tau\to\infty}\mathbb{E}\left[\left(\rho'\left(\mathsf{prox}_{b\rho}\left(\tau Z\right)\right)\right)^2\right] \quad = \quad \lim_{\tau\to\infty}\left\{\mathbb{P}\left\{Z>\frac{b}{\tau}\right\}+\frac{\tau^2}{b^2}\mathbb{E}\left[Z^2\mathbf{1}_{\left\{0<Z<\frac{b}{\tau}\right\}}\right]\right\}.$$

■

1.2 The probit case

The proof proceeds with the following 3 steps:

(i) Show that for any $b > 0$ and $\epsilon > 0$, there exist constants $c_{1,b}, c_{2,b}, c_3, c_4 > 0$, depending on ϵ, such that

$$\begin{cases} \sup_{z>c_{1,b}} \left|\mathsf{prox}_{b\rho}(z) - \frac{z}{b+1}\right| & \leq \epsilon, \\ \sup_{z<-c_{2,b}} \left|\mathsf{prox}_{b\rho}(z) - z\right| & \leq \epsilon, \end{cases} \quad \text{and} \quad \begin{cases} \sup_{z>c_3} \left|\rho''(z) - 1\right| & \leq \epsilon, \\ \sup_{z<-c_4} \left|\rho''(z)\right| & \leq \epsilon. \end{cases} \tag{16}$$

In particular, one can take

$$c_{1,b} := \max\left\{b\rho'(\sqrt{2}) + \sqrt{2},\ 2\sqrt{2}b,\ \frac{4}{\epsilon}b\right\} \quad \text{and} \quad c_{2,b} := \max\left\{2b\rho'(0),\ \sqrt{8\log\frac{b}{\epsilon}}\right\}. \tag{17}$$

(ii) Show that for any constant $\eta > 0$, for all τ sufficiently large, one has

$$\left|1 - \frac{1}{b(\tau)+1} - 2\kappa\right| \leq \eta. \tag{18}$$

(iii) Show that for any constant $0 < \eta < 1 - 2\kappa$ and for τ sufficiently large, one has

$$\left|\frac{\mathcal{V}(\tau^2)}{\tau^2} - 2\kappa\right| \leq \eta. \tag{19}$$

In the sequel, we elaborate on each of these three steps.

Step (i). Recall that for any $x > 0$, one has $\frac{\phi(x)}{x}\left(1 - \frac{1}{x^2}\right) \leq 1 - \Phi(x) \leq \frac{\phi(x)}{x}$. Since $\rho'(x) = \frac{\phi(x)}{1-\Phi(x)}$, this gives

$$\left|\rho'(x) - x\right| \leq \frac{1}{x - x^{-1}} \leq \frac{2}{x}, \qquad x \geq \sqrt{2}. \tag{20}$$

We start with the first inequality in (16). From the definition of $\mathsf{prox}(\cdot)$, we have the defining relation

$$b\rho'(\mathsf{prox}_{b\rho}(z)) + \mathsf{prox}_{b\rho}(z) = z. \tag{21}$$

Therefore, if we take $z_{b,1} := b\rho'(\sqrt{2}) + \sqrt{2}$, then this identity (21) indicates that $\mathsf{prox}_{b\rho}(z_{b,1}) = \sqrt{2}$. Moreover, $\mathsf{prox}_{b\rho}(z)$ is monotonically increasing in z (see [2, Eqn. (56)]), which tells us that

$$\mathsf{prox}_{b\rho}(z) \geq \mathsf{prox}_{b\rho}(z_{b,1}) = \sqrt{2}, \qquad \forall z > z_{b,1}. \tag{22}$$

