Supplementary Material: Robust Classification of HDLSS data using Data-adaptive Energy Distance

No Author Given

No Institute Given

A MATHEMATICAL DETAILS AND PROOFS

Lemma A.1. *Suppose* $\mathbf{U} \sim I, \mathbf{V} \sim J$, *and* $\mathbf{Z} \sim K$ *for* $I, J, K \in \{\mathbf{F}, \mathbf{G}\}$, *such that they are all independent. If assumptions 1 to 3 are satisfied, then*

$$\rho_0(\mathbf{U}, \mathbf{V}, \mathbf{Z}) \xrightarrow{\mathbb{P}} \frac{1}{\pi} \cos^{-1}\left(\frac{\mu_{IJK}}{\sqrt{\mu_{IK}\mu_{JK}}}\right) \text{ , as } d \to \infty$$

where $\mu_{IJK} = \lim_d \frac{1}{d}\mathbb{E}((\mathbf{U}-\mathbf{Z})^\top(\mathbf{V}-\mathbf{Z}))$ *and* $\mu_{IK} = \lim_d \frac{1}{d}\mathbb{E}(\|\mathbf{U}-\mathbf{Z}\|^2)$.

Proof of lemma A.1 Note that

$$\begin{aligned}
\frac{1}{d}\mathbb{E}((\mathbf{U}-\mathbf{Z})^\top(\mathbf{V}-\mathbf{Z})) &= \lim_d \frac{1}{d}\mathbb{E}[\mathbf{U}^\top\mathbf{V} - \mathbf{U}^\top\mathbf{Z} - \mathbf{V}^\top\mathbf{Z} + \mathbf{Z}^\top\mathbf{Z}] \\
&= \frac{1}{d}\left(\boldsymbol{\mu}_I^\top\boldsymbol{\mu}_J - \boldsymbol{\mu}_I^\top\boldsymbol{\mu}_{\boldsymbol{K}} - \boldsymbol{\mu}_J^\top\boldsymbol{\mu}_{\boldsymbol{K}} + \boldsymbol{\mu}_{\boldsymbol{K}}^\top\boldsymbol{\mu}_{\boldsymbol{K}} + \text{trace}(\Sigma_K)\right) \\
&= \frac{1}{d}(\boldsymbol{\mu}_I - \boldsymbol{\mu}_K)^\top(\boldsymbol{\mu}_J - \boldsymbol{\mu}_K) + \frac{1}{d}\text{trace}(\Sigma_K)
\end{aligned}$$

Since $I, J, K \in \{\mathbf{F}, \mathbf{G}\}$, at least two of these take same value. Hence, using assumption 2, we have

$$\mu_{IJK} = \begin{cases} \lambda_{IK} + \sigma_K^2, & \text{if } I = J \\ \sigma_K^2, & \text{otherwise.} \end{cases}$$

Now, using assumption 2, we shall show existence of the limiting constant μ_{IK}.

$$\begin{aligned}
\mu_{IK} &= \lim_d \frac{1}{d}\mathbb{E}[\mathbf{U}^\top\mathbf{U} - 2\mathbf{U}^\top\mathbf{Z} + \mathbf{Z}^\top\mathbf{Z}] \\
&= \lim_d \frac{1}{d}\left(\boldsymbol{\mu}_I^\top\boldsymbol{\mu}_I + \text{trace}(\Sigma_I) - 2\boldsymbol{\mu}_I^\top\boldsymbol{\mu}_{\boldsymbol{K}} + \boldsymbol{\mu}_{\boldsymbol{K}}^\top\boldsymbol{\mu}_{\boldsymbol{K}} + \text{trace}(\Sigma_K)\right) \\
&= \lambda_{IK} + \sigma_I^2 + \sigma_K^2.
\end{aligned}$$

Using Chebyshev's inequality, we have

$$\begin{aligned}
&\mathbb{P}\left[\left|\frac{1}{d}(\mathbf{U}-\mathbf{Z})^\top(\mathbf{V}-\mathbf{Z}) - \mathbb{E}\left(\frac{1}{d}(\mathbf{U}-\mathbf{Z})^\top(\mathbf{V}-\mathbf{Z})\right)\right| > \epsilon\right] \\
&\leq \frac{1}{\epsilon^2}\operatorname{var}\left(\frac{1}{d}\sum_{i=1}^{n}(U_i - Z_i)(V_i - Z_i)\right) \\
&= \frac{1}{\epsilon^2 d^2}\Bigg[\sum_{i=1}^{d}\operatorname{var}((U_i - Z_i)(V_i - Z_i)) \\
&\qquad + \sum_{1\leq i<j\leq d}\operatorname{cov}((U_i - Z_i)(V_i - Z_i), (U_j - Z_j)(V_j - Z_j))\Bigg] \\
&\leq \frac{1}{\epsilon^2 d^2}\Bigg[\sum_{i=1}^{d}\mathbb{E}((U_i - Z_i)^2(V_i - Z_i)^2) \\
&\qquad + \sum_{1\leq i<j\leq d}\operatorname{cov}((U_i - Z_i)(V_i - Z_i), (U_j - Z_j)(V_j - Z_j))\Bigg] \\
&\to 0 \ \text{, as } d \to \infty \ \ \text{[Using assumptions 1 and 3].}
\end{aligned}$$

So, as $d \to \infty$,

$$\begin{aligned}
\left|\frac{1}{d}(\mathbf{U}-\mathbf{Z})^\top(\mathbf{V}-\mathbf{Z}) - \mu_{IJK}\right| &\leq \left|\frac{1}{d}(\mathbf{U}-\mathbf{Z})^\top(\mathbf{V}-\mathbf{Z}) - \mathbb{E}\left(\frac{1}{d}(\mathbf{U}-\mathbf{Z})^\top(\mathbf{V}-\mathbf{Z})\right)\right| \\
&\quad + \left|\mathbb{E}\left(\frac{1}{d}(\mathbf{U}-\mathbf{Z})^\top(\mathbf{V}-\mathbf{Z})\right) - \mu_{IJK}\right| \xrightarrow{\mathbb{P}} 0.
\end{aligned}$$

Since μ_{IJK} is independent of d, we can write it as

$$\frac{1}{d}(\mathbf{U}-\mathbf{Z})^\top(\mathbf{V}-\mathbf{Z}) \xrightarrow{\mathbb{P}} \mu_{IJK}, \quad \text{, as } d \to \infty.$$

Similarly, we have

$$\frac{1}{d}\|\mathbf{U}-\mathbf{Z}\|^2 \xrightarrow{\mathbb{P}} \mu_{IK} \quad \text{and} \quad \frac{1}{d}\|\mathbf{V}-\mathbf{Z}\|^2 \xrightarrow{\mathbb{P}} \mu_{JK} \text{ , as } d \to \infty.$$

Using the continuous mapping theorem repeatedly, we get

$$\cos(\pi \cdot \rho_0(\mathbf{U},\mathbf{V},\mathbf{Z})) = \frac{\frac{1}{d}(\mathbf{U}-\mathbf{Z})^\top(\mathbf{V}-\mathbf{Z})}{\frac{1}{\sqrt{d}}\|\mathbf{U}-\mathbf{Z}\|.\frac{1}{\sqrt{d}}\|\mathbf{V}-\mathbf{Z}\|} \xrightarrow{\mathbb{P}} \frac{\mu_{IJK}}{\sqrt{\mu_{IK}\mu_{JK}}} \quad \text{, as } d \to \infty.$$

Therefore, we have

$$\rho_0(\mathbf{U},\mathbf{V},\mathbf{Z}) \xrightarrow{\mathbb{P}} \frac{1}{\pi}\cos^{-1}\left(\frac{\mu_{IJK}}{\sqrt{\mu_{IK}\mu_{JK}}}\right) \text{ , as } d \to \infty.$$

$\square$

Lemma A.2. *Suppose, assumptions 1 to 3 are satisfied. Then, for* $\mathbf{U} \sim I, \mathbf{V} \sim J$; $I, J \in \{\mathbf{F}, \mathbf{G}\}$ *, as* $d \to \infty$*,*

$$\hat{\rho}(\mathbf{U}, \mathbf{V}) \xrightarrow{\mathbb{P}} \frac{1}{\pi(m+n)} \left(m \cos^{-1} \left(\frac{\mu_{IJ\mathbf{F}}}{\sqrt{\mu_{I\mathbf{F}} \mu_{J\mathbf{F}}}} \right) + n \cos^{-1} \left(\frac{\mu_{IJ\mathbf{G}}}{\sqrt{\mu_{I\mathbf{G}} \mu_{J\mathbf{G}}}} \right) \right).$$

In particular, $\hat{\rho}(\mathbf{X_i}, \mathbf{X_j}) \xrightarrow{\mathbb{P}} \theta_{\mathbf{FF}}$, $\hat{\rho}(\mathbf{Y_i}, \mathbf{Y_j}) \xrightarrow{\mathbb{P}} \theta_{\mathbf{GG}}$, $\hat{\rho}(\mathbf{X_i}, \mathbf{Y_j}) \xrightarrow{\mathbb{P}} \theta_{\mathbf{FG}}$ *, as* $d \to \infty$*, where*

$$\theta_{\mathbf{FF}} = \frac{1}{\pi(m+n)} \left(\frac{m\pi}{3} + n \cos^{-1} \left(\frac{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{G}}}{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{G}} + \sigma^2_{\mathbf{F}}} \right) \right),$$

$$\theta_{\mathbf{GG}} = \frac{1}{\pi(m+n)} \left(m \cos^{-1} \left(\frac{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{F}}}{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{G}} + \sigma^2_{\mathbf{F}}} \right) + \frac{n\pi}{3} \right), \text{ and}$$

$$\begin{aligned}\theta_{\mathbf{FG}} = \frac{1}{2} - \frac{1}{2\pi(m+n)} \Bigg[& m \cos^{-1} \left(\frac{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{G}}}{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{G}} + \sigma^2_{\mathbf{F}}} \right) \\ & + n \cos^{-1} \left(\frac{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{F}}}{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{G}} + \sigma^2_{\mathbf{F}}} \right) \Bigg].\end{aligned}$$

Proof of lemma A.2 From lemma A.1 and definition of $\hat{\rho}$, we see that

$$\hat{\rho}(\mathbf{U}, \mathbf{V}) \xrightarrow{\mathbb{P}} \frac{1}{\pi(m+n)} \left(m \cos^{-1} \left(\frac{\mu_{IJ\mathbf{F}}}{\sqrt{\mu_{I\mathbf{F}} \mu_{J\mathbf{F}}}} \right) + n \cos^{-1} \left(\frac{\mu_{IJ\mathbf{G}}}{\sqrt{\mu_{I\mathbf{G}} \mu_{J\mathbf{G}}}} \right) \right),$$

as $d \to \infty$. Therefore, as $d \to \infty$,

$$\begin{aligned}\hat{\rho}(\mathbf{X}_i, \mathbf{X}_j) &\xrightarrow{\mathbb{P}} \frac{m \cos^{-1} \left(\frac{\mu_{FFF}}{\mu_{\mathbf{FF}}} \right) + n \cos^{-1} \left(\frac{\mu_{FFG}}{\mu_{\mathbf{FG}}} \right)}{\pi(m+n)} \\ &= \frac{1}{\pi(m+n)} \left(\frac{m\pi}{3} + n \cos^{-1} \left(\frac{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma_G}{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma_G + \sigma_F} \right) \right),\end{aligned}$$

$$\begin{aligned}\hat{\rho}(\mathbf{Y}_i, \mathbf{Y}_j) &\xrightarrow{\mathbb{P}} \frac{m \cos^{-1} \left(\frac{\mu_{\mathbf{GGF}}}{\mu_{\mathbf{GF}}} \right) + n \cos^{-1} \left(\frac{\mu_{GGG}}{\mu_{\mathbf{GG}}} \right)}{\pi(m+n)} \\ &= \frac{1}{\pi(m+n)} \left(m \cos^{-1} \left(\frac{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{F}}}{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{G}} + \sigma^2_{\mathbf{F}}} \right) + \frac{n\pi}{3} \right),\end{aligned}$$

and $\hat{\rho}(\mathbf{X}_i, \mathbf{Y}_j) \xrightarrow{\mathbb{P}} \frac{1}{\pi(m+n)} \left(m \cos^{-1} \left(\frac{\mu_{\mathbf{FGF}}}{\sqrt{\mu_{\mathbf{FF}} \mu_{\mathbf{GF}}}} \right) + n \cos^{-1} \left(\frac{\mu_{\mathbf{FGG}}}{\sqrt{\mu_{\mathbf{FG}} \mu_{\mathbf{GG}}}} \right) \right).$

Note that

$$\begin{aligned}
&\frac{m\cos^{-1}\left(\frac{\mu_{\mathbf{FGF}}}{\sqrt{\mu_{\mathbf{FF}}\mu_{\mathbf{GF}}}}\right)+n\cos^{-1}\left(\frac{\mu_{\mathbf{FGG}}}{\sqrt{\mu_{\mathbf{FG}}\mu_{\mathbf{GG}}}}\right)}{\pi(m+n)}\\
=&\frac{m\cos^{-1}\left(\frac{\sqrt{\sigma_{\mathbf{F}}^2}}{\sqrt{2(\lambda_{\mathbf{FF}}+\lambda_{\mathbf{GG}}-2\lambda_{\mathbf{FG}}+\sigma_{\mathbf{G}}^2+\sigma_{\mathbf{F}}^2)}}\right)+n\cos^{-1}\left(\frac{\sqrt{\sigma_{\mathbf{G}}^2}}{\sqrt{2(\lambda_{\mathbf{FF}}+\lambda_{\mathbf{GG}}-2\lambda_{\mathbf{FG}}+\sigma_{\mathbf{G}}^2+\sigma_{\mathbf{F}}^2)}}\right)}{\pi(m+n)}\\
=&\frac{\pi(m+n)-m\cos^{-1}\left(\frac{\lambda_{\mathbf{FF}}+\lambda_{\mathbf{GG}}-2\lambda_{\mathbf{FG}}+\sigma_{\mathbf{G}}^2}{\lambda_{\mathbf{FF}}+\lambda_{\mathbf{GG}}-2\lambda_{\mathbf{FG}}+\sigma_{\mathbf{G}}^2+\sigma_{\mathbf{F}}^2}\right)-n\cos^{-1}\left(\frac{\lambda_{\mathbf{FF}}+\lambda_{\mathbf{GG}}-2\lambda_{\mathbf{FG}}+\sigma_{\mathbf{F}}^2}{\lambda_{\mathbf{FF}}+\lambda_{\mathbf{GG}}-2\lambda_{\mathbf{FG}}+\sigma_{\mathbf{G}}^2+\sigma_{\mathbf{F}}^2}\right)}{2\pi(m+n)}.
\end{aligned}$$

