SUPPLEMENTARY MATERIAL FOR "OPTIMAL CLASSIFICATION IN SPARSE GAUSSIAN GRAPHIC MODEL"

By Yingying Fan, Jiashun Jin, and Zhigang Yao

University of Southern California, Carnegie Mellon University, and Ecole Polytechnique Fédérale de Lausanne

In this supplement we present the technical proofs for the main work [2]. Equation and theorem references made to the main document do not contain letters.

APPENDIX A: PROOFS FOR MAIN THEOREMS AND LEMMAS

In this section, we prove all key theorems and lemmas in the order they appear (except for Theorem 1.2-1.3 which are proved in Section 3.3). Secondary lemmas are proved in APPENDIX B.

A.1. Proof of Theorem 1.1. For short, write $n = n_p$. Recall that the training samples are $X_i \sim N(Y_i\mu, \Omega^{-1})$, $1 \leq i \leq n$, where $Y_i \in \{-1, 1\}$ are given. Consider an (independent) test sample $X \sim N(Y \cdot \mu, \Omega^{-1})$, where $Y = \pm 1$ with equal probabilities. Let $f_{\pm 1}$ be the joint of density of $(X_1, \ldots, X_n, X)$ in the case where $Y = 1$ and $Y = -1$, respectively, and let $H(f, g)$ be the Hellinger distance between two density functions f and g. To show the claim, it is sufficient to show $H(f_1, f_{-1}) \to 0$ as $p \to \infty$, uniformly for all $\Omega \in \mathcal{M}_p^*(a, K_p)$. Let f_0 be the joint density of $(X_1, \ldots, X_n, X)$ in the case where $X \sim N(0, \Omega^{-1})$ (but the distributions of X_i remain the same). By triangle inequality and symmetry, $H(f_1, f_{-1}) \leq H(f_1, f_0) + H(f_{-1}, f_0) = 2H(f_1, f_0)$. Therefore, it is sufficient to show

$$H(f_1, f_0) \to 0. \tag{A.1}$$

Since Ω is a K_p-sparse correlation matrix, by Lemma 1.1, there is a permutation matrix P and an integer $M_p = M_p(\Omega, K_p)$ such that $M_p \leq K_p$ and

$$P\Omega P' = \begin{pmatrix} \tilde{\Omega}_{11} & \ldots & \tilde{\Omega}_{1M_p} \\ \ldots & \ldots & \ldots \\ \tilde{\Omega}_{M_p 1} & \ldots & \tilde{\Omega}_{M_p M_p} \end{pmatrix}, \tag{A.2}$$

where on the diagonal, $\tilde{\Omega}_{11}$, ..., $\tilde{\Omega}_{M_p M_p}$ are identity matrices. Since permuting the coordinates of $X_1, X_2, \ldots, X$ simultaneously does not change the Hellinger distance $H(f_1, f_0)$, we assume $P = I_p$ for simplicity.

Now, corresponding to the partition of Ω in (A.2), we partition the mean-vector μ as $\mu = ((\mu^{(1)})', \ldots, (\mu^{(M_p)})')'$. For $0 \leq m \leq M_p$, let P_m be the projection matrix such that $P_m\mu = ((\mu^{(1)})', \ldots, (\mu^{(m)})', 0, \ldots, 0)'$, where generically, 0 denotes a row vector of zeros, and let $f^{(m)}$ be the joint density of $(X_1, \ldots, X_n, X)$ under the law that $X_i \sim N(Y_i\mu, \Omega^{-1})$ for all $1 \leq i \leq n$ and $X \sim N(P_m\mu, \Omega^{-1})$. Note that $f_0 = f^{(0)}$ and $f_1 = f^{(M_p)}$, and that by triangle inequality,

$$H(f^{(0)}, f^{(M_p)}) \leq \sum_{m=1}^{M_p} H(f^{(m-1)}, f^{(m)}). \tag{A.3}$$

Recalling $M_p \leq K_p$ and $K_p \leq L_p$, where L_p is a generic multi-log(p) term as in Definition 1.2, (A.1) follows by Lemma A.1 below. □

LEMMA A.1. *There is a constant $c_0 = c_0(\beta, r, \theta) > 0$ such that for any $1 \leq m \leq M_p - 1$,*

$$H(f^{(m-1)}, f^{(m)}) \leq L_p p^{-c_0}. \tag{A.4}$$

A.2. Proof of Lemma A.1. Denote $K = K_p$, $M = M_p$, and $n = n_p$ for short. Recall that each of $X, X_1, \ldots, X_n$ can be partitioned into M blocks. We simultaneously swap the first block and the m-th block of X and of each X_i, $1 \leq i \leq n$, but still denote the resultant vectors by X and X_i for notational simplicity. Denote $\tilde{\nu} = \mu^{(m)}$, $\tilde{\tilde{\nu}} = ((\mu^{(1)})', \ldots, (\mu^{(m-1)})', 0, \ldots, 0)'$, and $\tilde{\tilde{\mu}} = ((\mu^{(1)})', (\mu^{(2)})', \ldots, (\mu^{(m-1)})', (\mu^{(m+1)})', \ldots, (\mu^{(M)})')'$. After swapping, $f^{(m)}$ is the joint density of $(X_1, \ldots, X_n, X)$, where the common mean vector of $X_1, \ldots, X_n$ (which we still denote by μ for simplicity) is $\mu = (\tilde{\nu}', \tilde{\tilde{\mu}}')'$, the mean vector of X is $(\tilde{\nu}', \tilde{\tilde{\nu}}')'$, and the common precision matrix (still denote by Ω for simplicity) of $X_1, \ldots, X_n, X$ is

$$\Omega = \begin{pmatrix} I_k & B \\ B' & D \end{pmatrix}, \tag{A.5}$$

where I_k is a $k \times k$ identity matrix with $k = k(\Omega, m)$ equaling to the size of the m-th block (before swapping) and D is a correlation matrix. Similarly, $f^{(m-1)}$ is the joint density of $(X_1, \ldots, X_n, X)$, where the laws of $X_1, \ldots, X_n, X$ are the same as that of $f^{(m)}$ except for that the mean vector of X is $(0, \tilde{\tilde{\nu}}')'$ instead.

Denote for short $f_0 = f^{(m-1)}$, $f_1 = f^{(m)}$. Since (Y_i, X_i) are given and $Y_i \in \{-1, 1\}$, we assume $Y_i = 1$ for notational simplicity (when $Y_i = -1$, we can always multiply -1 to both Y_i and X_i). Consequently, $Z =$

$\frac{1}{\sqrt{n}}\sum_{i=1}^n Y_i X_i$ reduces to $Z = \frac{1}{\sqrt{n}}\sum_{i=1}^n X_i$. By definitions and elementary statistics, $f_0(x_1,\dots,x_n,x) = \phi(x,\Omega)\Pi_{i=1}^n \phi(x_i,\Omega)\cdot I$, and $f_1(x_1,\dots,x_n,x) = \phi(x,\Omega)\Pi_{i=1}^n\phi(x_i,\Omega)\cdot II$, where

$$I = \int e^{\sqrt{n}\mu'\Omega z - \frac{n}{2}\mu'\Omega\mu + (0,\tilde{\tilde{\nu}}')\Omega x - \frac{1}{2}\tilde{\tilde{\nu}}'D\tilde{\tilde{\nu}}} dF(\mu),$$

$$II = \int e^{\sqrt{n}\mu'\Omega z - \frac{n}{2}\mu'\Omega\mu + (\tilde{\nu}',\tilde{\tilde{\nu}}')\Omega x - \frac{1}{2}[\|\tilde{\nu}\|^2 + 2\tilde{\nu}'B\tilde{\tilde{\nu}} + \tilde{\tilde{\nu}}'D\tilde{\tilde{\nu}}]} dF(\mu),$$

and $F(\mu)$ denotes the cdf of μ. Here, x and x_i are $p\times 1$ vectors, $z = \frac{1}{\sqrt{n}}\sum_{i=1}^n x_i$, and $\phi(x,\Omega)$ is the joint density of $N(0,\Omega^{-1})$. For $1\le i\le k$, denote the i-th row of B in (A.5) by m_i'. Also, write $\Omega x = (\tilde{x}', \tilde{\tilde{x}}')'$ and $\Omega z = (\tilde{z}', \tilde{\tilde{z}}')'$ so that both the length of $\tilde{x}$ and the length of $\tilde{z}$ are k.

For simplicity, we assume H_p is a point mass at τ_p; the proof for general cases is similar since the support of H_p is contained in $[-\tau_p,\tau_p]$, but we need an extra layer of integral so the expression is much more cumbersome. Introduce $g = g(\tilde{z},\tilde{\tilde{\mu}})$, $h = h(\tilde{z},\tilde{x},\tilde{\tilde{\mu}},\tilde{\tilde{\nu}})$, and $w = w(\tilde{\tilde{z}},\tilde{\tilde{\mu}},\tilde{\tilde{\nu}})$ by

$$g = \Pi_{i=1}^k\big[(1-\epsilon_p) + \epsilon_p e^{\tau_p\tilde{z}_i - \frac{1}{2}\tau_p^2 - \sqrt{n}\tau_p(m_i,\tilde{\tilde{\mu}})}\big],$$

$$hg = \Pi_{i=1}^k\big[(1-\epsilon_p) + \epsilon_p e^{\tau_p\tilde{z}_i + (\tau_p/\sqrt{n})\tilde{x}_i - \frac{1}{2}\tau_p^2 - \frac{1}{2n}\tau_p^2 - \sqrt{n}\tau_p(m_i,\tilde{\tilde{\mu}}) - (\tau_p/\sqrt{n})(m_i,\tilde{\tilde{\nu}})}\big],$$

and

$$w = e^{\sqrt{n}\tilde{\tilde{\mu}}'\tilde{\tilde{z}} + \tilde{\tilde{v}}'\tilde{\tilde{x}} - \frac{n}{2}\tilde{\tilde{\mu}}'D\tilde{\tilde{\mu}} - \frac{1}{2}\tilde{\tilde{\nu}}'D\tilde{\tilde{\nu}}}.$$

Here, we have suppressed the expressions of g, h, and w as long as there is no confusion. Since that $\tilde{\nu}$ and $\tilde{\tilde{\mu}}$ are independent and that the entries of $\sqrt{n}\tilde{\nu}$ are iid samples from $(1-\epsilon_p)\nu_0 + \epsilon_p\nu_{\tau_p}$, integrating over $\tilde{\nu}$ gives

$$\begin{aligned} I &= \int e^{\sqrt{n}\tilde{\nu}'\tilde{z} + \sqrt{n}\tilde{\tilde{\mu}}'\tilde{\tilde{z}} + \tilde{\tilde{\nu}}'\tilde{\tilde{x}} - \frac{n}{2}\|\tilde{\nu}\|^2 - n\tilde{\nu}'B\tilde{\tilde{\mu}} - \frac{n}{2}\tilde{\tilde{\mu}}'D\tilde{\tilde{\mu}} - \frac{1}{2}\tilde{\tilde{\nu}}'D\tilde{\tilde{\nu}}} dF(\tilde{\nu})dF(\tilde{\tilde{\mu}}) \\ &= \int \left(\Pi_{i=1}^k\big[(1-\epsilon_p) + \epsilon_p e^{\tau_p\tilde{z}_i - \frac{1}{2}\tau_p^2 - \sqrt{n}\tau_p(m_i,\tilde{\tilde{\mu}})}\big]\right) e^{\sqrt{n}\tilde{\tilde{\mu}}'\tilde{\tilde{z}} + \tilde{\tilde{v}}'\tilde{\tilde{x}} - \frac{n}{2}\tilde{\tilde{\mu}}'D\tilde{\tilde{\mu}} - \frac{1}{2}\tilde{\tilde{\nu}}'D\tilde{\tilde{\nu}}} dF(\tilde{\tilde{\mu}}). \end{aligned}$$

By definitions, this implies that $I = \int gwdF(\tilde{\tilde{\mu}})$. Similarly, $II = \int hgwdF(\tilde{\tilde{\mu}})$.

Recall that $H(f_0,f_1)$ is the Hellinger distance between f_0 and f_1. Let E_0 be the expectation under the law that $X_1,\dots,X_n,X$ are iid from $N(0,\Omega^{-1})$. By Hölder inequality, $H(f_0,f_1)\le E_0[(\int(h-1)gwdF(\tilde{\tilde{\mu}}))^2/(\int gwdF(\tilde{\tilde{\mu}}))] \le E_0[\int(h-1)^2 gwdF(\tilde{\tilde{\mu}})]$. Since $E_0[\int hgwdF(\tilde{\tilde{\mu}})] = 1$ and $E_0[\int gwdF(\tilde{\tilde{\mu}})] = 1$, it is seen

$$H(f_0,f_1) \le E_0[\int h^2 gwdF(\tilde{\tilde{\mu}})] - 1. \tag{A.6}$$

Note that $h^2 g$ does not depend on $\tilde{\tilde{x}}$ and $\tilde{\tilde{z}}$. Also, note that $(\tilde{\tilde{x}}|\tilde{x})$ and $(\tilde{\tilde{z}}|\tilde{z})$ are the realizations of two (conditional) random vectors that are independent of each other and that distributed as $N(B'\tilde{x}, D - B'B)$ and $N(B'\tilde{z}, D - B'B)$, respectively. It follows that $E\big[w\big|(\tilde{x}, \tilde{z})\big] = \exp(\sqrt{n}\tilde{\tilde{\mu}}'B'\tilde{z} - \frac{n}{2}\tilde{\tilde{\mu}}'B'B\tilde{\tilde{\mu}} + \tilde{\tilde{\nu}}'B'\tilde{x} - \frac{1}{2}\tilde{\tilde{\nu}}'B'B\tilde{\tilde{\nu}})$, where w should also be interpreted as a random vector, not a realization of the random vector; we misuse the notation a little bit so that we don't have to introduce a new notation. Denote the right hand side by $v = v(\tilde{x}, \tilde{z}, \tilde{\tilde{\mu}}, \tilde{\tilde{\nu}})$. It follows that $E_0[\int h^2 gw dF(\tilde{\tilde{\mu}})] = E_0[\int h^2 gv dF(\tilde{\tilde{\mu}})]$. Combining this with (A.6) gives

$$H(f_0, f_1) \leq C\big(E_0[\int h^2 gv dF(\tilde{\tilde{\mu}})] - 1\big) \equiv C(IV - 1). \tag{A.7}$$

We now evaluate IV. Denote for short $a_i = (1 - \epsilon_p)$ and $b_i = \epsilon_p \exp(\tau_p \tilde{z}_i - \frac{\tau_p^2}{2} - \sqrt{n}\tau_p(m_i, \tilde{\tilde{\mu}}))$, $1 \leq i \leq k$. By direct calculations, IV equals to

$$E_0\Bigg[\int \Pi_{i=1}^k \Bigg(e^{\sqrt{n}(m_i,\tilde{\tilde{\mu}})\tilde{z}_i - \frac{n}{2}(m_i,\tilde{\tilde{\mu}})^2} \frac{\big[a_i + b_i e^{\frac{\tau_p}{\sqrt{n}}\tilde{x}_i - \frac{\tau_p^2}{2n} - \frac{\tau_p}{\sqrt{n}}(m_i,\tilde{\tilde{\nu}})}\big]^2}{a_i + b_i} e^{(\tilde{\tilde{\nu}},m_i)\tilde{x}_i - \frac{1}{2}(\tilde{\tilde{\nu}},m_i)^2}\Bigg) dF(\tilde{\tilde{\mu}})\Bigg]. \tag{A.8}$$