Rearranging the identity (21) and combining it with (20) and (22), we obtain

$$z - (b+1)\mathsf{prox}_{b\rho}(z) = b\rho'(\mathsf{prox}_{b\rho}(z)) - b\mathsf{prox}_{b\rho}(z)$$

$$\Longrightarrow \quad \left|\frac{z}{b+1} - \mathsf{prox}_{b\rho}(z)\right| = \frac{b}{b+1}\left|\rho'(\mathsf{prox}_{b\rho}(z)) - \mathsf{prox}_{b\rho}(z)\right| \leq \frac{2b/(b+1)}{\mathsf{prox}_{b\rho}(z)} \tag{23}$$

$$\leq \frac{\sqrt{2}b}{b+1}, \qquad \forall z > z_{b,1}. \tag{24}$$

This inequality provides a lower bound on $\mathsf{prox}_{b\rho}(z)$:

$$\mathsf{prox}_{b\rho}(z) \geq \frac{z - \sqrt{2}b}{b+1} \geq \frac{z}{2(b+1)}$$

for all z obeying $z > z_{b,1}$ and $z > 2\sqrt{2}b$. Substitution into (23) once again gives

$$\left|\frac{z}{b+1} - \mathsf{prox}_{b\rho}(z)\right| \leq \frac{2b/(b+1)}{\mathsf{prox}_{b\rho}(z)} \leq \frac{4b}{z} \leq \epsilon, \qquad \forall z > \max\left\{z_{b,1},\ 2\sqrt{2}b,\ \frac{4b}{\epsilon}\right\},$$

establishing the first bound in (16).

We now turn to the second result in (16). Similarly, it is seen from (21) that $\mathsf{prox}_{b\rho}(z_{b,2}) = 0$ with $z_{b,2} := b\rho'(0) > 0$. The monotonicity of $\mathsf{prox}_{b\rho}(\cdot)$ implies that

$$\mathsf{prox}_{b\rho}(z) \le \mathsf{prox}_{b\rho}(z_{b,2}) = 0, \qquad \forall z < z_{b,2}.$$

Recognizing that $\rho'(x) > 0$ and $\rho''(x) > 0$ for any x and using the relation (21), we arrive at

$$\left|z - \mathsf{prox}_{b\rho}(z)\right| = b\rho'(\mathsf{prox}_{b\rho}(z)) \le b\rho'(0), \qquad \forall z < z_{b,2}, \tag{25}$$

thus indicating that

$$\mathsf{prox}_{b\rho}(z) \le z + b\rho'(0) \le z/2, \qquad \forall z < -2z_{b,2} < 0.$$

Substituting it into (25) and using the fact that $\rho'(x) = \frac{\phi(x)}{1-\Phi(x)} \le 2\phi(x) \le e^{-x^2/2}$ for all $x < 0$, we get

$$\left|z - \mathsf{prox}_{b\rho}(z)\right| = b\rho'(\mathsf{prox}_{b\rho}(z)) \overset{\text{(a)}}{\le} b\rho'(z/2) \le be^{-z^2/8}, \qquad \forall z < -2z_{b,2} < 0, \tag{26}$$

where (a) follows since $\rho''(x) > 0$. The upper bound (26) will not exceed $\epsilon > 0$ as long as $z < -\max\left\{2z_{b,2}, \sqrt{8\log\frac{b}{\epsilon}}\right\}$. This establishes the second bound of (16).

The remaining two inequalities regarding ρ'' are rather straightforward and the proofs are thus omitted.

Step (ii). Recognizing that $\Psi'(z;b) = \left.\frac{b\rho''(x)}{1+b\rho''(x)}\right|_{x=\mathsf{prox}_{b\rho}(z)}$, we see that $b(\tau)$ is the solution to

$$1-\kappa = \mathbb{E}[g(\tau Z, b)] \qquad \text{with } g(x,b) := \frac{1}{1+b\rho''(\mathsf{prox}_{b\rho}(x))}. \tag{27}$$

As a result, everything boils down to quantifying $\mathbb{E}[g(\tau Z, b)]$.

Consider any sufficiently small $\epsilon > 0$. We first obtain an approximation of $\mathbb{E}[g(\tau Z, b)]$. Specifically, we claim that taking $c_\epsilon := \frac{1}{2}\tau\epsilon^2$ leads to

$$\mathbb{E}\left[g(\tau Z, b)\mathbf{1}_{\{|\tau Z| > c_\epsilon\}}\right] \le \mathbb{E}[g(\tau Z, b)] \le \mathbb{E}\left[g(\tau Z, b)\mathbf{1}_{\{|\tau Z| > c_\epsilon\}}\right] + \epsilon. \tag{28}$$

