Supplementary Material:
On Some Fast And Robust Classifiers For High Dimension, Low Sample Size Data

A MATHEMATICAL DETAILS AND PROOFS

We will use the following definitions while presenting the mathematical details.

1. $a_n = o(b_n)$ as $n \to \infty$ implies that for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $|a_n/b_n| < \epsilon$ for all $n \geq N$.

2. $a_n = O(b_n)$ as $n \to \infty$ implies that there exist $M > 0$ and $N \in \mathbb{N}$ such that $|a_n/b_n| < M$ for all $n \geq N$.

~~The~~ ~~R~~esults up to Theorem 3.3 on concentration of probabilities in the HDLSS asymptotic regime. They are derived for a fixed n and $p \to \infty$.

Proof of Lemma 2.1

Suppose $\mathbf{U} \sim \mathbf{F}_j$ and $\mathbf{V} \sim \mathbf{F}_{j'}$ for $j, j' \in \{1, 2\}$ and $\mathbf{U}, \mathbf{V}$ are independent. We have assumed in (b) that the limiting constants $\nu_{jj'}$ and σ_j^2 for $j, j' \in \{1, 2\}$ exist. Fix $\epsilon > 0$. Now, observe that

$$\begin{aligned}
\mathrm{P}\left[\left|\frac{1}{p}\mathbf{U}^\top\mathbf{V} - \nu_{jj'}\right| > \epsilon\right] &= \mathrm{P}\left[\left|\frac{1}{p}\mathbf{U}^\top\mathbf{V} - \frac{1}{p}\boldsymbol{\mu}_j^\top\boldsymbol{\mu}_{j'} + \frac{1}{p}\boldsymbol{\mu}_j^\top\boldsymbol{\mu}_{j'} - \nu_{jj'}\right| > \epsilon\right] \\
&\leq \mathrm{P}\left[\left|\frac{1}{p}\mathbf{U}^\top\mathbf{V} - \frac{1}{p}\boldsymbol{\mu}_j^\top\boldsymbol{\mu}_{j'}\right| > \frac{\epsilon}{2}\right] + \mathrm{I}\left[\left|\frac{1}{p}\boldsymbol{\mu}_j^\top\boldsymbol{\mu}_{j'} - \nu_{jj'}\right| > \frac{\epsilon}{2}\right] \text{ [using the union bound].}
\end{aligned}$$

~~Also,~~ Since $\lim_{p\to\infty} \boldsymbol{\mu}_j^\top\boldsymbol{\mu}_{j'} = \nu_{jj'}$, there exists $p_0 \in \mathbb{N}$ such that $\mathrm{I}\left[\left|\frac{1}{p}\boldsymbol{\mu}_j^\top\boldsymbol{\mu}_{j'} - \nu_{jj'}\right| > \frac{\epsilon}{2}\right] = 0$ for all $p \geq p_0$. So, we get

$$\mathrm{P}\left[\left|\frac{1}{p}\mathbf{U}^\top\mathbf{V} - \nu_{jj'}\right| > \epsilon\right] \leq \mathrm{P}\left[\left|\frac{1}{p}\mathbf{U}^\top\mathbf{V} - \frac{1}{p}\boldsymbol{\mu}_j^\top\boldsymbol{\mu}_{j'}\right| > \frac{\epsilon}{2}\right] \text{ for all } p \geq p_0.$$

Observe that

$$\begin{aligned}
&\mathrm{P}\left[\left|\frac{1}{p}\mathbf{U}^\top\mathbf{V} - \frac{1}{p}\boldsymbol{\mu}_j^\top\boldsymbol{\mu}_{j'}\right| > \frac{\epsilon}{2}\right] \qquad \text{(A.1)} \\
&= \mathrm{P}\left[\left|\frac{1}{p}\sum_{k=1}^{p} U_k V_k - \frac{1}{p}\sum_{k=1}^{p}\mathrm{E}[U_k]\mathrm{E}[V_k]\right| > \frac{\epsilon}{2}\right] \\
&\leq \frac{4}{\epsilon^2}\mathrm{Var}\left[\frac{1}{p}\sum_{k=1}^{p} U_k V_k\right] \text{ [using Chebyshev's inequality]} \\
&= \frac{4}{\epsilon^2 p^2}\sum_{k=1}^{p}\mathrm{Var}[U_k V_k] + \frac{8}{\epsilon^2 p^2}\sum\sum_{1\leq k<k'\leq p}\mathrm{Cov}\left(U_k V_k, U_{k'} V_{k'}\right) \\
&\leq \frac{4}{\epsilon^2 p^2}\sum_{k=1}^{p}\mathrm{E}[U_k^2 V_k^2] + \frac{8}{\epsilon^2 p^2}\sum\sum_{1\leq k<k'\leq p}\mathrm{Corr}\left(U_k V_k, U_{k'} V_{k'}\right)\sqrt{\mathrm{E}[U_k^2 V_k^2]\ \mathrm{E}[U_{k'}^2 V_{k'}^2]}
\end{aligned}$$

$$\leq \frac{4C}{\epsilon^2 p} + \frac{8C}{\epsilon^2 p^2} \sum\sum_{1\leq k<k'\leq p} \text{Corr}\,(U_kV_k, U_{k'}V_{k'}) \;\; [\text{for some } C<\infty \; (\cancel{\text{due to}} \text{ using } (\text{a}))]$$

$$= o(1) \text{ as } p\to\infty \; [\text{follows from (c)}]. \tag{A.2}$$

Therefore, $\mathrm{P}\left[\left|\frac{1}{p}\mathbf{U}^\top\mathbf{V}-\nu_{jj'}\right|>\epsilon\right] \leq \mathrm{P}\left[\left|\frac{1}{p}\mathbf{U}^\top\mathbf{V}-\frac{1}{p}\boldsymbol{\mu}_j^\top\boldsymbol{\mu}_{j'}\right|>\frac{\epsilon}{2}\right] = o(1)$ for $\mathbf{U}\sim\mathbf{F}_j$ and $\mathbf{V}\sim\mathbf{F}_{j'}$ with $j,j'\in\{1,2\}$ as $p\to\infty$.

Following similar arguments, one can also prove that (as $p\to\infty$),

$$\mathrm{P}\left[\left|\frac{1}{p}\|\mathbf{U}\|^2 - \frac{1}{p}\mathrm{E}[\|\mathbf{U}\|^2]\right|>\epsilon\right] \leq o(1)$$

$$\Rightarrow \mathrm{P}\left[\left|\frac{1}{p}\|\mathbf{U}\|^2 - \frac{1}{p}\{\|\boldsymbol{\mu}_i\|^2 + tr(\Sigma_j)\}\right|>\epsilon\right] \leq o(1)$$

$$\Rightarrow \mathrm{P}\left[\left|\frac{1}{p}\|\mathbf{U}\|^2 - \{\nu_{jj}+\sigma_j^2\}\right|>\epsilon\right] \leq o(1) \; [\cancel{\text{since}} \lim_{p\to\infty}\|\boldsymbol{\mu}_j\|^2/p=\nu_{jj} \text{ and } \lim_{p\to\infty} tr(\Sigma_j)=\sigma_j^2].$$

p?

Using the continuous mapping theorem (repeatedly), we obtain

$$\sin(2\pi h(\mathbf{U},\mathbf{V})) = \frac{1+\mathbf{U}^\top\mathbf{V}}{\sqrt{(1+\|\mathbf{U}\|^2)(1+\|\mathbf{V}\|^2)}} = \frac{\frac{1}{p}+\frac{\mathbf{U}^\top\mathbf{V}}{p}}{\sqrt{\left(\frac{1}{p}+\frac{\|\mathbf{U}\|^2}{p}\right)\left(\frac{1}{p}+\frac{\|\mathbf{V}\|^2}{p}\right)}} \xrightarrow{\mathrm{P}} \frac{\nu_{jj'}}{\sqrt{(\sigma_j^2+\nu_{jj})(\sigma_{j'}^2+\nu_{j'j'})}}$$

as $p\to\infty$. Hence, the proof. □

Note that $h(\mathbf{U},\mathbf{V}) \xrightarrow{\mathrm{P}} \frac{1}{2\pi}\sin^{-1}\left\{\frac{\nu_{jj'}}{\sqrt{(\sigma_j^2+\nu_{jj})(\sigma_{j'}^2+\nu_{j'j'})}}\right\}$ as $p\to\infty$. We define $\tau_{ii}=\frac{1}{2\pi}\sin^{-1}\left\{\frac{\nu_{ii}}{(\sigma_i^2+\nu_{ii})}\right\}$ for $i=1,2$ and $\tau_{12}=\frac{1}{2\pi}\sin^{-1}\left\{\frac{\nu_{12}}{\sqrt{(\sigma_1^2+\nu_{11})(\sigma_2^2+\nu_{22})}}\right\}$. Lemma 2.1 suggests that $h(\mathbf{U},\mathbf{V})\xrightarrow{\mathrm{P}}\tau_{jj'}$ as $p\to\infty$, where $\mathbf{U}\sim\mathbf{F}_j$, $\mathbf{V}\sim\mathbf{F}_{j'}$ for $j,j'\in\{1,2\}$ and $\mathbf{U},\mathbf{V}$ are independently distributed.

Corollary A.1 *For $j,j'\in\{1,2\}$, if assumptions* (a)-(c) *are satisfied, then*

(a) $|T_{jj'}-\tau_{jj'}|\xrightarrow{\mathrm{P}}0$ as $p\to\infty$, and

(b) if $\mathbf{Z}\sim\mathbf{F}_{j'}$, then $|T_j(\mathbf{Z})-\tau_{jj'}|\xrightarrow{\mathrm{P}}0$ as $p\to\infty$.

Proof of Corollary A.1

(a) Fix $\epsilon>0$. It follows from Lemma 2.1 that

$$\begin{aligned}
&\mathrm{P}[|T_{11}-\tau_{11}|>\epsilon] \\
&= \mathrm{P}\left[\left|\frac{1}{n_1(n_1-1)}\sum_{1\leq i\neq j\leq n_1}\{h(\mathbf{X}_i,\mathbf{X}_j)-\tau_{11}\}\right|>\epsilon\right] \\
&\leq \mathrm{P}\left[\frac{1}{n_1(n_1-1)}\sum_{1\leq i\neq j\leq n_1}|h(\mathbf{X}_i,\mathbf{X}_j)-\tau_{11}|>\epsilon\right] \\
&\leq \sum_{1\leq i\neq j\leq n_1}\mathrm{P}[|h(\mathbf{X}_i,\mathbf{X}_j)-\tau_{11}|>\epsilon] \\
&= n_1(n_1-1)o(1) = o(1) \text{ as } p\to\infty \; [\cancel{\text{since}} \; n_1 \text{ is fixed}].
\end{aligned} \tag{A.3}$$

ΣΣ ?

Therefore, $|T_{11}-\tau_{11}|\xrightarrow{\mathrm{P}}0$ as $p\to\infty$. Similarly, $|T_{12}-\tau_{12}|$ and $|T_{22}-\tau_{22}|$ also converge to 0 in probability as $p\to\infty$.

(b) Fix $\epsilon > 0$. Let $\mathbf{U} \in \mathcal{X}_i$ (i.e., $\mathbf{U} \sim \mathcal{X}_j$) and $\mathbf{Z} \sim \mathbf{F}_{j'}$ for $j, j' \in \{1, 2\}$. Since n_j is fixed for $j \in \{1, 2\}$, from Lemma 2.1, we have

$$\begin{aligned}
\mathrm{P}\big[|T_j(\mathbf{Z}) - \tau_{jj'}| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_{j'}\big] &= \mathrm{P}\left[\left|\left\{\frac{1}{n_j}\sum_{\mathbf{U}\in\mathcal{X}_j}\{h(\mathbf{U},\mathbf{Z}) - \mathrm{E}[h(\mathbf{U},\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}_{j'}]\}\right\}\right| > \epsilon \,\middle|\, \mathbf{Z} \sim \mathbf{F}_{j'}\right] \\
&\leq \mathrm{P}\left[\frac{1}{n_j}\sum_{\mathbf{U}\in\mathcal{X}_j}|h(\mathbf{U},\mathbf{Z}) - \mathrm{E}[h(\mathbf{U},\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}_{j'}]| > \epsilon \,\middle|\, \mathbf{Z} \sim \mathbf{F}_{j'}\right] \\
&\leq \sum_{\mathbf{U}\in\mathcal{X}_j}\mathrm{P}\left[|h(\mathbf{U},\mathbf{Z}) - \mathrm{E}[h(\mathbf{U},\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}_{j'}]| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_{j'}\right] \\
&\leq n_j o(1) = o(1) \text{ as } p \to \infty \text{ and } n_j \text{ is fixed.} \qquad \text{(A.4)}
\end{aligned}$$

twice?