Recall that $\tilde{x}$ and $\tilde{z}$ denote the realizations of $k \times 1$ sub-vectors of ΩX and ΩZ, respectively, where two random vectors are independent of each other, and each is normally distributed with the mean vector being 0 and the covariance matrix being the identity matrix. It follows

$$E_0\big[(a_i + b_i e^{\frac{\tau_p}{\sqrt{n}}\tilde{x}_i - \frac{1}{2n}\tau_p^2 - \frac{\tau_p}{\sqrt{n}}(m_i,\tilde{\tilde{\nu}})})^2 e^{(m_i,\tilde{\tilde{\nu}})\tilde{x}_i - \frac{1}{2}(m_i,\tilde{\tilde{\nu}})^2}\big] = (a_i + b_i)^2 + (e^{\frac{\tau_p^2}{n}} - 1)b_i^2. \tag{A.9}$$

Denote for short $\sqrt{n}(m_i, \tilde{\tilde{\mu}}) = d_i \tau_p$. By definitions and direct calculations,

$$E_0\big[e^{\sqrt{n}(m_i,\tilde{\tilde{\mu}})\tilde{z}_i - \frac{n}{2}(m_i,\tilde{\tilde{\mu}})^2}(a_i + b_i)\big] = 1, \tag{A.10}$$

and

$$E_0\big[e^{\sqrt{n}(m_i,\tilde{\tilde{\mu}})\tilde{z}_i - \frac{n}{2}(m_i,\tilde{\tilde{\mu}})^2} \frac{b_i^2}{a_i + b_i}\big] = \epsilon_p^2 e^{\tau_p^2} \cdot E\Bigg[\frac{e^{(2+d_i)\tau_p z_i - (2+d_i)^2\tau_p^2/2}}{(1 - \epsilon_p) + \epsilon_p e^{\tau_p \tilde{z}_i - \frac{\tau_p^2}{2} - b_i \tau_p^2}}\Bigg]. \tag{A.11}$$

Inserting (A.9)-(A.11) into (A.8) gives

$$\begin{aligned} IV &= \int \Pi_{i=1}^k \Bigg(e^{\sqrt{n}(m_i,\tilde{\tilde{\mu}})\tilde{z}_i - \frac{n}{2}(m_i,\tilde{\tilde{\mu}})^2}\big[a_i + b_i + (e^{\tau_p^2/n} - 1)\frac{b_i^2}{a_i + b_i}\big]\Bigg) dF(\tilde{\tilde{\mu}}) \\ &= \int \Pi_{i=1}^k \Bigg[1 + (e^{\frac{\tau_p^2}{n}} - 1)\epsilon_p^2 e^{\tau_p^2} E\big[\frac{e^{(2+d_i)\tau_p z_i - (2+d_i)^2\tau_p^2/2}}{1 - \epsilon_p + \epsilon_p e^{\tau_p \tilde{z}_i - \frac{\tau_p^2}{2} - b_i \tau_p^2}}\big]\Bigg] dF(\tilde{\tilde{\mu}}). \end{aligned} \tag{A.12}$$

Write $\frac{\tau_p^2}{n}\epsilon_p^2 e^{\tau_p^2} E\big[e^{(2\tau_p+d_i)z_i-(2\tau_p+d_i)^2/2}/[(1-\epsilon_p)+\epsilon_p e^{\tau_p\tilde{z}_i-\frac{\tau_p^2}{2}-d_i\tau_p}]\big] = A_i + B_i$, where

$$A_i = \left(\frac{\tau_p^2}{n}\epsilon_p^2 e^{\tau_p^2}\right) E\left[e^{(2\tau_p+b_i)z_i-(2\tau_p+b_i)^2/2} 1_{\{z\leq t_p+b_i\}}\right] = \left(\frac{\tau_p^2}{n}\epsilon_p^2 e^{\tau_p^2}\right)\Phi(t_p-2\tau_p),$$

$$B_i = \left(\frac{\tau_p^2}{n}\epsilon_p\right) E\left[e^{(\tau_p+b_i)z_i-(\tau_p+b_i)^2/2} 1_{\{z>t_p+b_i\}}\right] = \left(\frac{\tau_p^2}{n}\epsilon_p\right)\bar{\Phi}(t_p-\tau_p),$$

and $t_p = [(r+\beta)/(2r)]\tau_p$. First, by Mills' ratio [4], $A_i \leq L_p p^{-2\beta+2r-\theta}$. Second, for B_i, noting that $t_p/\tau_p > 1$ in the range of interest, so $B_i \leq L_p p^{-(\beta+r)^2/(4r)-\theta}$. By our assumptions, there is a constant $c_0 = c_0(\beta, r, \theta) > 0$ such that $\min\{2\beta - 2r + \theta, \frac{(\beta+r)^2}{4r} + \theta\} \geq 1 + c_0$. Combining these gives

$$\left(\frac{\tau_p^2}{n}\epsilon_p^2 e^{\tau_p^2}\right) E\left[\frac{e^{(2\tau_p+d_i)z_i-(2\tau_p+d_i)^2/2}}{1-\epsilon_p+\epsilon_p e^{\tau_p\tilde{z}_i-\frac{\tau_p^2}{2}-d_i\tau_p}}\right] \leq L_p p^{-(1+c_0)}. \tag{A.13}$$

Inserting (A.13) into (A.12), $IV \leq 1 + p^{-c_0}$. Inserting this into (A.7) gives the claim. ☐

A.3. Proof of Lemmas 2.1-2.2. Before we prove these two lemmas, we need some preparations. Recall that $D_j = \{k : 1 \leq k \leq p, \Omega(j,k) \neq 0\}$ for $1 \leq j \leq p$. Introduce events $A_{0j} = \{\mu(k) = 0, \forall k \in D_j\}$, $A_{1j} = \{\mu(k) \neq 0 \text{ for exactly one } k \in D_j\}$, and $A_{2j} = \{\mu(k) \neq 0 \text{ for some } k \in D_j, k \neq j\}$. Let $\tilde{\mu} = \Omega\mu$. It is seen that

- Over the event A_{0j}, $\tilde{\mu}(j) = 0$.
- Over the event $A_{1j} \cap \{\mu(j) \neq 0\}$, $\sqrt{n_p}\tilde{\mu}(j) = \sqrt{n_p}\mu(j) = \tau_p$.
- Over the event $A_{1j} \cap \{\mu(j) = 0\}$, $\sqrt{n_p}|\tilde{\mu}(j)| \leq a\tau_p$.

Let $h_0(t) = h_0(t, \epsilon_p, \tau_p, \Omega) = p^{-1}\sum_{j=1}^p P(|\tilde{Z}(j)| \geq t; A_{0j})$, $h_1^+(t) = h_1^+(t, \epsilon_p, \tau_p, \Omega) = p^{-1}\sum_{j=1}^p P(\tilde{Z}(j) \geq t; A_{1j} \cap \{\mu(j) \neq 0\})$, $h_1^-(t) = h_1^-(t, \epsilon_p, \tau_p, \Omega) = p^{-1}\sum_{j=1}^p P(\tilde{Z}(j) \leq -t; A_{1j} \cap \{\mu(j) \neq 0\})$, and $g_2(t) = \frac{\sqrt{n_p}}{p\tau_p}\sum_{j=1}^p E[\tilde{\mu}(j)\mathrm{sgn}(\tilde{Z}(j)) \cdot 1\{|\tilde{Z}(j)| \geq t\}|A_{2j}]P(A_{2j})$. Further, recall that $g_1(t) = \frac{1}{p}\sum_{j=1}^p P(|\tilde{Z}(j)| \geq t, A_{2j})$. By definitions, it follows that

$$\widetilde{F}(t) = h_0(t) + h_1^+(t) + h_1^-(t) + g_1(t), \;\; m_p(t) = n_p^{-1/2} p\tau_p\big(h_1^+(t) - h_1^-(t) + g_2(t)\big). \tag{A.14}$$

Lemma A.2 below summarizes some basic properties of these quantities, the proof of which is elementary so we omit it.

LEMMA A.2. *For any $t > 0$, we have (a)* $(1 - K_p\epsilon_p)\bar{\Psi}(t) \leq h_0(t) \leq \bar{\Psi}(t)$, *(b)* $(1 - K_p\epsilon_p)\epsilon_p\bar{\Phi}(t - \tau_p) \leq h_1^+(t) \leq \epsilon_p\bar{\Phi}(t - \tau_p)$, $\quad (1 - K_p\epsilon_p)\epsilon_p\bar{\Phi}(t + \tau_p) < h_1^-(t) \leq \epsilon_p\bar{\Phi}(t + \tau_p)$, *(c)* $0 < g_1(t) \leq K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t) + (K_p\epsilon_p)^2\bar{\Psi}_{(1+a)\tau_p}(t) + C(K_p\epsilon_p)^3$, *(d)* $0 \leq g_2(t) \leq K_p g_1(t)$, *and (e)* $(1 - K_p\epsilon_p)(\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)) \leq \widetilde{F}(t)$.

Next, the following lemma is proved in APPENDIX B.

LEMMA A.3. *Fix $a \in (0, 1)$ and $\tau \geq 0$. Let (X, Y) be a bivariate normal distribution with mean vector $(0, \tau)'$, variance one and correlation ρ. Then there is a constant $C = C(a) > 0$ such that for all $\rho \in [-a, a]$, $P(|X| \geq t\big||Y| \geq t) \leq C(1 + t)\exp\big(-\frac{(1-a)t^2}{2(1+a)}\big)$.*

By Lemma A.3, we have the following lemma which is proved in Section A.4.

LEMMA A.4. *For any $t > 0$, we can write $v_p(t) = p\big(\widetilde{F}(t) + rem(t)\big)$, where the reminder term $rem(t)/\widetilde{F}(t)$ can be bounded from above by*

$$\begin{cases} L_p p^{-\min\{r, \frac{\beta-r}{2}, (1-a)(\beta-ar)\}} + L_p(1+t)\exp\big(-\frac{(1-a)t^2}{2(1+a)}\big), & r < \beta \text{ and } t \leq \tau_p + \tilde{s}_p, \\ K_p, & r \geq \beta \text{ or } t > \tau_p + \tilde{s}_p, \end{cases} \tag{A.15}$$

where $\tilde{s}_p = \sqrt{\max\{2(\beta - r), (\beta + r)\}\log p}$. Moreover, when $r < \beta$ and $t \leq \tau_p + \tilde{s}_p$, we have $v_p(t)/(p\widetilde{F}(t)) \geq 1 - o(1)$. In addition, if the smallest eigenvalue of Ω is bounded from below by $b > 0$, then $v_p(t)/[p\widetilde{F}(t)] \geq b$.

Recall that in (2.17) and (2.8), we define $W_0(t)$ and its proxy $\widetilde{W_0}(t)$, respectively. Define $a(t) = \sqrt{p}(W_0(t))^{-1}[h_1^+(t) + h_1^-(t) + g_1(t)](v_p(t))^{-1/2}$ and $S_1(t) = (v_p(t))^{-1/2}[\sqrt{p}(g_2(t) - g_1(t) - 2h_1^-(t))]$. Then $\widetilde{Sep}(t, \epsilon_p, \tau_p, \Omega) = 2\tau_p\sqrt{p/n_p}[a(t)W_0(t) + S_1(t)]$. The following two lemmas are proved in Sections A.5 and A.6, respectively.

LEMMA A.5. *Fix $(\beta, r) \in (0, 1)^2$ and $\Omega \in \mathcal{M}_p^*(a, K_p)$. Then*

$$\sup_{\{0<t\leq\tau_p+\tilde{s}_p\}} |S_1(t)| \leq L_p\big(p^{-3\beta/2} + p^{-(\beta+r)}\big) + L_p p^{-c_0(\beta,r,a)} \sup_{\{0<t<\infty\}} \widetilde{W_0}(t), \tag{A.16}$$

where $c_0(\beta, r, a)$ is defined in (2.12) and $\tilde{s}_p$ is defined in Lemma A.4. If in addition $\Omega \in \widetilde{\mathcal{M}}_p^*(a, b, K_p)$, then the above inequality holds with the left hand side replaced with $\sup_{\{t>0\}} |S_1(t)|$.

Also, if $r < \beta$ *and* $t \leq \tau_p + \tilde{s}_p$, *then* $|a(t) - 1| \leq L_p p^{-\min\{r, \frac{\beta - r}{2}, (1-a)(\beta - ar)\}} + L_p(1+t)\exp\big(-\frac{(1-a)t^2}{2(1+a)}\big)$; *and if in addition* $\Omega \in \widetilde{\mathcal{M}}_p^*(a, b, K_p)$, *then* $K_p^{-1/2} \lesssim a(t) \lesssim b^{-1/2}$.

Lemma A.6. *Fix* $(r, \beta) \in (0,1)^2$. *Then*

$$
\sup_{\{t>0\}} |W_0(t) - \widetilde{W}_0(t)| \leq L_p p^{-3\beta/2} + 2 \sup_{\{t>0\}} \frac{K_p \epsilon_p \bar{\Psi}_{a\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + K_p \epsilon_p \bar{\Psi}_{a\tau_p}(t)}} = L_p p^{-3\beta/2} + L_p p^{-c_0(\beta, r, a)} \sup_{\{t>0\}} \widetilde{W}_0(t), \tag{A.17}
$$

where $c_0(\beta, r, a)$ *is defined in (2.12).*

We will also need the following lemma, which is proved in APPENDIX B.

Lemma A.7. *Let* $t_p(q) = \sqrt{2q\log p}$ *with* $q \in (0,1)$. *If* $r < \beta$, *then as* $p \to \infty$,

$$
\sup_{0<q<1} \Big\{ (1+t_p(q)) \exp\big(-\frac{(1-a)t_p(q)^2}{2(1+a)}\big) \widetilde{W}_0\big(t_p(q)\big) \Big\} \sim L_p p^{-\tilde{c}_0(\beta, r, a)} \sup_{0<q<1} \widetilde{W}_0(t_p(q)).
$$

We now prove Lemma 2.1 and Lemma 2.2 separately.

Consider Lemma 2.1 first. Write for short $\widetilde{Sep}(t) = \widetilde{Sep}(t, \epsilon_p, \tau_p, \Omega)$. We consider the two cases 1) $t > \tau_p + \tilde{s}_p$ and 2) $t \leq \tau_p + \tilde{s}_p$ separately, where $\tilde{s}_p$ is as in Lemma A.4.