□

We now define

$$t_{\mathbf{FG}}=\mathbb{E}\left[\rho\left(\mathbf{X}_1,\mathbf{Y}_1\right)\right] \text{ and } \hat{t}_{\mathbf{FG}}=\frac{1}{mn}\sum_{i,j}\hat{\rho}\left(\mathbf{X}_i,\mathbf{Y}_j\right).$$

Lemma A.3. *Suppose, assumptions 1 to 3 are satisfied. Then, as $d\to\infty$,*

$$\begin{aligned}
\mathcal{W}_{\mathbf{FG}}^*\to 2\theta_{\mathbf{FG}}-\theta_{\mathbf{FF}}-\theta_{\mathbf{GG}}=&\frac{2}{3}-\frac{1}{\pi}\Bigg[\cos^{-1}\left(\frac{\lambda_{\mathbf{FF}}+\lambda_{\mathbf{GG}}-2\lambda_{\mathbf{FG}}+\sigma_{\mathbf{G}}^2}{\lambda_{\mathbf{FF}}+\lambda_{\mathbf{GG}}-2\lambda_{\mathbf{FG}}+\sigma_{\mathbf{G}}^2+\sigma_{\mathbf{F}}^2}\right)\\
&+\cos^{-1}\left(\frac{\lambda_{\mathbf{FF}}+\lambda_{\mathbf{GG}}-2\lambda_{\mathbf{FG}}+\sigma_{\mathbf{F}}^2}{\lambda_{\mathbf{FF}}+\lambda_{\mathbf{GG}}-2\lambda_{\mathbf{FG}}+\sigma_{\mathbf{G}}^2+\sigma_{\mathbf{F}}^2}\right)\Bigg].
\end{aligned}$$

Proof of lemma A.3

$$\begin{aligned}
&\mathbb{E}[\hat{\rho}\left(\mathbf{X}_i,\mathbf{X}_j\right)]-t_{\mathbf{FF}}\\
&=\mathbb{E}[\hat{\rho}\left(\mathbf{X}_i,\mathbf{X}_j\right)]-\mathbb{E}\left[\rho\left(\mathbf{X}_i,\mathbf{X}_j\right)\right]\\
&=\frac{m}{m+n}\mathbb{E}\left[\rho_0(\mathbf{X}_i,\mathbf{X}_j;\mathbf{X}_k)\right]+\frac{n}{m+n}\mathbb{E}\left[\rho_0(\mathbf{X}_i,\mathbf{X}_j;\mathbf{Y}_k)\right]-\mathbb{E}\left[\rho\left(\mathbf{X}_i,\mathbf{X}_j\right)\right]\\
&=\frac{m}{m+n}\mathbb{E}\left[\rho_0(\mathbf{X}_i,\mathbf{X}_j;\mathbf{X}_k)\right]+\frac{n}{m+n}\mathbb{E}\left[\rho_0(\mathbf{X}_i,\mathbf{X}_j;\mathbf{Y}_k)\right]-\mathbb{E}\left[\mathbb{E}_{\mathbf{Q}}[\rho_0\left(\mathbf{X}_i,\mathbf{X}_j;\mathbf{Q}\right)]\right]\\
&\qquad\qquad\qquad\text{where } \mathbf{Q}\sim\frac{m}{m+n}\mathbf{F}+\frac{n}{m+n}\mathbf{G}\\
&=\frac{m}{m+n}\mathbb{E}\left[\rho_0(\mathbf{X}_i,\mathbf{X}_j;\mathbf{X}_3)\right]+\frac{n}{m+n}\mathbb{E}\left[\rho_0(\mathbf{X}_i,\mathbf{X}_j;\mathbf{Y}_3)\right]-\mathbb{E}\left[\rho_0\left(\mathbf{X}_i,\mathbf{X}_j;\mathbf{Q}\right)\right]\\
&=\frac{m}{m+n}\mathbb{E}\left[\rho_0(\mathbf{X}_i,\mathbf{X}_j;\mathbf{X}_3)\right]+\frac{n}{m+n}\mathbb{E}\left[\rho_0(\mathbf{X}_i,\mathbf{X}_j;\mathbf{Y}_3)\right]\\
&\quad-\frac{m}{m+n}\mathbb{E}\left[\rho\left(\mathbf{X}_i,\mathbf{X}_j;\mathbf{Q}\right)|\mathbf{Q}\sim\mathbf{F}\right]-\frac{n}{m+n}\mathbb{E}\left[\rho\left(\mathbf{X}_i,\mathbf{X}_j;\mathbf{Q}\right)|\mathbf{Q}\sim\mathbf{G}\right]\\
&=0.
\end{aligned}$$

Since, $\hat{\rho}(\mathbf{X_i},\mathbf{X_j})\xrightarrow{\mathbb{P}}\theta_{\mathbf{FF}}$ and $\hat{\rho}$ is a bounded function, using the Dominated Convergence Theorem, we have

$$t_{\mathbf{FF}}=\mathbb{E}[\hat{\rho}(\mathbf{X_i},\mathbf{X_j})]\to\theta_{\mathbf{FF}}\text{ , as } d\to\infty.$$

Similarly, we can show that $t_{\mathbf{GG}} \to \theta_{\mathbf{GG}}$ and $t_{\mathbf{FG}} \to \theta_{\mathbf{FG}}$, as $d \to \infty$. Hence, $\mathcal{W}^*_{\mathbf{FG}} \to 2\theta_{\mathbf{FG}} - \theta_{\mathbf{FF}} - \theta_{\mathbf{GG}}$, as $d \to \infty$. Substituting the values of $\theta_{\mathbf{FG}}, \theta_{\mathbf{FF}}$ and $\theta_{\mathbf{GG}}$, we get

$$\begin{aligned}\mathcal{W}^*_{\mathbf{FG}} \to & \frac{2}{3} - \frac{1}{\pi}\Bigg[\cos^{-1}\left(\frac{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{G}}}{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{G}} + \sigma^2_{\mathbf{F}}}\right) \\ & + \cos^{-1}\left(\frac{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{F}}}{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{G}} + \sigma^2_{\mathbf{F}}}\right)\Bigg] \text{, as } d \to \infty.\end{aligned}$$

$\square$

Lemma A.4. *Suppose, assumptions 1 to 3 are satisfied. Then,*

(a) i. $\hat{t}_{\mathbf{FF}} \xrightarrow{\mathbb{P}} \theta_{\mathbf{FF}}$ *, as* $d \to \infty$ *;*
ii. $\hat{t}_{\mathbf{GG}} \xrightarrow{\mathbb{P}} \theta_{\mathbf{GG}}$ *, as* $d \to \infty$ *;*
iii. $\hat{t}_{\mathbf{FG}} \xrightarrow{\mathbb{P}} \theta_{\mathbf{FG}}$ *, as* $d \to \infty$.

(b) i. if $\mathbf{Z} \sim \mathbf{F}$*, then* $\hat{t}_{\mathbf{F}}(\mathbf{Z}) \xrightarrow{\mathbb{P}} \theta_{\mathbf{FF}}$ *and* $\hat{t}_{\mathbf{G}}(\mathbf{Z}) \xrightarrow{\mathbb{P}} \theta_{\mathbf{FG}}$ *, as* $d \to \infty$ *;*
ii. if $\mathbf{Z} \sim \mathbf{G}$*, then* $\hat{t}_{\mathbf{F}}(\mathbf{Z}) \xrightarrow{\mathbb{P}} \theta_{\mathbf{FG}}$ *and* $\hat{t}_{\mathbf{G}}(\mathbf{Z}) \xrightarrow{\mathbb{P}} \theta_{\mathbf{GG}}$ *, as* $d \to \infty$ *.*

Proof of lemma A.4

(a) Since $\hat{\rho}(\mathbf{X_i}, \mathbf{X_j}) \xrightarrow{\mathbb{P}} \theta_{\mathbf{FF}}$,

$$\hat{t}_{\mathbf{FF}} = \frac{1}{m(m-1)} \sum_{i \neq j} \hat{\rho}\left(\mathbf{X}_i, \mathbf{X}_j\right) \xrightarrow{\mathbb{P}} \theta_{\mathbf{FF}} \text{ , as } d \to \infty.$$

Similarly, $\hat{t}_{\mathbf{GG}} \xrightarrow{\mathbb{P}} \theta_{\mathbf{GG}}$ and $\hat{t}_{\mathbf{FG}} \xrightarrow{\mathbb{P}} \theta_{\mathbf{FG}}$, as $d \to \infty$.

(b) If $\mathbf{Z} \sim \mathbf{F}$, $\hat{\rho}(\mathbf{X_i}, \mathbf{Z}) \xrightarrow{\mathbb{P}} \theta_{\mathbf{FF}}$ and $\hat{\rho}(\mathbf{Y_i}, \mathbf{Z}) \xrightarrow{\mathbb{P}} \theta_{\mathbf{FG}}$, as $d \to \infty$. So,

$$\hat{t}_{\mathbf{F}}(\mathbf{Z}) = \frac{1}{n} \sum_{i} \hat{\rho}\left(\mathbf{X}_i, \mathbf{Z}\right) \xrightarrow{\mathbb{P}} \theta_{\mathbf{FF}} \text{ and } \hat{t}_{\mathbf{G}}(\mathbf{Z}) = \frac{1}{n} \sum_{i} \hat{\rho}\left(\mathbf{Y}_i, \mathbf{Z}\right) \xrightarrow{\mathbb{P}} \theta_{\mathbf{FG}},$$

as $d \to \infty$. Similarly, if $\mathbf{Z} \sim \mathbf{G}$, $\hat{t}_{\mathbf{F}}(\mathbf{Z}) \xrightarrow{\mathbb{P}} \theta_{\mathbf{FG}}$ and $\hat{t}_{\mathbf{G}}(\mathbf{Z}) \xrightarrow{\mathbb{P}} \theta_{\mathbf{GG}}$, as $d \to \infty$.

$\square$

Proof of Theorem 4

From lemma A.3, we have $\theta^*_{\mathbf{FG}} = \lim_d \mathcal{W}^*_{\mathbf{FG}} = 2\theta_{\mathbf{FG}} - \theta_{\mathbf{FF}} - \theta_{\mathbf{GG}}$. Given $\mathbf{Z} \sim \mathbf{F}$,

$$\begin{aligned}& \left| l_{\mathbf{G}}(\mathbf{Z}) - l_{\mathbf{F}}(\mathbf{Z}) - \frac{1}{2}\theta^*_{\mathbf{FG}} \right| \\ &= \left| (\hat{t}_{\mathbf{G}}(\mathbf{Z}) - \theta_{\mathbf{FG}}) - (\hat{t}_{\mathbf{F}}(\mathbf{Z}) - \theta_{\mathbf{FF}}) + \frac{1}{2}(\hat{t}_{\mathbf{FF}} - \theta_{\mathbf{FF}}) - \frac{1}{2}(\hat{t}_{\mathbf{GG}} - \theta_{\mathbf{GG}}) \right| \\ &\leq |\hat{t}_{\mathbf{G}}(\mathbf{Z}) - \theta_{\mathbf{FG}}| + |\hat{t}_{\mathbf{F}}(\mathbf{Z}) - \theta_{\mathbf{FF}}| + \frac{1}{2}|\hat{t}_{\mathbf{FF}} - \theta_{\mathbf{FF}}| + \frac{1}{2}|\hat{t}_{\mathbf{GG}} - \theta_{\mathbf{GG}}| \\ &\xrightarrow{\mathbb{P}} 0, \text{ , as } d \to \infty \text{ [By using } \textit{lemma A.4}].\end{aligned}$$

Therefore, if $\mathbf{Z} \sim \mathbf{F}$, $l_{\mathbf{G}}(\mathbf{Z}) - l_{\mathbf{F}}(\mathbf{Z}) \xrightarrow{\mathbb{P}} \frac{1}{2}\theta^*_{\mathbf{FG}}$, as $d \to \infty$.

Part (b) can be shown in an exactly similar way.

□

Lemma A.5. *$2\theta_{\mathbf{FG}} - \theta_{\mathbf{FF}} - \theta_{\mathbf{GG}} = 0$ if and only if $\sigma^2_{\mathbf{G}} = \sigma^2_{\mathbf{F}}$ and $\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} = 0$.*

Proof of lemma A.5

$$\begin{aligned}2\theta_{\mathbf{FG}} - \theta_{\mathbf{FF}} - \theta_{\mathbf{GG}} =& \frac{2}{3} - \frac{1}{\pi}\Bigg[\cos^{-1}\left(\frac{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{G}}}{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{G}} + \sigma^2_{\mathbf{F}}}\right) \\ &+ \cos^{-1}\left(\frac{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma_{\mathbf{F}}}{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} + \sigma^2_{\mathbf{G}} + \sigma^2_{\mathbf{F}}}\right)\Bigg].\end{aligned}$$

Let, $x = \dfrac{\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}}}{\sigma^2_{\mathbf{G}} + \sigma^2_{\mathbf{F}}}$ and $\alpha = \dfrac{\sigma^2_{\mathbf{F}}}{\sigma^2_{\mathbf{G}} + \sigma^2_{\mathbf{F}}}$. Note that $x \geq 0$ and $\alpha \in (0, 1)$.