Hence, the proof. □

Recall the definition of τ_0 given as

$$\tau_0 = \frac{1}{2\pi}\sin^{-1}\left\{\frac{\nu_{11}}{(\sigma_1^2 + \nu_{11})}\right\} + \frac{1}{2\pi}\sin^{-1}\left\{\frac{\nu_{22}}{(\sigma_2^2 + \nu_{22})}\right\} - \frac{1}{\pi}\sin^{-1}\left\{\frac{\nu_{12}}{\sqrt{(\sigma_1^2 + \nu_{11})(\sigma_2^2 + \nu_{22})}}\right\}$$

i.e., $\tau_0 = \tau_{11} + \tau_{22} - 2\tau_{12}$.

~~Note that~~ If if $\nu_{11} = \nu_{12} = \nu_{22} = 0$, then $\tau_0 = 0$. Also, if $\nu_{11} = \nu_{12} = \nu_{22}$ and $\sigma_1^2 = \sigma_2^2$, then $\tau_0 = 0$.

Corollary A.2 *Suppose assumptions* (a)-(c) *are satisfied. Let* $\mathbf{Z} \in \mathbb{R}^p$ *be a test observation.*

$z \Leftrightarrow \mathbf{Z}$

(a) If $\mathbf{Z} \sim \mathbf{F}_1$, *then* $|L_2(\mathbf{Z}) - L_1(\mathbf{Z}) - \tau_0| \xrightarrow{\mathrm{P}} 0$ *as* $p \to \infty$.

(b) If $\mathbf{Z} \sim \mathbf{F}_2$, *then* $|L_2(\mathbf{Z}) - L_1(\mathbf{Z}) + \tau_0| \xrightarrow{\mathrm{P}} 0$ *as* $p \to \infty$.

$\tau_0 \Rightarrow$ free of p?

Proof of Corollary A.2

(a) Note that

$$L_2(\mathbf{Z}) - L_1(\mathbf{Z}) = \{T_{22} - 2T_2(\mathbf{Z})\} - \{T_{11} - 2T_1(\mathbf{Z})\}.$$

If $\mathbf{Z} \sim \mathbf{F}_1$, then it follows from Corollary A.1 that $L_2(\mathbf{Z}) - L_1(\mathbf{Z})$ converges in probabiliy to $\{\tau_{22} - 2\tau_{12}\} - \{\tau_{11} - 2\tau_{11}\} = \tau_{11} + \tau_{22} - 2\tau_{12} = \tau_0$ as $p \to \infty$.

(b) Similarly, if $\mathbf{Z} \sim \mathbf{F}_2$, $L_2(\mathbf{Z}) - L_1(\mathbf{Z})$ converges in probabiliy to $\{\tau_{22} - 2\tau_{22}\} - \{\tau_{11} - 2\tau_{12}\} = -\{\tau_{11} + \tau_{22} - 2\tau_{12}\} = -\tau_0$ as $p \to \infty$.

Hence, the proof. □

Proof of Theorem 2.2

The prior probability of an observation $\mathbf{Z}$ belonging to the first class is given by π_1 $(0 < \pi_1 < 1)$. The misclassification probability of δ_0 is as follows:

$$\begin{aligned}
\mathrm{P}[\delta_0(\mathbf{Z}) \neq \text{ true label of } \mathbf{Z}] &= \pi_1\mathrm{P}[\delta_0(\mathbf{Z}) = 2 \mid \mathbf{Z} \sim \mathbf{F}_1] + \pi_2\mathrm{P}[\delta_0(\mathbf{Z}) = 1 \mid \mathbf{Z} \sim \mathbf{F}_2] \\
&= \pi_1\mathrm{P}[L_2(\mathbf{Z}) \leq L_1(\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}_1] + \pi_2\mathrm{P}[L_2(\mathbf{Z}) > L_1(\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}_2]. \qquad \text{(A.5)}
\end{aligned}$$

We have assumed that either (a) $\nu_{11}, \nu_{12}, \nu_{22}$ are unequal, or (b) $\nu_{11} = \nu_{12} = \nu_{22} \neq \boxed{0}$, and $\sigma_1^2 = \sigma_2^2$. As a ~~result~~ consequence, τ_0 is strictly positive. Fix $0 < \epsilon < \tau_0$. Now, we have

$$\mathrm{P}\big[L_2(\mathbf{Z}) \leq L_1(\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}_1\big] \leq \mathrm{P}\big[L_2(\mathbf{Z}) - L_1(\mathbf{Z}) \leq \tau_0 - \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\big]$$

$\epsilon > \tau_0$?

follows from → using?

$$
\begin{aligned}
&\le \mathrm{P}\big[L_2(\mathbf{Z}) - L_1(\mathbf{Z}) - \tau_0 \le -\epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\big] \\
&\le \mathrm{P}\big[|L_2(\mathbf{Z}) - L_1(\mathbf{Z}) - \tau_0| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\big] \\
&= o(1) \text{ as } p \to \infty \text{ [follows from Corollary A.2(a))} \qquad \text{(A.6)}
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\mathrm{P}\big[L_2(\mathbf{Z}) > L_1(\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}_2\big] &\le \mathrm{P}\big[L_2(\mathbf{Z}) - L_1(\mathbf{Z}) > -\tau_0 + \epsilon \mid \mathbf{Z} \sim \mathbf{F}_2\big] \\
&\le \mathrm{P}\big[L_2(\mathbf{Z}) - L_1(\mathbf{Z}) + \tau_0 > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_2\big] \\
&\le \mathrm{P}\big[|L_2(\mathbf{Z}) - L_1(\mathbf{Z}) + \tau_0| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_2\big] \\
&= o(1) \text{ as } p \to \infty \text{ [follows from Corollary A.2(b))} \qquad \text{(A.7)}
\end{aligned}
$$

Combining Eqs. (A.5), (A.6) and (A.7), we get $\mathrm{P}[\delta_0(\mathbf{Z}) \neq$ true label of $\mathbf{Z}] = o(1)$ as $p \to \infty$. □

Lemma A.3 *For $j, j' \in \{1, 2\}$, if A1 is satisfied, then*

(a) $|\bar{T}_{jj'} - \bar{\tau}_p(j, j')| \xrightarrow{\mathrm{P}} 0$ *as* $p \to \infty$, *and*

(b) if $\mathbf{Z} \sim \mathbf{F}_j$, *then* $|\bar{T}_{j'}(\mathbf{Z}) - \bar{\tau}_p(j, j')| \xrightarrow{\mathrm{P}} 0$ *as* $p \to \infty$.

Proof of Lemma A.3

(a) Recall the definitions of $\bar{T}_{11}$ and $\bar{\tau}_p(1,1)$ given in (2.4) and (2.5), respectively. Fix $\epsilon > 0$. We have

$$
\begin{aligned}
&\mathrm{P}\big[|\bar{T}_{11} - \bar{\tau}_p(1,1)| > \epsilon\big] \\
&= \mathrm{P}\left[\left|\frac{1}{n_1(n_1-1)} \sum_{1 \le i \neq j \le n_1} \bar{h}_p(\mathbf{X}_i, \mathbf{X}_j) - \mathrm{E}[\bar{h}_p(\mathbf{X}_1, \mathbf{X}_2)]\right| > \epsilon\right] \\
&= \mathrm{P}\left[\left|\frac{1}{p}\sum_{k=1}^{p} \frac{1}{n_1(n_1-1)} \sum_{1 \le i \neq j \le n_1} h(X_{ik}, X_{jk}) - \frac{1}{p}\sum_{k=1}^{p} \mathrm{E}\big[h(X_{1k}, X_{2k})\big]\right| > \epsilon\right] \text{ (follows from the definition of } \bar{h}_p) \\
&= \mathrm{P}\left[\left|\frac{1}{n_1(n_1-1)} \sum_{1 \le i \neq j \le n_1} \frac{1}{p}\sum_{k=1}^{p} h(X_{ik}, X_{jk}) - \frac{1}{p}\sum_{k=1}^{p} \mathrm{E}\big[h(X_{1k}, X_{2k})\big]\right| > \epsilon\right] \\
&\le \mathrm{P}\left[\frac{1}{n_1(n_1-1)} \sum_{1 \le i \neq j \le n_1} \left|\frac{1}{p}\sum_{k=1}^{p} h(X_{ik}, X_{jk}) - \frac{1}{p}\sum_{k=1}^{p} \mathrm{E}\big[h(X_{1k}, X_{2k})\big]\right| > \epsilon\right] \text{ (using } \cancel{\text{the}} \text{ triangle ineq.)} \\
&\le \sum_{1 \le i \neq j \le n_1} \mathrm{P}\left[\left|\frac{1}{p}\sum_{k=1}^{p} \{h(X_{ik}, X_{jk}) - \mathrm{E}\big[h(X_{1k}, X_{2k})\big]\}\right| > \epsilon\right] \text{ (using the union bound)} \\
&\le \sum_{1 \le i \neq j \le n_1} \frac{1}{\epsilon^2} \mathrm{Var}\left[\frac{1}{p}\sum_{k=1}^{p} h(X_{ik}, X_{jk})\right] \text{ (}\overset{\text{using}}{\text{follows from}}\text{ Chebyshev's inequality).} \qquad \text{(A.8)}
\end{aligned}
$$

We will show that $\mathrm{Var}\,[\sum_{k=1}^{p} h(X_{ik}, X_{jk})/p]$ converges in probability to 0 for all $i \neq j$ as $p \to \infty$. Fix $1 \le i, j \le n_1$ with $i \neq j$. Observe that

$$
\mathrm{Var}\left[\frac{1}{p}\sum_{k=1}^{p} h(X_{ik}, X_{jk})\right] = \frac{1}{p^2}\sum_{k=1}^{p} \mathrm{Var}\big[h(X_{ik}, X_{jk})\big] + \frac{2}{p^2} \sum\sum_{1 \le k < k' \le p} \mathrm{Cov}\left(h(X_{ik}, X_{jk}), h(X_{ik'}, X_{jk'})\right). \qquad \text{(A.9)}
$$

Since, $0 \le h \le 1$, we have $\mathrm{Var}\big[h(X_{ik}, X_{jk})\big] \le 1$ for all $k = 1, \ldots, p$. Using the inequality $\mathrm{Cov}(X, Y) \le \mathrm{Corr}(X, Y)\sqrt{\mathrm{E}(X^2)\mathrm{E}(Y^2)}$ and the boundedness of h, we get

$$
\mathrm{Cov}\left(h(X_{ik}, X_{jk}), h(X_{ik'}, X_{jk'})\right) \le \mathrm{Corr}\left(h(X_{ik}, X_{jk}), h(X_{ik'}, X_{jk'})\right) \text{ for all } 1 \le k < k' \le p.
$$

Since A1 is satisfied, from (A.9) we obtain

$$\mathrm{Var}\left[\frac{1}{p}\sum_{k=1}^{p} h(X_{ik}, X_{jk})\right] \leq \frac{1}{p} + \frac{2}{p^2} \sum\sum_{1\leq k<k'\leq p} \mathrm{Corr}\left(h(X_{ik}, X_{jk}), h(X_{ik'}, X_{jk'})\right) = o(1) \text{ as } p \to \infty.$$

Since n_1 is finite, it now follows from (A.8) that $|\bar{T}_{11} - \bar{\tau}_p(1,1)| \xrightarrow{\mathrm{P}} 0$ as $p \to \infty$. Following similar arguments, one can show that if A1 is satisfied, then $|\bar{T}_{12} - \bar{\tau}_p(1,2)|$ and $|\bar{T}_{22} - \bar{\tau}_p(2,2)|$ also converge in probability to 0 as $p \to \infty$.

holds?