First consider case 1). We will show that (1a) $\widetilde{Sep}(t) \leq L_p p^{\frac{1-\theta}{2} - \max\{\beta - \frac{1}{2}r, \frac{3\beta + r}{4}\}}$ and (1b) $\widetilde{W}_0(t) \leq L_p p^{-\max\{\beta - \frac{1}{2}r, \frac{3\beta+r}{4}\}}$. Then combining (1a) and (1b) completes the proof of the lemma in case 1). We now proceed to prove (1a) and (1b). The result (1b) follows immediately from the definition of $\widetilde{W}_0(t)$ and the inequalities $\widetilde{W}_0(t) \leq \sqrt{\epsilon_p \bar{\Psi}_{\tau_p}(t)} \leq L_p p^{-\max\{4\beta - 2r, 3\beta + r\}/4}$. It remains to prove (1a). Let η be a $p \times 1$ vector such that $\eta(j) = 1\{(\Omega \hat{\mu}_t^{\tilde{Z}})(j) \neq 0\}$, $1 \leq j \leq p$. Also, for any $p \times 1$ vectors x and y, let $x \circ y$ be the $p \times 1$ vector such that $(x \circ y)(j) = x(j)y(j)$, $1 \leq j \leq p$. By definition, it is seen that $m_p(t) = E[\tilde{M}_p(t)] = E[(\hat{\mu}_t^{\tilde{Z}})'\Omega\mu] = E[(\hat{\mu}_t^{\tilde{Z}})'\Omega(\mu \circ \eta)]$. Using Cauchy-Schwartz inequality, $m_p(t) \leq \big(E[(\hat{\mu}_t^{\tilde{Z}})'\Omega\hat{\mu}_t^{\tilde{Z}}]\big)^{1/2}\big(E[(\mu\circ\eta)'\Omega(\mu\circ\eta)]\big)^{1/2}$. Recalling that $v_p(t) = E[\tilde{V}_p(t)] = E[(\hat{\mu}_t^{\tilde{Z}})'\Omega\hat{\mu}_t^{\tilde{Z}}]$, it follows that

$$
|\widetilde{Sep}(t)| = 2m_p(t)(v_p(t))^{-1/2} \leq 2\big(E[(\mu\circ\eta)'\Omega(\mu\circ\eta)]\big)^{1/2}. \tag{A.18}
$$

Since the largest eigenvalue of Ω is no greater than K_p, the last term above $\leq 2K_p^{1/2}(E\|\mu\circ\eta\|^2)^{1/2}$ and so $|\widetilde{Sep}(t)| \leq 2K_p^{1/2}(E\|\mu\circ\eta\|^2)^{1/2}$. It remains to study $E\|\mu\circ\eta\|^2$. By definition,

$$E\|\mu\circ\eta\|^2 = \sum_{i=1}^p \frac{\tau_p^2}{n_p} P(\mu(i)\neq 0, (\Omega\hat{\mu}_t^{\tilde{Z}})(i)\neq 0) \leq \frac{\tau_p^2}{n_p}\sum_{i=1}^p\sum_{j\in D_i} P(\mu(i)\neq 0, \hat{\mu}_t^{\tilde{Z}}(j)\neq 0)$$

$$= \frac{\tau_p^2}{n_p}\sum_{i=1}^p\sum_{j\in D_i} P(\mu(i)\neq 0, |\tilde{Z}(j)|\geq t) \leq L_p p^{1-\theta}(\epsilon_p\bar{\Psi}_{\tau_p}(t) + \epsilon_p\bar{\Psi}_{a\tau_p}(t) + CK_p\epsilon_p^2).$$

Since we consider the range $t > \tau_p + \tilde{s}_p$, the above expectation can be bounded as $E\|\mu\circ\eta\|^2 \leq L_p p^{1-\theta-\max\{4\beta-2r, 3\beta+r\}/2}$. Inserting this into (A.18) we complete the proof of (1a).

Now we consider the case 2). Recall that $\widetilde{Sep}(t) = 2\tau_p\sqrt{p/n_p}[a(t)W_0(t) + S_1(t)]$. Noting that $n_p = p^\theta$, the key is to show

$$\begin{aligned}\text{(A.19)}\quad &\sup_{\{0<t\leq\tau_p+\tilde{s}_p\}}\left|(2\tau_p)^{-1}p^{(\theta-1)/2}\widetilde{Sep}(t) - W_0(t)\right| \leq L_p p^{-3\beta/2} + L_p p^{-\beta-r}\\ &+ L_p\Big(p^{-\min\{r, \frac{\beta-r}{2}, (1-a)(\beta-ar)\}} + p^{-c_0(\beta,r,a)} + p^{-\tilde{c}_1(\beta,r,a)}\Big)\sup_{\{t>0\}}\widetilde{W}_0(t)\Big).\end{aligned}$$

In fact, once this is proved, the claim follows by using Lemma A.6.

We now show (A.19). By Lemma A.5,

$$\begin{aligned}\text{(A.20)}\qquad &\sup_{\{0<t\leq\tau_p+\tilde{s}_p\}}|p^{(\theta-1)/2}(2\tau_p)^{-1}\widetilde{Sep}(t) - W_0(t)|\\ &\leq \sup_{\{0<t\leq\tau_p+\tilde{s}_p\}}|a(t)-1|W_0(t) + \sup_{\{0<t\leq\tau_p+\tilde{s}_p\}}|S_1(t)|.\end{aligned}$$

The second term on the right was studied in Lemma A.5 inequality (A.16). We now study the first term on the right. By lemma A.5,

$$\text{(A.21)}\qquad \sup_{\{0<t\leq\tau_p+\tilde{s}_p\}}|a(t)-1|W_0(t) \leq \sup_{\{t\geq 0\}} I_1(t) + \sup_{\{t\geq 0\}} I_2(t),$$

where $I_1(t) = L_p\big(p^{-\min\{r, \frac{\beta-r}{2}, (1-a)(\beta-ar)\}} + C(1+t)\exp(-\frac{1-a}{2(1+a)}t^2)\big)\widetilde{W}_0(t)$, and $I_2(t) = L_p|W_0(t) - \widetilde{W}_0(t)|$.

Consider $I_2(t)$ first. By Lemma A.6,

$$\text{(A.22)}\qquad \sup_{\{t\geq 0\}} I_2(t) \leq L_p\Big(p^{-3\beta/2} + p^{-c_0(\beta,r,a)}\sup_{\{0<t<\infty\}}\widetilde{W}_0(t)\Big).$$

Consider $I_1(t)$ next. Write $I_1(t) = I_{1a}(t) + I_{1b}(t)$, where $I_{1a}(t) = L_p\cdot p^{-\min\{r, \frac{\beta-r}{2}, (1-a)(\beta-ar)\}}\widetilde{W}_0(t)$ and $I_{1b}(t) = L_p(1+t)\exp(-\frac{1-a}{2(1+a)}t^2)\widetilde{W}_0(t)$.

We first study $I_{1b}(t)$. By Lemma A.7,

$$\sup_{\{0<t<\infty\}}\{(1+t)\exp(-\frac{(1-a)}{2(1+a)}t^2)\widetilde{W}_0(t)\} = L_p p^{-\tilde{c}_0(\beta,r,a)} \sup_{\{0<t<\infty\}} \widetilde{W}_0(t),$$

where $\tilde{c}_0(\beta, r, a)$ is defined in (2.12). Combining these results and comparing terms yields

$$\text{(A.23)}\quad \sup_{t>0} I_1(t) \le L_p\Big(p^{-\min\{r,\frac{\beta-r}{2},(1-a)(\beta-ar)\}}+p^{-\tilde{c}_0(\beta,r,a)}\Big) \sup_{\{0<t<\infty\}} \widetilde{W}_0(t).$$

Combing (A.23) and (A.22) with (A.21) yields

$$\sup_{\{0<t\le\tau_p+\tilde{s}_p\}} |a(t)-1|W_0(t) \le L_p p^{-3\beta/2}$$
$$+\Big(p^{-\min\{r,\frac{\beta-r}{2},(1-a)(\beta-ar)\}}+p^{-c_0(\beta,r,a)}+p^{-\tilde{c}_0(\beta,r,a)}\Big) \sup_{\{0<t<\infty\}} \widetilde{W}_0(t).$$

Inserting this and (A.16) into (A.20) shows the claim for the case where $t \le \tau_p + \tilde{s}_p$. This completes the proof of Lemma 2.1.

We now show Lemma 2.2. First, we consider (a)-(b). By Lemma A.5, $(2\tau_p)^{-1}\sqrt{n_p/p}\widetilde{Sep}(t) \le b^{-1/2}W_0(t)+S_1(t)$, where $W_0(t)$ is defined in (2.17), and $S_1(t)$ is as in Lemma A.5. The key is to prove that there is a constant $d_0 > 0$ such that for any fixed t satisfying either $0 \le t \le \sqrt{2\beta\log p} - d_0\log\log p/\sqrt{\log p}$ or $t > \tau_p + 2\sqrt{\log(K_p\log p)}$,

$$\text{(A.24)}\qquad W_0(t) \lesssim \tfrac{2\sqrt{b\epsilon_p}}{3K_p}, \qquad S_1(t) \lesssim \tfrac{\sqrt{b\epsilon_p}}{3K_p(\log p)}.$$

In fact, once these are proved, then

$$\text{(A.25)}\qquad \widetilde{Sep}(t) \le 2\tau_p p^{(1-\theta)/2}[b^{-1/2}W_0(t)+S_1(t)] \lesssim \frac{5}{3}\tau_p K_p^{-1} p^{(1-\theta-\beta)/2},$$

and parts (a)-(b) of the lemma follow.

We now show (A.24). Recall that by the proof of Lemmas A.5-A.6,

$$\text{(A.26)}\qquad |S_1(t)| \le L_p(p^{-3\beta/2}+p^{-\beta-r}) + \frac{CK_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}{\sqrt{\bar{\Psi}(t)+K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}},$$

$$\text{(A.27)}\qquad 0 < W_0(t) - \widetilde{W}_0(t) \le L_p p^{-3\beta/2} + \frac{CK_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}{\sqrt{\bar{\Psi}(t)+K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}};$$

note that the last terms in the above two inequalities are the same. We now consider the case $t \le \sqrt{2\beta\log p} - d_0\log\log p/\sqrt{\log p}$ and the case $t > \tau_p + 2\sqrt{\log(K_p\log p)}$ separately.

In the first case, by Mills's ratio [4], with the constant $d_0 > 0$ being appropriately chosen, $\bar{\Psi}(t) + K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t) \geq 9C^2b^{-1}K_p^4(\log p)^2\epsilon_p$ and $\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t) \geq 9b^{-1}K_p^2\epsilon_p$. As a result,

$$\frac{CK_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}} \leq \frac{\sqrt{b\epsilon_p}}{3K_p\log p}, \quad \widetilde{W}_0(t) = \frac{\epsilon_p\bar{\Psi}_{\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)}} \leq \frac{\sqrt{b\epsilon_p}}{3K_p}.$$

Inserting these into (A.26) and (A.27), the claim follows by noting that $\epsilon_p = p^{-\beta}$.

Consider the second case. In this case, $\epsilon_p\bar{\Psi}_{a\tau_p}(t) = o(\epsilon_p p^{-(1-a)^2 r})$. Thus,

$$\frac{K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}} \leq \sqrt{K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)} = o(K_p^{-1}(\log p)^{-1}\sqrt{\epsilon_p}),$$

and

$$\widetilde{W}_0(t) = \frac{\epsilon_p\bar{\Psi}_{\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)}} \leq \sqrt{\epsilon_p\bar{\Psi}_{\tau_p}(t)} \lesssim \sqrt{b\epsilon_p}/(3K_p).$$

Inserting these into (A.26) and (A.27) proves (A.24), the claim follows by similar reasons.

Next, consider (c). Write for short $s_p = \sqrt{2\beta\log p} - d_0\log\log p/\sqrt{\log p}$. Since the eigenvalue of Ω is bounded from above by K_p, by definition we have $v_p(t) \leq K_p p\widetilde{F}(t)$. Thus, $\widetilde{Sep}(t) = 2m_p(t)/\sqrt{v_p(t)} \geq 2K_p^{-1/2}m_p(t)/\sqrt{p\widetilde{F}(t)}$. By definitions in (A.14) and Lemma A.2 we can further obtain that

$$\widetilde{Sep}(t) \geq \frac{2\tau_p p^{\frac{1-\theta}{2}}(h_1^+(t) - h_1^-(t))}{\sqrt{K_p\widetilde{F}(t)}} \geq \frac{2\tau_p p^{\frac{1-\theta}{2}}[(1-K_p\epsilon_p)\epsilon_p\bar{\Phi}(t-\tau_p) - \epsilon_p\bar{\Phi}(t+\tau_p)]}{\sqrt{K_p\big(\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t) + K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t) + C(K_p\epsilon_p)^2\big)}}.$$

When $s_p \leq t \leq \tau_p$, the numerator above $\sim 2\tau_p p^{\frac{1-\theta}{2}-\beta}$, and the denominator above $\leq K_p p^{-\frac{\beta}{2}}$. Thus, $\widetilde{Sep}(t) \geq 2\tau_p K_p^{-1}p^{(1-\theta-\beta)/2}$. On the other hand, recall that $\sup_{t>0}\widetilde{W}_0(t) = L_p p^{-\beta/2}$ when $r \geq \beta$, which together with Lemmas A.5-A.6 ensures $\sup_{t>0}W_0(t) \leq L_p p^{-\beta/2}$ and $\sup_{t>0}S_1(t) \leq L_p p^{-\beta/2}$. Since $(2\tau_p)^{-1}\sqrt{n_p/p}\widetilde{Sep}(t) \leq b^{-1/2}W_0(t) + S_1(t)$, combining these entails $\widetilde{Sep}(t) \leq L_p p^{(1-\theta-\beta)/2}$. This shows part (c) and completes the proof of Lemma 2.2. □

A.4. Proof of Lemma A.4. The last claim follows trivially from the assumption on the minimum eigenvalue of Ω. And in the case of $r \geq \beta$, by definition of $v_p(t)$ and noting that the maximum eigenvalue of Ω is bounded by K_p, we obtain that $v_p(t) \leq K_p p\widetilde{F}(t)$. So we only need to prove the first claim in the case of $r < \beta$ and the second claim.

Consider the first claim. Let $D_i = \{j : \Omega(i,j) \neq 0\}$ and $\tilde{D}_i = D_i \setminus \{i\}$. Write $h(t) = \tilde{h}_0(t) + \tilde{h}_1(t)$, where $h(t) = p^{-1}\sum_{i=1}^p \sum_{j\in\tilde{D}_i} P(|\tilde{Z}(i)| \geq t, |\tilde{Z}(j)| \geq t)$, $\tilde{h}_0(t) = p^{-1}\sum_{i,j\in\tilde{D}_i} P(|\tilde{Z}(i)| \geq t, |\tilde{Z}(j)| \geq t, \tilde{\mu}(i) = 0 \text{ or } \tilde{\mu}(j) = 0)$, $\tilde{h}_1(t) = p^{-1}\sum_{i,j\in\tilde{D}_i} P(|\tilde{Z}(i)| \geq t, |\tilde{Z}(j)| \geq t, \tilde{\mu}(i) \neq 0 \text{ and } \tilde{\mu}(j) \neq 0)$. By definitions, it is seen that

$$v_p(t) = p(\widetilde{F}(t) + rem(t)), \text{ where } |rem(t)| \leq h(t) = \tilde{h}_0(t) + \tilde{h}_1(t). \tag{A.28}$$

To show the claim, it is sufficient to show that the ratio $[\tilde{h}_0(t) + \tilde{h}_1(t)]/\widetilde{F}(t)$ does not exceed the right hand side of (A.15).