We need to solve $\cos^{-1}\left(\frac{x+\alpha}{x+1}\right) + \cos^{-1}\left(\frac{x+1-\alpha}{x+1}\right) = \frac{2\pi}{3}$, for $x \geq 0$ and $\alpha \in (0, 1)$, which is equivalent to solve

$$\begin{aligned}&\cos^{-1}\left(\frac{x+\alpha}{x+1} \cdot \frac{x+1-\alpha}{x+1} - \sqrt{\left(1 - \left(\frac{x+\alpha}{x+1}\right)^2\right)\left(1 - \left(\frac{x+1-\alpha}{x+1}\right)^2\right)}\right) = \frac{2\pi}{3} \\ &\implies \frac{x+\alpha}{x+1} \cdot \frac{x+1-\alpha}{x+1} - \sqrt{\left(1 - \left(\frac{x+\alpha}{x+1}\right)^2\right)\left(1 - \left(\frac{x+1-\alpha}{x+1}\right)^2\right)} = -\frac{1}{2} \\ &\implies x^2 + \alpha(1-\alpha) + x - \sqrt{\alpha(1-\alpha)(4x^2 + 2 + 6x + \alpha(1-\alpha))} = -\frac{1}{2}(x+1)^2 \\ &\implies \left(\frac{1}{2}(x+1)(3x+1) + \beta\right)^2 = \beta(4x^2 + 2 + 6x + \beta) \text{ ; substituting, } \beta = \alpha(1-\alpha) \\ &\implies \beta = \frac{(3x+1)^2}{4} \geq \frac{1}{4} \text{ since } x \geq 0.\end{aligned}$$

But, since $\alpha \in (0, 1)$, by AM-GM inequality, $\beta = \alpha(1-\alpha) \leq \frac{1}{4}$, equality holds iff $\alpha = \frac{1}{2}$. So, we must have $\beta = \frac{1}{4}$ which will imply $\alpha = \frac{1}{2}$ and $x = 0$, i.e, $\sigma_G = \sigma_F$ and $\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} = 0$.

Also, $\sigma_{\mathbf{G}} = \sigma_{\mathbf{F}}$ and $\lambda_{\mathbf{FF}} + \lambda_{\mathbf{GG}} - 2\lambda_{\mathbf{FG}} = 0$ implies $2\theta_{\mathbf{FG}} - \theta_{\mathbf{FF}} - \theta_{\mathbf{GG}} = 0$.

□

Proof of Theorem 5

The misclassification probability of the classifier δ_0 can be written as

$$\begin{aligned}\Delta_0 &= \mathbb{P}\left[\delta_0(\mathbf{Z}) = 2, \mathbf{Z} \sim \mathbf{F}\right] + \mathbb{P}\left[\delta_0(\mathbf{Z}) = 1, \mathbf{Z} \sim \mathbf{G}\right] \\ &= \frac{m}{m+n}\mathbb{P}\left[\delta_0(\mathbf{Z}) = 2 \mid \mathbf{Z} \sim \mathbf{F}\right] + \frac{n}{m+n}\mathbb{P}\left[\delta_0(\mathbf{Z}) = 1 \mid \mathbf{Z} \sim \mathbf{G}\right] \\ &= \frac{m}{m+n}\mathbb{P}\left[l_{\mathbf{G}}(\mathbf{Z}) - l_{\mathbf{F}}(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \frac{n}{m+n}\mathbb{P}\left[l_{\mathbf{G}}(\mathbf{Z}) - l_{\mathbf{F}}(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right].\end{aligned}$$

By lemma A.5, we know that if either $\lim_{d\to\infty} \frac{1}{d}\|\boldsymbol{\mu}_{\mathbf{F}} - \boldsymbol{\mu}_{\mathbf{G}}\|^2 \neq 0$ or $\sigma^2_{\mathbf{F}} \neq \sigma^2_{\mathbf{G}}$ holds, $2\theta_{\mathbf{FG}} - \theta_{\mathbf{FF}} - \theta_{\mathbf{GG}} > 0$. We can choose $\epsilon > 0$ such that $\epsilon < \frac{1}{2}(2\theta_{\mathbf{FG}} - \theta_{\mathbf{FF}} - \theta_{\mathbf{GG}})$. Therefore, we have:

$$\begin{aligned}
&\mathbb{P}\left[l_{\mathbf{G}}(\mathbf{Z}) - l_{\mathbf{F}}(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\leq \mathbb{P}\left[l_{\mathbf{G}}(\mathbf{Z}) - l_{\mathbf{F}}(\mathbf{Z}) \leq \frac{1}{2}(2\theta_{\mathbf{FG}} - \theta_{\mathbf{FF}} - \theta_{\mathbf{GG}}) - \epsilon \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\leq \mathbb{P}\left[l_{\mathbf{G}}(\mathbf{Z}) - l_{\mathbf{F}}(\mathbf{Z}) - \frac{1}{2}(2\theta_{\mathbf{FG}} - \theta_{\mathbf{FF}} - \theta_{\mathbf{GG}}) \leq -\epsilon \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\leq \mathbb{P}\left[\left|l_{\mathbf{G}}(\mathbf{Z}) - l_{\mathbf{F}}(\mathbf{Z}) - \frac{1}{2}(2\theta_{\mathbf{FG}} - \theta_{\mathbf{FF}} - \theta_{\mathbf{GG}})\right| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}\right] \quad \to 0 \text{ , as } d \to \infty \text{ .} \\
&\hfill \text{[Using Theorem 4]}
\end{aligned}$$

Similarly, one can show that

$$\mathbb{P}\left[l_{\mathbf{G}}(\mathbf{Z}) - l_{\mathbf{F}}(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right] \to 0 \text{ , as } d \to \infty.$$

Thus, we conclude that $\Delta_0 \to 0$ as $d \to \infty$.

□

Lemma A.6. *It follows from assumption 6 that*

$$\sum_{1\leq k_1<k_2\leq d} \operatorname{cov}(\hat{\rho}(U_{k_1}, V_{k_1}), \hat{\rho}(U_{k_2}, V_{k_2})) = o(d^2).$$

Proof of lemma A.6 Let, $\mathcal{X} = \{\mathbf{X}_1, \cdots, \mathbf{X}_m\}$ and $\mathcal{Y} = \{\mathbf{Y}_1, \cdots, \mathbf{Y}_n\}$. Then,

$$\begin{aligned}
&\operatorname{cov}(\hat{\rho}(U_{k_1}, V_{k_1}), \hat{\rho}(U_{k_2}, V_{k_2})) \\
&= \frac{1}{(m+n)^2} \sum_{\mathbf{Q},\mathbf{Q}^*\in\mathcal{X}\cup\mathcal{Y}} \operatorname{cov}(\rho_0(U_{k_1}, V_{k_1}; Q_{k_1}), \rho_0(U_{k_2}, V_{k_2}; Q^*_{k_2})).
\end{aligned}$$

Hence, it follows that

$$\begin{aligned}
&\sum_{1\leq k_1<k_2\leq d} \operatorname{cov}(\hat{\rho}(U_{k_1}, V_{k_1}), \hat{\rho}(U_{k_2}, V_{k_2})) \\
&= \frac{1}{(m+n)^2} \sum_{1\leq k_1<k_2\leq d} \sum_{\mathbf{Q},\mathbf{Q}^*\in\mathcal{X}\cup\mathcal{Y}} \operatorname{cov}(\rho_0(U_{k_1}, V_{k_1}; Q_{k_1}), \rho_0(U_{k_2}, V_{k_2}; Q^*_{k_2})) \\
&= \frac{1}{(m+n)^2} \sum_{\mathbf{Q},\mathbf{Q}^*\in\mathcal{X}\cup\mathcal{Y}} \sum_{1\leq k_1<k_2\leq d} \operatorname{cov}(\rho_0(U_{k_1}, V_{k_1}; Q_{k_1}), \rho_0(U_{k_2}, V_{k_2}; Q^*_{k_2})) \\
&\hfill \text{[Since the sum over } \mathbf{Q}, \mathbf{Q}^* \text{ is finite sum.]} \\
&= o(d^2) \text{ [By using assumption 6].}
\end{aligned}$$

□

We now define

$$T_{\mathbf{FG}} = \mathbb{E}\left[\bar{\rho}\left(\mathbf{X}_1, \mathbf{Y}_1\right)\right] \quad \text{and} \quad \hat{T}_{\mathbf{FG}} = \frac{1}{nm} \sum_{i,j} \hat{\bar{\rho}}\left(\mathbf{X}_i, \mathbf{Y}_j\right).$$

Lemma A.7. *Suppose assumption 6 is satisfied. Then,*

(a) Irrespective of whether $\mathbf{U}, \mathbf{V}$ *are coming from* $\mathbf{F}$ *and/or* $\mathbf{G}$*,*

$$\hat{\rho}(\mathbf{U}, \mathbf{V}) - \mathbb{E}[\hat{\rho}(\mathbf{U}, \mathbf{V})] \xrightarrow{\mathbb{P}} 0 \ , \ as\ d \to \infty$$

(b) i. $\hat{T}_{\mathbf{FF}} - T_{\mathbf{FF}} \xrightarrow{\mathbb{P}} 0$ *, as* $d \to \infty$ *;*
ii. $\hat{T}_{\mathbf{GG}} - T_{\mathbf{GG}} \xrightarrow{\mathbb{P}} 0$ *, as* $d \to \infty$ *;*
iii. $\hat{T}_{\mathbf{FG}} - T_{\mathbf{FG}} \xrightarrow{\mathbb{P}} 0$ *, as* $d \to \infty$ *.*

(c) i. If $\mathbf{Z} \sim \mathbf{F}$*, then as* $d \to \infty$*,* $\hat{T}_{\mathbf{F}}(\mathbf{Z}) - T_{\mathbf{FF}} \xrightarrow{\mathbb{P}} 0$ *and* $\hat{T}_{\mathbf{G}}(\mathbf{Z}) - T_{\mathbf{FG}} \xrightarrow{\mathbb{P}} 0$ *;*
ii. If $\mathbf{Z} \sim \mathbf{G}$*, then as* $d \to \infty$*,* $\hat{T}_{\mathbf{F}}(\mathbf{Z}) - T_{\mathbf{FG}} \xrightarrow{\mathbb{P}} 0$ *and* $\hat{T}_{\mathbf{G}}(\mathbf{Z}) - T_{\mathbf{GG}} \xrightarrow{\mathbb{P}} 0$ *.*

Proof of lemma A.7

(a) For any $\epsilon > 0$, using *Chebyshev's Inequality*, we have,

$$\mathbb{P}[|\hat{\hat{\rho}}(\mathbf{U}, \mathbf{V}) - \mathbb{E}[\hat{\hat{\rho}}(\mathbf{U}, \mathbf{V})]| > \epsilon] \leq \frac{\mathrm{var}(\hat{\hat{\rho}}(\mathbf{U}, \mathbf{V}))}{\epsilon^2}.$$

Note that $\rho_0 \in [-1, 1]$. Thus, $\hat{\rho}$ being a convex combination of some ρ_0, lies between $[-1, 1]$. Also, since $\rho(a, b) = \mathbb{E}[\rho_0(a, b)]$, we have $\hat{\rho} \in [-1, 1]$. So, $\mathbb{E}(\hat{\rho}^2(U_i, V_i)) \leq 1$.

$$\begin{aligned}
&\mathbb{P}[|\hat{\hat{\rho}}(\mathbf{U}, \mathbf{V}) - \mathbb{E}[\hat{\hat{\rho}}(\mathbf{U}, \mathbf{V})]| > \epsilon] \\
&\leq \frac{1}{\epsilon^2} \mathrm{var}(\frac{1}{d} \sum_{i=1}^{d} \hat{\rho}(U_i, V_i)) \\
&= \frac{1}{\epsilon^2 d^2} \left[\sum_{i=1}^{d} \mathrm{var}(\hat{\rho}(U_i, V_i)) + \sum_{1 \leq i < j \leq d} \mathrm{cov}(\hat{\rho}(U_i, V_i), \hat{\rho}(U_j, V_j)) \right] \\
&\leqslant \frac{1}{\epsilon^2 d^2} \left[\sum_{i=1}^{d} \mathbb{E}(\hat{\rho}^2(U_i, V_i)) + \sum_{1 \leq i < j \leq d} \mathrm{cov}(\hat{\rho}(U_i, V_i), \hat{\rho}(U_j, V_j)) \right] \\
&\leqslant \frac{1}{\epsilon^2 d^2} \left[d + o(d^2) \right] \ \to 0 \ \ , \text{ as } d \to \infty.
\end{aligned}$$

The last assertion holds good due to lemma A.6 and the fact that $\mathbb{E}(\hat{\rho}^2(U_i, V_i)) \leq 1$. Therefore, we have:

$$|\hat{\hat{\rho}}(\mathbf{U}, \mathbf{V}) - \mathbb{E}[\hat{\hat{\rho}}(\mathbf{U}, \mathbf{V})]| \xrightarrow{\mathbb{P}} 0, \text{ as } d \to \infty.$$

(b) Once we have proved part **(a)**, we have the following:

$$
\begin{aligned}
&\hat{T}_{\mathbf{FF}} - T_{\mathbf{FF}} \\
&= \frac{1}{m(m-1)} \sum_{i \neq j} \left[\hat{\bar{\rho}}\,(\mathbf{X}_i, \mathbf{X}_j) - \mathbb{E}[\hat{\bar{\rho}}\,(\mathbf{X}_i, \mathbf{X}_j)] + \mathbb{E}[\bar{\rho}\,(\mathbf{X}_i, \mathbf{X}_j)] - \mathbb{E}\left[\hat{\bar{\rho}}\,(\mathbf{X}_i, \mathbf{X}_j)\right] \right] \\
&= \frac{1}{m(m-1)} \sum_{i \neq j} \left[\hat{\bar{\rho}}\,(\mathbf{X}_i, \mathbf{X}_j) - \mathbb{E}[\hat{\bar{\rho}}\,(\mathbf{X}_i, \mathbf{X}_j)] \right] \\
&\qquad + \frac{1}{m(m-1)} \sum_{i \neq j} \left[\mathbb{E}[\hat{\bar{\rho}}\,(\mathbf{X}_i, \mathbf{X}_j)] - \mathbb{E}\left[\bar{\rho}\,(\mathbf{X}_i, \mathbf{X}_j)\right] \right].
\end{aligned}
$$