(b) Fix $\epsilon > 0$, and recall the definitions of $\bar{T}_1(\mathbf{Z})$ and $\bar{\tau}_p(1,1)$. We have

$$\begin{aligned}
&\mathrm{P}\left[|\bar{T}_1(\mathbf{Z}) - \bar{\tau}_p(1,1)| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\right] \\
&= \mathrm{P}\left[\left|\frac{1}{p}\sum_{k=1}^{p} T_{1k}(Z_k) - \frac{1}{p}\sum_{k=1}^{p} \mathrm{E}\left[h(X_{1k}, X_{2k})\right]\right| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\right] \\
&= \mathrm{P}\left[\left|\frac{1}{p}\sum_{k=1}^{p} \frac{1}{n_1}\sum_{i=1}^{n_1} \left\{h(X_{ik}, Z_k) - \mathrm{E}[h(X_{1k}, Z_k) \mid \mathbf{Z} \sim \mathbf{F}_1]\right\}\right| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\right] \quad ? \\
&= \mathrm{P}\left[\left|\frac{1}{n_1}\sum_{i=1}^{n_1} \frac{1}{p}\sum_{k=1}^{p} \left\{h(X_{ik}, Z_k) - \mathrm{E}[h(X_{1k}, Z_k) \mid \mathbf{Z} \sim \mathbf{F}_1]\right\}\right| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\right] \\
&\leq \mathrm{P}\left[\frac{1}{n_1}\sum_{i=1}^{n_1} \left|\frac{1}{p}\sum_{k=1}^{p} \left\{h(X_{ik}, Z_k) - \mathrm{E}[h(X_{1k}, Z_k) \mid \mathbf{Z} \sim \mathbf{F}_1]\right\}\right| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\right] \text{ [follows from } \cancel{\text{the}} \text{ triangle ineq.]} \\
&\leq \sum_{i=1}^{n_1} \mathrm{P}\left[\left|\frac{1}{p}\sum_{k=1}^{p} \left\{h(X_{ik}, Z_k) - \mathrm{E}[h(X_{1k}, Z_k) \mid \mathbf{Z} \sim \mathbf{F}_1]\right\}\right| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\right] \text{ [using the union bound]} \\
&\leq \sum_{i=1}^{n_1} \frac{1}{\epsilon^2} \mathrm{Var}\left[\frac{1}{p}\sum_{k=1}^{p} h(X_{ik}, Z_k) \mid \mathbf{Z} \sim \mathbf{F}_1\right] = \sum_{i=1}^{n_1} \frac{1}{\epsilon^2} \mathrm{Var}\left[\frac{1}{p}\sum_{k=1}^{p} h(X_{ik}, X'_k)\right], \text{[Chebyshev's]} \qquad \text{(A.10)}
\end{aligned}$$

where $\mathbf{X}' = (X'_1, \ldots, X'_p)^\top \sim \mathbf{F}_1$ and 2 independent from X_1. We have already shown in the proof of Lemma 3.1(a) that $\mathrm{Var}\left[\frac{1}{p}\sum_{k=1}^{p} h(X_{ik}, X'_k)\right] = o(1)$ as $p \to \infty$ ~~due to~~ boundedness of h and assumption A1. Since n_1 is fixed, $\sum_{i=1}^{n_1} \mathrm{Var}\left[\frac{1}{p}\sum_{k=1}^{p} h(X_{ik}, X'_k)\right] = o(1)$ as $p \to \infty$. Therefore, it follows from (A.10) that $|\bar{T}_1(\mathbf{Z}) - \bar{\tau}_p(1,1)|$ converges in probability to 0 as $p \to \infty$ (when $\mathbf{Z} \sim \mathbf{F}_1$). using the

Following similar arguments, one can prove that for every $\epsilon > 0$,

$$\begin{aligned}
\lim_{p\to\infty} \mathrm{P}\left[|\bar{T}_2(\mathbf{Z}) - \bar{\tau}_p(1,2)| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\right] &= \lim_{p\to\infty} \mathrm{P}\left[|\bar{T}_1(\mathbf{Z}) - \bar{\tau}_p(1,2)| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_2\right] \\
&= \lim_{p\to\infty} \mathrm{P}\left[|\bar{T}_2(\mathbf{Z}) - \bar{\tau}_p(2,2)| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_2\right] = 0.
\end{aligned}$$

looks odd!

Hence, the proof. □

Proof of Lemma 3.1 → part (a)/(b) ?

Recall that $\bar{L}_1(\mathbf{Z}) = \bar{T}_{11} - 2\bar{T}_1(\mathbf{Z})$, $\bar{L}_2(\mathbf{Z}) = \bar{T}_{22} - 2\bar{T}_2(\mathbf{Z})$, and

$$\begin{aligned}
\bar{\theta}(\mathbf{Z}) &= \frac{1}{2}\bar{T}(\bar{L}_2(\mathbf{Z}) - \bar{L}_1(\mathbf{Z})) + \frac{1}{2}(\bar{T}_{22} - \bar{T}_{11})(\bar{L}_2(\mathbf{Z}) + \bar{L}_1(\mathbf{Z}) + 2\bar{T}_{12}) \\
&= \frac{1}{2}\left\{(\bar{T}_{11} - 2\bar{T}_{12} + \bar{T}_{22}) \times (\bar{L}_2(\mathbf{Z}) - \bar{L}_1(\mathbf{Z}))\right\} \\
&\quad + \frac{1}{2}\left\{(\bar{T}_{22} - \bar{T}_{11}) \times (\bar{T}_{22} - 2\bar{T}_2(\mathbf{Z}) + \bar{T}_{11} - 2\bar{T}_1(\mathbf{Z}) + 2\bar{T}_{12})\right\}. \qquad \text{(A.11)}
\end{aligned}$$

Let us denote $\bar{L}_2(\mathbf{Z}) - \bar{L}_1(\mathbf{Z})$ by $\bar{L}(\mathbf{Z})$ and $\bar{T}_{22} - 2\bar{T}_2(\mathbf{Z}) + \bar{T}_{11} - 2\bar{T}_1(\mathbf{Z}) + 2\bar{T}_{12}$ by $\bar{S}(\mathbf{Z})$.

~~Therefore~~ Now, we can write $\bar{\theta}(\mathbf{Z}) = \frac{1}{2}\left\{(\bar{T}_{11} - 2\bar{T}_{12} + \bar{T}_{22}) \times \bar{L}(\mathbf{Z})\right\} + \frac{1}{2}\left\{(\bar{T}_{22} - \bar{T}_{11}) \times \bar{S}(\mathbf{Z})\right\}$. (A.12)

(a) Fix $\epsilon > 0$. Now,

$$
\begin{aligned}
& \mathrm{P}\left[|\bar{L}_2(\mathbf{Z}) - \bar{L}_1(\mathbf{Z}) - \bar{\tau}_p| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\right] \\
&= \mathrm{P}\left[|\{\bar{T}_{22} - 2\bar{T}_2(\mathbf{Z}) - \bar{T}_{11} + 2\bar{T}_1(\mathbf{Z})\} - \{\bar{\tau}_p(1,1) - 2\bar{\tau}_p(1,2) + \bar{\tau}_p(2,2)\}| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\right] \\
&\leq \mathrm{P}\left[|\{\bar{T}_{22} - 2\bar{T}_2(\mathbf{Z}) - \bar{T}_{11} + 2\bar{T}_1(\mathbf{Z})\} - \{2\bar{\tau}_p(1,1) - \bar{\tau}_p(1,1) - 2\bar{\tau}_p(1,2) + \bar{\tau}_p(2,2)\}| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\right] \\
&\leq \mathrm{P}\left[|\bar{T}_{11} - \bar{\tau}_p(1,1)| > \frac{\epsilon}{4}\right] + \mathrm{P}\left[|\{\bar{T}_{22} - \bar{\tau}_p(2,2)| > \frac{\epsilon}{4}\right] \\
&\quad + \mathrm{P}\left[2|\bar{T}_2(\mathbf{Z}) - \bar{\tau}_p(1,2)| > \frac{\epsilon}{4} \mid \mathbf{Z} \sim \mathbf{F}_1\right] + \mathrm{P}\left[2|\bar{T}_1(\mathbf{Z}) - \bar{\tau}_p(1,1)| > \frac{\epsilon}{4} \mid \mathbf{Z} \sim \mathbf{F}_1\right] \\
&= o(1) \text{ as } p \to \infty \text{ [~~following~~ using Lemma 3.1].}
\end{aligned}
\tag{A.13}
$$

Therefore, if $\mathbf{Z} \sim \mathbf{F}_1$, then $|\bar{L}_2(\mathbf{Z}) - \bar{L}_1(\mathbf{Z}) - \bar{\tau}_p| \xrightarrow{\mathrm{P}} 0$ as $p \to \infty$. This proves the first part of Lemma 3.1(a). Next, we use the continuous mapping theorem and Lemma 3.1 to obtain the following:

?,

$$
\begin{aligned}
&|\{\bar{T}_{11} - 2\bar{T}_{12} + \bar{T}_{22}\} - \bar{\tau}_p| \xrightarrow{\mathrm{P}} 0,\ |\{\bar{T}_{22} - \bar{T}_{11}\} - \{\bar{\tau}_p(2,2) - \bar{\tau}_p(1,1)\}| \xrightarrow{\mathrm{P}} 0 \text{ and} \\
&|\bar{S}(\mathbf{Z}) - \{\bar{\tau}_p(2,2) - \bar{\tau}_p(1,1)\}| \xrightarrow{\mathrm{P}} 0 \text{ (if } \mathbf{Z} \sim \mathbf{F}_1) \text{ as } p \to \infty.
\end{aligned}
$$

looks odd!

Using the continuous mapping theorem once again, we conclude from (A.12) that if $\mathbf{Z} \sim \mathbf{F}_1$, then

$$
\left|\bar{\theta}(\mathbf{Z}) - \left\{\frac{1}{2}\bar{\tau}_p^2 + \frac{1}{2}(\bar{\tau}_p(2,2) - \bar{\tau}_p(1,1))^2\right\}\right| \xrightarrow{\mathrm{P}} 0 \text{ as } p \to \infty
$$

$$
\Rightarrow \cancel{\text{i.e.}}\ |\bar{\theta}(\mathbf{Z}) - \bar{\psi}_p| \xrightarrow{\mathrm{P}} 0\ \cancel{(\text{if } \mathbf{Z} \sim \mathbf{F}_1)} \text{ as } p \to \infty. \tag{A.14}
$$

(b) The arguments for this part of the proof are similar to part (a), and we skip it. □

Proof of Theorem 3.2

Drop?

[Recall that the prior probability of an observation $\mathbf{Z}$ belonging to the first class is given by π_1 $(0 < \pi_1 < 1)$.] The misclassification probability of the classifier δ_1 can be written as

$$
\begin{aligned}
\mathrm{P}[\delta_1(\mathbf{Z}) \neq \text{ true label of } \mathbf{Z}] &= \pi_1 \mathrm{P}[\delta_1(\mathbf{Z}) = 2 \mid \mathbf{Z} \sim \mathbf{F}_1] + \pi_2 \mathrm{P}[\delta_1(\mathbf{Z}) = 1 \mid \mathbf{Z} \sim \mathbf{F}_2] \\
&= \pi_1 \mathrm{P}[\bar{L}_2(\mathbf{Z}) \leq \bar{L}_1(\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}_1] + \pi_2 \mathrm{P}[\bar{L}_2(\mathbf{Z}) > \bar{L}_1(\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}_2].
\end{aligned}
\tag{A.15}
$$

Since A2 is satisfied (i.e., $\liminf_p \bar{\tau}_p > 0$), we can choose $\epsilon > 0$ such that $\epsilon < \bar{\tau}_p$ for all $p \geq p_0$ for some $p_0 \in \mathbb{N}$. Therefore,

$$
\begin{aligned}
\mathrm{P}\left[\bar{L}_2(\mathbf{Z}) \leq \bar{L}_1(\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}_1\right] &\leq \mathrm{P}\left[\bar{L}_2(\mathbf{Z}) - \bar{L}_1(\mathbf{Z}) \leq \bar{\tau}_p - \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\right] \\
&\leq \mathrm{P}\left[\bar{L}_2(\mathbf{Z}) - \bar{L}_1(\mathbf{Z}) - \bar{\tau}_p \leq -\epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\right] \leq \mathrm{P}\left[|\bar{L}_2(\mathbf{Z}) - \bar{L}_1(\mathbf{Z}) - \bar{\tau}_p| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\right]
\end{aligned}
$$

for all $p \geq p_0$. Then, it follows from part (a) of Lemma 3.1~~(a)~~ that $\mathrm{P}\left[\bar{L}_2(\mathbf{Z}) \leq \bar{L}_1(\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}_1\right] = o(1)$ as $p \to \infty$. Similarly,

$$
\begin{aligned}
\mathrm{P}\left[\bar{L}_2(\mathbf{Z}) > \bar{L}_1(\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}_2\right] &\leq \mathrm{P}\left[\bar{L}_2(\mathbf{Z}) - \bar{L}_1(\mathbf{Z}) > -\bar{\tau}_p + \epsilon \mid \mathbf{Z} \sim \mathbf{F}_2\right] \\
&\leq \mathrm{P}\left[\bar{L}_2(\mathbf{Z}) - \bar{L}_1(\mathbf{Z}) + \bar{\tau}_p > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_2\right] \leq \mathrm{P}\left[|\bar{L}_2(\mathbf{Z}) - \bar{L}_1(\mathbf{Z}) + \bar{\tau}_p| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_2\right]
\end{aligned}
$$

for all $p \geq p_0$. Since $\mathrm{P}\left[|\bar{L}_2(\mathbf{Z}) - \bar{L}_1(\mathbf{Z}) + \bar{\tau}_p| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_2\right] = o(1)$ as $p \to \infty$ (~~due to~~ using part (b) of Lemma 3.1~~(b)~~), $\mathrm{P}\left[\bar{L}_2(\mathbf{Z}) > \bar{L}_1(\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}_2\right] = o(1)$ as $p \to \infty$. Consequently, it follows from Eq. (A.15) that $\mathrm{P}[\delta_1(\mathbf{Z}) \neq \text{ true label of } \mathbf{Z}] = \pi_1 o(1) + \pi_2 o(1) = o(1)$ as $p \to \infty$. Hence, the proof. □

new page?