First, consider $\tilde{h}_0(t)$. If at least one of $\tilde{Z}(i)$ and $\tilde{Z}(j)$ has mean 0, by Lemma A.3 and definitions, $P(|\tilde{Z}(i)| \geq t, |\tilde{Z}(j)| \geq t, \tilde{\mu}(i) = 0 \text{ or } \tilde{\mu}(j) = 0) \leq CK_p(1+t)\exp(-\frac{(1-a)t^2}{2(1+a)})\big(P(|\tilde{Z}(i)| \geq t) + P(|\tilde{Z}(j)| \geq t)\big)$. Since $\tilde{D}_i$ has at most K_p components, it follows from the definition of $\widetilde{F}(t)$ that

(A.29)
$$\begin{aligned}\tilde{h}_0(t) &\leq CK_p(1+t)\exp(-\frac{(1-a)t^2}{2(1+a)})p^{-1}\sum_{i,j\in\tilde{D}_i}\big(P(|\tilde{Z}(i)| \geq t) + P(|\tilde{Z}(j)| \geq t)\big)\\ &\leq CK_p^2(1+t)\exp\big(-\frac{(1-a)t^2}{2(1+a)}\big)\widetilde{F}(t).\end{aligned}$$

Next, consider $\tilde{h}_1(t)$. Define events $A_{1,ij} = \{\mu(k) \neq 0 \text{ for some } k \in D_i \setminus D_j\}$, $A_{2,ij} = \{\mu(k) \neq 0 \text{ for exactly one } k, \text{ which is in } D_i \cap D_j\}$, and $A_{3,ij} = \{\mu(k) \neq 0 \text{ for two or more } k, \text{ all of which are in } D_i \cap D_j\}$. It is seen that

$$\begin{aligned}\tilde{h}_1(t) &= p^{-1}\sum_{i,j\in\tilde{D}_i} P(|\tilde{Z}(i)| \geq t, |\tilde{Z}(j)| \geq t, \tilde{\mu}(i) \neq 0 \text{ and } \tilde{\mu}(j) \neq 0)\\ &= \tilde{h}_{1,1}(t) + \tilde{h}_{1,2}(t) + \tilde{h}_{1,3}(t),\end{aligned}$$

where $\tilde{h}_{1,1}(t) = p^{-1}\sum_{i,j\in\tilde{D}_i} P(|\tilde{Z}(i)| \geq t, |\tilde{Z}(j)| \geq t, A_{1,ij} \cap A_{1,ji})$, $\tilde{h}_{1,2}(t) = p^{-1}\sum_{i,j\in\tilde{D}_i} P(|\tilde{Z}(i)| \geq t, |\tilde{Z}(j)| \geq t, A_{2,ij})$, and $\tilde{h}_{1,3}(t) = p^{-1}\sum_{i,j\in\tilde{D}_i} P(|\tilde{Z}(i)| \geq t, |\tilde{Z}(j)| \geq t, A_{3,ij})$.

We first consider $\tilde{h}_{1,1}(t)$. Note that

$$\begin{aligned}&P(|\tilde{Z}(i)| \geq t, |\tilde{Z}(j)| \geq t, A_{1,ij} \cap A_{1,ji})\\ &\leq P(|\tilde{Z}(i)| \geq t, A_{1,ij})K_p\epsilon_p \leq K_p\epsilon_p P(|\tilde{Z}(i)| \geq t).\end{aligned}$$

Thus, $\tilde{h}_{1,1}(t) \le \epsilon_p K_p^2 p^{-1} \sum_{i=1}^p P(|\tilde{Z}(i)| \ge t) = L_p \epsilon_p \widetilde{F}(t)$.

Now we consider $\tilde{h}_{1,2}(t)$. For any $(i,j) \in A_{2,ij}$, we use $(\tilde{Z}^*(i), \tilde{Z}^*(j))$ to denote the demeaned pair of $(\tilde{Z}(i), \tilde{Z}(j))$. By definition there exists a k such that $\sqrt{n_p}\mu(k) = \tau_p$, $\tilde{\mu}(i) = \Omega(i,k)\mu(k)$ and $\tilde{\mu}(j) = \Omega(j,k)\mu(k)$. Thus, $|\sqrt{n_p}\tilde{\mu}(i)| \le a\tau_p$ or $|\sqrt{n_p}\tilde{\mu}(j)| \le a\tau_p$ and

$$P(|\tilde{Z}(i)| \ge t, |\tilde{Z}(j)| \ge t, A_{2,ij}) \le K_p\epsilon_p P(|\tilde{Z}^*(i)| \ge t - a\tau_p) = K_p\epsilon_p \bar{\Psi}_{a\tau_p}(t).$$

Then $\tilde{h}_{1,2}(t) \le K_p^2 \epsilon_p \bar{\Psi}_{a\tau_p}(t)$. Direct calculations yield

$$\frac{\tilde{h}_{1,2}(t)}{\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)} \le \frac{K_p^2\epsilon_p\bar{\Psi}_{a\tau_p}(t)}{\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)} \le L_p p^{-(1-a)(\beta - ar)}, \text{ for all } t \le \tau_p + \tilde{s}_p. \tag{A.30}$$

By Lemma A.2, $\widetilde{F}(t) \gtrsim \bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)$. It follows that $\tilde{h}_{1,2}(t) \le L_p p^{-(1-a)(\beta-ar)}\widetilde{F}(t)$.

Now, consider $\tilde{h}_{1,3}(t)$. Observe that $\tilde{h}_{1,3}(t) \le p^{-1}\sum_{i,j\in\tilde{D}_i} P(A_{3,ij}) \le K_p(K_p\epsilon_p)^2$. By Lemma A.2,

$$\frac{\tilde{h}_{1,3}(t)}{\widetilde{F}(t)} \le \frac{1}{1 - K_p\epsilon_p}\frac{K_p(K_p\epsilon_p)^2)}{\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)} \le \frac{CK_p^3\epsilon_p^2}{\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)}.$$

When $r < \beta$ and $t \le \tau_p + \tilde{s}_p$, we have $\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t) \ge L_p p^{-\max\{4\beta-2r, 3\beta+r\}/2}$, and thus $CK_p^3\epsilon_p^2/[\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)] \le L_p(p^{-(\beta-r)/2} + p^{-r})$. When $t > \tau_p + \tilde{s}_p$, by the definition of $v_p(t)$ and recalling that the largest eigenvalue of Ω is bounded by K_p, we have $v_p(t) \le K_p p\widetilde{F}(t)$. Combining these together and noting that $\widetilde{F}(t) \gtrsim \bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)$, we obtain $\tilde{h}_{1,3}(t)/\widetilde{F}(t) \le K_p$ if $t > \tau_p + \tilde{s}_p$, and $\tilde{h}_{1,3}(t)/\widetilde{F}(t) \le L_p(p^{-(\beta-r)/2} + p^{-r})$ if $t \le \tau_p + \tilde{s}_p$.

Combining the bounds on $\tilde{h}_{1,1}(t), \tilde{h}_{1,2}(t)$ and $\tilde{h}_{1,3}(t)$ entails that when $r < \beta$, $\tilde{h}_1(t)/\widetilde{F}(t) \le p^{-(\beta-r)/2} + p^{-r} + p^{-(1-a)(\beta-ar)}$ if $t \le \tau_p + \tilde{s}_p$ and $\tilde{h}_1(t)/\widetilde{F}(t) \le K_p$ if $t > \tau_p + \tilde{s}_p$. These together with (A.28) and (A.29) completes the proof of the first claim when $r < \beta$.

Next, we consider the second claim. The goal is to show that $v_p(t)/(p\widetilde{F}(t)) \gtrsim 1$, assuming $r < \beta$ and $t \le \tau_p + \tilde{s}_p$. We consider the cases (a) $d_3 \log\log p \le t \le \tau_p + \tilde{s}_p$ and (b) $t < d_3 \log\log p$ separately, where $d_3 > 0$ is a large constant.

In Case (a), using (A.15), it is seen that $|rem(t)|/\widetilde{F}(t)| = o(1)$, uniformly for all $d_3 \log\log p \le t \le \tau_p + \tilde{s}_p$. Using (A.28), $|v_p(t)/(p\widetilde{F}(t)) - 1| = o(1)$ and the claim follows.

In Case (b), recall that $v_p(t) = E[(\hat{\mu}_t^{\tilde{Z}})'\Omega\mu_t^{\tilde{Z}}]$, where $\hat{\mu}_t^{\tilde{Z}}(j) = \mathrm{sgn}(\tilde{Z}(j))1\{|\tilde{Z}(j)| \ge t\}$ and $\tilde{Z} = \Omega Z$. Write $\tilde{Z} = \sqrt{n_p}\tilde{\mu} + W$, where $\tilde{\mu} = \Omega\mu$ and $W \sim N(0, \Omega)$.

Let $\hat{\mu}_t$ be the counterpart of $\hat{\mu}_t^{\tilde{Z}}$ defined by $\hat{\mu}_t(j) = \mathrm{sgn}(W(j))1\{|W(j)| \geq t\}$. We claim (b1) $E[(\hat{\mu}_t^{\tilde{Z}})'\Omega\mu_t^{\tilde{Z}}] = E[(\hat{\mu}_t)'\Omega\hat{\mu}_t] + O(L_p p^{1-\beta/2})$ and (b2) $E[(\hat{\mu}_t)'\Omega\hat{\mu}_t] \geq p\widetilde{F}(t)$. The claim follows by combining (b1) and (b2) and noting that $p\widetilde{F}(t) \geq L_p p(1 - K_p\epsilon_p)$ when $t \leq d_3 \log\log p$.

Consider (b1). Let $S = \{1 \leq i \leq p : \hat{\mu}_t^{\tilde{Z}}(i) \neq \hat{\mu}_t(i)\}$. Note that for all $p \times 1$ vectors ξ and η, by Schwartz inequality and that the spectral norm of $\Omega \leq K_p$, $|(\xi+\eta)'\Omega(\xi+\eta) - \eta'\Omega\eta| \leq \xi'\Omega\xi + 2[(\xi'\Omega\xi)\cdot(\eta'\Omega\eta)]^{1/2} \leq L_p[\|\xi\|^2 + \|\xi\|\|\eta\|]$. Applying this with $\eta = \hat{\mu}_t$, $\xi = \hat{\mu}_t^{\tilde{Z}} - \hat{\mu}_t$, and noting that each coordinate of $\hat{\mu}_t^{\tilde{Z}} - \hat{\mu}_t$ has magnitude no greater than 2, we claim that $|E[(\hat{\mu}_t^{\tilde{Z}})'\Omega\mu_t^{\tilde{Z}}] - E[(\hat{\mu}_t)'\Omega\hat{\mu}_t]| \leq L_p E[|S| + \sqrt{p|S|}] \leq L_p E[\sqrt{p|S|}]$. Note that for any $i \in S$, we must have $\tilde{\mu}(i) \neq 0$. Therefore, by definitions, $|S| \leq \sum_{i=1}^p 1\{(\Omega\mu)(i) \neq 0\} \leq \sum_{i=1}^p \sum_{j:\Omega(i,j)\neq 0} 1\{\mu(j) \neq 0\} \leq K_p \sum_{i=1}^p 1\{\mu(i) \neq 0\}$, where we have used the assumption that Ω is K_p-sparse. Note that $\sum_{i=1}^p 1\{\mu(i) \neq 0\} \sim \mathrm{Binomial}(p, \epsilon_p)$, where $\epsilon_p = p^{-\beta}$, so $E[\sqrt{p|S|}] \sim p^{1-\beta/2}$. Combining these gives (b1).

Consider (b2). Denoting $B = E[\hat{\mu}_t\hat{\mu}_t']$, we have $E[(\hat{\mu}_t)'\Omega\hat{\mu}_t] = E[\Omega\hat{\mu}_t\hat{\mu}_t'] = tr(\Omega B)$. We claim that for any $i \neq j$ such that $\Omega(i,j) \neq 0$, $B(i,j)$ has the same sign as that of $\Omega(i,j)$. To see the point, write $B(i,j) = E[\mathrm{sgn}(\tilde{Z}(i))\mathrm{sgn}(\tilde{Z}(j))\cdot 1\{|\tilde{Z}(i| > t, |\tilde{Z}(j)| > t\}$. By symmetry and basic statistics, $B(i,j) = 2[P(\tilde{Z}(j) > t, \tilde{Z}(j) > t|\Omega(i,j)) - P(\tilde{Z}(i) > t, \tilde{Z}(j) > t| - \Omega(i,j))]$, where for any $\rho \in (-1,1)$, $P(\tilde{Z}(i) > t, \tilde{Z}(j) > t|\rho)$ is evaluated at the law that $\mathrm{corr}(\tilde{Z}(i), \tilde{Z}(j)) = \rho$. The claim follows by noting that for any $\rho > 0$, $P(\tilde{Z}(j) > t, \tilde{Z}(j) > t|\rho) > P(\tilde{Z}(i) > t)P(\tilde{Z}(j) > t) > P(\tilde{Z}(i) > t, \tilde{Z}(j) > t| - \rho)$. As a result, $tr(\Omega B) \geq tr(B) \equiv p\widetilde{F}(t)$, where we have used the fact that the diagonals of Ω are ones. This proves (b2). □

A.5. Proof of Lemma A.5. Consider the first claim. By Lemma A.2 (part (d)), $|g_2(t)| \leq K_p g_1(t)$. So by definitions,

$$|S_1(t)| \leq (K_p+1)\frac{\sqrt{p}g_1(t)}{\sqrt{v_p(t)}} + \frac{2\sqrt{p}h_1^-(t)}{\sqrt{v_p(t)}} \equiv (K_p+1)B_0(t) + B_1(t). \tag{A.31}$$

Consider $B_0(t)$ first. Rewrite $B_0(t) = [g_1(t)/\sqrt{\widetilde{F}(t)}]\sqrt{p\widetilde{F}(t)/v_p(t)}$. Note that when $r < \beta$ and $t \leq \tau_p + \tilde{s}_p$, $p\widetilde{F}(t)/v_p(t) \lesssim 1$, and when $r \geq \beta$ and $\Omega \in \widetilde{M}^*(a,b,K_p)$, by the last claim of Lemma A.4, $p\widetilde{F}(t)/v_p(t) \leq b^{-1}$. This says that $p\widetilde{F}(t)/v_p(t) \leq C$ and so $B_0(t) \leq Cg_1(t)/\sqrt{\widetilde{F}(t)}$, where $C > 0$ is a generic constant. At the same time, by definitions and Lemma A.2, $\widetilde{F}(t) = h_0(t) + h_1^+(t) + h_1^-(t) + g_1(t) \geq (1 - K_p\epsilon_p)[\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)] + g_1(t)$, so we have

$$B_0(t) \leq C g_1(t) / \sqrt{\bar{\Psi}(t) + \epsilon_p \bar{\Psi}_{\tau_p}(t) + g_1(t)}.$$