We can show that as $d \to \infty$, the first summand goes to 0 in probability using part (a). Now,

$$
\begin{aligned}
&\sum_{i \neq j} \left[\mathbb{E}[\hat{\bar{\rho}}\,(\mathbf{X}_i, \mathbf{X}_j)] - \mathbb{E}\left[\bar{\rho}\,(\mathbf{X}_i, \mathbf{X}_j)\right] \right] \\
&= \frac{1}{d} \sum_{k=1}^{d} \left[\mathbb{E}[\hat{\rho}\,(X_{1k}, X_{2k})] - \mathbb{E}\left[\rho\,(X_{1k}, X_{2k})\right] \right] \\
&= \frac{1}{d} \sum_{k=1}^{d} \left[\frac{m}{m+n} \mathbb{E}\left[\rho_0(X_{1k}, X_{2k}; X_{3k})\right] + \frac{n}{m+n} \mathbb{E}\left[\rho_0(X_{1k}, X_{2k}; Y_{3k})\right] \right. \\
&\qquad \left. - \mathbb{E}\left[\rho\,(X_{1k}, X_{2k})\right] \right] \\
&= \frac{1}{d} \sum_{k=1}^{d} \left[\frac{m}{m+n} \mathbb{E}\left[\rho_0(X_{1k}, X_{2k}; X_{3k})\right] \right. \\
&\qquad \left. + \frac{n}{m+n} \mathbb{E}\left[\rho_0(X_{1k}, X_{2k}; Y_{3k})\right] - \mathbb{E}\left[E_{\mathbf{Q}}[\rho_0\,(X_{1k}, X_{2k}; Q_k)]\right] \right] \\
&\qquad\qquad\qquad\qquad \text{where } \mathbf{Q} \sim \frac{m}{m+n}\mathbf{F} + \frac{n}{m+n}\mathbf{G} \\
&= \frac{1}{d} \sum_{k=1}^{d} \left[\frac{m}{m+n} \mathbb{E}\left[\rho_0(X_{1k}, X_{2k}; X_{3k})\right] + \frac{n}{m+n} \mathbb{E}\left[\rho_0(X_{1k}, X_{2k}; Y_{3k})\right] \right. \\
&\qquad \left. - \mathbb{E}\left[\rho_0\,(X_{1k}, X_{2k}; Q_k)\right] \right] \\
&= \frac{1}{d} \sum_{k=1}^{d} \left[\frac{m}{m+n} \mathbb{E}\left[\rho_0(X_{1k}, X_{2k}; X_{3k})\right] + \frac{n}{m+n} \mathbb{E}\left[\rho_0(X_{1k}, X_{2k}; Y_{3k})\right] \right. \\
&\qquad\qquad \left. - \frac{m}{m+n} \mathbb{E}\left[\rho\,(X_{1k}, X_{2k}; Q_k)\,|\mathbf{Q} \sim \mathbf{F}\right] - \frac{n}{m+n} \mathbb{E}\left[\rho\,(X_{1k}, X_{2k}; Q_k)\,|\mathbf{Q} \sim \mathbf{G}\right] \right] \\
&= 0.
\end{aligned}
$$

ii. The proof is the same as above.

iii. The proof is the same as above.

(c) Once again, we shall use the result from part **(a)**.

i. If $\mathbf{Z} \sim \mathbf{F}$, in that case, $T_{\mathbf{FF}} = \mathbb{E}\left[\bar{\rho}\left(\mathbf{X}_1, \mathbf{X}_2\right)\right] = \mathbb{E}[\bar{\rho}\left(\mathbf{X_i}, \mathbf{Z}\right)]$. So,

$$
\begin{aligned}
&\hat{T}_{\mathbf{F}}(\mathbf{Z}) - T_{\mathbf{FF}} \\
&= \frac{1}{m}\sum_i (\hat{\bar{\rho}}\left(\mathbf{X_i}, \mathbf{Z}\right) - \mathbb{E}[\bar{\rho}\left(\mathbf{X_i}, \mathbf{Z}\right)]) \\
&= \frac{1}{m}\sum_i \left[(\hat{\bar{\rho}}\left(\mathbf{X_i}, \mathbf{Z}\right) - \mathbb{E}[\hat{\bar{\rho}}\left(\mathbf{X_i}, \mathbf{Z}\right)]) + (\mathbb{E}[\hat{\bar{\rho}}\left(\mathbf{X_i}, \mathbf{Z}\right)] - \mathbb{E}[\bar{\rho}\left(\mathbf{X_i}, \mathbf{Z}\right)])\right].
\end{aligned}
$$

$$
\begin{aligned}
&\frac{1}{m}\sum_i (\mathbb{E}[\hat{\bar{\rho}}\left(\mathbf{X_i}, \mathbf{Z}\right)] - \mathbb{E}[\bar{\rho}\left(\mathbf{X_i}, \mathbf{Z}\right)]) \\
&= \mathbb{E}[\hat{\bar{\rho}}\left(\mathbf{X_1}, \mathbf{Z}\right)] - \mathbb{E}[\bar{\rho}\left(\mathbf{X_1}, \mathbf{Z}\right)] \\
&= \frac{1}{d}\sum_{k=1}^{d}\Bigg[\mathbb{E}[\hat{\rho}\left(\mathbf{X_1}, \mathbf{Z}\right)] - \mathbb{E}[\rho\left(\mathbf{X_1}, \mathbf{Z}\right)]\Bigg] \\
&= \frac{1}{d}\sum_{k=1}^{d}\Bigg[\frac{m}{m+n}\mathbb{E}\left[\rho_0(X_{1k}, Z_k; X_{2k})\right] + \frac{n}{m+n}\mathbb{E}\left[\rho_0(X_{1k}, Z_k; Y_{2k})\right] \\
&\qquad - \mathbb{E}\left[\rho\left(X_{1k}, Z_k\right)\right]\Bigg] \\
&= \frac{1}{d}\sum_{k=1}^{d}\Bigg[\frac{m}{m+n}\mathbb{E}\left[\rho_0(X_{1k}, Z_k; X_{2k})\right] + \frac{n}{m+n}\mathbb{E}\left[\rho_0(X_{1k}, Z_k; Y_{2k})\right] - \\
&\qquad \mathbb{E}\left[E_{Q_k}\left[\rho_0\left(X_{1k}, Z_k; Q_k\right)\right]\right]\Bigg], \text{ where } \mathbf{Q} \sim \frac{m}{m+n}\mathbf{F} + \frac{n}{m+n}\mathbf{G} \\
&= \frac{1}{d}\sum_{k=1}^{d}\Bigg[\frac{m}{m+n}\mathbb{E}\left[\rho_0(X_{1k}, Z_k; X_{2k})\right] + \frac{n}{m+n}\mathbb{E}\left[\rho_0(X_{1k}, Z_k; Y_{2k})\right] \\
&\qquad - \mathbb{E}\left[\rho_0\left(X_{1k}, Z_k; Q_k\right)\right]\Bigg] \\
&= \frac{1}{d}\sum_{k=1}^{d}\Bigg[\frac{m}{m+n}\mathbb{E}\left[\rho_0(X_{1k}, Z_k; X_{2k})\right] + \frac{n}{m+n}\mathbb{E}\left[\rho_0(X_{1k}, Z_k; Y_{2k})\right] \\
&\qquad - \frac{m}{m+n}\mathbb{E}\left[\rho\left(X_{1k}, Z_k; Q_k\right)|\mathbf{Q} \sim \mathbf{F}\right] - \frac{n}{m+n}\mathbb{E}\left[\rho\left(X_{1k}, Z_k; Q_k\right)|\mathbf{Q} \sim \mathbf{G}\right]\Bigg] \\
&= 0 \quad \square
\end{aligned}
$$

Proof of Theorem 7

We shall prove part (a) only, and the next part will follow analogously.

$$
\begin{aligned}
&|\mathscr{D}_1(\mathbf{Z}) - \frac{1}{2}\overline{\mathcal{W}}^*_{\mathbf{FG}}| \\
&= \left|(\hat{T}_{\mathbf{G}}(\mathbf{Z}) - T_{\mathbf{FG}}) - (\hat{T}_{\mathbf{F}}(\mathbf{Z}) - T_{\mathbf{FF}}) - \frac{1}{2}(\hat{T}_{\mathbf{GG}} - T_{\mathbf{GG}}) + \frac{1}{2}(\hat{T}_{\mathbf{FF}} - T_{\mathbf{FF}})\right| \\
&\leq |\hat{T}_{\mathbf{G}}(\mathbf{Z}) - T_{\mathbf{FG}}| + |\hat{T}_{\mathbf{F}}(\mathbf{Z}) - T_{\mathbf{FF}}| + \frac{1}{2}|\hat{T}_{\mathbf{GG}} - T_{\mathbf{GG}}| + \frac{1}{2}|\hat{T}_{\mathbf{FF}} - T_{\mathbf{FF}}| \\
&\xrightarrow{\mathbb{P}} 0, \text{ as } d \to \infty,\ \text{given } \mathbf{Z} \sim \mathbf{F}.
\end{aligned}
$$

Here, the last assertion follows from lemma A.7.Now,

$$
\begin{aligned}
&\left|S(\mathbf{Z}) - \frac{1}{2}\left(T_{\mathbf{FF}} - T_{\mathbf{GG}}\right)\right| \\
&= \left|\left(\hat{T}_{\mathbf{F}}(\mathbf{Z}) - \hat{T}_{\mathbf{FF}}\right) + \left(\hat{T}_{\mathbf{G}}(\mathbf{Z}) - \hat{T}_{\mathbf{FG}}\right) + \frac{1}{2}\left(\hat{T}_{\mathbf{FF}} - T_{\mathbf{FF}}\right) - \frac{1}{2}\left(\hat{T}_{\mathbf{GG}} - T_{\mathbf{GG}}\right)\right| \\
&\leq |\hat{T}_{\mathbf{F}}(\mathbf{Z}) - \hat{T}_{\mathbf{FF}}| + |\hat{T}_{\mathbf{G}}(\mathbf{Z}) - \hat{T}_{\mathbf{FG}}| + \frac{1}{2}|\hat{T}_{\mathbf{FF}} - T_{\mathbf{FF}}| + \frac{1}{2}|\hat{T}_{\mathbf{GG}} - T_{\mathbf{GG}}|.
\end{aligned}
$$

From here, it follows from lemma A.7 that

$$
\left|S(\mathbf{Z}) - \frac{1}{2}\left(T_{\mathbf{FF}} - T_{\mathbf{GG}}\right)\right| \xrightarrow{\mathbb{P}} 0, \text{ as } d \to \infty,\ \text{given } \mathbf{Z} \sim \mathbf{F}. \tag{1}
$$

Also, from theorem 1, given $\mathbf{Z} \sim \mathbf{F}$,

$$
\left|\mathscr{D}_1(\mathbf{Z}) - \frac{1}{2}\overline{\mathcal{W}}^*_{\mathbf{FG}}\right| \xrightarrow{\mathbb{P}} 0, \text{ as } d \to \infty.
$$

Next, we have:

$$
\begin{aligned}
|\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} - \overline{\mathcal{W}}^*_{\mathbf{FG}}| &= \left|2\left(\hat{T}_{\mathbf{FG}} - T_{\mathbf{FG}}\right) - \left(\hat{T}_{\mathbf{FF}} - T_{\mathbf{FF}}\right) - \left(\hat{T}_{\mathbf{GG}} - T_{\mathbf{GG}}\right)\right| \\
&\leq 2|\hat{T}_{\mathbf{FG}} - T_{\mathbf{FG}}| + |\hat{T}_{\mathbf{FF}} - T_{\mathbf{FF}}| + |\hat{T}_{\mathbf{GG}} - T_{\mathbf{GG}}| \\
&\xrightarrow{\mathbb{P}} 0, \text{ as } d \to \infty.
\end{aligned}
$$

And also:

$$
\begin{aligned}
\left|S_{\mathbf{FG}} - \left(T_{\mathbf{FF}} - T_{\mathbf{GG}}\right)\right| &= \left|\left(\hat{T}_{\mathbf{FF}} - T_{\mathbf{FF}}\right) - \left(\hat{T}_{\mathbf{GG}} - T_{\mathbf{GG}}\right)\right| \\
&\leq |\hat{T}_{\mathbf{FF}} - T_{\mathbf{FF}}| + |\hat{T}_{\mathbf{GG}} - T_{\mathbf{GG}}| \xrightarrow{\mathbb{P}} 0, \text{ as } d \to \infty.
\end{aligned}
$$

Combining the 4 results stated above, we can see that given $\mathbf{Z} \sim \mathbf{F}$,

$$
\left|\frac{1}{2}\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \frac{1}{2}S_{\mathbf{FG}} \cdot S(\mathbf{Z}) - \bar{\tau}_{\mathbf{FG}}\right| \xrightarrow{\mathbb{P}} 0, \text{ as } d \to \infty.
$$

□

Proof of Theorem 10

We shall prove part (a) only, and the next part will follow analogously. We have already shown that given $\mathbf{Z} \sim \mathbf{F}$,

$$\left|\mathscr{D}_1(\mathbf{Z}) - \frac{1}{2}\overline{\mathcal{W}}^*_{\mathbf{FG}}\right| \xrightarrow{\mathbb{P}} 0, \text{ as } d \to \infty.$$

Let, $\liminf_d \overline{\mathcal{W}}^*_{\mathbf{FG}} = \delta$. By $\delta > 0$. So, there exists some $d_0 \in \mathbb{N}$, such that $\overline{\mathcal{W}}^*_{\mathbf{FG}} > \frac{\delta}{2}$, for all $d \geq d_0$.
We have, $\mathbb{P}\left[\left|\mathscr{D}_1(\mathbf{Z}) - \frac{1}{2}\overline{\mathcal{W}}^*_{\mathbf{FG}}\right| < \frac{\delta}{4} \mid \mathbf{Z} \sim \mathbf{F}\right] \to 1$, as $d \to \infty$.
For all $d \geq d_0$,

$$\begin{aligned}\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{F}\right] &\geq \mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) > \frac{1}{2}\overline{\mathcal{W}}^*_{\mathbf{FG}} - \frac{\delta}{4} \mid \mathbf{Z} \sim \mathbf{F}\right] \\ &\geq \mathbb{P}\left[|\mathscr{D}_1(\mathbf{Z}) - \frac{1}{2}\overline{\mathcal{W}}^*_{\mathbf{FG}}| < \frac{\delta}{4} \mid \mathbf{Z} \sim \mathbf{F}\right] \\ \implies \mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{F}\right] &\to 1, \text{ as } d \to \infty.\end{aligned}$$

Given $\mathbf{Z} \sim \mathbf{F}$,

$$\operatorname{sign}\left(\mathscr{D}_1(\mathbf{Z})\right) \xrightarrow{\mathbb{P}} 1 \text{ , as } d \to \infty$$

Also, as we have seen before:

$$\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} - \overline{\mathcal{W}}^*_{\mathbf{FG}} = 2\left(\hat{T}_{\mathbf{FG}} - T_{\mathbf{FG}}\right) - \left(\hat{T}_{\mathbf{FF}} - T_{\mathbf{FF}}\right) - \left(\hat{T}_{\mathbf{GG}} - T_{\mathbf{GG}}\right) \xrightarrow{\mathbb{P}} 0 \text{ , as } d \to \infty.$$

Hence, we have

$$\frac{1}{2}\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign}\left(\mathscr{D}_1(\mathbf{Z})\right) - \frac{1}{2}\overline{\mathcal{W}}^*_{\mathbf{FG}} \xrightarrow{\mathbb{P}} 0 \text{ , as } d \to \infty. \tag{2}$$

Now, assumption 9 implies, for any fixed $\epsilon > 0$, we can choose some d_1, such that $|T_{\mathbf{FF}} - T_{\mathbf{GG}}| > \epsilon$, $\forall d \geq d_1$. From (1),

$$\mathbb{P}\left[\ \left|S(\mathbf{Z}) - \frac{1}{2}\left(T_{\mathbf{FF}} - T_{\mathbf{GG}}\right)\right| \leq \epsilon/4 \mid \mathbf{Z} \sim \mathbf{F}\right] \to 1. \tag{3}$$

So, if $\operatorname{sign}(T_{\mathbf{FF}} - T_{\mathbf{GG}}) = 1$, we have, $T_{\mathbf{FF}} - T_{\mathbf{GG}} > \epsilon$, $\forall d \geq d_0$. Then, $\left|S(\mathbf{Z}) - \frac{1}{2}\left(T_{\mathbf{FF}} - T_{\mathbf{GG}}\right)\right| \leq \epsilon/4$ implies that $S(\mathbf{Z}) > \frac{3\epsilon}{4}$ and so, $\operatorname{sign}(S(\mathbf{Z})) = 1$
Similarly, if $\operatorname{sign}(T_{\mathbf{FF}} - T_{\mathbf{GG}}) = -1$, we can show that $\left|S(\mathbf{Z}) - \frac{1}{2}\left(T_{\mathbf{FF}} - T_{\mathbf{GG}}\right)\right| \leq \epsilon/4$ implies $\operatorname{sign}(S(\mathbf{Z})) = -1$.