Proof of Theorem 3.3

First observe that

$$\liminf_p \bar{\tau}_p > 0 \Rightarrow \liminf_p \frac{1}{2}\bar{\tau}_p^2 > 0 \Rightarrow \liminf_p \frac{1}{2}\left\{\bar{\tau}_p^2 + (\bar{\tau}_p(2,2) - \bar{\tau}_p(1,1))^2\right\} > 0.$$

Thus, if A2 is satisfied, then $\liminf_p \bar{\psi}_p > 0$. Now, let us consider the misclassfication probability of δ_2.

$$\begin{aligned} \mathrm{P}[\delta_2(\mathbf{Z}) \neq \text{ true label of } \mathbf{Z}] &= \pi_1 \mathrm{P}[\delta_2(\mathbf{Z}) = 2 \mid \mathbf{Z} \sim \mathbf{F}_1] + \pi_2 \mathrm{P}[\delta_2(\mathbf{Z}) = 1 \mid \mathbf{Z} \sim \mathbf{F}_2] \\ &= \pi_1 \mathrm{P}[\bar{\theta}(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}_1] + \pi_2 \mathrm{P}[\bar{\theta}(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{F}_2]. \end{aligned} \tag{A.16}$$

Since $\liminf_p \bar{\psi}_p > 0$, we can choose $\epsilon > 0$ such that $\epsilon < \bar{\psi}_p$ for all $p \geq p_1$ for some $p_1 \in \mathbb{N}$. Therefore,

$\epsilon > \bar{\psi}_p$?

$$\begin{aligned} \mathrm{P}\big[\bar{\theta}(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}_1\big] &\leq \mathrm{P}\big[\bar{\theta}(\mathbf{Z}) \leq \bar{\psi}_p - \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\big] \\ &\leq \mathrm{P}\big[\bar{\theta}(\mathbf{Z}) - \bar{\psi}_p \leq -\epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\big] \leq \mathrm{P}\big[|\bar{\theta}(\mathbf{Z}) - \bar{\psi}_p| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\big] \text{ for all } p \geq p_1. \end{aligned}$$

Therefore, it follows from part (a) of Lemma 3.1(a) that $\mathrm{P}\big[\bar{\theta}(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}_1\big] = o(1)$ as $p \to \infty$. Similarly,

$$\begin{aligned} \mathrm{P}\big[\bar{\theta}(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{F}_2\big] &\leq \mathrm{P}\big[\bar{\theta}(\mathbf{Z}) > -\bar{\psi}_p + \epsilon \mid \mathbf{Z} \sim \mathbf{F}_1\big] \\ &\leq \mathrm{P}\big[\bar{\theta}(\mathbf{Z}) + \bar{\psi}_p > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_2\big] \leq \mathrm{P}\big[|\bar{\theta}(\mathbf{Z}) + \bar{\psi}_p| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}_2\big] \text{ for all } p \geq p_1, \end{aligned}$$

implying first $\mathrm{P}\big[\bar{\theta}(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{F}_2\big] = o(1)$ as $p \to \infty$ (~~due to~~ using part (b) of Lemma 3.1(b)). As a result, Using Eq. (A.16), we get $\mathrm{P}[\delta_2(\mathbf{Z}) \neq \text{ true label of } \mathbf{Z}] = \pi_1 o(1) + \pi_2 o(1) = o(1)$ as $p \to \infty$. Hence, the proof. □

Space!

Proof of Theorem 3.4

?

We have assumed in assumption (A3) that there exists a p_0 such that $\bar{\tau}_p(1,2)$ lies between $\bar{\tau}_p(1,1)$ and $\bar{\tau}_p(2,2)$ for all $p \geq p_0$. With out loss of generality, let us assume that $\bar{\tau}_p(1,1) < \bar{\tau}_p(2,2)$. As a result,

$$\bar{\tau}_p < \bar{\tau}_p(2,2) - \bar{\tau}_p(1,1) \text{ for all } p \geq p_0. \tag{A.17}$$

Recall that

$$\Delta_1 = \mathrm{P}[\delta_1(\mathbf{Z}) \neq \text{ true label of } \mathbf{Z}] = \pi_1\ \mathrm{P}[\bar{L}(\mathbf{Z}) \leq 0 | \mathbf{Z} \sim \mathbf{F}_1] + \pi_2\ \mathrm{P}[\bar{L}(\mathbf{Z}) > 0 | \mathbf{Z} \sim \mathbf{F}_2], \text{ and}$$

$$\Delta_2 = \mathrm{P}[\delta_2(\mathbf{Z}) \neq \text{ true label of } \mathbf{Z}] = \pi_1\ \mathrm{P}[\bar{\theta}(\mathbf{Z}) \leq 0 | \mathbf{Z} \sim \mathbf{F}_1] + \pi_2\ \mathrm{P}[\bar{\theta}(\mathbf{Z}) > 0 | \mathbf{Z} \sim \mathbf{F}_2].$$

It follows from (A.17) that

$$\begin{aligned} \mathrm{P}[\bar{\theta}(\mathbf{Z}) \leq 0 | \mathbf{Z} \sim \mathbf{F}_1] &= \mathrm{P}[\bar{\tau}_p \bar{L}(\mathbf{Z}) + \{\bar{\tau}_p(2,2) - \bar{\tau}_p(1,1)\}\bar{S}(\mathbf{Z}) \leq 0 | \mathbf{Z} \sim \mathbf{F}_1] \\ &\leq \mathrm{P}[\bar{\tau}_p\{\bar{L}(\mathbf{Z}) + \bar{S}(\mathbf{Z})\} \leq 0 | \mathbf{Z} \sim \mathbf{F}_1] \text{ for all } p \geq p_0. \end{aligned}$$

Consequently, for all $p \geq p_0$, we have the following:

$$\begin{aligned} &\mathrm{P}[\bar{\theta}(\mathbf{Z}) \leq 0 | \mathbf{Z} \sim \mathbf{F}_1] \\ &\leq \mathrm{P}[\bar{L}(\mathbf{Z}) + \bar{S}(\mathbf{Z}) \leq 0 | \mathbf{Z} \sim \mathbf{F}_1] \quad (\text{since } \bar{\tau}_p > 0) \\ &= \mathrm{P}[\bar{L}(\mathbf{Z}) \leq -\bar{S}(\mathbf{Z}) | \mathbf{Z} \sim \mathbf{F}_1] \\ &= \mathrm{P}[\bar{L}(\mathbf{Z}) \leq -\bar{S}(\mathbf{Z}), \bar{S}(\mathbf{Z}) \geq 0 | \mathbf{Z} \sim \mathbf{F}_1] + \mathrm{P}[\bar{L}(\mathbf{Z}) \leq -\bar{S}(\mathbf{Z}), \bar{S}(\mathbf{Z}) < 0 | \mathbf{Z} \sim \mathbf{F}_1] \\ &\leq \mathrm{P}[\bar{L}(\mathbf{Z}) \leq 0 | \mathbf{Z} \sim \mathbf{F}_1] + \mathrm{P}[\bar{S}(\mathbf{Z}) < 0 | \mathbf{Z} \sim \mathbf{F}_1]\ . \end{aligned} \tag{A.18}$$

Similarly, one can show that

$$\mathrm{P}[\bar{\theta}(\mathbf{Z}) > 0 | \mathbf{Z} \sim \mathbf{F}_2] \leq \mathrm{P}[\bar{L}(\mathbf{Z}) > 0 | \mathbf{Z} \sim \mathbf{F}_2] + \mathrm{P}[\bar{S}(\mathbf{Z}) > 0 | \mathbf{Z} \sim \mathbf{F}_2] \text{ for all } p \geq p_0. \tag{A.19}$$

Adding the two inequalities (A.18) and (A.19), we obtain

$$\Delta_2 \leq \Delta_1 + \pi_1 \mathrm{P}[\bar{S}(\mathbf{Z}) > 0 | \mathbf{Z} \sim \mathbf{F}_1] + \pi_2 \mathrm{P}[\bar{S}(\mathbf{Z}) > 0 | \mathbf{Z} \sim \mathbf{F}_2] \text{ for all } p \geq p_0. \tag{A.20}$$

(a)/(b) ?

Now, it follows from Lemma 3.1 that for $\mathbf{Z} \sim \mathbf{F}_1$, $|\bar{S}(\mathbf{Z}) - \{\bar{\tau}_p(2,2) - \bar{\tau}_p(1,1)\}| \xrightarrow{\mathrm{P}} 0$ as $p \to \infty$. Therefore, for any $\epsilon_1 > 0$ and $\epsilon_2 > 0$, there exists a $\tilde{p}_1(\epsilon_1, \epsilon_2)$ such that for all $p \geq \tilde{p}_1(\epsilon_1, \epsilon_2)$

$$\begin{aligned}
&\mathrm{P}\left[|\bar{S}(\mathbf{Z}) - \{\bar{\tau}_p(2,2) - \bar{\tau}_p(1,1)\}| > \epsilon_1 \big| \mathbf{Z} \sim F_1\right] < \epsilon_2 \\
\Rightarrow\ &\mathrm{P}\left[\bar{S}(\mathbf{Z}) - \{\bar{\tau}_p(2,2) - \bar{\tau}_p(1,1)\} < -\epsilon_1 \big| \mathbf{Z} \sim F_1\right] < \epsilon_2 \\
\Rightarrow\ &\mathrm{P}\left[\bar{S}(\mathbf{Z}) < \{\bar{\tau}_p(2,2) - \bar{\tau}_p(1,1)\} - \epsilon_1 \big| \mathbf{Z} \sim F_1\right] < \epsilon_2.
\end{aligned}$$

We have already assumed that $\bar{\tau}_p(2,2) > \bar{\tau}_p(1,1)$ for all $p \geq p_0$. ~~Let~~ & define $\lambda_0 = \liminf_p \{\bar{\tau}_p(2,2) - \bar{\tau}_p(1,1)\}$, ~~and we have~~ ~~Note that~~ we $\lambda_0 \geq \liminf_p \bar{\tau}_p$. Since assumption A2 ~~is satisfied~~ holds, we have $\lambda_0 > 0$. Therefore, it follows from the above inequality that for any $0 < \epsilon_1 < \lambda_0$,

$$\mathrm{P}[\bar{S}(\mathbf{Z}) < 0 \big| \mathbf{Z} \sim F_1] < \epsilon_2 \text{ for all } p \geq \max\{\tilde{p}_1(\epsilon_1, \epsilon_2), p_0\}.$$

Following similar arguments, one can show that for any $0 < \epsilon < \lambda_0$,

$$\mathrm{P}[\bar{S}(\mathbf{Z}) > 0 \big| \mathbf{Z} \sim F_2] < \epsilon_2 \text{ for all } p \geq \max\{\tilde{p}_1(\epsilon_1, \epsilon_2), p_0\}.$$

Therefore, it follows from (A.20) that for any $0 < \epsilon_1 < \lambda_0$, $\epsilon \geq \lambda_0$?

$$\Delta_2 \leq \Delta_1 + \epsilon_2 \text{ for all } p \geq \max\{\tilde{p}_2(\epsilon_1, \epsilon_2), p_0\}.$$

Since $\epsilon_2 > 0$ is chosen arbitrarily, we conclude that for any $0 < \epsilon_1 < \lambda_0$,

$$\Delta_2 \leq \Delta_1 \text{ for all } p \geq p_0' = \max\{\tilde{p}_2(\epsilon_1, \epsilon_2), p_0\}.$$

This completes the proof. □

Check p. 1

The following results are derived for ultrahigh-dimensional settings, where n goes to infinity and $p(= p_n)$ is assumed to ~~be an increasing function of~~ increase with n. In particular, we assume $\log p_n = O(n^\beta)$ for some $0 \leq \beta < 1$.