The lower bound is trivial since $0 \le g(x,b) \le 1$. To see why the upper bound holds, we invoke Cauchy-Schwarz to derive

$$\mathbb{E}\left[g(\tau Z, b)\mathbf{1}_{\{|\tau Z| \le c_\epsilon\}}\right] \le \sqrt{\mathbb{E}\left[g^2(\tau Z, b)\right]}\sqrt{\mathbb{P}\left(|Z| \le \frac{c_\epsilon}{\tau}\right)} \overset{\text{(b)}}{\le} \sqrt{\mathbb{P}\left(|Z| \le \frac{c_\epsilon}{\tau}\right)} \le \sqrt{2\frac{c_\epsilon}{\tau}} = \epsilon, \tag{29}$$

where (b) arises since $0 \le g(x,b) \le 1$. This inequality (29) matches the upper bound in (28). In short, we see that $\mathbb{E}\left[g(\tau Z, b)\mathbf{1}_{\{|\tau Z| > c_\epsilon\}}\right]$ is a reasonably tight approximation of $\mathbb{E}\left[g(\tau Z, b)\right]$, and it suffices to look at

$$\mathbb{E}\left[g(\tau Z, b)\mathbf{1}_{\{|\tau Z| > c_\epsilon\}}\right] = \mathbb{E}\left[g(\tau Z, b)\mathbf{1}_{\{\tau Z < -c_\epsilon\}}\right] + \mathbb{E}\left[g(\tau Z, b)\mathbf{1}_{\{\tau Z > c_\epsilon\}}\right]. \tag{30}$$

We first control the second term in the right-hand side of (30). Suppose for the moment that

$$c_\epsilon > \max\left\{c_{1,b},\ (c_3+\epsilon)(b+1),\ c_{2,b},\ c_4+\epsilon\right\}.$$

According to (16), on the event $\{\tau Z > c_\epsilon\}$ one has

$$\frac{\tau Z}{b+1} - \epsilon \le \mathsf{prox}_{b\rho}(\tau Z) \le \frac{\tau Z}{b+1} + \epsilon \quad \text{and} \quad 1-\epsilon \le \rho''(\mathsf{prox}_{b\rho}(\tau Z)) \le 1+\epsilon,$$

where the second inequality holds since $\mathsf{prox}_{b\rho}(\tau Z) \ge \frac{\tau Z}{b+1} - \epsilon > \frac{c_\epsilon}{b+1} - \epsilon \ge c_3$. Plugging these inequalities into (27) gives

$$\frac{1}{1+b(1+\epsilon)} \le g(\tau Z, b) \le \frac{1}{1+b(1-\epsilon)}.$$

In addition, similar to (29) we get

$$\frac{1}{2} \geq \mathbb{P}\left(\tau Z > c_\epsilon\right) = \mathbb{P}\left(\tau Z < -c_\epsilon\right) = \frac{1}{2}\left\{1 - \mathbb{P}\left(|Z| \leq \frac{c_\epsilon}{\tau}\right)\right\} \geq \frac{1}{2}\left\{1 - \frac{2c_\epsilon}{\tau}\right\} = \frac{1}{2}(1-\epsilon^2).$$

The above bounds taken collectively reveal that

$$\frac{1}{1+b(1+\epsilon)} \cdot \frac{1}{2}(1-\epsilon^2) \leq \mathbb{E}\left[g(\tau Z, b)\mathbf{1}_{\{\tau Z > c_\epsilon\}}\right] \leq \frac{1}{1+b(1-\epsilon)} \cdot \frac{1}{2}. \tag{31}$$

We can employ similar arguments to control the first term in the right-hand side of (28) as well. Since $c_\epsilon > \max\{c_{2,b}, c_4 + \epsilon\}$, on the event $\{\tau Z < -c_\epsilon\}$ we have

$$\tau Z - \epsilon \leq \mathsf{prox}_{b\rho}(\tau Z) \leq \tau Z + \epsilon \quad \text{and} \quad -\epsilon \leq \rho''(\mathsf{prox}_{b\rho}(\tau Z)) \leq \epsilon,$$

a direct consequence of (16). This implies that

$$\frac{1}{1+b\epsilon} \leq g(\tau Z, b) \leq \frac{1}{1-b\epsilon}$$

and, therefore,

$$\frac{1}{1+b\epsilon} \cdot \frac{1}{2}(1-\epsilon^2) \leq \mathbb{E}\left[g(\tau Z, b)\mathbf{1}_{\{\tau Z < -c_\epsilon\}}\right] \leq \frac{1}{1-b\epsilon} \cdot \frac{1}{2}. \tag{32}$$