Hence, for all $d \geq d_1$,

$$\mathbb{P}\Big[\operatorname{sign}(S(\mathbf{Z})) - \operatorname{sign}(T_{\mathbf{FF}} - T_{\mathbf{GG}}) = 0 \mid \mathbf{Z} \sim \mathbf{F}\Big] \geq \mathbb{P}\left[\ \left|S(\mathbf{Z}) - \frac{1}{2}\left(T_{\mathbf{FF}} - T_{\mathbf{GG}}\right)\right| \leq \epsilon/4 \mid \mathbf{Z} \sim \mathbf{F}\right]. \tag{4}$$

From (3) and (4) we get, $\mathbb{P}\left[\text{sign}(S(\mathbf{Z})) - \text{sign}(T_{\mathbf{FF}} - T_{\mathbf{GG}}) = 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \to 1$.

$$\implies \text{sign}(S(\mathbf{Z})) - \text{sign}(T_{\mathbf{FF}} - T_{\mathbf{GG}}) \xrightarrow{\mathbb{P}} 0.$$

And also,

$$\hat{S}_{\mathbf{FG}} - (T_{\mathbf{FF}} - T_{\mathbf{GG}}) = \left(\hat{T}_{\mathbf{FF}} - T_{\mathbf{FF}}\right) - \left(\hat{T}_{\mathbf{GG}} - T_{\mathbf{GG}}\right) \xrightarrow{\mathbb{P}} 0 \text{ , as } d \to \infty.$$

Hence, given $\mathbf{Z} \sim \mathbf{F}$,

$$\frac{1}{2}\hat{S}_{\mathbf{FG}}\,\text{sign}(S(\mathbf{Z})) - \frac{1}{2}\left|T_{\mathbf{FF}} - T_{\mathbf{GG}}\right| \xrightarrow{\mathbb{P}} 0 \text{ , as } d \to \infty. \tag{5}$$

From (2) and (5), we obtain that given $\mathbf{Z} \sim \mathbf{F}$,

$$\left|\frac{1}{2}\hat{\overline{\mathcal{W}}}^{*}_{\mathbf{FG}}\,\text{sign}\left(\mathscr{D}_1(\mathbf{Z})\right) + \frac{1}{2}\hat{S}_{\mathbf{FG}}\,\text{sign}(S(\mathbf{Z})) - \bar{\psi}_{\mathbf{FG}}\right| \xrightarrow{\mathbb{P}} 0 \text{ , as } d \to \infty.$$

□

Proof of Theorem 11

(1) The misclassification probability of the classifier δ_1 can be written as

$$\begin{aligned}
\Delta_1 &= \mathbb{P}\left[\delta_1(\mathbf{Z}) = 2, \mathbf{Z} \sim \mathbf{F}\right] + \mathbb{P}\left[\delta_1(\mathbf{Z}) = 1, \mathbf{Z} \sim \mathbf{G}\right] \\
&= \frac{m}{m+n}\mathbb{P}\left[\delta_1(\mathbf{Z}) = 2 \mid \mathbf{Z} \sim \mathbf{F}\right] + \frac{n}{m+n}\mathbb{P}\left[\delta_1(\mathbf{Z}) = 1 \mid \mathbf{Z} \sim \mathbf{G}\right] \\
&= \frac{m}{m+n}\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \le 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \frac{n}{m+n}\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right].
\end{aligned}$$

Since $\liminf_d \overline{\mathcal{W}}^{*}_d > 0$ (assumption 8), we can choose $\epsilon > 0$ such that $\epsilon < \overline{\mathcal{W}}^{*}$ for all $d \ge d_0$ for some $d_0 \in \mathbb{N}$. Therefore, we have:

$$\begin{aligned}
\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \le 0 \mid \mathbf{Z} \sim \mathbf{F}\right] &\le \mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \le \overline{\mathcal{W}}^{*}_d - \epsilon \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\le \mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) - \overline{\mathcal{W}}^{*}_d \le -\epsilon \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\le \mathbb{P}\left[\left|\mathscr{D}_1(\mathbf{Z}) - \overline{\mathcal{W}}^{*}_d\right| > \epsilon \mid \mathbf{Z} \sim \mathbf{F}\right] \quad \to 0 \text{ , as } d \to \infty.
\end{aligned}$$

Similarly, one can show that

$$\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right] \to 0 \text{ , as } d \to \infty.$$

Thus, we conclude that $\Delta_1 \to 0$ as $d \to \infty$.

(2) The misclassification probability of the classifier δ_2 can be written as

$$\begin{aligned}
\Delta_2 &= \mathbb{P}\left[\delta_2(\mathbf{Z}) = 2, \mathbf{Z} \sim \mathbf{F}\right] + \mathbb{P}\left[\delta_2(\mathbf{Z}) = 1, \mathbf{Z} \sim \mathbf{G}\right] \\
&= \frac{m}{m+n}\mathbb{P}\left[\delta_2(\mathbf{Z}) = 2 \mid \mathbf{Z} \sim \mathbf{F}\right] + \frac{n}{m+n}\mathbb{P}\left[\delta_2(\mathbf{Z}) = 1 \mid \mathbf{Z} \sim \mathbf{G}\right] \\
&= \frac{m}{m+n}\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^{*}_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \cdot S(\mathbf{Z}) \le 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad + \frac{n}{m+n}\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^{*}_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \cdot S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right].
\end{aligned}$$

Assumption 8 implies, we have:

$$\liminf_d \bar{\tau} = \liminf_d \left[\frac{1}{2}\overline{\mathcal{W}}^{*2}_{\mathbf{FG}} + \frac{1}{2}\left(T_{\mathbf{FF}} - T_{\mathbf{GG}}\right)^2\right] \geq \frac{1}{2}\liminf_d \overline{\mathcal{W}}^{*2}_{\mathbf{FG}} > 0.$$

From here onward, we can proceed exactly similar to the proof of part **(1)**, and we can show that $\Delta_2 \to 0$ as $d \to \infty$.

(3) The misclassification probability of the classifier δ_3 can be written as

$$\begin{aligned}
\Delta_3 &= \mathbb{P}\left[\delta_3(\mathbf{Z}) = 2, \mathbf{Z} \sim \mathbf{F}\right] + \mathbb{P}\left[\delta_3(\mathbf{Z}) = 1, \mathbf{Z} \sim \mathbf{G}\right] \\
&= \frac{m}{m+n}\mathbb{P}\left[\delta_3(\mathbf{Z}) = 2 \mid \mathbf{Z} \sim \mathbf{F}\right] + \frac{n}{m+n}\mathbb{P}\left[\delta_3(\mathbf{Z}) = 1 \mid \mathbf{Z} \sim \mathbf{G}\right]. \\
&= \frac{m}{m+n}\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^{*}_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad + \frac{n}{m+n}\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^{*}_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right].
\end{aligned}$$

Assumption 8 implies, we have:

$$\liminf_d \bar{\psi} = \liminf_d \left[\frac{1}{2}\overline{\mathcal{W}}^{*}_{\mathbf{FG}} + \frac{1}{2}|T_{\mathbf{FF}} - T_{\mathbf{GG}}|\right] \geq \frac{1}{2}\liminf_d \overline{\mathcal{W}}^{*}_{\mathbf{FG}} > 0.$$

The argument for rest of the proof is similar to what has been shown for part **(1)**. Finally, we shall conclude that $\Delta_3 \to 0$ as $d \to \infty$.

□

Lemma A.8. *Suppose assumptions 6 and 8 are satisfied.*

(a) *If* $\liminf_d \left(\max\{T_{\mathbf{FF}}, T_{\mathbf{GG}}\} - T_{\mathbf{FG}}\right) > 0$ *there exists* $d_0' \in \mathbb{N}$*, such that* $\Delta_2 \leq \Delta_1$ *for all* $d \geq d_0'$.

(b) *If* $\liminf_d \left(T_{\mathbf{FG}} - \max\{T_{\mathbf{FF}}, T_{\mathbf{FG}}\}\right) > 0$*, there exists* $d_0' \in \mathbb{N}$*, such that* $\Delta_2 \geq \Delta_1$ *for all* $d \geq d_0'$.

Proof of lemma A.8

(a) Since, $\overline{\mathcal{W}}^{*}_{\mathbf{FG}} - \hat{\overline{\mathcal{W}}}^{*}_{\mathbf{FG}} \xrightarrow{\mathbb{P}} 0$ as $d \to \infty$. Therefore, for any $\epsilon_0 > 0$ and $\delta > 0$, there exists a d_0 such that for all $d \geq d_0$,

$$\begin{aligned}
&\mathbb{P}\left[\left|\hat{\overline{\mathcal{W}}}^{*}_{\mathbf{FG}} - \overline{\mathcal{W}}^{*}_{\mathbf{FG}}\right| > \delta\right] < \epsilon_0 \\
\Rightarrow &\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^{*}_{\mathbf{FG}} - \overline{\mathcal{W}}^{*}_{\mathbf{FG}} < -\delta\right] < \epsilon_0 \\
\Rightarrow &\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^{*}_{\mathbf{FG}} < \overline{\mathcal{W}}^{*}_{\mathbf{FG}} - \delta\right] < \epsilon_0.
\end{aligned}$$

Using assumption 8, we have $\lambda_0 = \liminf_d \overline{\mathcal{W}}^{*}_{\mathbf{FG}} > 0$. Hence, for any $0 < \delta < \lambda_0$,

$$\mathbb{P}[\hat{\overline{\mathcal{W}}}^{*}_{\mathbf{FG}} < 0] < \epsilon_0, \text{ for all } d \geq d_0. \tag{6}$$

WLOG, we assume that $T_{\mathbf{FF}} > T_{\mathbf{GG}}$. Since, $\liminf_d \left(\max\{T_{\mathbf{FF}}, T_{\mathbf{GG}}\} - T_{\mathbf{FG}}\right) > 0$, there exists $d' \in \mathbb{N}$ such that $T_{\mathbf{FF}} > T_{\mathbf{FG}}$ for all $d \geq d^{'}$. Also, $\overline{\mathcal{W}}^*_{\mathbf{FG}} = 2T_{\mathbf{FG}} - T_{\mathbf{FF}} - T_{\mathbf{GG}} > 0$. So, we have, $T_{\mathbf{FG}} > T_{\mathbf{GG}}$ for all $d \geq d^{'}$. Hence,

$$\overline{\mathcal{W}}^*_{\mathbf{FG}} = (T_{\mathbf{FG}} - T_{\mathbf{FF}}) + (T_{\mathbf{FG}} - T_{\mathbf{GG}}) < T_{\mathbf{FF}} - T_{\mathbf{GG}} = S_{\mathbf{FG}}$$

Since $\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} - \hat{S}_{\mathbf{FG}} - \left(\overline{\mathcal{W}}^*_{\mathbf{FG}} - S_{\mathbf{FG}}\right) \xrightarrow{\mathbb{P}} 0$ as $d \to \infty$ and $\liminf_d \left(S_{\mathbf{FG}} - \overline{\mathcal{W}}^*_{\mathbf{FG}}\right) = \liminf_d 2\left(T_{\mathbf{FF}} - T_{\mathbf{FG}}\right) > 0$, in a similar way, we can show that for every $\epsilon_1 > 1$, there exists a d_1, such that

$$\mathbb{P}[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} > \hat{S}_{\mathbf{FG}}] < \epsilon_1 \text{ for all } d \geq d_1 \tag{7}$$

Now, it follows from (1) that for $\mathbf{Z} \sim \mathbf{F}, \left|\, S(\mathbf{Z}) - \frac{1}{2}\left(T_{\mathbf{FF}} - T_{\mathbf{GG}}\right)\right\} \,\Big| \xrightarrow{\mathbb{P}} 0$ as $d \to \infty$. We have already assumed that $T_{\mathbf{FF}} > T_{\mathbf{GG}}$. Define $\lambda_0 = \liminf_d \frac{1}{2}\left(T_{\mathbf{FF}} - T_{\mathbf{GG}}\right)$.