Let us define the following random variables:

$$\begin{aligned}
T_{11k} &= \frac{1}{n_1(n_1-1)} \sum\sum_{1 \leq i \neq j \leq n_1} h(X_{ik}, X_{jk}), \quad T_{12k} = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} h(X_{ik}, Y_{jk}), \text{ and} \\
T_{22k} &= \frac{1}{n_2(n_2-1)} \sum\sum_{1 \leq i \neq j \leq n_2} h(Y_{ik}, Y_{jk}) \text{ for } k = 1, \ldots, p_n.
\end{aligned} \tag{A.21}$$

Also, for $\mathbf{z} = (z_1, \ldots, z_{p_n})^\top \in \mathbb{R}^{p_n}$, we define

$$\begin{aligned}
&T_{1k}(z_k) = \frac{1}{n_1} \sum_{i=1}^{n_1} h(X_{ik}, z_k), \quad T_{2k}(z_k) = \frac{1}{n_2} \sum_{j=1}^{n_2} h(Y_{jk}, z_k), \quad L_{1k}(Z_k) = T_{11k} - 2T_{1k}(z_k) \text{ and} \\
&L_{2k}(z_k) = T_{22k} - 2T_{2k}(z_k) \text{ for } k = 1, \ldots, p_n.
\end{aligned} \tag{A.22}$$

Observe that the estimators of $\bar{\tau}_{11}, \bar{\tau}_{12}$ and $\bar{\tau}_{22}$ defined in (2.4) can be expressed as follows:

$$\begin{aligned}
\bar{T}_{11} &= \frac{1}{n_1(n_1-1)p_n} \sum_{k=1}^{p_n} \sum\sum_{1 \leq i \neq j \leq n_1} h(X_{ik}, X_{jk}), \quad \bar{T}_{12} = \frac{1}{n_1 n_2 p_n} \sum_{k=1}^{p_n} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} h(X_{ik}, Y_{jk}) \text{ and} \\
\bar{T}_{22} &= \frac{1}{n_2(n_2-1)p_n} \sum_{k=1}^{p_n} \sum\sum_{1 \leq i \neq j \leq n_1} h(Y_{ik}, Y_{jk}), \\
\text{i.e., } \bar{T}_{11} &= \frac{1}{p_n} \sum_{k=1}^{p_n} T_{11k}, \ \bar{T}_{12} = \frac{1}{p_n} \sum_{k=1}^{p_n} T_{12k} \text{ and } \bar{T}_{22} = \frac{1}{p_n} \sum_{k=1}^{p_n} T_{22k}.
\end{aligned}$$

Similarly, for $\mathbf{z} \in \mathbb{R}^{p_n}$, we can write

$$\bar{T}_1(\mathbf{z}) = \frac{1}{p_n}\sum_{k=1}^{p_n} T_{1k}(z_k), \text{ and } \bar{T}_2(\mathbf{z}) = \frac{1}{p_n}\sum_{k=1}^{p_n} T_{2k}(z_k).$$

Recall the definitions of $\bar{L}_1(\mathbf{z})$, $\bar{L}_2(\mathbf{z})$ and $\bar{\theta}(\mathbf{z})$ given in Eq. (A.11). We now derive ~~the~~ upper bounds on the rates of convergence of these random variables.

First, let us consider the bounded differences inequality that will be used to derive ~~the~~ concentration bounds. Given vectors $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ and an index $l \in \{1, \ldots, n\}$, we define a new vector $\mathbf{x}^{\setminus l} \in \mathbb{R}^n$ as follows:

$$\mathbf{x}^{\setminus l} = \begin{cases} x_j, & \text{if } j \neq l, \\ x'_l, & \text{if } j = l. \end{cases} \tag{A.23}$$

With this notation, we say that $f : \mathbb{R}^n \to \mathbb{R}$ satisfies the bounded difference inequality with parameters $(M_1, \ldots, M_n)^\top$ if

$$|f(\mathbf{x}) - f(\mathbf{x}^{\setminus l})| \leq M_l \text{ for each } l = 1, \ldots, n \text{ and for all } \mathbf{x}, \mathbf{x}' \in \mathbb{R}^n.$$

Lemma A.4 *(Wainwright, 2019, page 37) Suppose that f satisfies the bounded difference property* (A.23) *with parameters $(M_1, \ldots, M_n)^\top$ and that the random vector $\mathbf{U} = (U_1, \ldots, U_n)^\top$ has independent components. Then,*

$$\mathrm{P}\left[|f(\mathbf{U}) - \mathrm{E}[f(\mathbf{U})]| > \epsilon\right] \leq 2e^{-\frac{2\epsilon^2}{\sum_{l=1}^{n} M_l^2}} \text{ for all } \epsilon > 0.$$

Using Lemma A.4, we first derive the rates of convergence of $\bar{T}_{ij}$ and $\bar{T}_i(\mathbf{z})$ for $j, j' \in \{1, 2\}$ and $\mathbf{z} \in \mathbb{R}^{p_n}$.

Lemma A.5 *Fix $0 < \gamma < 1/2$. There exist~~s~~ positive constants a_{ij}, b_i for $j, j' \in \{1, 2\}$ such that*

$i \in ?$

(a) $\mathrm{P}\left[|\bar{T}_{ij} - \tau_p(i,j)| > n^{-\gamma}\right] \leq O(p_n e^{-a_{ij} n^{1-2\gamma}})$ *and*

(b) $\mathrm{P}\left[|\bar{T}_i(\mathbf{z}) - E[\bar{T}_i(\mathbf{z})]| > n^{-\gamma}\right] \leq O(p_n e^{-b_i n^{1-2\gamma}})$ *for all* $\mathbf{z} \in \mathbb{R}^{p_n}$.

Proof of Lemma A.5

(a) Fix $k \in \{1, \ldots, p_n\}$. Recall the definitions of T_{11k}, T_{22k} and T_{12k} in Eq. (A.21) and note that the first two random variables are one sample U-statistics with kernel h (of order 2), while the third random variable is a two sample U-statistic with kernel h (of order (1,1)).

The random vector $\mathcal{X}_k = (X_{1k}, \ldots, X_{n_1 k})^\top$ has independent components. Since $|h| \leq 1$, Lemma A.4 can be used to estalish the concentration of T_{11k} around its mean. Viewing T_{11k} as a function $f(X_{1k}, \ldots, X_{n_1 k})$, for any given coordinate $l \in \{1, \ldots, n_1\}$, we have

$$|f(\mathcal{X}_k) - f(\mathcal{X}_k^{\setminus l})| \leq \frac{2}{n_1(n_1-1)}\sum_{j \neq l}|h(X_{jk}, X_{lk}) - h(X_{jk}, X'_{lk})| \leq 2(n_1 - 1)\frac{2}{n_1(n_1-1)} = \frac{4}{n_1}.$$

So, the bounded difference property holds with parameter $M_l = 4/n_1$ in each coordinate. We conclude from Lemma A.4 that

$$\mathrm{P}\left[|T_{11k} - \mathrm{E}[T_{11k}]| > n^{-\gamma}\right] \leq 2e^{-\frac{n_1 n^{-2\gamma}}{8}}. \tag{A.24}$$

Since $\lim_{n\to\infty} n_1/n = \pi_1 < 1$, there exist constants $a_{11} > 0$ and $N \in \mathbb{N}$ such that

$$P[|T_{11k} - \mathrm{E}[T_{11k}]| \geq n^{-\gamma}] \leq 2e^{-a_{11} n^{1-2\gamma}} \text{ for all } n \geq N. \tag{A.25}$$

Clearly, Eq. (A.25) is true for all $k = 1, \ldots, p_n$, ~~i.e.,~~. So, we have

$$
\begin{aligned}
& P[|T_{11k} - \mathrm{E}[T_{11k}]| \geq n^{-\gamma}] \leq O\left(e^{-a_{11}n^{1-2\gamma}}\right) \quad \text{for all } k = 1, \ldots, p_n \\
\Rightarrow & \sum_{k=1}^{p_n} P[|T_{11k} - \mathrm{E}[T_{11k}]| \geq n^{-\gamma}] \leq O\left(p_n e^{-a_{11}n^{1-2\gamma}}\right) \\
\Rightarrow & P\left[\frac{1}{p_n}\sum_{k=1}^{p_n} |T_{11k} - \mathrm{E}[T_{11k}]| \geq n^{-\gamma}\right] \leq O\left(p_n e^{-a_{11}n^{1-2\gamma}}\right) \\
\Rightarrow & P\left[\left|\frac{1}{p_n}\sum_{k=1}^{p_n} (T_{11k} - \mathrm{E}[T_{11k}])\right| \geq n^{-\gamma}\right] \leq O\left(p_n e^{-a_{11}n^{1-2\gamma}}\right) \\
\Rightarrow & P\left[|\bar{T}_{11} - \bar{\tau}_p(1,1))| \geq n^{-\gamma}\right] \leq O\left(p_n e^{-a_{11}n^{1-2\gamma}}\right) \left[\text{since } \sum_{k=1}^{p_n} \mathrm{E}[T_{11k}]/p_n = \bar{\tau}_{p_n}(1,1)\right]. \qquad \text{(A.26)}
\end{aligned}
$$

Following similar arguments, it can be shown that there exist positive constants a_{12} and a_{22} such that

$$
\mathrm{P}\left[|\bar{T}_{12} - \bar{\tau}_p(1,2)| > n^{-\gamma}\right] \leq O(p_n e^{-a_{12}n^{1-2\gamma}}) \text{ and } \mathrm{P}\left[|\bar{T}_{22} - \bar{\tau}_p(2,2)| > n^{-\gamma}\right] \leq O(p_n e^{-a_{22}n^{1-2\gamma}}). \quad \text{(A.27)}
$$

(b) Recall the definition of $\bar{T}_1(\mathbf{z})$ ~~in~~ from (A.22), and observe that for each $\mathbf{z} \in \mathbb{R}^{p_n}$, we have the following:

$$
\begin{aligned}
& \mathrm{P}\left[|\bar{T}_1(\mathbf{z}) - E[\bar{T}_1(\mathbf{z})]| > n^{-\gamma}\right] \\
= & \mathrm{P}\left[\left|\frac{1}{p_n}\sum_{k=1}^{p_n} T_{1k}(z_k) - \frac{1}{p_n}\sum_{k=1}^{p_n} E[T_{1k}(z_k)]\right| > n^{-\gamma}\right] \\
\leq & \mathrm{P}\left[\frac{1}{p_n}\sum_{k=1}^{p_n} |T_{1k}(z_k) - E[T_{1k}(z_k)]| > n^{-\gamma}\right] \\
\leq & \sum_{k=1}^{p_n} \mathrm{P}\left[|T_{1k}(z_k) - E[T_{1k}(z_k)]| > n^{-\gamma}\right] \\
\leq & \sum_{k=1}^{p_n} \mathrm{P}\left[\left|\frac{1}{n_1}\sum_{i=1}^{n_1} h(X_{ik}, z_k) - \frac{1}{n_1}\sum_{i=1}^{n_1} E[h(X_{ik}, z_k)]\right| > n^{-\gamma}\right] \\
= & \sum_{k=1}^{p_n} \mathrm{P}\left[\left|\frac{1}{n_1}\sum_{i=1}^{n_1} \{h(X_{ik}, z_k) - E[h(X_{ik}, z_k)]\}\right| > n^{-\gamma}\right]. \qquad \text{(A.28)}
\end{aligned}
$$

Note that $\sum_{i=1}^{n_1} h(X_{ik}, z_k)/n_1$ is an average of independently distributed random variables for each $\mathbf{z} \in \mathbb{R}^{p_n}$. Using Hoeffding's inequality, we obtain the following:

$$
\begin{aligned}
& P\left[\left|\frac{1}{n_1}\sum_{i=1}^{n_1} \{h(X_{ik}, z_k) - E[h(X_{ik}, z_k)]\}\right| > n^{-\gamma}\right] \leq 2e^{-2n_1 n^{-2\gamma}} \quad \text{for all } 1 \leq k \leq p_n \\
\Rightarrow & \sum_{k=1}^{p_n} P\left[\left|\frac{1}{n_1}\sum_{i=1}^{n_1} \{h(X_{ik}, z_k) - E[h(X_{ik}, z_k)]\}\right| > n^{-\gamma}\right] \leq 2p_n e^{-2n_1 n^{-2\gamma}} \\
\Rightarrow & \sum_{k=1}^{p_n} P\left[\left|\frac{1}{n_1}\sum_{i=1}^{n_1} \{h(X_{ik}, z_k) - E[h(X_{1k}, z_k)]\}\right| > n^{-\gamma}\right] = O\left(p_n e^{-b_1 n^{1-2\gamma}}\right) \quad \text{for some } b_1 > 0. \qquad \text{(A.29)}
\end{aligned}
$$

Combining (A.28) and (A.29), for every $\mathbf{z} \in \mathbb{R}^{p_n}$, we get

$$
\mathrm{P}\left[|\bar{T}_1(\mathbf{z}) - E[\bar{T}_1(\mathbf{z})]| > n^{-\gamma}\right] \leq O\left(p_n e^{-b_1 n^{1-2\gamma}}\right) \quad \text{for some } b_1 > 0.
$$

Similarly, one can show that there exists a constant $b_2 > 0$ such that

$$
\mathrm{P}\left[|\bar{T}_2(\mathbf{z}) - E[\bar{T}_2(\mathbf{z})]| > n^{-\gamma}\right] \leq O\left(p_n e^{-b_2 n^{1-2\gamma}}\right).
$$

previous page?