Finally, using Lemma A.2 and noting that $x/\sqrt{A+x}$ is an increasing function in $x \in (0,\infty)$ for any number $A > 0$, we obtain

$$B_0(t) \leq \frac{C\big(K_p\epsilon_p \bar{\Psi}_{a\tau_p}(t) + (K_p\epsilon_p)^2 \bar{\Psi}_{(1+a)\tau_p}(t) + (K_p\epsilon_p)^3\big)}{\sqrt{\bar{\Psi}(t) + \epsilon_p \bar{\Psi}_{\tau_p}(t) + K_p\epsilon_p \bar{\Psi}_{a\tau_p}(t) + (K_p\epsilon_p)^2 \bar{\Psi}_{(1+a)\tau_p}(t) + (K_p\epsilon_p)^3}}.$$

where the right hand side $\leq I + II + C(K_p\epsilon_p)^{3/2}$, with

$$I = \frac{C K_p \epsilon_p \bar{\Psi}_{a\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + K_p\epsilon_p \bar{\Psi}_{a\tau_p}(t)}}, \; II = \frac{C(K_p\epsilon_p)^2 \bar{\Psi}_{(1+a)\tau_p}(t)}{\sqrt{\epsilon_p \bar{\Psi}_{\tau_p}(t) + K_p\epsilon_p \bar{\Psi}_{a\tau_p}(t) + (K_p\epsilon_p)^2 \bar{\Psi}_{(1+a)\tau_p}(t)}}.$$

The above two terms have been considered in Lemma A.6 (see the last two terms of (A.34)). Using the results over there we can show that

$$\text{(A.32)} \qquad \sup_{\{0<t\leq \tilde{s}_p\}} B_0(t) \leq L_p p^{-3\beta/2} + L_p p^{-c_0(\beta,r,a)} \sup_{\{0<t<\infty\}} \widetilde{W}_0(t).$$

Next we consider $B_1(t)$. Write $B_1(t) = 2\cdot[(p\widetilde{F}(t)/v_p(t))^{1/2}]\cdot[h_1^-(t)(\widetilde{F}(t))^{-1/2}]$. We have just proved $p\widetilde{F}(t)/v_p(t) \leq C$ when $r \geq \beta$ or $0 < t \leq \tau_p + \tilde{s}_p$ with $C > 0$ some generic constant. At the same time, using (A.14) and parts (a)-(b) of Lemma A.2, first, $h_1^-(t) \leq \epsilon_p \bar{\Phi}(t+\tau_p)$, and second, $\widetilde{F}(t) \geq h_0(t) + h_1^+(t) + h_1^-(t) \geq (1 - K_p\epsilon_p)[\bar{\Psi}(t) + \epsilon_p \bar{\Psi}_{\tau_p}(t)]$. Combining these gives $h_1^-(t)(\widetilde{F}(t))^{-1/2} \leq C\epsilon_p \bar{\Phi}(t+\tau_p)/\sqrt{\bar{\Psi}(t) + \epsilon_p \bar{\Psi}_{\tau_p}(t)}$. It follows that $B_1(t) \leq C\epsilon_p \bar{\Phi}(t+\tau_p)/\sqrt{\bar{\Psi}(t) + \epsilon_p \bar{\Psi}_{\tau_p}(t)}$. This together with direct calculations yields

$$\text{(A.33)} \qquad \sup_{0<t\leq \tilde{s}_p} B_1(t) \leq C\epsilon_p \sup_{0<t<\infty} \frac{\bar{\Phi}(t+\tau_p)}{\sqrt{\bar{\Psi}(t) + \epsilon_p \bar{\Psi}_{\tau_p}(t)}} = C p^{-(\beta+r)}.$$

Inserting (A.32) and (A.33) into (A.31) completes the proof.

Consider the last two claims. Write $a(t) = A_1 \cdot A_2 \cdot A_3$, where

$$A_1 = \frac{h_1^+(t) + h_1^-(t) + g_1(t)}{\epsilon_p \bar{\Psi}_{\tau_p}(t) + g_1(t)}, \quad A_2 = \Big(\frac{\bar{\Psi}(t) + \epsilon_p \bar{\Psi}_{\tau_p}(t) + g_1(t)}{\widetilde{F}(t)}\Big)^{1/2},$$

and $A_3 = \big(p\widetilde{F}(t)/v_p(t)\big)^{1/2}$. First, by Lemma A.2 (part (b)), $\epsilon_p(1-K_p\epsilon_p)\bar{\Psi}_{\tau_p}(t) \leq h_1^+(t) + h_1^-(t) \leq \epsilon_p \bar{\Psi}_{\tau_p}(t)$ and thus $1 - K_p\epsilon_p \leq A_1 \leq 1$. Second, similarly, by Lemma A.2, $1 \leq A_2 \leq (1 - K_p\epsilon_p)^{-1/2}$. Since by basis algebra,

$|AB-1| \leq |A-1| + |B-1| + |A-1||B-1|$ for any numbers A and B, we have $|a(t)-1| \leq CK_p\epsilon_p(1+|A_3-1|) + |A_3-1|$. Now, by Lemma A.4, $|A_3-1| \leq L_p\Big(p^{-\min\{r, \frac{\beta-r}{2}, (1-a)(\beta-ar)\}} + (1+t)\exp\big(-\frac{(1-a)t^2}{2(1+a)}\big)\Big)$ when $r<\beta$ and $0<t\leq \tau_p+\tilde{s}_p$, and $K_p^{-1/2} \leq A_3(t) \leq b^{-1/2}$ when $\Omega \in \widetilde{\mathcal{M}}_p^*(a,b,K_p)$, and so the claim follows. □

A.6. Proof of Lemma A.6. Recall that $W_0(t) = \frac{\epsilon_p\bar{\Psi}_{\tau_p}(t)+g_1(t)}{\sqrt{\bar{\Psi}(t)+\epsilon_p\bar{\Psi}_{\tau_p}(t)+g_1(t)}}$, where $g_1(t)$ is as in Lemma A.2. We will compare $W_0(t)$ with $\widetilde{W}_0(t)$ defined in (2.8). On one hand, since $(A+x)/\sqrt{B+x}$ is an increasing function of x when $0\leq A<B$, it is seen that $W_0(t) \geq \widetilde{W}_0(t)$. On the other hand, writing for short $b(t) = K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t) + (K_p\epsilon_p)^2\bar{\Psi}_{(1+a)\tau_p}(t)$, it follows from Lemma A.2(c) that

(A.34)
$$
\begin{aligned}
W_0(t) &\leq \frac{\epsilon_p\bar{\Psi}_{\tau_p}(t) + b(t) + CK_p^3\epsilon_p^3}{\sqrt{\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t) + b(t) + CK_p^3\epsilon_p^3}} \\
&\leq \widetilde{W}_0(t) + CK_p^{3/2}p^{-3\beta/2} + \frac{K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}} + \frac{K_p^2\epsilon_p^2\bar{\Psi}_{(1+a)\tau_p}(t)}{\sqrt{\epsilon_p\bar{\Psi}_{\tau_p}(t) + b(t)}}.
\end{aligned}
$$

Combining these and recalling $\epsilon_p = p^{-\beta}$, we have

$$
\sup_{0<t<\infty} |W_0(t) - \widetilde{W}_0(t)| \leq L_p p^{-3\beta/2} + I + II,
$$

where

$$
I = \sup_{0<t<\infty} \frac{K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}}, \qquad II = \sup_{0<t<\infty} \frac{K_p^2\epsilon_p^2\bar{\Psi}_{(1+a)\tau_p}(t)}{\sqrt{\epsilon_p\bar{\Psi}_{\tau_p}(t) + b(t)}}.
$$

To show the first inequality of claim, it is sufficient to show
(A.35)
$$
II \leq L_p p^{-3\beta/2} + L_p p^{-\beta/2} \sup_{0<t<\infty} \frac{K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}} \equiv L_p p^{-3\beta/2} + L_p p^{-\beta/2}\cdot I.
$$

Towards this end, we write $II \leq IIa + IIb$, where IIa and IIb are the supremums of $K_p^2\epsilon_p^2\bar{\Psi}_{(1+a)\tau_p}(t)/\sqrt{\epsilon_p\bar{\Psi}_{\tau_p}(t)+b(t)}$ over the intervals $0<t<\tau_p$ and $\tau_p \leq t < \infty$, respectively. Consider IIa. When $0\leq t\leq \tau_p$, $\bar{\Psi}_{\tau_p}(t) \geq$

$1/2$, and so $IIa \leq K_p^2 \epsilon_p^2 \sup_{\{0<t<\tau_p\}} \frac{\bar{\Psi}_{(1+a)\tau_p}(t)}{\sqrt{\epsilon_p \bar{\Psi}_{\tau_p}(t)}} \leq L_p \epsilon_p^{3/2}$. Consider IIb. By definitions and change-of-variable, and recalling $\epsilon_p = p^{-\beta}$,

$$IIb \leq \sup_{\{\tau_p \leq t < \infty\}} \frac{K_p^2 \epsilon_p^2 \bar{\Psi}_{(1+a)\tau_p}(t)}{\sqrt{\epsilon_p \bar{\Psi}_{\tau_p}(t) + K_p^2 \epsilon_p^2 \bar{\Psi}_{(1+a)\tau_p}(t)}} = \sup_{\{0 \leq t < \infty\}} \frac{K_p^2 \epsilon_p^{3/2} \bar{\Psi}_{a\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + K_p^2 \epsilon_p \bar{\Psi}_{a\tau_p}(t)}}$$
$$\leq L_p \epsilon_p^{1/2} \cdot I = L_p p^{-\beta/2} \cdot I.$$

Combining these proves (A.35). Consequently, the first inequality of the claim follows.

To show the second inequality in the claim, we use similar calculations as in [1] and get

$$\sup_{\{0 \leq t < \infty\}} \{\widetilde{W}_0(t)\} = L_p p^{-\delta(r,\beta)}, \; I = L_p p^{-\delta(a^2 r, \beta)} \equiv L_p p^{-c_0(\beta, r, a)} \sup_{0<t<\infty} \widetilde{W}_0(t),$$

where we have used $c_0(\beta, r, a) = \delta(\beta, a^2 r) - \delta(\beta, r)$ as in (2.10). □

A.7. Proof of Lemmas 2.3-2.4. Write for short $W(t) = p^{-1/2} HC(t, \widetilde{F})$. Recalling $W_0(t) = [\epsilon_p \bar{\Psi}_{\tau_p}(t) + g_1(t)] / \sqrt{\bar{\Psi}(t) + \epsilon_p \bar{\Psi}_{\tau_p}(t) + g_1(t)}$ as defined in (2.17), where $g_1(t)$ is as in Lemma A.2, we let $a_1(t) = \big(W_0(t)\big)^{-1} [\widetilde{F}(t) - h_0(t)] \cdot (\widetilde{F}(t)(1 - \widetilde{F}(t))^{-1/2}$, and $W_1(t) = [\bar{\Psi}(t) - h_0(t)] \cdot (\widetilde{F}(t)(1 - \widetilde{F}(t))^{-1/2}$, where $h_0(t)$ is as in Lemma A.2. By these notations, $W(t) = a_1(t) W_0(t) - W_1(t)$. The following lemma is proved in Section A.8.

Lemma A.8. *Fix a sufficiently large p. There is a universal constant $C > 0$ such that for all $\Omega \in \mathcal{M}_p^*(a, K_p)$,*

(A.36)
$$0 < W_1(t) \leq C K_p \epsilon_p \bar{\Psi}(t) / \sqrt{\bar{\Psi}(t) + \epsilon_p \bar{\Psi}_{\tau_p}(t)}, \; \textit{for all } t \geq \bar{\Psi}^{-1}(1/2)$$

(A.37)
$$1 - C K_p \epsilon_p \leq a_1(t) \leq (1 + C K_p \epsilon_p)(1 - \bar{\Psi}(t) - K_p \epsilon_p)^{-1/2}, \; \textit{for all } t \geq 0.$$

Consider Lemma 2.3. Using Lemma A.8, $|a_1(t) - 1| \leq C(K_p \epsilon_p + \bar{\Psi}(t))$ for all $t \geq 0$. Recalling $W(t) = a_1(t) W_0(t) - W_1(t)$, we have

(A.38)
$$\sup_{\{t \geq \bar{\Psi}^{-1}(\frac{1}{2})\}} |W(t) - W_0(t)| \leq \sup_{\{t \geq 0\}} \{|a_1(t) - 1| W_0(t)\} + \sup_{\{t \geq \bar{\Psi}^{-1}(\frac{1}{2})\}} W_1(t)$$
$$\leq L_p (I + II + III),$$

where $I = K_p\epsilon_p \sup_{\{t\geq 0\}}\{W_0(t)\}$, $II = \sup_{\{t\geq 0\}}\{\bar{\Psi}(t)W_0(t)\}$, and $III = \sup_{\{t\geq \bar{\Psi}^{-1}(\frac{1}{2})\}}\{W_1(t)\}$.

First, consider I. By basic algebra and Lemma A.6,

$$I \leq L_p\epsilon_p[\sup_{\{t\geq 0\}} \widetilde{W}_0(t) + \sup_{t\geq 0} |W_0(t) - \widetilde{W}_0(t)|] \leq L_p p^{-\beta}[p^{-3\beta/2} + \sup_{\{t\geq 0\}}\{\widetilde{W}_0(t)\}].$$

Next, consider II. Write

$$II \leq \sup_{\{t\geq 0\}}[\bar{\Psi}(t)\widetilde{W}_0(t)] + \sup_{\{t\geq 0\}}[\bar{\Psi}(t)|W_0(t) - \widetilde{W}_0(t)|] \equiv IIa + IIb. \tag{A.39}$$

On one hand, elementary calculus shows that $IIa \leq p^{-\beta}$. On the other hand, by similar argument as in the proof of Lemma A.6, $IIb \leq L_p(p^{-\beta} + p^{-\frac{a^2 r}{3}-\beta} + p^{-3\beta/2})$. Combining these, $II \leq L_p(p^{-\beta} + p^{-\frac{a^2 r}{3}-\beta} + p^{-3\beta/2})$. Last, consider III. By (A.36) and direct calculations,

$$III \leq CK_p\epsilon_p \sup_{\{t\geq 0\}}\{\bar{\Psi}(t)/\sqrt{\bar{\Psi}(t) + \epsilon_p\bar{\Psi}(t-\tau_p)}\} \leq L_p p^{-\beta}.$$

Inserting these into (A.38) gives the claim.

Next, we show Lemma 2.4. The first claim has already been proved in Lemma A.6. So we only need to prove claims (a)–(c) in the case of $r \geq \beta$.