Combining (28), (31) and (32), we conclude that for any $\epsilon > 0$,

$$\frac{1-\epsilon^2}{2}\left\{\frac{1}{1+b(1+\epsilon)} + \frac{1}{1+b\epsilon}\right\} \leq \mathbb{E}\left[g(\tau Z, b)\right] \leq \frac{1}{2}\left\{\frac{1}{1+b(1-\epsilon)} + \frac{1}{1-b\epsilon}\right\} + \epsilon,$$

as long as $c_\epsilon = \frac{1}{2}\tau\epsilon^2 > \max\left\{c_{1,b},\ (c_3+\epsilon)(b+1),\ c_{2,b}, c_4+\epsilon\right\}$, or equivalently,

$$\tau > \frac{2\max\left\{c_{1,b},\ (c_3+\epsilon)(b+1),\ c_{2,b},\ c_4+\epsilon\right\}}{\epsilon^2},$$

where the lower bound is on the order of b/ϵ^3. Effectively, we have established that for any given b and any sufficiently small $\epsilon > 0$ (so that $b\epsilon < 1$ and $\epsilon < 1$), if τ is sufficiently large (as specified above) one has

$$\left|\mathbb{E}\left[g(\tau Z, b)\right] - \frac{1}{2}\left(\frac{1}{1+b} + 1\right)\right| \leq \tilde{c}_4\left(\epsilon + b\epsilon\right) \tag{33}$$

for some universal constant $\tilde{c}_4 > 0$ independent of b, ϵ, τ.

We can then combine this result (33) with the constraint (27) to derive an estimate on $b(\tau)$. Fix any $\eta > 0$. Let b_1 and b_2 be two constants such that

$$\frac{1}{2}\left(\frac{1}{1+b_1} + 1\right) = 1 - \kappa - \frac{\eta}{4}, \qquad \frac{1}{2}\left(\frac{1}{1+b_2} + 1\right) = 1 - \kappa + \frac{\eta}{4}.$$

Picking $\epsilon > 0$ sufficiently small so that $\max\{\tilde{c}_4(1+b_1)\epsilon, \tilde{c}_4(1+b_2)\epsilon\} < \eta/4$ and $\tau \gg \max\left\{b_1, b_2\right\}/\epsilon^3$, we can ensure that

$$\mathbb{E}\left[g(\tau Z, b_1)\right] < 1 - \kappa < \mathbb{E}\left[g(\tau Z, b_2)\right].$$

Recall that for any $\tau > 0$, the function $G(b) := 1 - \mathbb{E}\left[g(\tau Z, b)\right]$ is strictly increasing in b (see [1, Lemma 5]) and, hence,

$$b_2 \ \leq\ b(\tau)\ \leq\ b_1, \qquad \Longrightarrow \qquad \frac{1}{2(1+b_1)} \leq \frac{1}{2(1+b(\tau))} \leq \frac{1}{2(1+b_2)}.$$

Combining these together, we obtain

$$\left|\left(1 - \frac{1}{b(\tau)+1}\right) - 2\kappa\right| \leq \eta, \tag{34}$$

for any $\eta > 0$ with the proviso that τ is sufficiently large. This finishes Step (ii). In particular, this yields

$$\lim_{\tau\to\infty} b(\tau) = \frac{2\kappa}{1-2\kappa}. \tag{35}$$

Step (iii). Now we move on to the variance map

$$\mathcal{V}(\tau^2) = \frac{b(\tau)^2}{\kappa}\,\mathbb{E}\left[\rho'(\mathsf{prox}_{b(\tau)\rho}(\tau Z))^2\right]. \tag{36}$$

For notational convenience, we set

$$h(x) := \rho'(\mathsf{prox}_{b\rho}(x))^2,$$

a key mapping in the definition (36). Before proceeding, we remark that from the properties of ρ', for any $\epsilon > 0$, there exist constants $c_5, c_6 > 0$, depending on ϵ, such that

$$\sup_{z>c_5} |\rho'(z) - z| \le \epsilon, \ \sup_{z<-c_6} |\rho'(z)| \le \epsilon. \tag{37}$$