Hence, the proof. □

fine?

Lemma A.6 *Suppose* $P[|X_n - a_0| > \epsilon] = O(p_n e^{-M_1 n\epsilon^2})$ *and* $P[|Y_n - b_0| > \epsilon] = O(p_n e^{-M_2 n\epsilon^2})$ *for all* $\epsilon > 0$ *where* $\max\{|a_0|, |b_0|\} > 0$ *and* M_1, M_2 *are positive constants. Then, there exists a positive constant* M_3 *such that* $P[|X_n Y_n - a_0 b_0| > \epsilon] = O(p_n e^{-M_3 n\epsilon^2})$ *for all* $\epsilon > 0$.

Proof: define ~~Let~~ $c_0 = \max\{|a_0|, |b_0|\}$. using ~~From the~~ triangle inequality, we ~~have~~ get

$$
\begin{aligned}
&|X_nY_n - a_0b_0| \le |X_nY_n - b_0X_n - a_0Y_n + a_0b_0| + |b_0||X_n - a_0| + |a_0||Y_n - b_0| \\
\Rightarrow\ &|X_nY_n - a_0b_0| \le |X_n - a_0||Y_n - b_0| + |b_0||X_n - a_0| + |a_0||Y_n - b_0| \\
\Rightarrow\ &|X_nY_n - a_0b_0| \le |X_n - a_0||Y_n - b_0| + c_0(|X_n - a_0| + |Y_n - b_0|).
\end{aligned}
$$

Therefore, $|X_n - a_0| \le \epsilon$ and $|Y_n - b_0| \le \epsilon$ implies that $|X_nY_n - a_0b_0| \le \epsilon^2 + 2c_0\epsilon$. We choose M such that $M > 2 + \epsilon/c_0$. Therefore, $\epsilon^2 + 2c_0\epsilon \le Mc_0\epsilon$. Now,

$$
\begin{aligned}
&P[|X_n - a_0| \le \epsilon, |Y_n - b_0| \le \epsilon] \le P[|X_nY_n - a_0b_0| \le \epsilon^2 + 2c_0\epsilon] \\
\Rightarrow\ &P[|X_n - a_0| \le \epsilon, |Y_n - b_0| \le \epsilon] \le P[|X_nY_n - a_0b_0| \le Mc_0\epsilon] \\
\Rightarrow\ &P[|X_nY_n - a_0b_0| > Mc_0\epsilon] \le P[|X_n - a_0| > \epsilon] + P[|Y_n - b_0| > \epsilon] \\
\Rightarrow\ &P[|X_nY_n - a_0b_0| > Mc_0\epsilon] \le O(p_n e^{-M_1 n\epsilon^2}) + O(p_n e^{-M_2 n\epsilon^2}) \\
\Rightarrow\ &P[|X_nY_n - a_0b_0| > Mc_0\epsilon] \le O(p_n e^{-\min\{M_1,M_2\} n\epsilon^2}) \\
\Rightarrow\ &P[|X_nY_n - a_0b_0| > \epsilon] \le O(p_n e^{-\frac{\min\{M_1,M_2\}}{Mc_0} n\epsilon^2}).
\end{aligned}
$$

Therefore, $P[|X_n - a_0| \le \epsilon, |Y_n - b_0| \le \epsilon] \le O(p_n e^{-\frac{\min\{M_1,M_2\}}{Mc_0} n\epsilon^2})$ for all $\epsilon > 0$ with ~~where~~ $M > 2 + c_0/\epsilon$. Hence, the proof. □

Proof of Lemma 3.5

(a) For a $\mathbf{z} \in \mathbb{R}^{p_n}$, recall the definitions of $\bar{L}(\mathbf{z})$ and $\bar{L}_0(\mathbf{z})$ given in Section 3.2. For any $0 < \gamma < 1/2$, we have

$$
\begin{aligned}
&P\left[|\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| > n^{-\gamma}\right] \\
&= P\left[|\bar{L}_2(\mathbf{z}) - \bar{L}_1(\mathbf{z}) - \bar{L}_2^0(\mathbf{z}) + \bar{L}_1^0(\mathbf{z})| > n^{-\gamma}\right] \\
&= P\left[|\bar{T}_{22} - 2\bar{T}_2(\mathbf{z}) - \bar{T}_{11} + 2\bar{T}_1(\mathbf{z}) - \bar{\tau}_{p_n}(2,2) + 2E[\bar{h}_{p_n}(\mathbf{Y}_1, \mathbf{z})] - \bar{\tau}_{p_n}(1,1) + 2E[\bar{h}_{p_n}(\mathbf{X}_1, \mathbf{z})]| > n^{-\gamma}\right] \\
&\le P\left[|\bar{T}_{11} - \bar{\tau}_{p_n}(1,1)| > \frac{n^{-\gamma}}{4}\right] + P\left[|\bar{T}_{22} - \bar{\tau}_{p_n}(2,2)| > \frac{n^{-\gamma}}{4}\right] \\
&\quad + P\left[|\bar{T}_1(\mathbf{z}) - E[\bar{h}_{p_n}(\mathbf{X}_1, \mathbf{z})]| > \frac{n^{-\gamma}}{2}\right] + P\left[|\bar{T}_2(\mathbf{z}) - E[\bar{h}_{p_n}(\mathbf{Y}_1, \mathbf{z})]| > \frac{n^{-\gamma}}{2}\right] \\
&= P_1 + P_2 + P_3 + P_4. \qquad (A.30)
\end{aligned}
$$

Space?

We have already proved in part (a) of Lemma A.5 that $P_1 \le O\left(p_n e^{-a_{11}^* n^{1-2\gamma}}\right)$ and $P_2 \le O\left(p_n e^{-a_{22}^* n^{1-2\gamma}}\right)$ for some positive constants a_{11}^* and a_{22}^*. Now, let us consider the term P_3. Observe that

$$
P_3 = P\left[|\bar{T}_2(\mathbf{z}) - E[\bar{h}_{p_n}(\mathbf{X}_1, \mathbf{z})]| > \frac{n^{-\gamma}}{2}\right] = P\left[|\bar{T}_1(\mathbf{z}) - E[\bar{T}_1(\mathbf{z})]| > \frac{n^{-\gamma}}{1}\right] \quad ?
$$

It is shown in part (b) of Lemma A.5 that

$$
P\left[|\bar{T}_1(\mathbf{z}) - E[\bar{T}_1(\mathbf{z})]| > \frac{n^{-\gamma}}{2}\right] \le O\left(p_n e^{-b_1^* n^{1-2\gamma}}\right) \text{ for some positive constant } b_1^*.
$$

Therefore, $P_3 \leq O\left(p_n e^{-b_1^* n^{1-2\gamma}}\right)$. Similarly, $P_4 \leq O\left(p_n e^{-b_2^* n^{1-2\gamma}}\right)$ for some positive constant b_2^*. It follows from (A.30) that

$$
\begin{aligned}
&\mathrm{P}\left[|\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| > n^{-\gamma}\right] \\
&\leq O\left(p_n e^{-a_{11}^* n^{1-2\gamma}}\right) + O\left(p_n e^{-a_{22}^* n^{1-2\gamma}}\right) + O\left(p_n e^{-b_1^* n^{1-2\gamma}}\right) + O\left(p_n e^{-b_2^* n^{1-2\gamma}}\right) \\
&= O\left(p_n e^{-B_0^* n^{1-2\gamma}}\right), \text{where } B_0^* = \min\{a_{11}^*, a_{22}^*, b_1^*, b_2^*\}.
\end{aligned}
$$

Recall that there exist $M > 0$ and $N \in \mathbb{N}$ such that

$$
\begin{aligned}
&p_n \leq e^{Mn^\beta} \text{ for all } n \geq N \\
\Rightarrow\ & p_n e^{-B_0^* n^{1-2\gamma}} \leq e^{-\{B_0^* n^{1-2\gamma} - Mn^\beta\}} \text{ for all } n \geq N \\
\Rightarrow\ & p_n e^{-B_0^* n^{1-2\gamma}} \leq e^{-B_0\{n^{1-2\gamma} - n^\beta\}} \text{ for all } n \geq N,
\end{aligned}
$$

where $B_0 = \min\{B_0^*, M\}$. Therefore, $\mathrm{P}\left[|\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| > n^{-\gamma}\right] \leq O\left(e^{-B_0\{n^{1-2\gamma} - n^\beta\}}\right)$.

(b) Now, we derive a rate of convergence ~~of~~ for the random variable $\bar{\theta}(\mathbf{z}) - \bar{\theta}^0(\mathbf{z})$ for $\mathbf{z} \in \mathbb{R}^{p_n}$. As defined in (A.11), we have

$$
\bar{\theta}(\mathbf{z}) = \frac{1}{2}\{(\bar{T}_{11} - 2\bar{T}_{12} + \bar{T}_{22}) \times \bar{L}(\mathbf{z})\} + \frac{1}{2}\{(\bar{T}_{22} - \bar{T}_{11}) \times \bar{S}(\mathbf{z})\},
$$

where $\bar{L}(\mathbf{z}) = \bar{T}_{22} - 2\bar{T}_2(\mathbf{z}) - \bar{T}_{11} + 2\bar{T}_1(\mathbf{z})$ and $\bar{S}(\mathbf{z}) = \bar{T}_{22} - 2\bar{T}_2(\mathbf{z}) + \bar{T}_{11} - 2\bar{T}_1(\mathbf{z}) + 2\bar{T}_{12}$. Further, $\bar{\theta}^0(\mathbf{z})$ is defined as

$$
\begin{aligned}
\bar{\theta}^0(\mathbf{z}) = &\frac{\bar{\tau}_{p_n}}{2}\{\bar{\tau}_{p_n}(2,2) - 2E[\bar{h}_{p_n}(\mathbf{Y}_1, \mathbf{z}) - \bar{\tau}_{p_n}(1,1) + 2E[\bar{h}_{p_n}(\mathbf{X}_1, \mathbf{z})]\} \\
&+ \frac{1}{2}(\bar{\tau}_{p_n}(2,2) - \bar{\tau}_{p_n}(1,1))\{\bar{\tau}_{p_n}(2,2) - 2E[\bar{h}_{p_n}(\mathbf{Y}_1, \mathbf{z})] + \bar{\tau}_{p_n}(1,1) - 2E[\bar{h}_{p_n}(\mathbf{X}_1, \mathbf{z})] + 2\bar{\tau}_{p_n}(1,2)\}
\end{aligned}
$$

$$
\Rightarrow \bar{\theta}^0(\mathbf{z}) = \frac{\bar{\tau}_{p_n}}{2}E[\bar{L}(\mathbf{z})] + \frac{1}{2}(\bar{\tau}_{p_n}(2,2) - \bar{\tau}_{p_n}(1,1))E[\bar{S}(\mathbf{z})].
$$

Note that $\mathrm{E}[\bar{T}_{11} - 2\bar{T}_{12} + \bar{T}_{22}] = \bar{\tau}_{p_n}$ and $\mathrm{E}[\bar{T}_{22} - \bar{T}_{11}] = \bar{\tau}_{p_n}(2,2) - \bar{\tau}_{p_n}(1,1)$. It follows from part (a) of Lemma A.5 that there exist positive constants c_1 and c_2 such that

$$
\begin{aligned}
&\mathrm{P}\left[|\{\bar{T}_{11} - 2\bar{T}_{12} + \bar{T}_{22}\} - \bar{\tau}_{p_n}| > n^{-\gamma}\right] \leq O\left(p_n e^{-c_1 n^{1-2\gamma}}\right) \text{ and} \\
&\mathrm{P}\left[|\{\bar{T}_{22} - \bar{T}_{11}\} - \{\bar{\tau}_{p_n}(2,2) - \bar{\tau}_{p_n}(1,1)\}| > n^{-\gamma}\right] \leq O\left(p_n e^{-c_2 n^{1-2\gamma}}\right).
\end{aligned} \tag{A.31}
$$

Part (b) of Lemma A.5 suggests that there exist positive constants c_3 and c_4 such that

$$
\begin{aligned}
&\mathrm{P}\left[|\bar{L}(\mathbf{z}) - E[\bar{L}(\mathbf{z})]| > n^{-\gamma}\right] \leq O\left(p_n e^{-c_3 n^{1-2\gamma}}\right) \text{ and} \\
&\mathrm{P}\left[|\bar{S}(\mathbf{z}) - E[\bar{S}(\mathbf{z})]| > n^{-\gamma}\right] \leq O\left(p_n e^{-c_4 n^{1-2\gamma}}\right) \text{ for all } \mathbf{z} \in \mathbb{R}^{p_n}.
\end{aligned} \tag{A.32}
$$