Since $\overline{\mathcal{W}}^*_{\mathbf{FG}} < T_{\mathbf{FF}} - T_{\mathbf{GG}}$, $\liminf_d \left(T_{\mathbf{FF}} - T_{\mathbf{GG}}\right) \geq \liminf_d \overline{\mathcal{W}}^*_{\mathbf{FG}} > 0$, by assumption 8. Hence, we have $\lambda_0 > 0$. Following similar arguments, one can show that for any $\epsilon_2 > 0$, there exists d_2, such that

$$\mathbb{P}[S(\mathbf{Z}) < 0 \mid \mathbf{Z} \sim \mathbf{F}] < \epsilon_2 \text{ for all } d \geq d_2 \tag{8}$$

Recall that

$$\Delta_1 = \frac{m}{m+n}\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \frac{n}{m+n}\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right]$$

and

$$\begin{aligned}\Delta_2 = &\frac{m}{m+n}\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \cdot S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\ &+ \frac{n}{m+n}\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \cdot S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right]\end{aligned}$$

It follows that

$$\begin{aligned}&\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}}S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\ &= \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}}S(\mathbf{Z}) \leq 0, S(\mathbf{Z}) \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\ &\quad + \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}}S(\mathbf{Z}) > 0, S(\mathbf{Z}) < 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\ &\leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}}S(\mathbf{Z}) \leq 0, S(\mathbf{Z}) \geq 0, \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \leq \hat{S}_{\mathbf{FG}} \mid \mathbf{Z} \sim \mathbf{F}\right] \\ &\quad + \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}}S(\mathbf{Z}) \leq 0, S(\mathbf{Z}) \geq 0, \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} > \hat{S}_{\mathbf{FG}} \mid \mathbf{Z} \sim \mathbf{F}\right] \\ &\quad + \mathbb{P}[S(\mathbf{Z}) < 0 \mid \mathbf{Z} \sim \mathbf{F}] \\ &\leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}\left\{\mathscr{D}_1(\mathbf{Z}) + S(\mathbf{Z})\right\} \leq 0, S(\mathbf{Z}) \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right]\end{aligned}$$

$$
\begin{aligned}
&+ \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} > \hat{S}_{\mathbf{FG}}\right] + \epsilon_2 \text{ for all } d \geq d_2 \\
\leq\ & \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \{\mathscr{D}_1(\mathbf{Z}) + S(\mathbf{Z})\} \leq 0, S(\mathbf{Z}) \geq 0, \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&+ \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} < 0\right] + \epsilon_1 + \epsilon_2 \text{ for all } d \geq \max\{d_1, d_2\} \\
\leq\ & \mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) + S(\mathbf{Z}) \leq 0, S(\mathbf{Z}) \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_0 + \epsilon_1 + \epsilon_2 \\
&\qquad\qquad\qquad\qquad \text{for all } d \geq \max\{d_0, d_1, d_2\} \\
\leq\ & \mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_0 + \epsilon_1 + \epsilon_2, \text{ for all } d \geq \max\{d_0, d_1, d_2\}
\end{aligned}
$$

Similarly, one can show that, for all $d \geq \max\{d_0, d_1, d_2\}$

$$
\begin{aligned}
\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right] &\leq \mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right] \\
&+ \epsilon_0 + \epsilon_1 + \epsilon_2
\end{aligned}
$$

Adding the two inequalities, we obtain

$$
\Delta_2 \leq \Delta_1 + \epsilon_0 + \epsilon_1 + \epsilon_2 \text{ for all } d \geq \max\{d_0, d_1, d_2\}
$$

Let, $d_0' = \max\{d_0, d_1, d_2\}$. Since, $\epsilon_0, \epsilon_1, \epsilon_2 > 0$ are arbitrary, we have $\Delta_2 \leq \Delta_1$ for all $d \geq d_0'$

(b) We have already proved that for every $\epsilon_1 > 0$, there exists a d_1, such that $\mathbb{P}[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} < 0] < \epsilon_1$, for all $d \geq d_1$.

$$
\begin{aligned}
&\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
=\ & \mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \leq 0, \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \leq 0, \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} < 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
\leq\ & \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} < 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
\leq\ & \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) \leq 0, \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&+ \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) \leq 0, \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_1 \\
\leq\ & \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&+ \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) - \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) < 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_1
\end{aligned}
$$

Now, we know, if $\mathbf{Z} \sim \mathbf{F}$,

$$
\left| \mathscr{D}_1(\mathbf{Z}) - \frac{1}{2} \overline{\mathcal{W}}^*_{\mathbf{FG}} \right| \xrightarrow{\mathbb{P}} 0 \text{ , as } d \to \infty
$$

$$
\left| S(\mathbf{Z}) - \frac{1}{2} \left(T_{\mathbf{FF}} - T_{\mathbf{GG}}\right) \right| \xrightarrow{\mathbb{P}} 0 \text{ , as } d \to \infty
$$

Also, we know,

$$\left|\hat{S}_{\mathbf{FG}} - (T_{\mathbf{FF}} - T_{\mathbf{GG}})\right| \xrightarrow{\mathbb{P}} 0 \text{ , as } d \to \infty$$

$$\left|\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} - \overline{\mathcal{W}}^*_{\mathbf{FG}}\right| \xrightarrow{\mathbb{P}} 0 \text{ , as } d \to \infty$$

Combining these, we get that given $\mathbf{Z} \sim \mathbf{F}$,

$$\left|\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) - \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) - \frac{1}{2}\left(\overline{\mathcal{W}}^{*2}_{\mathbf{FG}} - (T_{\mathbf{FF}} - T_{\mathbf{GG}})^2\right)\right| \xrightarrow{\mathbb{P}} 0 \text{ , as } d \to \infty$$

Now, if there exists a $d^{'} \in \mathbb{N}$ such that $T_{\mathbf{FG}} > \max\{T_{\mathbf{FF}}, T_{\mathbf{FG}}\}$ for all $d \geq d^{'}$,

$$\begin{aligned}
\overline{\mathcal{W}}^{*2}_{\mathbf{FG}} - (T_{\mathbf{FF}} - T_{\mathbf{GG}})^2 &= \overline{\mathcal{W}}^*_{\mathbf{FG}} - (T_{\mathbf{FF}} - T_{\mathbf{GG}})) \cdot (\overline{\mathcal{W}}^*_{\mathbf{FG}} + (T_{\mathbf{FF}} - T_{\mathbf{GG}})) \\
&= 4(T_{\mathbf{FG}} - T_{\mathbf{FF}})(T_{\mathbf{FG}} - T_{\mathbf{GG}}) \\
&> 4(T_{\mathbf{FG}} - \max\{T_{\mathbf{FF}}, T_{\mathbf{GG}}\})^2 \\
&> 0, \text{ for all } d \geq d^{'}
\end{aligned}$$

$$\begin{aligned}
&\liminf_d \, (T_{\mathbf{FG}} - \max\{T_{\mathbf{FF}}, T_{\mathbf{GG}}\}) > 0 \\
&\implies \liminf_d \, (T_{\mathbf{FG}} - \max\{T_{\mathbf{FF}}, T_{\mathbf{GG}}\})^2 > 0 \\
&\implies \liminf_d \left(\overline{\mathcal{W}}^{*2}_{\mathbf{FG}} - (T_{\mathbf{FF}} - T_{\mathbf{GG}})^2\right) > 4 \cdot \liminf_d \, (T_{\mathbf{FG}} - \max\{T_{\mathbf{FF}}, T_{\mathbf{GG}}\})^2 > 0
\end{aligned}$$

Using this, we can show that for every $\epsilon_2 > 0$, there exists a d_0', such that

$$\mathbb{P}[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) - \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) < 0 \mid \mathbf{Z} \sim \mathbf{F}] < \epsilon_2, \text{ for all } d \geq d_0'$$

So for all $d \geq d_0'$, we have,

$$\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_1 + \epsilon_2$$

Since $\epsilon_1, \epsilon_2 > 0$ are arbitrary, we can say that, for all $d \geq d_0'$,

$$\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right]$$

Similarly, one can show that, for all $d \geq d_0'$

$$\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right] \leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right]$$

Adding the last two inequalities, we obtain $\Delta_2 \geq \Delta_1$ for all $d \geq d_0'$.

□

Lemma A.9. *Suppose assumptions 6,8 and 9 are satisfied.*

(a) If $\liminf_d (\max\{T_{\mathbf{FF}}, T_{\mathbf{GG}}\} - T_{\mathbf{FG}}) > 0$ *there exists* $d_0' \in \mathbb{N}$*, such that* $\Delta_3 \leq \Delta_1$ *for all* $d \geq d_0'$

(b) Suppose assumption 9 holds in addition. If $\liminf_d (T_{\mathbf{FG}} - \max\{T_{\mathbf{FF}}, T_{\mathbf{FG}}\}) > 0$*, there exists* $d_0' \in \mathbb{N}$*, such that* $\Delta_3 \geq \Delta_1$ *for all* $d \geq d_0'$*.*

Proof of lemma A.9

(a) WLOG, we assume that $T_{\mathbf{FF}} > T_{\mathbf{GG}}$.
We have shown in (6), (7) and (8) that for every $\epsilon_0 > 0$, there exists a d_0, such that $\mathbb{P}[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} < 0] < \epsilon_0$ for all $d \geq d_0$, for every $\epsilon_1 > 0$, there exists a d_1, such that $\mathbb{P}[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} > \hat{S}_{\mathbf{FG}}] < \epsilon_1$, for all $d \geq d_1$ and for any $\epsilon_2 > 0$, there exists d_2, such that $\mathbb{P}[S(\mathbf{Z}) < 0 \mid \mathbf{Z} \sim \mathbf{F}] < \epsilon_2$ for all $d \geq d_2$ respectively. It follows that

$$
\begin{aligned}
&\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} L_(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&= \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0, S(\mathbf{Z}) \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\qquad + \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) > 0, S(\mathbf{Z}) < 0, \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0, S(\mathbf{Z}) \geq 0, \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \leq \hat{S}_{\mathbf{FG}} \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad + \mathbb{P}[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} > \hat{S}_{\mathbf{FG}}] + \mathbb{P}[S(\mathbf{Z}) < 0 \mid \mathbf{Z} \sim \mathbf{F}] \\
&\leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \{\operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \operatorname{sign} S(\mathbf{Z})\} \leq 0, S(\mathbf{Z}) \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_1 + \epsilon_2 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for all } d \geq \max\{d_1, d_2\} \\
&\leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \{\operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \operatorname{sign} S(\mathbf{Z})\} \leq 0, S(\mathbf{Z}) \geq 0, \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad + \mathbb{P}[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} < 0] + \epsilon_1 + \epsilon_2 \\
&\leq \mathbb{P}\left[\operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \operatorname{sign} S(\mathbf{Z}) \leq 0, S(\mathbf{Z}) \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_0 + \epsilon_1 + \epsilon_2 \\
&\leq \mathbb{P}\left[\operatorname{sign} \mathscr{D}_1(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_0 + \epsilon_1 + \epsilon_2 \\
&= \mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_0 + \epsilon_1 + \epsilon_2
\end{aligned}
$$

Similarly, one can show that, for all $d \geq \max\{d_0, d_1, d_2, d^{'}\}$

$$
\begin{aligned}
&\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign}(S(\mathbf{Z})) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right] \\
&\qquad \leq \mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right] + \epsilon_0 + \epsilon_1 + \epsilon_2
\end{aligned}
$$

Combining the two inequalities, we obtain

$$\Delta_3 \leq \Delta_1 + \epsilon_0 + \epsilon_1 + \epsilon_2 \text{ for all } d \geq \max\{d_0, d_1, d_2, d^{'}\}$$

Let, $d_0' = \max\left\{d_0, d_1, d_2, d^{'}\right\}$. Since, $\epsilon_0, \epsilon_1, \epsilon_2 > 0$ are arbitrary, we have $\Delta_3 \leq \Delta_1$ for all $d \geq d_0'$

(b) In (6), we have shown that for every $\epsilon_0 > 0$, there exists a d_0, such that $\mathbb{P}[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} < 0] < \epsilon_0$, for all $d \geq d_0$.

$$\begin{aligned}
&\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&= \mathbb{P}\left[\operatorname{sign} \mathscr{D}_1(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&= \mathbb{P}\left[\operatorname{sign} \mathscr{D}_1(\mathbf{Z}) \leq 0, \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad + \mathbb{P}\left[\operatorname{sign} \mathscr{D}_1(\mathbf{Z}) \leq 0, \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} < 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} < 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) \leq 0, \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad + \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) \leq 0, \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_0 \\
&\leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad + \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) - \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) < 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_0, \text{ for all } d \geq d_0
\end{aligned}$$

In the proof of theorem 10, we have shown that given $\mathbf{Z} \sim \mathbf{F}$,

$$\hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) - |T_{\mathbf{FF}} - T_{\mathbf{GG}}| \xrightarrow{\mathbb{P}} 0 \text{ , as } d \to \infty$$

$$\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) - \overline{\mathcal{W}}^*_{\mathbf{FG}} \xrightarrow{\mathbb{P}} 0 \text{ , as } d \to \infty$$

Hence, as $d \to \infty$,

$$\left| \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) - \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) - \left(\overline{\mathcal{W}}^*_{\mathbf{FG}} - |T_{\mathbf{FF}} - T_{\mathbf{GG}}| \right) \right| \xrightarrow{\mathbb{P}} 0.$$

Now,

$$\begin{aligned}
&\overline{\mathcal{W}}^*_{\mathbf{FG}} - |T_{\mathbf{FF}} - T_{\mathbf{GG}}| \\
&= 2T_{\mathbf{FG}} - T_{\mathbf{FF}} - T_{\mathbf{GG}} - (\max\{T_{\mathbf{FF}}, T_{\mathbf{GG}}\} - \min\{T_{\mathbf{FF}}, T_{\mathbf{GG}}\}) \\
&= 2(T_{\mathbf{FG}} - \max\{T_{\mathbf{FF}}, T_{\mathbf{GG}}\}).
\end{aligned}$$

So, $\liminf_d \left(\overline{\mathcal{W}}^*_{\mathbf{FG}} - |T_{\mathbf{FF}} - T_{\mathbf{GG}}| \right) > 0$ and using this, one can show that for every $\epsilon_1 > 0$, there exists a d_0', such that $\mathbb{P}[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) - \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) < 0] < \epsilon_1$, for all $d \geq d_0'$.