As before, we decompose the function $\mathcal{V}(\tau^2)$ as follows:

$$\left|\mathcal{V}(\tau^2) - \frac{b(\tau)^2}{\kappa}\,\mathbb{E}\left[h(\tau Z)\mathbf{1}_{\{|\tau Z|>\alpha_\epsilon\}}\right]\right| = \frac{b(\tau)^2}{\kappa}\,\mathbb{E}\left[h(\tau Z)\mathbf{1}_{\{|\tau Z|\le\alpha_\epsilon\}}\right]$$

for some point $\alpha_\epsilon > 0$ to be specified later. This gives

$$\mathbb{E}[h(\tau Z)\mathbf{1}_{\{|\tau Z|<\alpha_\epsilon\}}] \le \sqrt{\mathbb{E}[h^2(\tau Z)\mathbf{1}_{\{|\tau Z|<\alpha_\epsilon\}}]}\sqrt{\mathbb{P}\left(|\tau Z| < \alpha_\epsilon\right)} \le C(\alpha_\epsilon, b)\sqrt{2\Phi\left(\frac{\alpha_\epsilon}{\tau}\right) - 1}, \tag{38}$$

where

$$C(\alpha_\epsilon, b) = \rho'(\mathsf{prox}_{b\rho}(\alpha_\epsilon))^2.$$

The last inequality of (38) holds since (1) $\rho'(z) \ge 0$ is an increasing function of z; (2) $\mathsf{prox}_{b\rho}(x)$ is an increasing function of x (see [2, Eqn. (56)]). For any given $\epsilon > 0$, one can pick τ sufficiently large so that the above bound $C(\alpha_\epsilon, b)\sqrt{2\Phi\left(\frac{\alpha_\epsilon}{\tau}\right) - 1}$ is below ϵ. The particular choice of τ will be made clear later. Under these conditions,

$$\mathbb{E}[h(\tau Z)\mathbf{1}_{\{|\tau Z|>\alpha_\epsilon\}}] \le \mathbb{E}[h(\tau Z)] \le \mathbb{E}[h(\tau Z)\mathbf{1}_{\{\tau Z<-\alpha_\epsilon\}}] + \mathbb{E}[h(\tau Z)\mathbf{1}_{\{\tau Z>\alpha_\epsilon\}}] + \epsilon. \tag{39}$$

We first control the second term in the right-hand side of (39). To this end, we choose

$$\alpha_\epsilon > \max\left\{c_{1,b}, c_{2,b}, \left(c_5 + \epsilon\right)\left(b+1\right), c_6 + 2\epsilon\right\}$$

as before. Then from (16) and (37), on the event $\{\tau Z > \alpha_\epsilon\}$ we have

$$\frac{\tau Z}{b+1} - \epsilon \le \mathsf{prox}_{b\rho}(\tau Z) \le \frac{\tau Z}{b+1} + \epsilon \qquad \text{and} \qquad \frac{\tau Z}{b+1} - 2\epsilon \le \rho'(\mathsf{prox}_{b\rho}(\tau Z)) \le \frac{\tau Z}{b+1} + 2\epsilon.$$

This yields

$$\left(\frac{\tau Z}{b+1} - 2\epsilon\right)^2 \le h(\tau Z) \le \left(\frac{\tau Z}{b+1} + 2\epsilon\right)^2$$

on the event $\{\tau Z > \alpha_\epsilon\}$, and hence

$$\mathbb{E}\left[\left(\frac{\tau Z}{b+1} - 2\epsilon\right)^2 \mathbf{1}_{\{\tau Z>\alpha_\epsilon\}}\right] \le \mathbb{E}[h(\tau Z)\mathbf{1}_{\{\tau Z>\alpha_\epsilon\}}] \le \mathbb{E}\left[\left(\frac{\tau Z}{b+1} + 2\epsilon\right)^2 \mathbf{1}_{\{\tau Z>\alpha_\epsilon\}}\right]. \tag{40}$$