Now, for $\mathbf{z} \in \mathbb{R}^{p_n}$, we have

$$
\begin{aligned}
\mathrm{P}\left[|\bar{\theta}(\mathbf{z}) - \bar{\theta}^0(\mathbf{z})| > n^{-\gamma}\right] \leq\ &\mathrm{P}\left[\left|\frac{1}{2}\{(\bar{T}_{11} - 2\bar{T}_{12} + \bar{T}_{22}) \times \bar{L}(\mathbf{z})\} - \frac{\bar{\tau}_{p_n}}{2}E[\bar{L}(\mathbf{z})]\right| > \frac{n^{-\gamma}}{2}\right] \\
&+ \mathrm{P}\left[\left|\frac{1}{2}\{(\bar{T}_{22} - \bar{T}_{11}) \times \bar{S}(\mathbf{z})\} - \frac{1}{2}\{\bar{\tau}_{p_n}(2,2) - \bar{\tau}_{p_n}(1,1)\}E[\bar{S}(\mathbf{z})]\right| > \frac{n^{-\gamma}}{2}\right].
\end{aligned} \tag{A.33}
$$

Combining (A.31) and (A.32) with Lemma A.6, we conclude that there exists a constant c_{10} such that

$$
\mathrm{P}\left[\left|\frac{1}{2}\{(\bar{T}_{11} - 2\bar{T}_{12} + \bar{T}_{22}) \times \bar{L}(\mathbf{z})\} - \frac{\bar{\tau}_{p_n}}{2}E[\bar{L}(\mathbf{z})]\right| > \frac{n^{-\gamma}}{2}\right] \leq O(p_n e^{-c_{10} n^{1-2\gamma}}). \tag{A.34}
$$

Similarly, there exists a constant $c_{11} > 0$ such that

$$\mathrm{P}\left[\left|\frac{1}{2}\{(\bar{T}_{22} - \bar{T}_{11}) \times \bar{S}(\mathbf{z})\} - \frac{1}{2}\{\bar{\tau}_{p_n}(2,2) - \bar{\tau}_{p_n}(1,1)\}E[\bar{S}(\mathbf{z})]\right| > \frac{n^{-\gamma}}{2}\right] \leq O(p_n e^{-c_{11} n^{1-2\gamma}}). \tag{A.35}$$

Define $B_1^* = \min\{c_{10}, c_{11}\}$. Then, it follows from ~~Eqs.~~ (A.33), (A.34) and (A.35) that

$$\mathrm{P}\left[|\bar{\theta}(\mathbf{z}) - \bar{\theta}^0(\mathbf{z})| > n^{-\gamma}\right] \leq O(p_n e^{-B_1^* n^{1-2\gamma}}) \text{ for all } \mathbf{z} \in \mathbb{R}^{p_n}.$$

alignment?

Since there exist $M > 0$ and $N \in \mathbb{N}$ such that

$$p_n \leq e^{Mn^\beta} \text{ for all } n \geq N$$
$$\Rightarrow p_n e^{-B_1^* n^{1-2\gamma}} \leq e^{-B_1\{n^{1-2\gamma} - n^\beta\}} \text{ for all } n \geq N,$$

where $B_1 = \min\{B_1^*, M\}$. Therefore, $\mathrm{P}\left[|\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| > n^{-\gamma}\right] \leq O\left(e^{-B_1\{n^{1-2\gamma} - n^\beta\}}\right)$ for all $\mathbf{z} \in \mathbb{R}^{p_n}$.

Hence, the proof. □

Proof of Theorem 3.6

Let $l_{\mathbf{Z}}$ denote the true class label of $\mathbf{Z}$ with $P[l_{\mathbf{Z}} = j] = \pi_j$ where $\pi_1 + \pi_2 = 1$. Therefore, $\mathbf{Z} \mid l_{\mathbf{Z}} = 1 \sim \mathbf{F}_1$ and $\mathbf{Z} \mid l_{\mathbf{Z}} = 2 \sim \mathbf{F}_2$. The unconditional distribution of $\mathbf{Z}$ is defined as $\mathbf{H}(\mathbf{z}) = \pi_1 \mathbf{F}_1(\mathbf{z}) + \pi_2 \mathbf{F}_2(\mathbf{z})$, for $\mathbf{z} \in \mathbb{R}^{p_n}$.

(a) Recall that the misclassification probabilities of δ_1 and δ_1^0 are defined as $\Delta_1 = \mathrm{P}[\delta_1(\mathbf{Z}) \neq l_{\mathbf{Z}}]$ and $\Delta_1^0 = \mathrm{P}[\delta_1^0(\mathbf{Z}) \neq l_{\mathbf{Z}}]$, respectively. ~~Therefore,~~ Now,

$$\begin{aligned}
&\Delta_1 - \Delta_1^0 \\
&= \mathrm{P}[\delta_1(\mathbf{Z}) \neq l_{\mathbf{Z}}] - \mathrm{P}[\delta_1^0(\mathbf{Z}) \neq l_{\mathbf{Z}}] \\
&= \int \left\{\mathrm{P}[\delta_1(\mathbf{z}) \neq l_{\mathbf{z}}] - \mathrm{P}[\delta_1^0(\mathbf{z}) \neq l_{\mathbf{z}}]\right\} \, d\mathbf{H}(\mathbf{z}) \\
&= \int \left\{\mathrm{P}[\delta_1^0(\mathbf{z}) = l_{\mathbf{z}}] - \mathrm{P}[\delta_1(\mathbf{z}) = l_{\mathbf{z}}]\right\} \, d\mathbf{H}(\mathbf{z}) \\
&= \int \left\{I[\delta_1^0(\mathbf{z}) = 1]\mathrm{P}[l_{\mathbf{z}} = 1] + I[\delta_1^0(\mathbf{z}) = 0]\mathrm{P}[l_{\mathbf{z}} = 0] - \mathrm{P}[\delta_1(\mathbf{z}) = 1]\mathrm{P}[l_{\mathbf{z}} = 1] + \mathrm{P}[\delta_1(\mathbf{z}) = 0]\mathrm{P}[l_{\mathbf{z}} = 0]\right\} \, d\mathbf{H}(\mathbf{z}) \\
&= \int \left\{(I[\delta_1^0(\mathbf{z}) = 1] - \mathrm{P}[\delta_1(\mathbf{z}) = 1])\mathrm{P}[l_{\mathbf{z}} = 1] + (I[\delta_1^0(\mathbf{z}) = 0] - \mathrm{P}[\delta_1(\mathbf{z}) = 0])\mathrm{P}[l_{\mathbf{z}} = 0]\right\} \, d\mathbf{H}(\mathbf{z}) \\
&= \int (I[\delta_1^0(\mathbf{z}) = 1] - E\left[I[\delta_1(\mathbf{z}) = 1]\right])(2\mathrm{P}[l_{\mathbf{z}} = 1] - 1) \, d\mathbf{H}(\mathbf{z}) \\
&\leq \int \left|E\left[I[\delta_1^0(\mathbf{z}) = 1] - I[\delta_1(\mathbf{z}) = 1]\right]\right| \, |2\mathrm{P}[l_{\mathbf{z}} = 1] - 1| \, d\mathbf{H}(\mathbf{z}) \\
&= \int E\left[|I[\delta_1^0(\mathbf{z}) = 1] - I[\delta_1(\mathbf{z}) = 1]|\right] \, d\mathbf{H}(\mathbf{z}) \\
&= \int E\left[I[\delta_1^0(\mathbf{z}) \neq \delta_1(\mathbf{z})]\right] \, d\mathbf{H}(\mathbf{z}) \\
&= \int \mathrm{P}[\delta_1^0(\mathbf{z}) \neq \delta_1(\mathbf{z})] \, d\mathbf{H}(\mathbf{z}) \\
&= \int \mathrm{P}[\bar{L}(\mathbf{z}) \leq 0, \bar{L}^0(\mathbf{z}) > 0] \, d\mathbf{H}(\mathbf{z}) + \int \mathrm{P}[\bar{L}(\mathbf{z}) > 0, \bar{L}^0(\mathbf{z}) \leq 0] \, d\mathbf{H}(\mathbf{z}) \\
&= P_1 + P_2.
\end{aligned} \tag{A.36}$$

Consider the first term. We have the following:

$$P_1 = \int \mathrm{P}[\bar{L}(\mathbf{z}) \leq 0, \bar{L}^0(\mathbf{z}) > 0] \, d\mathbf{H}(\mathbf{z})$$

$\gamma \geq ?$

$$
\begin{aligned}
&= \int \mathrm{P}[\bar{L}(\mathbf{z}) \leq 0, \bar{L}^0(\mathbf{z}) > 0, |\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| \leq n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) \\
&\qquad + \int \mathrm{P}[\bar{L}(\mathbf{z}) \leq 0, \bar{L}^0(\mathbf{z}) > 0, |\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) \\
&\leq \int \mathrm{P}[\bar{L}(\mathbf{z}) \leq 0, \bar{L}^0(\mathbf{z}) > 0, |\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| \leq n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) + \int \mathrm{P}[|\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) \\
&= P_{11}(\gamma) + \int \mathrm{P}[|\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}). \qquad \text{(A.37)}
\end{aligned}
$$

Note that

$$
\begin{aligned}
P_{11}(\gamma) &= \int \mathrm{P}[\bar{L}(\mathbf{z}) \leq 0, \bar{L}^0(\mathbf{z}) > 0, |\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| \leq n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) \\
&= \int \mathrm{P}[\bar{L}(\mathbf{z}) \leq 0, \bar{L}^0(\mathbf{z}) > 0, -\bar{L}(\mathbf{z}) + \bar{L}^0(\mathbf{z}) \leq n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) \\
&\leq \int \mathrm{P}[\bar{L}^0(\mathbf{z}) \leq n^{-\gamma}, \bar{L}^0(\mathbf{z}) > 0, \bar{L}(\mathbf{z}) \leq 0]\, d\mathbf{H}(\mathbf{z}) \\
&\leq \int \mathrm{P}[\bar{L}^0(\mathbf{z}) \leq n^{-\gamma}, \bar{L}^0(\mathbf{z}) > 0]\, d\mathbf{H}(\mathbf{z}) = \mathrm{P}[0 < \bar{L}^0(\mathbf{Z}) \leq n^{-\gamma}]. \qquad \text{(A.38)}
\end{aligned}
$$

Combining (A.37) and (A.38), we observe that

$$
P_1 \leq \mathrm{P}[0 < \bar{L}^0(\mathbf{Z}) \leq n^{-\gamma}] + \int \mathrm{P}[|\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}). \qquad \text{(A.39)}
$$

Following similar arguments, we can write P_2 as

$$
\begin{aligned}
P_2 &= \int \mathrm{P}[\bar{L}^0(\mathbf{z}) \leq 0, \bar{L}(\mathbf{z}) > 0]\, d\mathbf{H}(\mathbf{z}) \\
&\leq \int \mathrm{P}[\bar{L}^0(\mathbf{z}) \leq 0, \bar{L}(\mathbf{z}) > 0, |\bar{L}^0(\mathbf{z}) - \bar{L}(\mathbf{z})| \leq n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) + \int \mathrm{P}[|\bar{L}^0(\mathbf{z}) - \bar{L}(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) \\
&= \int \mathrm{P}[\bar{L}^0(\mathbf{z}) \leq 0, \bar{L}(\mathbf{z}) > 0, |\bar{L}^0(\mathbf{z}) - \bar{L}(\mathbf{z})| \leq n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) + \int \mathrm{P}[|\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| > n^{-\gamma}] d\mathbf{H}(\mathbf{z}) \\
&\leq \int \mathrm{P}[\bar{L}^0(\mathbf{z}) \leq 0, \bar{L}(\mathbf{z}) > 0, -\bar{L}^0(\mathbf{z}) + \bar{L}(\mathbf{z}) \leq n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) + \int \mathrm{P}[|\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| > n^{-\gamma}] d\mathbf{H}(\mathbf{z}) \\
&\leq \int \mathrm{P}[-n^{\gamma} < \bar{L}^0(\mathbf{z}) \leq 0]\, d\mathbf{H}(\mathbf{z}) + \int \mathrm{P}[|\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| > n^{-\gamma}] d\mathbf{H}(\mathbf{z}) \\
&= \mathrm{P}[-n^{-\gamma} < \bar{L}^0(\mathbf{Z}) \leq 0] + \int \mathrm{P}[|\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| > n^{-\gamma}] d\mathbf{H}(\mathbf{z}). \qquad \text{(A.40)}
\end{aligned}
$$

Combining (A.36), (A.39) and (A.40), we obtain

$$
\Delta_1 - \Delta_1^0 \leq \mathrm{P}[|\bar{L}^0(\mathbf{Z})| < n^{-\gamma}] + 2 \int \mathrm{P}[|\bar{L}(\mathbf{z}) - \bar{L}^0(\mathbf{z})| > n^{-\gamma}] d\mathbf{H}(\mathbf{z}).
$$