First consider claims (a) and (b) in Lemma 2.4. Comparing Lemma A.6 and the desired claim, it is sufficient to verify that

$$W_0(t) \leq p^{-\beta/2}/\sqrt{2}, \text{ if } t \leq \sqrt{2\beta \log p} - \Delta_1 \text{ or } t > \tau_p, \tag{A.40}$$

where $\Delta_1 = d_0(\log\log p)/\sqrt{\log p}$ is as defined in the statement of Lemma 2.4. Once this is proved, recalling that $W(t) = a_1(t)W_0(t) - W_1(t)$ and we have just proved $\sup_{t\geq \bar{\Psi}^{-1}(1/2)}\{\bar{\Psi}(t)W_0(t)\} \leq L_p p^{-\beta}$, then by Lemma A.8 we have

$$W(t) \leq a(t)W_0(t) \lesssim (1 + C\bar{\Psi}(t) + CK_p\epsilon_p)W_0(t) \lesssim p^{-\beta/2}/\sqrt{2}.$$

We now proceed to prove (A.40). By the proof of Lemma A.6 (inequality (A.34)), we have

$$0 \leq W_0(t) - \widetilde{W}_0(t) \leq L_p p^{-\beta} + K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)/\sqrt{\bar{\Psi}(t) + K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}, \tag{A.41}$$

where we have noted that the last term in (A.34) is bounded by $K_p\epsilon_p\sqrt{\bar{\Psi}_{(1+a)\tau_p}(t)} \leq L_p p^{-\beta}$. First consider the case when $t \leq \sqrt{2\beta\log p} - \Delta_1$. By Mills's ratio, for appropriately chosen d_0 in $\Delta_1 = d_0(\log\log p)/\sqrt{\log p}$, we have $\bar{\Psi}(t) + K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t) \geq 8K_p^2\epsilon_p$, and $\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{a\tau_p}(t) \geq 8\epsilon_p$. As a result,

$$\frac{K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}} \leq \sqrt{2\epsilon_p}/4, \qquad \widetilde{W}_0(t) \leq \frac{\epsilon_p\bar{\Psi}_{\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)}} \leq \sqrt{2\epsilon_p}/4.$$

Inserting these into (A.41), we complete the proof of (A.40) when $t \leq \sqrt{2\beta\log p} - \Delta_1$. Now we consider the case of $t > \tau_p$. Since $\epsilon_p\bar{\Psi}_{a\tau_p}(t) = o(\epsilon_p p^{-(1-a)^2 r})$, it follows that

$$\frac{K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)}} \leq \sqrt{K_p\epsilon_p\bar{\Psi}_{a\tau_p}(t)} = o(p^{-\beta/2})$$

and

$$\widetilde{W}_0(t) \leq \frac{\epsilon_p\bar{\Psi}_{\tau_p}(t)}{\sqrt{\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)}} \leq \sqrt{\epsilon_p\bar{\Psi}_{\tau_p}(t)} \leq \sqrt{\epsilon_p/2}.$$

Inserting these into (A.41) proves (A.40) when $t > \tau_p$.

Finally we prove part (c). Write for short $s_p = \sqrt{2\beta\log p} - \Delta_1$. By (A.37) and recalling that we have just proved $\sup_{t>0} W_1(t) \leq L_p p^{-\beta}$, we obtain that $W(t) = a_1(t)W_0(t) - W_1(t) \geq (1 - K_p\epsilon_p)W_0(t) - \sup_{t>0} W_1(t) \geq (1 - CK_p\epsilon_p)W_0(t) - L_p p^{-\beta}$. Further recall that in Lemma A.6, we have shown that $W_0(t) \geq \widetilde{W}_0(t)$ for all $t \geq 0$. Thus, $W(t) \geq (1 - CK_p\epsilon_p)\widetilde{W}_0(t) - L_p p^{-\beta}$. Taking $t_p^* = \frac{\beta+r}{2r}\tau_p$, it is seen that for sufficiently large p, $s_p \leq t_p^* \leq \tau_p$. Therefore, $\sup_{\{s_p \leq t \leq \tau_p\}} W(t) \geq (1 - CK_p\epsilon_p)\sup_{\{s_p \leq t \leq \tau_p\}} \widetilde{W}_0(t) \geq (1 - CK_p\epsilon_p)\widetilde{W}_0(t_p^*)$, and the first inequality of part c) follows from $\widetilde{W}_0(t_p^*) \sim p^{-\beta/2}$ and $\epsilon_p = p^{-\beta}$ for large enough p. On the other hand, by Lemma A.6 and recall $r \geq \beta$, we have $\sup_{t>0} W_0(t) \leq L_p \sup_{t>0} \widetilde{W}_0(t) \sim L_p p^{-\beta/2}$. Further, by (A.37) and the expression $W(t) = a_1(t)W_0(t) - W_1(t)$, we have $\sup_{s_p \leq t \leq \tau_p} W(t) \leq \sup_{s_p \leq t \leq \tau_p}\{a_1(t)W_0(t)\} \leq C\sup_{s_p \leq t \leq \tau_p} W_0(t) \sim L_p p^{-\beta/2}$. Thus, the second inequality in the claim follows. $\square$

A.8. Proof of Lemma A.8. Let $h_0(t)$, $h_1^{\pm}(t)$ and $g_1(t)$ be as in Lemma A.4. Consider the first claim. By Lemma A.2 parts (a) and (e), we have

$$0 \leq \bar{\Psi}(t) - h_0(t) \leq K_p\epsilon_p\bar{\Psi}(t), \qquad \widetilde{F}(t) \geq (1 - K_p\epsilon_p)[\bar{\Psi}(t) + \epsilon_p\bar{\Psi}_{\tau_p}(t)]. \tag{A.42}$$

At the same time, note that $\widetilde{F}(t) \leq \bar{\Psi}(t) + K_p\epsilon_p$. Combining these ensures that

$$1 \leq (1 - \widetilde{F}(t))^{-1/2} \leq [1 - \bar{\Psi}(t) - K_p\epsilon_p]^{-1/2}. \tag{A.43}$$

Inserting (A.42) and (A.43) into the definition of $W_1(t)$ gives

$$0 \leq W_1(t) \leq \frac{K_p \epsilon_p \bar{\Psi}(t)}{\sqrt{(1-\bar{\Psi}(t)-K_p\epsilon_p)(1-K_p\epsilon_p)[\bar{\Psi}(t)+\epsilon_p\bar{\Psi}_{\tau_p}(t)]}}.$$

Thus the first claim follows by noting $(1-\bar{\Psi}(t)-K_p\epsilon_p) \geq 1/2 - K_p\epsilon_p$ for all $t \geq \bar{\Psi}^{-1}(\frac{1}{2})$.

Consider the second claim. Recall that $\widetilde{F}(t) = h_0(t)+h_1^+(t)+h_1^-(t)+g_1(t)$. By definitions,

$$a_1(t) = (1-\widetilde{F}(t))^{-1/2} \cdot I \cdot II, \tag{A.44}$$

where $I = [h_1^+(t)+h_1^-(t)+g_1(t)]/[\epsilon_p\bar{\Psi}_{\tau_p}(t)+g_1(t)]$, and

$$II = \sqrt{\bar{\Psi}(t)+\epsilon_p\bar{\Psi}_{\tau_p}(t)+g_1(t)}/\sqrt{h_0(t)+h_1^+(t)+h_1^-(t)+g_1(t)}.$$

By (a) and (b) in Lemma A.2, we have

$$(1-K_p\epsilon_p) \leq I \leq 1, \qquad 1 \leq II \leq (1-K_p\epsilon_p)^{-1/2}. \tag{A.45}$$

Inserting (A.43) and (A.45) into (A.44), we obtain that there is a universal constant $C > 0$ such that (A.37) holds. $\square$

A.9. Proof of Theorems 2.1-2.2. The following lemma is proved in Section A.10.

LEMMA A.9. *Fix $(\beta, r) \in (0,1)^2$ and a sufficiently large p. When t ranges in $(0,\infty)$, $\widetilde{W}_0(t)$ first strictly increases and reaches the maximum at $t = t_p^{**} \sim \min\{2, \frac{r+\beta}{2r}\}\tau_p\, (\equiv t_p^*)$, and then strictly decreases. Additionally, if $r < \beta$, then there are positive constants $c_4 = c_4(\beta,r)$ and $c_5 = c_5(\beta,r)$ such that for all $|t - t_p^{**}| \leq c_4\tau_p^{-1}$, $\widetilde{W}_0''(t) \leq -2c_5\widetilde{W}_0(t)$.*

Denote by $W(t) = p^{-1/2}HC(t,\widetilde{F})$. By the first claim in Lemma 2.3 and Lemma A.6, and noting that $\beta > c_0(\beta,r,a)$, we obtain

$$\sup_{\{t\geq 0\}} |W(t) - \widetilde{W}_0(t)| \leq L_p[p^{-\beta} + p^{-c_0(\beta,r,a)}\sup_{\{t\geq 0\}}\widetilde{W}_0(t)]. \tag{A.46}$$

First, we show Theorem 2.1, where we assume $r < \beta$. Once the first claim is proved, the second claim follows by combining Taylor expansion with Lemmas 2.3, 2.4, and A.9, so we only show the first claim. The idea is to prove T_{HC} and T_{ideal} are both close to t_p^{**}, then they are close to each other.

We first prove that T_{HC} and t_p^{**} are close. We will show that (i) $W(t_p^{**}+u)-W(t_p^{**})<0$ for all $p^{-c_1}\le|u|\le c_4/\tau_p$, and (ii) $W(t_p^{**}+u)-W(t_p^{**})<0$ for all $|u|>c_4/\tau_p$. Then combining these proves

$$|T_{HC}(\widetilde{F})-t_p^{**}|\le p^{-c_1}, \tag{A.47}$$

with $c_1=c_1(\beta,r,a)>0$ some constant to be specified later.

We now prove the first case (i). Recall that t_p^{**} is the maximizer of $\widetilde{W}_0(t)$ and $\widetilde{W}_0(t_p^{**})=L_p p^{-\delta(\beta,r)}$, where $\delta(\beta,r)$ is as in (2.10). Thus, $\widetilde{W}_0'(t_p^{**})=0$. By Taylor expansion, $\widetilde{W}_0(t_p^{**}+u)-\widetilde{W}_0(t_p^{**})=\frac{u^2}{2}\widetilde{W}_0''(\tilde{t}_p)$, where $\tilde{t}_p$ lies between t_p^{**} and $t_p^{**}+u$. Next, by Lemma A.9, for $|u|\le\frac{c_4}{\tau_p}$ we can further write $\widetilde{W}_0(t_p^{**}+u)-\widetilde{W}_0(t_p^{**})\le-c_5u^2\widetilde{W}_0(\tilde{t}_p)=-c_5u^2\widetilde{W}_0(t_p^{**})-c_5u^2(\widetilde{W}_0(\tilde{t}_p)-\widetilde{W}_0(t_p^{**}))\le-c_5u^2\widetilde{W}_0(t_p^{**})-c_5u^2(\widetilde{W}_0(t_p^{**}+u)-\widetilde{W}_0(t_p^{**}))$, where the last step is because of $\widetilde{W}_0(t_p^{**}+u)\le\widetilde{W}_0(\tilde{t}_p)$. Thus, the inequality can be further written as $\widetilde{W}_0(t_p^{**}+u)-\widetilde{W}_0(t_p^{**})\le-c_5u^2\widetilde{W}_0(t_p^{**})/(1+c_5u^2)$. Then by (A.46) we obtain that

$$\begin{aligned}W(t_p^{**}+u)-W(t_p^{**})&=\big(W(t_p^{**}+u)-\widetilde{W}_0(t_p^{**}+u)\big)-\big(W(t_p^{**})-\widetilde{W}_0(t_p^{**})\big)\\&+\big(\widetilde{W}_0(t_p^{**}+u)-\widetilde{W}_0(t_p^{**})\big)\le L_p(p^{-\beta}+p^{-c_0(\beta,r,a)}\widetilde{W}_0(t_p^{**}))+\big(\widetilde{W}_0(t_p^{**}+u)-\widetilde{W}_0(t_p^{**})\big)\\&\le L_pp^{-\beta}+\big(L_pp^{-c_0(\beta,r,a)}-c_5u^2/(1+c_5u^2)\big)\widetilde{W}_0(t_p^{**}).\end{aligned} \tag{A.48}$$

It is easy to check that $p^{-c_0(\beta,r,a)}\widetilde{W}_0(t_p^{**})\gg L_pp^{-\beta}$ when $\rho_\theta^*(\beta)<r<\beta$. By Lemma A.9, we obtain that if $|u|\ge p^{-c_1}$ with $c_1=c_1(\beta,r,a)\in(0,\frac{1}{3}c_0(\beta,r,a))$, then for all $p^{-c_1}\le|u|\le c_4/\tau_p$,

$$W(t_p^{**}+u)-W(t_p^{**})\le-L_pp^{-2c_1(\beta,r,a)}\widetilde{W}_0(t^*)(1+o(1))<0,$$

which completes the proof of case (i). It remains to prove case (ii). Direct calculations yield $\widetilde{W}_0(t_p^{**}\pm c_4/\tau_p)\lesssim e^{-c_5}\widetilde{W}_0(t_p^{**})$, where $c_5>0$ is a constant depending on whether $r<\beta/3$ or $r\ge\beta/3$. By Lemma A.9, $\widetilde{W}_0(t)\le\widetilde{W}_0(t_p^{**}\pm c_4/\tau_p)\lesssim e^{-c_5}\widetilde{W}_0(t_p^{**})$ for all $|t-t_p^{**}|>c_4/\tau_p$. Thus, similar to (A.48) we have $W(t)-W(t_p^{**})\le L_p(p^{-\beta}+p^{-c_0(\beta,r,a)}\widetilde{W}_0(t_p^{**}))+(\widetilde{W}_0(t)-\widetilde{W}_0(t_p^{**}))\lesssim L_pp^{-\beta}+(e^{-c_5}-1+L_pp^{-c_0(\beta,r,a)})\widetilde{W}_0(t_p^{**})=L_pp^{-\beta}+(e^{-c_5}-1+L_pp^{-c_0(\beta,r,a)})p^{-\delta(\beta,r)}<0$, where the last step is because $\beta>\delta(\beta,r)$. This proves case (ii). Consequently, we have proved (A.47).

Using similar method as above and in view of Lemma 2.1 we can also prove that for appropriately chosen $c_1>0$,

$$|T_{ideal}(\epsilon_p,\tau_p,\Omega)-t_p^{**}|\le p^{-c_1}. \tag{A.49}$$

Thus the claim in Theorem 2.1 follows when $r < \beta$.

We now show Theorem 2.2, where we assume $r \geq \beta$. In this range $\widetilde{W}_0(t)$ is maximized at $t_p^{**} = \frac{\beta + r}{2r}\tau_p$ and $\widetilde{W}_0(t_p^{**}) \sim p^{-\frac{\beta}{2}}$. By Lemma 2.4 we see that the maximizer of $W_0(t)$ is in the range $[\sqrt{2\beta \log p} - \Delta_1, \tau_p)$. By (A.40) and Lemma 2.3 we obtain that if $0 \leq t < \sqrt{2\beta \log p} - \Delta_1$ or $\tau_p \leq t < \infty$,

$$W(t) = W_0(t) + (W(t) - W_0(t)) \leq \frac{1}{\sqrt{2}} p^{-\beta/2} + L_p p^{-\beta} = \frac{1}{\sqrt{2}} p^{-\beta/2}(1 + o(1)),$$

and if $\sqrt{2\beta \log p} - \Delta_1 \leq t < \tau_p$,

$$W(t) = W_0(t) + (W(t) - W_0(t)) \geq p^{-\beta/2} - L_p p^{-\beta} = p^{-\beta/2}(1 - o(1)).$$

Thus, the maximizer $T_{HC}(\widetilde{F})$ is in the interval $[\sqrt{2\beta \log p} - \Delta_1, \tau_p)$.