Therefore, for all $d \geq \max\{d_0, d_0'\}$, we have,

$$\begin{aligned}
\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \leq\ & \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) - \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
& + \epsilon_1 + \epsilon_0
\end{aligned}$$

Since $\epsilon_1, \epsilon_0 > 0$ are arbitrary, we can say that, for all $d \geq \max\{d_0, d_0'\}$,

$$\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) - \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right]$$

Similarly, one can show that, for all $d \geq \max\{d_0, d_0'\}$

$$\mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right] \leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) - \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right]$$

Adding the two inequalities, we obtain $\Delta_3 \geq \Delta_1$ for all $d \geq \max\{d_0, d_0'\}$.

□

Lemma A.10. *Suppose assumptions 6,8 and 9 are satisfied.*

(a) If $\liminf_d (\max\{T_{\mathbf{FF}}, T_{\mathbf{GG}}\} - T_{\mathbf{FG}}) > 0$ *there exists* $d_0' \in \mathbb{N}$*, such that* $\Delta_2 \leq \Delta_3$ *for all* $d \geq d_0'$

(b) Suppose assumption 9 holds in addition. If $\liminf_d (T_{\mathbf{FG}} - \max\{T_{\mathbf{FF}}, T_{\mathbf{FG}}\}) > 0$*, there exists* $d_0' \in \mathbb{N}$*, such that* $\Delta_3 \leq \Delta_2$ *for all* $d \geq d_0'$.

Proof of lemma A.10

(a) WLOG, we assume that $T_{\mathbf{FF}} > T_{\mathbf{GG}}$.
We have shown in (6) and (7) that for every $\epsilon_0 > 0$, there exists a d_0, such that $\mathbb{P}[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} < 0] < \epsilon_0$ for all $d \geq d_0$, for every $\epsilon_1 > 0$, there exists a d_1, such that $\mathbb{P}[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} > \hat{S}_{\mathbf{FG}}] < \epsilon_1$.

$$\begin{aligned}
&\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0, \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \mathbb{P}[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} < 0] \\
&\leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0, 0 \leq \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \leq \hat{S}_{\mathbf{FG}} \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad + \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} > \hat{S}_{\mathbf{FG}}\right] + \epsilon_0 \\
&\leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0, 0 \leq \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \leq \hat{S}_{\mathbf{FG}}, S(\mathbf{Z}) < 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad + \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0, 0 \leq \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \leq \hat{S}_{\mathbf{FG}}, S(\mathbf{Z}) \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad + \epsilon_0 + \epsilon_1
\end{aligned}$$

Now, we see that $0 \leq \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \leq \hat{S}_{\mathbf{FG}}$ implies that $\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) - \hat{S}_{\mathbf{FG}} \leq 0$. Therefore,

$$\begin{aligned}
&\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0, 0 \leq \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \leq \hat{S}_{\mathbf{FG}}, S(\mathbf{Z}) < 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&= \mathbb{P}\Big[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0, 0 \leq \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \leq \hat{S}_{\mathbf{FG}}, \operatorname{sign} S(\mathbf{Z}) = -1, \\
&\qquad \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) - \hat{S}_{\mathbf{FG}} \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\Big] \\
&\leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right]
\end{aligned}$$

Now,

$$\begin{aligned}
&\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}}S(\mathbf{Z}) \leq 0, 0 \leq \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \leq \hat{S}_{\mathbf{FG}}, S(\mathbf{Z}) \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&= \mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \leq -\frac{\hat{S}_{\mathbf{FG}}}{\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}} S(\mathbf{Z}), 0 \leq \hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \leq \hat{S}_{\mathbf{FG}}, S(\mathbf{Z}) \geq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&= \mathbb{P}\left[\mathscr{D}_1(\mathbf{Z}) \leq -S(\mathbf{Z}) \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&= \mathbb{P}\left[2T_{\mathbf{G}}(\mathbf{Z}) - \hat{T}_{\mathbf{FG}} - \hat{T}_{\mathbf{GG}} \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right]
\end{aligned}$$

Given $\mathbf{Z} \sim \mathbf{F}$,

$$\left|2T_{\mathbf{G}}(\mathbf{Z}) - \hat{T}_{\mathbf{FG}} - \hat{T}_{\mathbf{GG}} - (T_{\mathbf{FG}} - T_{\mathbf{GG}})\right| \xrightarrow{\mathbb{P}} 0, \text{ , as } d \to \infty$$

$$\liminf_d (T_{\mathbf{FG}} - T_{\mathbf{GG}}) = \liminf_d \left(\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} + (T_{\mathbf{FF}} - T_{\mathbf{FG}})\right) > 0$$

Hence, for any $\epsilon_2 > 0$, there exists d_2, such that

$$\mathbb{P}\left[2T_{\mathbf{G}}(\mathbf{Z}) - \hat{T}_{\mathbf{FG}} - \hat{T}_{\mathbf{GG}} \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] < \epsilon_2, \text{ for all } d \geq d_2$$

Combining all these, we get that for $d \geq \max\{d_0, d_1, d_2\} = d_0'$(let),

$$\begin{aligned}
&\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}}S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad \leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_0 + \epsilon_1 + \epsilon_2
\end{aligned}$$

Since $\epsilon_0, \epsilon_1, \epsilon_2 > 0$ are arbitrary, we can say that, for all $d \geq d_0'$,

$$\begin{aligned}
&\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}}S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad \leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right]
\end{aligned}$$

Similarly, one can show that, for all $d \geq d_0'$,

$$\begin{aligned}
&\mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}}\mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}}S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right] \\
&\quad \leq \mathbb{P}\left[\hat{\overline{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right]
\end{aligned}$$

Combing the above two inequalities, we obtain $\Delta_2 \leq \Delta_3$ for all $d \geq d_0'$.

(b) WLOG, assume that $T_{\mathbf{FF}} > T_{\mathbf{GG}}$. W e know that $\hat{S}_{\mathbf{FG}} \xrightarrow{\mathbb{P}} T_{\mathbf{FF}} - T_{\mathbf{GG}}$, as $d \to \infty$. Then assumption 9 implies that for every $\epsilon_0 > 0$, there exists a d_0, such that

$$\mathbb{P}[\hat{S}_{\mathbf{FG}} \leq 0] < \epsilon_0 \text{ for all } d \geq d_0$$

Since, $\liminf_d (T_{\mathbf{FG}} - \max\{T_{\mathbf{FF}}, T_{\mathbf{FG}}\}) > 0$, there exists $d' \in \mathbb{N}$ such that $T_{\mathbf{FG}} > T_{\mathbf{FF}}$ for all $d \geq d^{'}$. Hence, for all $d \geq d^{'}$,

$$\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} = (T_{\mathbf{FG}} - T_{\mathbf{FF}}) + (T_{\mathbf{FG}} - T_{\mathbf{GG}}) > T_{\mathbf{FF}} - T_{\mathbf{GG}} = S_{\mathbf{FG}}$$

Since $\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} - \hat{S}_{\mathbf{FG}} - \left(\overline{\mathcal{W}}^*_{\mathbf{FG}} - S_{\mathbf{FG}}\right) \xrightarrow{\mathbb{P}} 0$ as $d \to \infty$.
Since $\liminf_d \left(\overline{\mathcal{W}}^*_{\mathbf{FG}} - S_{\mathbf{FG}}\right) = \liminf_d 2\,(T_{\mathbf{FG}} - T_{\mathbf{FF}}) > 0$, hence, we can show that for every $\epsilon_1 > 0$, there exists a d_1, such that

$$\mathbb{P}[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \leq \hat{S}_{\mathbf{FG}}] < \epsilon_1 \text{ for all } d \geq d_1$$

It follows that

$$\begin{aligned}
&\mathbb{P}\left[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\leq \mathbb{P}\left[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0, \hat{S}_{\mathbf{FG}} > 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \mathbb{P}[\hat{S}_{\mathbf{FG}} \leq 0] \\
&= \mathbb{P}\left[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0, \overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} > \hat{S}_{\mathbf{FG}} > 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad + \mathbb{P}[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \leq \hat{S}_{\mathbf{FG}}] + \epsilon_0 \\
&\leq \mathbb{P}\Bigg[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0, \\
&\qquad\qquad \overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0, \overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} > \hat{S}_{\mathbf{FG}} > 0 \mid \mathbf{Z} \sim \mathbf{F}\Bigg] \\
&\quad + \mathbb{P}\Bigg[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0, \overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} > \hat{S}_{\mathbf{FG}} > 0, \\
&\qquad\qquad \overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{F}\Bigg] + \epsilon_0 + \epsilon_1 \\
&\leq \mathbb{P}\left[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_0 + \epsilon_1 + p_0
\end{aligned}$$

where $p_0 = \mathbb{P}(A(\mathbf{Z}))$ with

$$\begin{aligned}
A(\mathbf{Z}) = &\{\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0, \\
&\qquad \overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) > 0, \overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} > \hat{S}_{\mathbf{FG}} > 0 \mid \mathbf{Z} \sim \mathbf{F}\}
\end{aligned}$$

We consider the four mutually disjoint and exclusive events here,

1. $L_{\mathbf{G}}(\mathbf{Z}) > L_{\mathbf{F}}(\mathbf{Z}), S(\mathbf{Z}) > 0 \implies A(\mathbf{Z})$ cannot occur.
2. $L_{\mathbf{G}}(\mathbf{Z}) > L_{\mathbf{F}}(\mathbf{Z}), S(\mathbf{Z}) < 0 \implies A(\mathbf{Z})$ cannot occur.
3. $L_{\mathbf{G}}(\mathbf{Z}) < L_{\mathbf{F}}(\mathbf{Z}), S(\mathbf{Z}) > 0$
4. $L_{\mathbf{G}}(\mathbf{Z}) < L_{\mathbf{F}}(\mathbf{Z}), S(\mathbf{Z}) < 0, \implies A(\mathbf{Z})$ cannot occur.

Hence, $A(\mathbf{Z})$ implies $L_{\mathbf{G}}(\mathbf{Z}) < L_{\mathbf{F}}(\mathbf{Z}), S(\mathbf{Z}) > 0$. So,

$$\begin{aligned}
p_0 &= \mathbb{P}[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0, \overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) > 0, \\
&\quad \overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} > \hat{S}_{\mathbf{FG}} > 0, L_{\mathbf{G}}(\mathbf{Z}) < L_{\mathbf{F}}(\mathbf{Z}), S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{F}] \\
&= \mathbb{P}[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} > \hat{S}_{\mathbf{FG}} > 0, L_{\mathbf{G}}(\mathbf{Z}) < L_{\mathbf{F}}(\mathbf{Z}), S(\mathbf{Z}) > 0, \\
&\quad \overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{F}] \\
&\leq \mathbb{P}\left[S(\mathbf{Z}) - (\mathscr{D}_1(\mathbf{Z})) > 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&= \mathbb{P}\left[2L_{\mathbf{F}}(\mathbf{Z}) - T_{\mathbf{FG}} > 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&= \mathbb{P}\left[T_{\mathbf{FF}} + T_{\mathbf{FG}} - 2T_{\mathbf{F}}(\mathbf{Z}) < 0 \mid \mathbf{Z} \sim \mathbf{F}\right]
\end{aligned}$$

Now, given $\mathbf{Z} \sim \mathbf{F}$,

$$T_{\mathbf{FF}} + T_{\mathbf{FG}} - 2T_{\mathbf{F}}(\mathbf{Z}) - (T_{\mathbf{FG}} - T_{\mathbf{FF}}) \xrightarrow{\mathbb{P}} 0$$

Now, $\liminf_d (T_{\mathbf{FG}} - T_{\mathbf{FF}}) > 0$, because we have assumed that $T_{\mathbf{FF}} = \max\{T_{\mathbf{FF}}, T_{\mathbf{GG}}\}$. So, for every $\epsilon_2 > 0$, there exists a d_2 such that

$$\mathbb{P}\left[T_{\mathbf{FF}} + T_{\mathbf{FG}} - 2T_{\mathbf{F}}(\mathbf{Z}) < 0 \mid \mathbf{Z} \sim \mathbf{F}\right] < \epsilon_2 \text{ for all } d \geq d_2.$$

Hence, $p_0 \leq \epsilon_2$ for all $d \geq d_2$. And for all d $\geq \max\{d_0, d_1, d_2\}$,

$$\begin{aligned}
&\mathbb{P}\left[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad \leq \mathbb{P}\left[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] + \epsilon_0 + \epsilon_1 + \epsilon_2
\end{aligned}$$

Let, $d_0' = \max\{d_0, d_1, d_2\}$. Since $\epsilon_0, \epsilon_1, \epsilon_2 > 0$ are arbitrary, we can say that, for all $d \geq d_0'$,

$$\begin{aligned}
&\mathbb{P}\left[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right] \\
&\quad \leq \mathbb{P}\left[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) \leq 0 \mid \mathbf{Z} \sim \mathbf{F}\right]
\end{aligned}$$

Similarly, one can show that, for all $d \geq d_0'$,

$$\begin{aligned}
&\mathbb{P}\left[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \operatorname{sign} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} \operatorname{sign} S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right] \\
&\quad \leq \mathbb{P}\left[\overline{\hat{\mathcal{W}}}^*_{\mathbf{FG}} \mathscr{D}_1(\mathbf{Z}) + \hat{S}_{\mathbf{FG}} S(\mathbf{Z}) > 0 \mid \mathbf{Z} \sim \mathbf{G}\right]
\end{aligned}$$

Combing the above two inequalities, we obtain $\Delta_3 \leq \Delta_2$ for all $d \geq d_0'$.

□

Proof of Theorem 12

The proof of this theorem follows directly from lemma A.8, lemma A.9 and lemma A.10.

B ADDITIONAL NUMERICAL RESULTS and DETAILS

B.1 SIMULATION DETAILS

GLMNET: The R-package `glmnet` is used for implementing GLMNET. The tuning parameter α in the elastic-net penalty term is fixed at the default value 1. The weight λ of the penalty term is chosen via cross-validation using the function `cv.glmnet` with default values of its parameters.
NN-RP: The function `classify` from the package `RandPro` is used with default values of its parameters.
SVM: The R package `e1071` is used for implementing SVM with linear (SVM-LIN) and radial basis function (SVF-RBF) kernels. For the RBF kernel, i.e., $K_\theta(\mathbf{x}, \mathbf{y}) = \exp\{-\theta\|\mathbf{x} - \mathbf{y}\|^2\}$, the default value of the tuning parameter θ was chosen, i.e., $\theta = \frac{1}{d}$.
N-NET: We used the `nnet` function from the R package `nnet` to fit a single-hidden-layer artificial neural network with default parameters. The number of units in the hidden layer was allowed to vary from 1 up to 10. Among them, the one with the minimum misclassification rate was reported as N-NET.
1-NN: The `knn1` function from the R-package `class` is used for implementing the 1-nearest neighbor classifier.