Similarly for the first term in the right-hand side of (39), as $\alpha_\epsilon > \max\{c_2, c_6 + 2\epsilon\}$, on the event $\{\tau Z < -\alpha_\epsilon\}$, we have

$$\tau Z - \epsilon \le \mathsf{prox}_{b\rho}(\tau Z) \le \tau Z + \epsilon \qquad \text{and} \qquad -\epsilon \le \rho'(\mathsf{prox}_{b\rho}(\tau Z) \le \epsilon.$$

Note that $\mathbb{P}(\tau Z > \alpha_\epsilon) = \mathbb{P}(\tau Z < -\alpha_\epsilon) = \frac{1}{2}(1-\delta_\epsilon)$ for some δ_ϵ small which is a function of ϵ and which vanishes as $\epsilon \to 0$. This yields

$$0 \le \mathbb{E}[h(\tau Z)\mathbf{1}_{\{\tau Z < -\alpha_\epsilon\}}] \le \frac{\epsilon^2}{2}(1-\delta_\epsilon). \tag{41}$$

Combining the relations (39), (40) and (41) we obtain that

$$\frac{b^2}{\kappa}\mathbb{E}\left[\left(\frac{\tau Z}{b+1} - 2\epsilon\right)^2 \mathbf{1}_{\{\tau Z > \alpha_\epsilon\}}\right] \le \mathcal{V}(\tau^2) \le \frac{b^2}{\kappa}\left\{\mathbb{E}\left[\left(\frac{\tau Z}{b+1} + 2\epsilon\right)^2 \mathbf{1}_{\{\tau Z > \alpha_\epsilon\}}\right] + \frac{\epsilon^2}{2}(1-\delta_\epsilon) + \epsilon\right\}. \tag{42}$$

We still need to evaluate $\mathbb{E}\left[\left(\frac{\tau Z}{b+1} - 2\epsilon\right)^2 \mathbf{1}_{\{\tau Z > \alpha_\epsilon\}}\right]$. To this end, we define two quantities

$$\alpha_1 := \mathbb{E}\left[Z\mathbf{1}_{\{\tau Z > \alpha_\epsilon\}}\right] \qquad \text{and} \qquad \alpha_2 := \mathbb{E}\left[Z^2\mathbf{1}_{\{\tau Z > \alpha_\epsilon\}}\right].$$

Using the properties of the normal CDF, one can show that

$$\frac{\tau}{\sqrt{2\pi}} - \alpha_\epsilon\frac{\delta_\epsilon}{2} \le \tau\alpha_1 \le \frac{\tau}{\sqrt{2\pi}} \qquad \text{and} \qquad \frac{\tau^2}{2} - \alpha_\epsilon^2\frac{\delta_\epsilon}{2} \le \tau^2\alpha_2 \le \frac{\tau^2}{2}. \tag{43}$$

Using the above relations and rearranging, the bounds in (42) can be rewritten as

$$\begin{aligned}
\mathcal{V}(\tau^2) &\ge \frac{b^2}{\kappa}\left[\frac{\tau^2}{2(b+1)^2} - \frac{\alpha_\epsilon^2\delta_\epsilon}{2(b+1)^2} - 4\epsilon\frac{\tau}{\sqrt{2\pi}(b+1)} + 2\epsilon^2(1-\delta_\epsilon)\right];\\
\mathcal{V}(\tau^2) &\le \frac{b^2}{\kappa}\left[\frac{\tau^2}{2(b+1)^2} + \epsilon\left(\frac{4\tau}{\sqrt{2\pi}(b+1)} + 1\right) + \frac{5}{2}\epsilon^2(1-\delta_\epsilon)\right].
\end{aligned}$$

Finally, observing that $b \ge 0$, we arrive at

$$\left|\mathcal{V}(\tau^2) - \frac{b^2}{2\kappa}\frac{\tau^2}{(b+1)^2}\right| \le \frac{b^2}{\kappa}\left\{\epsilon\left(\frac{8\tau}{\sqrt{2\pi}} + 1\right) + \frac{\delta_\epsilon\alpha_\epsilon^2}{2} + \frac{\epsilon^2}{2}(1-\delta_\epsilon)\right\},$$

which is equivalent to

$$\left|\frac{\mathcal{V}(\tau^2)}{\tau^2} - \frac{1}{2\kappa}\left(1 - \frac{1}{b+1}\right)^2\right| \le \frac{b^2}{\kappa}\left\{\epsilon\left(\frac{8}{\sqrt{2\pi}\tau} + \frac{1}{\tau^2}\right) + \frac{\delta_\epsilon\alpha_\epsilon^2}{2\tau^2} + \frac{\epsilon^2}{2\tau^2}(1-\delta_\epsilon)\right\}. \tag{44}$$