By Lemma 2.2, the maximizer of $\widetilde{Sep}(t, \epsilon_p, \tau_p, \Omega)$ is in the interval $[\sqrt{2\beta \log p} - \Delta_1, \tau_p + \Delta_2)$, and Theorem 2.2 follows immediately. $\square$

A.10. Proof of Lemma A.9. Let $\psi_{\tau_p}(t) = \phi(t - \tau_p) + \phi(t + \tau_p)$ and $\psi(t) = 2\phi(t)$. Introduce $m_0(t) = \psi(t)/\bar{\Psi}(t)$, $m_1(t) = \bar{\Psi}_{\tau_p}(t)/\psi_{\tau_p}(t)$, $d(t) = -\psi'_{\tau_p}(t)/\psi_{\tau_p}(t)$, $a(t) = \epsilon_p \psi_{\tau_p}(t)/\psi(t)$, $R(t) = m_1(t)/m_0(t)$, and $g(t) = (1/2)(1 + a(t))/(R^{-1}(t) + a(t))$. The following lemma is proved in APPENDIX B.

LEMMA A.10. *Fix a sufficiently large p, $R(t) > 1$ and is strictly decreasing for all $t > 0$.*

Consider the first claim. By direct calculations and our notations,

$$\widetilde{W}_0'(t)/\widetilde{W}_0(t) = \frac{1}{2}\Big[\frac{\psi(t) + \epsilon_p \psi_{\tau_p}(t)}{\bar{\Psi}(t) + \epsilon_p \bar{\Psi}_{\tau_p}(t)}\Big] - \frac{\psi_{\tau_p}(t)}{\bar{\Psi}_{\tau_p}(t)} \equiv [g(t) - 1]/m_1(t). \tag{A.50}$$

To show the claim, it suffices to show that equation $g(t) = 1$ has exactly one solution. Recall that $g(t) = (1/2)(1 + a(t))/(R^{-1}(t) + a(t))$, where $R(t) > 1$ and both $a(t)$ and $R^{-1}(t)$ are strictly increasing in t. It follows from basic calculus that $g(t)$ is strictly decreasing in $(0, \infty)$, and the equation $g(t) = 1$ has at most one solution.

The equation also has at least one solution. Note that $g(0) \geq Ce^{\tau_p^2/2}$ which > 1 for sufficiently large p, it suffices to show that there is a t such that $g(t) < 1$. We show this for the case of $r < \beta/3$ and $r > \beta/3$ separately. In the first case, for all t such that $|t - 2\tau_p| \leq 4\tau_p^{-1}$, $a(t)$ is algebraically small, and so by Mills' ratio [4], for any fixed b,

$$g(2\tau_p + b\tau_p^{-1}) \leq \frac{1}{2}\Big[\frac{1}{2} - \frac{3b}{2}\tau_p^{-2} + O(\tau_p^{-4})\Big],$$

and the claim follows. Note that this shows that the solution t_p^{**} of the equation $g(t) = 1$ satisfies $|t_p^{**} - 2\tau_p| \le 2\tau_p^{-1}$. In the second case, $a(\sqrt{2\log(p)}) = L_p p^{1-\beta-(1-\sqrt{r})^2}$, where the the exponent > 0 since $r > \beta/3$ and $r > \rho(\beta)$ (recall that $\rho(\beta)$ is the standard phase function). Therefore, $g(t_0) \sim 1/2$ and the claim follows. This completes the proof of the first claim.

Consider the second claim. We discuss for the case $0 < r < \beta/3$ and $\beta/3 < r < \beta$ separately.

Consider the first case. Recalling that $|t_p^{**} - 2\tau_p| \le 2\tau_p^{-1}$, it is sufficient to show that for all t such that $|t_p - 2\tau_p| \le 4\tau_p^{-1}$, $\widetilde{W}_0''(t)/\widetilde{W}_0(t) \lesssim -1/2$. Introduce $s(t) = [t\psi(t) + d(t)\psi_{\tau_p}(t)] \cdot [\bar{\Psi}(t) + \epsilon_p \bar{\Psi}_{\tau_p}(t)]/[\psi(t) + \epsilon_p \psi_{\tau_p}(t)]^2$. By direct calculations,

$$\widetilde{W}''/\widetilde{W}(t) = I + II - \frac{1}{2} III, \tag{A.51}$$

where

$$I = (g(t) - 1)^2/m_1^2(t),\; II = d(t)/m_1(t) - m_1^{-2}(t),\; III = (s(t) - 1)g^2(t)m_1^{-2}(t).$$

Consider I first. When $|t - 2\tau_p| \le 4\tau_p^{-1}$, on one hand, by Mills' ratio, $m_1^{-1}(t) \sim (t - \tau_p) \sim \tau_p$. On the other hand, by similar argument, $|g(t) - 1| \le C\tau_p^{-2}$. It follows that $I \le C\tau_p^{-2}$. Consider II next. By Mills' ratio, $m_1^{-1}(t) = (t - \tau_p) + \frac{1}{t - \tau_p} + O(\tau_p^{-3})$. Since $|d(t) - (t - \tau_p)|$ is algebraically small, it follows from basic algebra that $II \sim -1$. Consider III. Note that both the ratio $\epsilon_p \psi_{\tau_p}(t)/\psi(t)$ and the ratio $\epsilon_p \bar{\Psi}_{\tau_p}(t)/\bar{\Psi}_{\tau_p}(t)$ are algebraically small. Combining this with $\bar{\Psi}(t)/\psi(t) = (1/t) - (1/t^3) + O(t^{-5})$ gives

$$s(t) = \frac{t\psi(t)\bar{\Psi}(t)}{(\psi(t))^2} + O(\tau_p^{-3}) = 1 - \frac{1}{t^2} + O(\tau_p^{-3}),$$

Recall that $m_1^{-1}(t) \sim \tau_p$ and $g(t) \sim 1$, it follows that $III \sim -4\tau_p^2/t^2 \sim -1$. Inserting these into (A.51) gives that for all $|t - 2\tau_p| \le 4\tau_p^{-1}$, $\widetilde{W}_0''(t)/\widetilde{W}_0(t) \lesssim -1/2$ and the second claim follows.

Consider the second case, where $r \ge \beta$. For a constant $\eta_0 \in (0, 1)$ to be determined, choose t_0 and $t_p^{\pm}$ such that $a(t_0) = \frac{3r-\beta}{\beta+r}$, and $a(t_p^{\pm}) = (1 \pm \eta_0)a(t_0)$. It is seen that $|t_p^{\pm} - \frac{\beta+r}{2r}\tau_p| \le C\tau_p^{-1}$, and $|t_0 - \frac{\beta+r}{2r}\tau_p| \le C\tau_p^{-1}$. Combining these with definitions and Mills' ratio, for $t_p^- \le t \le t_p^+$, $R^{-1}(t) \sim (t - \tau_p)/t \sim (\beta - r)/(\beta + r)$, and that

$$g(t) \sim \frac{1}{2} \cdot \frac{1 + a(t)}{[(\beta - r)/(\beta + r)] + a(t)}. \tag{A.52}$$

By direct calculations, $g(t_p^-) > 1$ and $g(t_p^+) < 1$. Since $g(t_p^{**}) = 1$, we have $t_p^- < t_p^{**} < t_p^+$.

We now use (A.51) to calculate $\widetilde{W}_0''(t)/\tilde{W}_0(t)$ with. First, recall that $II \sim -1$. Second, by similar argument, $m_1^{-1}(t) \sim (t - \tau_p) \sim (\beta - r)/(2r)\tau_p$. Combining this with (A.52),

$$I = m_1^{-2}(t)[g(t) - 1]^2 = (\frac{\beta - r}{2r})^2 \tau_p^2 \cdot \big([\frac{1}{2}\frac{1 + a(t)}{[(\beta - r)/(\beta + r) + a(t)} - 1]^2 + o(1)\big).$$

Last, by similar argument,

$$\frac{t\psi(t) + \epsilon_p d(t)\psi_{\tau_p}(t)}{\bar{\Psi}(t) + \epsilon_p \bar{\Psi}_{\tau_p}(t)} \sim \frac{(\beta + r)/(\beta - r) + a(t)}{(\beta - r)/(\beta + r) + a(t)}(\frac{\beta - r}{2r})^2 \tau_p^2.$$

$$s(t) \sim \frac{[(\beta + r)/(\beta - r) + a(t)] \cdot [(\beta - r)/(\beta + r) + a(t)]}{(1 + a(t))^2}.$$

Combining this with (A.52), III equals to $(\frac{\beta-r}{2r})^2\tau_p^2 \cdot [\frac{1+a(t)}{(\beta-r)/(\beta+r)+a(t)}]^2$ times

$$\big[\frac{[(\beta + r)/(\beta - r) + a(t)] \cdot [(\beta - r)/(\beta + r) + a(t)]}{(1 + a(t))^2} - 1 + o(1)\big].$$

Inserting these into (A.51) and recalling that $a(t_0) = (3r - \beta)/(\beta + r)$, it follows from basic algebra that $\widetilde{W}''(t_0)/\widetilde{W}(t_0) \lesssim -\frac{3r-\beta}{2(\beta-r)} \cdot (\frac{\beta-r}{2r})^2\tau_p^2$, Recall that $a(t_p^{\pm}) = (1 \pm \eta_0)a(t_0)$. By the continuity of I and III on $a(t)$, if we choose η_0 sufficiently small, then for all $t_p^- \leq t \leq t_p^+$,

$$\widetilde{W}_0''(t)/\widetilde{W}_0(t) \leq -\frac{3r - \beta}{4(\beta - r)} \cdot (\frac{\beta - r}{2r})^2 \tau_p^2,$$

and the claim follows. $\square$

A.11. Proof of Theorem 2.3. We write $\widetilde{Sep}(t) = \widetilde{Sep}(t, \epsilon_p, \tau_p, \Omega)$, $Sep(t) = Sep(t, \tilde{Z}, \mu, \Omega)$, and $T_{ideal} = T_{ideal}(\epsilon_p, \tau_p, \Omega)$ for short. The following lemmas are proved in Section A.12 and Section A.13 respectively.

LEMMA A.11. *Fix a constant $\kappa > 0$. As $p \to \infty$, for any sequence $t_p \in (0, \tau_p + \tilde{s}_p]$ with $\tilde{s}_p$ defined in Lemma A.4 such that $\widetilde{Sep}(t_p) \geq L_p p^\kappa$, we have $P(YL_t(X, \Omega) < 0 | t = t_p) = \bar{\Phi}\big((1 + o(1))\frac{1}{2}\widetilde{Sep}(t_p)\big)$.*

LEMMA A.12. *For any sequence of closed subset $A_p \subset [0, \tau_p + \tilde{s}_p]$ with $\tilde{s}_p$ defined in lemma A.4, if there exists a constant $\kappa > 0$ such that $\sup_{t \in A_p}\{\widetilde{Sep}(t)\} \geq p^\kappa$ for sufficiently large p, then with probability at least $1 - o(1/p)$,*

$$\sup_{t \in A_p} Sep(t) \leq \big(1 + o(1/\sqrt{\log p})\big)\frac{1}{2}\sup_{0 < t \leq \tau_p + \tilde{s}_p} \widetilde{Sep}(t).$$

Now we proceed to prove the theorem. The key is to show

$$\min_{t>\tau_p+\tilde{s}_p} P(YL_t(X,\Omega)<0|t) > \bar{\Phi}\Big(\big(1+o(1)\big)\frac{1}{2}\widetilde{Sep}\big(T_{ideal}\big)\Big), \tag{A.53}$$

and

$$\min_{0<t\le\tau_p+\tilde{s}_p} P(YL_t(X,\Omega)<0|t) = \bar{\Phi}\Big(\big(1+o(1)\big)\frac{1}{2}\widetilde{Sep}\big(T_{ideal}\big)\Big). \tag{A.54}$$

Then combining the above results completes the proof of the theorem.

We first prove (A.53). When $r<\beta$, by proof (1b) in Lemma 3.4 we have $Sep(t,\tilde{Z},\mu,\Omega)\le L_p p^{\frac{1-\theta}{2}-\frac{1}{4}\max\{4\beta-2r,3\beta+4\}}$ for all $t>\tau_p+\tilde{s}_p$ with probability at least $1-o(p^{-1})$. When $r\ge\beta$, by Lemma 3.4 we have

$$Sep(t)=\widetilde{Sep}(t)+(Sep(t)-\widetilde{Sep}(t))\le\widetilde{Sep}(t)+L_p p^{-\theta/2}.$$

Following the same line as that in the proof of Lemma 2.2 we can show that for $r\ge\beta$, $\widetilde{Sep}(t)\le L_p p^{\frac{1-\theta}{2}-c_8}$ with $c_8=c_8(\beta,r)>\delta(\beta,r)$ for all $t>\tau_p+\tilde{s}_p$. Combining these and recalling that $r>\rho^*_\theta(\beta)$ and $\beta\in(\frac{1-\theta}{2},1-\theta)$, we have $Sep(t)\le L_p p^{\frac{1-\theta}{2}-c_9(\beta,r)}$ with $c_9(\beta,r)$ some constant whose value depends on whether $r<\beta$ or $r\ge\beta$ and satisfies $c_9(\beta,r)>\delta(\beta,r)$, for all $t>\tau_p+\tilde{s}_p$, with probability at least $1-o(p^{-1})$. Recall that $\widetilde{Sep}\big(T_{ideal}\big)=L_p p^{\frac{1-\theta}{2}-\delta(\beta,r)}$. Thus,

$$\begin{aligned}P(YL_t(X,\Omega)<0|t)&=E\Big(\bar{\Phi}\Big(\frac{1}{2}Sep(t)\Big)\Big)\ge\bar{\Phi}\Big(L_p p^{\frac{1-\theta}{2}-c_9(\beta,r)}\Big)(1-o(p^{-1}))\\&\gg\bar{\Phi}\Big(\big(1+o(1)\big)\frac{1}{2}\widetilde{Sep}\big(T_{ideal}\big)\Big).\end{aligned}$$

This completes the proof of (A.53).

Next we prove (A.54). We only need to prove that uniformly over all $0<t\le\tau_p+\tilde{s}_p$,

$$P(YL_t(X,\Omega)<0|t)\ge\bar{\Phi}\Big((1+o(1))\frac{1}{2}\sup_{0<t\le\tau_p+\tilde{s}_p}\widetilde{Sep}(t)\Big). \tag{A.55}$$

Then, taking $t_p=T_{ideal}$ in Lemma A.11 and noting that $T_{ideal}\in(0,\tau_p+\tilde{s}_p]$ shows $P(YL_t(X,\Omega)<0|T_{ideal})=\bar{\Phi}\big((1+o(1))\frac{1}{2}\sup_{0<t\le\tau_p+\tilde{s}_p}\widetilde{Sep}(t)\big)$. Combining this with (A.55) yields (A.54).