B.2 TABLES

Table 1. Estimated values of $T_{\mathbf{FF}}, T_{\mathbf{FG}}$ and $T_{\mathbf{GG}}$ with standard errors (in parentheses) for δ_1, δ_2, δ_3, for simulated examples at $d = 1000$

Example	$\hat{T}_{\mathbf{FF}}$	$\hat{T}_{\mathbf{FG}}$	$\hat{T}_{\mathbf{GG}}$	$\hat{T}_{\mathbf{FG}} > \max(\hat{T}_{\mathbf{FF}}, \hat{T}_{\mathbf{GG}})$
1	0.28362 (0.00008)	0.31839 (0.000023)	0.3461 (0.000064)	FALSE
2	0.34206 (0.00007)	0.31841 (0.000025)	0.2876 (0.00008)	FALSE
3	0.3004 (0.000112)	0.33211 (0.000058)	0.30041 (0.000104)	TRUE
4	0.27954 (0.00009)	0.31944 (0.000024)	0.34796 (0.000077)	FALSE
5	0.26462 (0.000078)	0.32142 (0.000032)	0.35871 (0.000076)	TRUE

Table 2. Estimated misclassification probabilities (in percentages) with standard errors (in parentheses) for δ_1, δ_2, δ_3, and popular classifiers **Simulated Example 1**.

d	Popular Classifiers							Proposed Classifiers		
	Bayes	GLMNET	NN-RP	SVM-LIN	SVM-RBF	N-NET	1-NN	δ_1	δ_2	δ_3
5	30.36 (0.0031)	49.30 (0.0038)	45.10 (0.0030)	48.49 (0.0037)	36.29 (0.0039)	44.76 (0.0042)	45.68 (0.0038)	45.23 (0.0039)	33.66 (0.0032)	33.62 (0.0033)
10	22.4 (0.0027)	49.64 (0.0035)	46.51 (0.0021)	49.21 (0.0033)	29.90 (0.0040)	44.51 (0.0042)	44.74 (0.0035)	43.74 (0.0041)	26.83 (0.0033)	26.71 (0.0032)
25	11.44 (0.0023)	48.36 (0.0033)	48.90 (0.0011)	48.30 (0.0038)	18.53 (0.0032)	45.12 (0.0039)	45.89 (0.0024)	40.62 (0.0036)	16.24 (0.0028)	16.36 (0.0028)
50	4.16 (0.0014)	48.05 (0.0038)	49.93 (0.0000)	47.76 (0.0036)	9.45 (0.0024)	44.82 (0.0041)	47.98 (0.0013)	35.98 (0.0038)	8.04 (0.0020)	7.96 (0.0020)
100	0.73 (0.0000)	48.33 (0.0037)	50.00 (0.0000)	47.08 (0.0032)	2.90 (0.0013)	44.78 (0.0041)	49.64 (0.0000)	30.61 (0.0038)	2.38 (0.0013)	2.40 (0.0013)
250	0.01 (0.0000)	47.60 (0.0038)	50.00 (0.0000)	46.50 (0.0028)	0.13 (0.0000)	45.92 (0.0035)	49.99 (1e-04)	21.39 (0.0032)	0.10 (0.0000)	0.11 (0.0000)
500	0.00 (0.0000)	47.95 (0.0037)	50.00 (0.0000)	46.48 (0.0025)	0.00 (0.0000)	45.08 (0.0038)	50.00 (0.0000)	13.30 (0.0024)	0.00 (0.0000)	0.00 (0.0000)
1000	0.00 (0.0000)	47.14 (0.0038)	50.00 (0.0000)	47.11 (0.0017)	0.00 (0.0000)	44.46 (0.0040)	50.00 (0.0000)	5.74 (0.0017)	0.00 (0.0000)	0.00 (0.0000)

Table 3. Estimated misclassification probabilities (in percentages) with standard errors (in parentheses) for δ_1, δ_2, δ_3, and popular classifiers **Simulated Example 2**.

d	Popular Classifiers							Proposed Classifiers		
	Bayes	GLMNET	NN-RP	SVM-LIN	SVM-RBF	N-NET	1-NN	δ_1	δ_2	δ_3
5	30.64 (0.0030)	49.98 (0.0039)	46.78 (0.0003)	49.58 (0.0039)	43.48 (0.0051)	47.24 (0.0042)	45.15 (0.0039)	45.18 (0.0045)	34.79 (0.0032)	35.05 (0.0041)
10	24.17 (0.0031)	49.60 (0.0035)	47.84 (0.0029)	49.66 (0.0035)	42.19 (0.0042)	48.01 (0.0035)	46.56 (0.0031)	42.91 (0.0036)	28.73 (0.0032)	28.61 (0.0033)
25	13.05 (0.0024)	49.94 (0.0037)	48.91 (0.0020)	50.04 (0.0039)	41.54 (0.0052)	49.63 (0.0036)	47.84 (0.0034)	39.53 (0.0045)	19.22 (0.0030)	19.40 (0.0029)
50	5.72 (0.0016)	49.48 (0.0038)	49.65 (0.0016)	50.07 (0.0040)	40.86 (0.0041)	49.54 (0.0038)	48.68 (0.0027)	35.29 (0.0039)	11.34 (0.0021)	11.48 (0.0020)
100	1.20 (0.0000)	49.01 (0.0031)	50.00 (0.0012)	49.40 (0.0034)	39.50 (0.0044)	49.56 (0.0029)	49.84 (0.0019)	30.14 (0.0034)	4.09 (0.0015)	4.15 (0.0015)
250	0.00 (0.0000)	48.80 (0.0037)	50.15 (0.0011)	49.28 (0.0032)	37.24 (0.0033)	49.12 (0.0037)	49.63 (0.0014)	20.71 (0.0031)	0.31 (0.0000)	0.30 (0.0000)
500	0.00 (0.0000)	49.25 (0.0034)	50.08 (0.0000)	49.73 (0.0035)	35.91 (0.0033)	49.54 (0.0033)	50.19 (0.0000)	12.66 (0.0022)	0.00 (0.0000)	0.00 (0.0000)
1000	0.00 (0.0000)	49.06 (0.0033)	49.96 (0.0000)	49.60 (0.0032)	34.84 (0.0030)	49.06 (0.0033)	50.06 (0.0000)	5.19 (0.0018)	0.00 (0.0000)	0.00 (0.0000)

Table 4. Estimated misclassification probabilities (in percentages) with standard errors (in parentheses) for δ_1, δ_2, δ_3, and popular classifiers **Simulated Example 3**.

d	Popular Classifiers							Proposed Classifiers		
	Bayes	GLMNET	NN-RP	SVM-LIN	SVM-RBF	N-NET	1-NN	$\boldsymbol{\delta_1}$	$\boldsymbol{\delta_2}$	$\boldsymbol{\delta_3}$
5	21.94 (0.0030)	40.09 (0.0060)	39.20 (0.0047)	43.21 (0.0059)	41.50 (0.0058)	38.29 (0.0050)	40.65 (0.0041)	28.28 (0.0041)	30.41 (0.0058)	30.30 (0.0074)
10	14.18 (0.0025)	39.50 (0.0053)	41.54 (0.0050)	41.64 (0.0058)	41.52 (0.0054)	39.58 (0.0044)	41.82 (0.0043)	21.40 (0.0037)	22.92 (0.0043)	22.00 (0.0057)
25	4.64 (0.0013)	37.41 (0.0045)	44.6 (0.0048)	40.71 (0.0050)	41.69 (0.0049)	41.92 (0.0042)	44.31 (0.0040)	10.18 (0.0025)	10.57 (0.0028)	10.18 (0.0025)
50	0.64 (0.0000)	37.48 (0.0042)	46.80 (0.0037)	41.62 (0.0044)	43.43 (0.0041)	44.16 (0.0042)	46.61 (0.0040)	3.86 (0.0014)	4.04 (0.0015)	3.86 (0.0014)
100	0.04 (0.0000)	36.07 (0.0038)	47.88 (0.0036)	39.93 (0.0042)	43.17 (0.0032)	44.97 (0.0035)	47.54 (0.0037)	0.69 (0.0000)	0.76 (0.0000)	0.69 (0.0000)
250	0.00 (0.0000)	35.82 (0.0036)	49.42 (0.0023)	40.45 (0.0035)	46.26 (0.0023)	46.44 (0.0039)	48.62 (0.0032)	0.01 (0.0000)	0.01 (0.0000)	0.01 (0.0000)
500	0.00 (0.0000)	35.78 (0.0032)	49.83 (0.0022)	39.76 (0.0033)	48.05 (0.0016)	46.74 (0.0040)	49.66 (0.0025)	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0000)
1000	0.00 (0.0000)	35.27 (0.0035)	50.24 (0.0022)	39.69 (0.0031)	49.64 (0.0000)	47.78 (0.0041)	49.53 (0.0029)	0.00 (0.0000)	0.00 (0.0000)	0.00 (0.0000)

Table 5. Estimated misclassification probabilities (in percentages) with standard errors (in parentheses) for δ_1, δ_2, δ_3, and popular classifiers **Simulated Example 4**.

d	Popular Classifiers							Proposed Classifiers		
	Bayes	GLMNET	NN-RP	SVM-LIN	SVM-RBF	N-NET	1-NN	$\boldsymbol{\delta_1}$	$\boldsymbol{\delta_2}$	$\boldsymbol{\delta_3}$
5	28.90 (0.0025)	45.85 (0.0023)	44.24 (0.0023)	46.37 (0.0017)	40.02 (0.0033)	41.82 (0.0023)	40.68 (0.0028)	43.39 (0.0046)	31.5 (0.0033)	31.52 (0.0034)
10	22.44 (0.0026)	45.54 (0.0020)	45.88 (0.0021)	45.32 (0.0017)	37.28 (0.0024)	41.90 (0.0020)	40.46 (0.0024)	39.29 (0.0045)	25.16 (0.0026)	25.20 (0.0028)
25	11.17 (0.0019)	45.43 (0.0016)	48.18 (0.0014)	45.24 (0.0015)	34.90 (0.0021)	42.38 (0.0020)	41.53 (0.0026)	34.06 (0.0041)	16.56 (0.0030)	16.94 (0.0029)
50	4.13 (0.0014)	44.99 (0.0014)	49.01 (0.0000)	44.54 (0.0015)	35.38 (0.0027)	43.53 (0.0020)	43.92 (0.0021)	27.92 (0.0038)	10.53 (0.0028)	11.00 (0.0029)
100	0.76 (0.0000)	44.97 (0.0017)	49.68 (0.0000)	43.85 (0.0018)	38.31 (0.0028)	44.98 (0.0025)	45.96 (0.0016)	21.30 (0.0035)	6.22 (0.0021)	6.76 (0.0022)
250	0.00 (0.0000)	45.06 (0.0015)	49.85 (0.0000)	44.80 (0.0014)	44.62 (0.0021)	45.44 (0.0026)	47.43 (0.0013)	10.96 (0.0025)	1.76 (0.0012)	2.06 (0.0012)
500	0.00 (0.0000)	44.97 (0.0012)	49.94 (0.0000)	44.38 (0.0017)	48.02 (0.0012)	45.48 (0.0030)	48.44 (0.0010)	4.19 (0.0018)	0.20 (0.0000)	0.27 (0.0000)
1000	0.00 (0.0000)	44.78 (0.0017)	49.92 (0.0000)	45.10 (0.0014)	49.78 (0.0000)	46.08 (0.0019)	49.02 (0.0000)	1.18 (0.0000)	0.00 (0.0000)	0.00 (0.0000)

Table 6. Estimated misclassification probabilities (in percentages) with standard errors (in parentheses) for δ_1, δ_2, δ_3, and popular classifiers **Simulated Example 5**.

d	Popular Classifiers							Proposed Classifiers		
	Bayes	**GLMNET**	**NN-RP**	**SVM-LIN**	**SVM-RBF**	**N-NET**	**1-NN**	$\boldsymbol{\delta_1}$	$\boldsymbol{\delta_2}$	$\boldsymbol{\delta_3}$
5	20.78 (0.0027)	48.63 (0.0041)	42.47 (0.0032)	46.85 (0.0039)	32.21 (0.0047)	40.90 (0.0053)	39.40 (0.0039)	38.14 (0.0048)	24.97 (0.0033)	25.32 (0.0034)
10	11.77 (0.0022)	48.20 (0.0041)	45.82 (0.0026)	47.60 (0.0037)	28.02 (0.0048)	40.71 (0.0045)	40.31 (0.0029)	33.56 (0.0047)	17.26 (0.0033)	17.71 (0.0033)
25	2.86 (0.0012)	47.45 (0.0038)	49.40 (0.0017)	48.30 (0.0035)	24.94 (0.0044)	44.00 (0.0044)	46.39 (0.0022)	25.09 (0.0043)	7.98 (0.0023)	8.33 (0.0023)
50	0.36 (0.0000)	46.66 (0.0036)	49.91 (0.0015)	47.84 (0.0035)	23.55 (0.0035)	45.47 (0.0038)	48.87 (0.0021)	17.22 (0.0033)	3.08 (0.0015)	3.43 (0.0016)
100	0.00 (0.0000)	47.26 (0.0038)	49.83 (0.0014)	47.59 (0.0032)	24.30 (0.0028)	46.46 (0.0036)	49.77 (0.0019)	9.60 (0.0025)	0.67 (0.0000)	0.84 (0.0000)
250	0.00 (0.0000)	46.08 (0.0040)	50.00 (0.0017)	48.01 (0.0026)	29.72 (0.0026)	48.06 (0.0035)	50.33 (0.0020)	2.56 (0.0013)	0.02 (0.0000)	0.04 (0.0000)
500	0.00 (0.0000)	45.47 (0.0030)	50.08 (0.0016)	48.43 (0.0030)	37.14 (0.0025)	48.22 (0.0037)	49.88 (0.0019)	0.44 (0.0000)	0.00 (0.0000)	0.00 (0.0000)
1000	0.00 (0.0000)	45.25 (0.0028)	49.84 (0.0023)	48.59 (0.0032)	44.47 (0.0015)	49.34 (0.0042)	49.90 (0.0021)	0.01 (0.0000)	0.00 (0.0000)	0.00 (0.0000)