Note that in the bound above α_ϵ also depends on b. Henceforth we denote α_ϵ as $\alpha_\epsilon(b)$. Next, we invoke the result from Step (ii) to ensure that $b(\tau)$ is bounded for all sufficiently large values of τ.

Fix $\eta' > 0$ such that $0 < \eta' < 1 - 2\kappa$. Let τ_0 be the threshold above which for all values of τ the relation (34) holds with $\eta = \eta'/2$. Then $\forall \tau \ge \tau_0$, one has

$$b(\tau) \le \frac{2\kappa + \eta'}{1 - 2\kappa - \eta'} =: a(\eta').$$

For all $\tau \ge \tau_0$, we have

$$\left|\frac{\mathcal{V}(\tau^2)}{\tau^2} - \frac{1}{2\kappa}\left(1 - \frac{1}{b+1}\right)^2\right| \le \frac{a(\eta')^2}{\kappa}\left\{\epsilon\left(\frac{8}{\sqrt{2\pi}\tau} + \frac{1}{\tau^2}\right) + \frac{\delta_\epsilon(\alpha_\epsilon(a(\eta)))^2}{2\tau^2} + \frac{\epsilon^2}{2\tau^2}(1-\delta_\epsilon)\right\},$$

where $\alpha_\epsilon(a(\eta))$ is any constant above $\max\{c_{1,a(\eta)}, c_{2,a(\eta)}, (c_5+\epsilon)(a(\eta)+1), c_6 + 2\epsilon\}$. We choose $\tau > \tau_0$ so that $C(\alpha_\epsilon(a(\eta)), a(\eta))\sqrt{2\Phi(\alpha_\epsilon) - 1}$ is below ϵ, and the above bound in the RHS is below $\eta = \eta'/2$. This gives

$$\left|\frac{\mathcal{V}(\tau^2)}{\tau^2} - 2\kappa\right| \le \left|\frac{\mathcal{V}(\tau^2)}{\tau^2} - \frac{1}{2\kappa}\left(1 - \frac{1}{b+1}\right)^2\right| + \left|2\kappa - \frac{1}{2\kappa}\left(1 - \frac{1}{b+1}\right)^2\right| \le \eta'.$$

Hence, for any such τ

$$\frac{\mathcal{V}(\tau^2)}{\tau^2} \leq 2\kappa + \eta' < 1,$$

from the choice of η'. In particular, we have established that

$$\lim_{\tau\to\infty} \frac{\mathcal{V}(\tau^2)}{\tau^2} = 2\kappa.$$

Remark 1. In fact, the above analysis works for a broader class of link functions beyond the probit case. Specifically, more general sufficient conditions for the above result to hold are the following: in addition to conditions mentioned in [1, Section 2.3.3].

- $\rho'(x) \to 0$ when $x \to -\infty$, and $\rho'(x)/x \to 1$, when $x \to \infty$; further, $|\rho'(x) - x| \leq f(x)$ for all x positive, where $f(x)$ is some function obeying $f(x) \to 0$ when $x \to \infty$.
- ρ'' is bounded, converges to 1 when $x \to \infty$ and converges to 0 when $x \to -\infty$. $-\infty$ are swapped.
- In addition, for any given z, $b\rho''(\mathsf{prox}_{b\rho}(z)) \to \infty$ when $b \to \infty$.

References

[1] Pragya Sur, Yuxin Chen, and Emmanuel Candès. The likelihood ratio test in high-dimensional logistic regression is asymptotically a *rescaled* chi-square. 2017.

[2] David Donoho and Andrea Montanari. High dimensional robust M-estimation: Asymptotic variance via approximate message passing. *Probability Theory and Related Fields*, pages 1–35, 2013.