Using now

It follows ~~from~~ Lemma 3.5 that

$$
\Delta_1 - \Delta_1^0 \leq \mathrm{P}[|\bar{L}^0(\mathbf{Z})| < n^{-\gamma}] + O\left(e^{-B_0\{n^{1-2\gamma} - n^{\beta}\}}\right).
$$

(b) The misclassification probabilities of δ_2 and δ_2^0 are defined as $\Delta_2 = \mathrm{P}[\delta_2(\mathbf{Z}) \neq l_{\mathbf{Z}}]$ and $\Delta_2^0 = \mathrm{P}[\delta_2^0(\mathbf{Z}) \neq l_{\mathbf{Z}}]$, respectively. Similar to (A.36), we have the following:

$$
\Delta_2 - \Delta_2^0 \leq \int \mathrm{P}[\bar{\theta}(\mathbf{z}) \leq 0, \bar{\theta}^0(\mathbf{z}) > 0]\, d\mathbf{H}(\mathbf{z}) + \int \mathrm{P}[\bar{\theta}(\mathbf{z}) > 0, \bar{\theta}^0(\mathbf{z}) \leq 0]\, d\mathbf{H}(\mathbf{z}) = P_1 + P_2. \qquad \text{(A.41)}
$$

Consider the first term. Similar to (A.37), we have the following:

$$\begin{aligned}
P_1 &= \int \mathrm{P}[\bar{\theta}(\mathbf{z}) \le 0, \bar{\theta}^0(\mathbf{z}) > 0]\, d\mathbf{H}(\mathbf{z}) \\
&\le \int \mathrm{P}[\bar{\theta}(\mathbf{z}) \le 0, \bar{\theta}^0(\mathbf{z}) > 0, |\bar{\theta}(\mathbf{z}) - \bar{\theta}^0(\mathbf{z})| \le n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) + \int \mathrm{P}[|\bar{\theta}(\mathbf{z}) - \bar{\theta}^0(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) \\
&= P_{11}(\gamma) + \int \mathrm{P}[|\bar{\theta}(\mathbf{z}) - \bar{\theta}^0(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}). \qquad \text{(A.42)}
\end{aligned}$$

Note that

$$\begin{aligned}
P_{11}(\gamma) &= \int \mathrm{P}[\bar{\theta}(\mathbf{z}) \le 0, \bar{\theta}^0(\mathbf{z}) > 0, |\bar{\theta}(\mathbf{z}) - \bar{\theta}^0(\mathbf{z})| \le n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) \\
&= \int \mathrm{P}[\bar{\theta}(\mathbf{z}) \le 0, \bar{\theta}^0(\mathbf{z}) > 0, -\bar{\theta}(\mathbf{z}) + \bar{\theta}^0(\mathbf{z}) \le n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) \\
&\le \int \mathrm{P}[\bar{\theta}^0(\mathbf{z}) \le n^{-\gamma}, \bar{\theta}^0(\mathbf{z}) > 0, \bar{\theta}(\mathbf{z}) \le 0]\, d\mathbf{H}(\mathbf{z}) \\
&\le \int \mathrm{P}[\bar{\theta}^0(\mathbf{z}) \le n^{-\gamma}, \bar{\theta}^0(\mathbf{z}) > 0]\, d\mathbf{H}(\mathbf{z}) = \mathrm{P}[0 < \bar{\theta}^0(\mathbf{Z}) \le n^{-\gamma}]. \qquad \text{(A.43)}
\end{aligned}$$

Combining (A.42) and (A.43) we observe that

$$P_1 \le \mathrm{P}[0 < \bar{\theta}^0(\mathbf{Z}) \le n^{-\gamma}] + \int \mathrm{P}[|\bar{\theta}(\mathbf{z}) - \bar{\theta}^0(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}). \qquad \text{(A.44)}$$

Same proof?

Following similar arguments, we can write P_2 as

$$\begin{aligned}
P_2 &= \int \mathrm{P}[\bar{\theta}^0(\mathbf{z}) \le 0, \bar{\theta}(\mathbf{z}) > 0]\, d\mathbf{H}(\mathbf{z}) \\
&\le \int \mathrm{P}[\bar{\theta}^0(\mathbf{z}) \le 0, \bar{\theta}(\mathbf{z}) > 0, |\bar{\theta}^0(\mathbf{z}) - \bar{\theta}(\mathbf{z})| \le n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) + \int \mathrm{P}[|\bar{\theta}^0(\mathbf{z}) - \bar{\theta}(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) \\
&= \int \mathrm{P}[\bar{\theta}^0(\mathbf{z}) \le 0, \bar{\theta}(\mathbf{z}) > 0, |\bar{\theta}^0(\mathbf{z}) - \bar{\theta}(\mathbf{z})| \le n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) + \int \mathrm{P}[|\bar{\theta}(\mathbf{z}) - \bar{\theta}^0(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) \\
&\le \int \mathrm{P}[\bar{\theta}^0(\mathbf{z}) \le 0, \bar{\theta}(\mathbf{z}) > 0, -\bar{\theta}^0(\mathbf{z}) + \bar{\theta}(\mathbf{z}) \le n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) + \int \mathrm{P}[|\bar{\theta}(\mathbf{z}) - \bar{\theta}^0(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) \\
&\le \int \mathrm{P}[-n^{\gamma} < \bar{\theta}^0(\mathbf{z}) \le 0]\, d\mathbf{H}(\mathbf{z}) + \int \mathrm{P}[|\bar{\theta}(\mathbf{z}) - \bar{\theta}^0(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}) \\
&= \mathrm{P}[-n^{-\gamma} < \bar{\theta}^0(\mathbf{Z}) \le 0] + \int \mathrm{P}[|\bar{\theta}(\mathbf{z}) - \bar{\theta}^0(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}). \qquad \text{(A.45)}
\end{aligned}$$

Combining (A.41), (A.44) and (A.45) we obtain

$$\Delta_2 - \Delta_2^0 \le \mathrm{P}[|\bar{\theta}^0(\mathbf{z})| < n^{-\gamma}] + 2\int \mathrm{P}[|\bar{\theta}(\mathbf{z}) - \bar{\theta}^0(\mathbf{z})| > n^{-\gamma}]\, d\mathbf{H}(\mathbf{z}).$$

It follows from Lemma 3.5 that

$$\Delta_2 - \Delta_2^0 \le \mathrm{P}[|\bar{\theta}^0(\mathbf{Z})| < n^{-\gamma}] + O\left(e^{-B_1\{n^{1-2\gamma} - n^{\beta}\}}\right).$$

Hence, the proof. ☐

?

Proof of Corollary 3.7

Observe that $\beta = 0$ implies $p_n(= p)$ is fixed. Then, it follows from Theorem 3.6 that for $0 < \gamma < 1/2$,

$$\Delta_1 - \Delta_1^0 \le \mathrm{P}[|\bar{L}^0(\mathbf{Z})| < n^{-\gamma}] + O\left(e^{-B_0 n^{1-2\gamma}}\right) \text{ and } \Delta_2 - \Delta_2^0 \le \mathrm{P}[|\bar{\theta}^0(\mathbf{Z})| < n^{-\gamma}] + O\left(e^{-B_1 n^{1-2\gamma}}\right)$$

where B_0 and B_1 are positive constants. Clearly, we have $e^{-B_0 n^{1-2\gamma}} = o(1)$ and $e^{-B_1 n^{1-2\gamma}} = o(1)$ for all $0 < \gamma < 1/2$. Now, if $\mathbf{F}_1$, and $\mathbf{F}_2$ are absolutely continuous distribution functions, then the unconditional distribution of $\mathbf{Z}$, given by $\mathbf{H}(\mathbf{z}) = \pi_1 \mathbf{F}_1(\mathbf{z}) + \pi_2 \mathbf{F}_2(\mathbf{z})$ is also absolutely continuous. Consequently, $P[|\bar{L}^0(\mathbf{Z})| < n^{-\gamma}]$ and $P[|\bar{\theta}^0(\mathbf{Z})| < n^{-\gamma}]$ converge to 0 as $n \to \infty$. Hence, $\Delta_i - \Delta_i^0 \to 0$ as $n \to \infty$ for $i = 1, 2$. □

Remark : If $\beta > 0$, then p_n can grow with n. Theorem 3.6 suggests that

i. ~~If~~ $P[|\bar{L}^0(\mathbf{Z})| < n^{-\gamma}] \to 0$ and $\Delta_1^0 \to 0$, then $\Delta_1 \to 0$ as $n \to \infty,\ p_n \to \infty$.

ii. ~~If~~ $P[|\bar{\theta}^0(\mathbf{Z})| < n^{-\gamma}] \to 0$ and $\Delta_2^0 \to 0$, then $\Delta_2 \to 0$ as $n \to \infty,\ p_n \to \infty$.

$\min(n, p_n) \to \infty$.

B SIMULATION DETAILS

- GLMNET: The R-package `glmnet` ~~is~~ was used for the implementation of GLMNET. The tuning parameter α in the elastic-net penalty term ~~is~~ was kept fixed at the default value 1. The weight λ of the penalty term ~~is~~ was chosen by cross-validation using the function `cv.glmnet` with default values of its arguments.

- 1NN: The `knn1` function from the R-package `class` ~~is~~ was used for implementation of the usual 1-nearest neighbor classifier.

- NN-RAND: The function `classify` from the package `RandPro` ~~is~~ was used with default values of the arguments.

- NNET: ~~For NNET~~, We used `nnet` from the package `nnet` to fit a single-hidden-layer neural network with default parameters. The number of units in the hidden layer ~~was~~ allowed to vary in the set $\{1, 3, 5, 10\}$, and the minimum misclassification rate was reported as NNET. were

- SVM: The R package `e1071` ~~is~~ was used for implementing SVM with linear and RBF kernel. For the RBF kernel, i.e., $K_\theta(\mathbf{x}, \mathbf{y}) = \exp\{-\theta\|\mathbf{x} - \mathbf{y}\|^2\}$, we took the default value of the tuning parameter θ, i.e., $\theta = 1/p$. considered.

Table 2: Estimates of $\bar{\tau}_p(1,1), \bar{\tau}_p(1,2)$ and $\bar{\tau}_p(2,2)$ in the simulated examples (along with the standard errors in parentheses) based on 100 replications. → space.

Example

Ex	$\bar{T}_{11}$	$\bar{T}_{12}$	$\bar{T}_{22}$	$\bar{T}_{12} \geq \min\{\bar{T}_{11}, \bar{T}_{22}\}$
1	0.1562 (0.0019)	0.1446 (0.0020)	0.1273 (0.0022)	True
2	0.0909 (0.0014)	0.0984 (0.0010)	0.1109 (0.0015)	True
3	0.0857 (0.0018)	0.0821 (0.0016)	0.1018 (0.0027)	False
4	0.0857 (0.0018)	0.0748 (0.0016)	0.0545 (0.0016)	True
5	0.2077 (0.0005)	0.2106 (0.0004)	0.2136 (0.0004)	True

If $\bar{T}_{12} \geq \min\{\bar{T}_{11}, \bar{T}_{22}\}$, then we expect the classifier δ_2 to yield better misclassification rates than δ_1. We have observed this while studying the performance of the proposed classifiers in the simulated examples (see Figure 3 of the main article).

lower? ref?

20

Table 3: Average time (in seconds) taken by the classifiers to classify 100+100 test observations in Example 1

p	δ_0	δ_1	δ_2	GLM NET	1NN	NN RAND	NNET* 1	NNET* 3	NNET* 5	NNET* 10	SVM LIN	SVM RBF
50	0.0149	0.0189	0.0188	0.094	0.0008	2.7834	0.009	0.0162	0.0328	0.111	0.0052	0.006
100	0.0156	0.0236	0.0238	0.0978	0.0024	3.4872	0.013	0.0454	0.107	0.4012	0.0104	0.0102
250	0.0185	0.039	0.0389	0.105	0.0048	4.6608	0.0382	0.2232	0.5982	4.2194	0.0224	0.024
500	0.0209	0.0551	0.0549	0.1132	0.007	5.3308	0.1104	0.8512	3.924	19.7896	0.0398	0.0402
1000	0.0263	0.0807	0.0808	0.153	0.012	6.7963	0.3883	6.3370	19.1236	100.7417	0.0713	0.0797

* 1,3,5,10 represent the numbers of units in the single-hidden-layer of the neural network.

Note that the weight λ of the elastic-net penalty in GLMNET is tuned using cross-validation method. The rest of the existing classifiers in Table 3 are implemented with their respective default arguments. The proposed classifiers are free from parameter tuning.

Codes: The R-codes for implementation of the proposed classifiers are uploaded along with the supplementary material as 'RCodes.zip'.

Link!

Font!

digits?