We now proceed to prove (A.55). Define $A_p=\{t:t\in(0,\tau_p+\tilde{s}_p],\ \widetilde{Sep}(t)\le\frac{1}{2}\sup_{0<t\le\tau_p+\tilde{s}_p}\{\widetilde{Sep}(t)\}\}$. Then by Lemma A.12, with probability at least $1-o(p^{-1})$,

$$\sup\nolimits_{t\in A_p}Sep(t)\le(1+o(\frac{1}{\sqrt{\log p}}))\frac{1}{2}\sup\nolimits_{0<t\le\tau_p+\tilde{s}_p}\{\widetilde{Sep}(t)\}.$$

We claim that it remains to show with probability at least $1 - o(1/p)$, uniformly for all $t \in A_p^c \equiv (0, \tau_p + \tilde{s}_p] \setminus A_p$,

$$Sep(t) \leq (1 + L_p p^{-\kappa}) \frac{1}{2} \sup_{0<t\leq \tau_p+\tilde{s}_p} \widetilde{Sep}(t), \tag{A.56}$$

where $\kappa = (1-\theta)/2 - \delta(\beta, r) > 0$. Then, combining the above two inequalities yields that uniformly for all $t \in (0, \tau_p + \tilde{s}_p]$,

$$\begin{aligned} P(YL_t(X, \Omega) < 0|t) &\geq \bar{\Phi}\Big((1 + o(\frac{1}{\sqrt{\log p}}))\frac{1}{2} \sup_{0<t\leq \tau_p+\tilde{s}_p} \widetilde{Sep}(t)\Big)(1 - o(\frac{1}{p})) \\ &= \bar{\Phi}\Big((1 + o(1))\frac{1}{2} \sup_{0<t\leq \tau_p+\tilde{s}_p} \widetilde{Sep}(t)\Big), \end{aligned}$$

which completes the proof of (A.55).

We proceed to prove the above claim (A.56). Introduce the event

$$B_p = \{\sup_{t\in A_p^c} \frac{|\tilde{M}_p(t) - m_p(t)|}{m_p(t)} \leq L_p p^{-\kappa}, \qquad \sup_{t\in A_p^c} \frac{|\tilde{V}_p(t) - v_p(t)|}{p\widetilde{F}(t)} \leq L_p p^{-\kappa}\},$$

where κ is as introduced in (A.56). Then on the event B_p,

$$\tilde{M}_p(t)/\sqrt{\tilde{V}_p(t)} \leq (1 + L_p p^{-\kappa}) m_p(t)/\sqrt{v_p(t)} \leq (1 + L_p p^{-\kappa})\frac{1}{2} \sup_{0<t\leq \tau_p+\tilde{s}_p} \widetilde{Sep}(t), \tag{A.57}$$

and the above claim (A.56) holds by noting $Sep(t) = \tilde{M}_p(t)/\sqrt{\tilde{V}_p(t)}$. Next we show that $P(B_p) \geq 1 - o(1/p)$. Recall that we have proved in (2.13) that $\sup_{0<t<\sqrt{2\log p}} \widetilde{Sep}(t) = L_p p^{\kappa}$. By Lemma A.4 and (A.87), $v_p(t) \geq Cp\widetilde{F}(t)$ with some constant $C > 0$, where the value of C depends on whether $r < \beta$ or $r \geq \beta$. Moreover, by definition of A_p^c, $\widetilde{Sep}(t) \geq \frac{1}{2} L_p p^{-\kappa}$ for $t \in A_p^c$. It follows that $m_p(t) = \frac{1}{2}\sqrt{v_p(t)}\widetilde{Sep}(t) \geq \sqrt{Cp\widetilde{F}(t) L_p p^{\kappa}}$. On the other hand, by Lemma A.15 $m_p(t) \leq L_p p^{1-\theta/2}\widetilde{F}(t)$, so we can derive $p\widetilde{F}(t) \geq L_p p^{2\kappa+\theta}$ and consequently, $\sqrt{n_p} m_p(t) \geq L_p p^{2\kappa+\theta}$ and $v_p(t) \geq L_p p^{2\kappa+\theta}$. By Lemma A.17 and using similar arguments as those in Lemma A.16, we can prove that for each $t \in A_p^c$,

$$P\Big(\frac{|\tilde{M}_p(t) - m_p(t)|}{m_p(t)} \geq L_p p^{-\kappa}\Big) \leq o(\frac{1}{p^3}), \quad P\Big(\frac{|\tilde{V}_p(t) - v_p(t)|}{v_p(t)} \geq L_p p^{-\kappa}\Big) \leq o(\frac{1}{p^3}).$$

Using the grid point method as that in the proof of Lemma 3.1 shows that $P(B_p) \geq 1 - o(1/p)$. This completes the proof of (A.56) and the results in the theorem follow immediately.

$\square$

A.12. Proof of Lemma A.11. Write for short $\widetilde{Sep}(t) = \widetilde{Sep}(t, \epsilon_p, \tau_p, \Omega)$, $\tilde{M}_p(t) = M_p(t, \tilde{Z}, \Omega, \mu)$, $\tilde{V}_p(t) = V_p(t, \tilde{Z}, \Omega)$, $m_p(t) = m_p(t, \epsilon_p, \tau_p, \Omega)$, and $v_p(t) = v_p(t, \epsilon_p, \tau_p, \Omega)$. Define event

$$B_p = \{|\tilde{V}_p(t_p) - v_p(t_p)| \leq L_p p^{-\theta/2} p\widetilde{F}(t_p),\ |\tilde{M}_p(t_p) - m_p(t_p)| \leq L_p p^{-\theta/2} m_p(t_p)\}.$$

The key is to first show that (a)

$$P(B_p^c) \leq \exp\Big(-\frac{1}{2}\log(p)(\widetilde{Sep}(t_p))^2 \cdot (1 + o(1))\Big), \tag{A.58}$$

and then show that (b) the desired claim in the lemma holds on the event B_p. Combining (a) and (b) proves that the desired claim holds.

We first prove claim (a). Note that by Lemma A.4, $v_p(t) \geq Cp\widetilde{F}(t)$ with some constant $C > 0$, where the value of C depends on whether $r \geq \beta$ or $r < \beta$. Further by Lemma A.15, $0 < \sqrt{n_p} m_p(t) \leq K_p^2 (\log p)^{3/2} p\widetilde{F}(t) \leq CK_p^2 (\log p)^{3/2} v_p(t)$, and so that $\sqrt{n_p} m_p(t) \geq Cn_p m_p^2(t)/[K_p^2(\log p)^{3/2} v_p(t)] = \frac{Cn_p}{K_p^2(\log p)^{3/2}}(\widetilde{Sep}(t))^2$. Taking $\lambda_p = K_p(\log p)\Big(\frac{\sqrt{n_p}\widetilde{Sep}^2(t_p)}{c_2 m_p(t_p)}\Big)^{1/2} m_p(t_p)$, then $\lambda_p \leq L_p m_p(t_p)$. It follows that $P\big(\sqrt{n_p}|\tilde{M}_p(t_p) - m_p(t_p)| \geq K_p^3 \cdot L_p m_p(t_p)\big) \leq P(\sqrt{n_p}|\tilde{M}_p(t_p) - m_p(t_p)| \geq K_p^3 \lambda_p)$, where by Lemma A.17, the right hand side

$$\leq K_p^3 \exp\Big(- \big(\widetilde{Sep}(t_p)\big)^2 (\log p)\Big).$$

Since $\widetilde{Sep}(t_p) \geq L_p p^{\kappa} \to \infty$, it follows easily that

$$P\big(|\tilde{M}_p(t_p) - m_p(t_p)| \geq L_p p^{-\theta/2} m_p(t_p)\big) \leq \exp\Big(-\big(\widetilde{Sep}(t_p)\big)^2 (\log p)(1 + o(1))\Big). \tag{A.59}$$

Next we consider $\tilde{V}_p(t)$. Let $\lambda_p = \widetilde{Sep}(t_p)\sqrt{(\log p) K_p p\widetilde{F}(t_p)}$. Using the same technique as for proving (A.59) we obtain that $\lambda_p \leq L_p p^{-\theta/2} p\widetilde{F}(t)$. Further, by Lemma A.17 we have

$$P(|\tilde{V}_p(t_p) - v_p(t_p)| \geq L_p p^{-\theta/2} p\widetilde{F}(t_p)) \leq \exp\Big(- \big(\widetilde{Sep}(t_p)\big)^2 (\log p)\big(1 + o(1)\big)\Big). \tag{A.60}$$

Combing (A.59) with (A.60) proves (A.58).

On the set B_p, since $v_p(t_p) \geq Cp\widetilde{F}(t_p)$ by Lemma A.4, we have $\frac{\tilde{V}_p(t_p)}{v_p(t_p)} = 1 + o(1)$, $\frac{\tilde{M}_p}{m_p(t_p)} = 1 + o(1)$. Therefore,

$$\frac{\tilde{M}_p(t_p)}{\sqrt{\tilde{V}_p(t_p)}} = \frac{m_p(t_p)}{\sqrt{v_p(t_p)}}\Big(1 + o(1)\Big) = \widetilde{Sep}(t_p)\big(1 + o(1)\big). \tag{A.61}$$

Combining (A.58) with (A.61), the misclassification rate can be bounded as

$$P(YL_t(X,\Omega)<0|t_p) \le \bar\Phi\left(\frac{1}{2}\widetilde{Sep}(t_p)\big(1+o(1)\big)\right) + P(B_p^c) \lesssim \bar\Phi\left(\frac{1}{2}\widetilde{Sep}(t_p)\big(1-o(1)\big)\right),$$

and

$$P(YL_t(X,\Omega)<0|t_p) \ge \bar\Phi\left(\frac{1}{2}\widetilde{Sep}(t_p)\big(1+o(1)\big)\right) P(B_p) \gtrsim \bar\Phi\left(\frac{1}{2}\widetilde{Sep}(t_p)\big(1+o(1)\big)\right).$$

Thus the claim follows easily. □

A.13. Proof of Lemma A.12. Write $\widetilde{Sep}(t) = \widetilde{Sep}(t,\epsilon_p,\tau_p,\Omega)$ for short. We consider the cases (a) $p\widetilde{F}(t) \ge K_p^8(\log p)^7$, $\sqrt{n_p}m_p(t) \ge K_p^8(\log p)^7$, (b) $\sqrt{n_p}m_p(t) \le K_p^8(\log p)^7$, $p\widetilde{F}(t) \ge K_p^8(\log p)^7$, and (c) $\sqrt{n_p}m_p(t) \ge K_p^8(\log p)^7$, $p\widetilde{F}(t) \le K_p^8(\log p)^7$ separately.

For case (a), define the event

$$B_p = \{\sup_{t\in A_p}\frac{|\tilde M_p(t)-m_p(t)|}{m_p(t)} \le \frac{1}{\sqrt{\log p}}, \qquad \sup_{t\in A_p}\frac{|\tilde V_p(t)-v_p(t)|}{p\widetilde{F}(t)} \le \frac{1}{\sqrt{\log p}}\}.$$

We will first prove $P(B_p^c) \le o(1/p)$. Let $\lambda = \lambda_p = CK_p^{-3}(\log p)^{-1/2}\sqrt{n_p}m_p(t)$ with $C>0$ some constant. Then by Lemma A.17, using similar arguments as those in Lemma A.16 we obtain that with probability at least $1-o(p^{-3})$, $|\tilde M_p(t)-m_p(t)| \le (\log p)^{-1/2}m_p(t)$. Using the grid points method as that in Lemma 3.1, we can prove that except for a probability of $o(1/p)$,

$$\sup_{t\in A_p,\sqrt{n_p}m_p(t)\ge K_p^7(\log p)^3}\frac{|\tilde M_p(t)-m_p(t)|}{m_p(t)} \le (\log p)^{-1/2}.$$

As for $\tilde V_p(t)$, using similar arguments and Lemma A.16 we obtain that with probability at least $1-o(1/p)$,

$$\sup_{t\in A_p,p\widetilde{F}(t)\ge K_p^7(\log p)^3}\frac{|\tilde V_p(t)-v_p(t)|}{p\widetilde{F}(t)} \le (\log p)^{-1/2}. \tag{A.62}$$

Thus we have proved the desired claim that $P(B_p) \ge 1-o(1/p)$.

Next by Lemma A.4, $p\widetilde{F}(t)/v_p(t) \le C$ for all $0<t\le\tau_p+\tilde s_p$. Then on the event B_p,

$$\frac{\tilde M_p(t)}{m_p(t)} = 1+o\big(\frac{1}{\sqrt{\log p}}\big) \qquad \frac{\tilde V_p(t)}{v_p(t)} = 1+o\big(\frac{1}{\sqrt{\log p}}\big)\frac{p\widetilde{F}(t)}{v_p(t)} = 1+o\big(\frac{1}{\sqrt{\log p}}\big),$$

where the $o(1)$ is uniformly over all t. Therefore, for any $t \in A_p$,

$$Sep(t) = \tilde{M}_p(t)/\sqrt{\tilde{V}_p(t)} = (1 + o(\frac{1}{\sqrt{\log p}}))m_p(t)/\sqrt{v_p(t)}$$
$$\leq (1 + o(\frac{1}{\sqrt{\log p}}))\frac{1}{2}\sup_{t\in A_p} \widetilde{Sep}(t),$$

and the desired claim in the lemma has been proved.

Now we consider case (b). By the proof of Lemma A.16 we obtain that except for a probability of $o(1/p)$, for any $t \in A_p$, $\tilde{M}_p(t) \leq m_p(t)+L_p n_p^{-1/2} \leq L_p n_p^{-1/2}$. Since we assumed that $p\widetilde{F}(t) \geq K_p^8(\log p)^7$, by (A.62) and the same arguments as that for (A.93), we have $\frac{\tilde{V}_p(t)}{v_p(t)} = 1 + o(\frac{1}{\sqrt{\log p}})$ except for a probability of $o(1/p)$. Since by Lemma A.4, $v_p(t) \geq Cp\widetilde{F}(t) \geq C(\log p)^{-1/2}$ with some constant $C > 0$ whose value depends on whether $r \geq \beta$ or $r < \beta$. Thus, with probability at least $1 - o(1/p)$, for any $t \in A_p$,

$$Sep(t) = \tilde{M}_p(t)/\sqrt{\tilde{V}_p(t)} \leq L_p n_p^{-1/2}/\sqrt{v_p(t)} \leq L_p n_p^{-1/2}. \tag{A.63}$$

Thus, the claim in the lemma follows automatically by the assumption that $\sup_{t\in A_p}\{\widetilde{Sep}(t)\} \geq p^\kappa$ with $\kappa > 0$.

Finally we consider case (c). By Lemma 3.1, $p\widetilde{F}_p(t) \leq L_p$ with probability at least $1 - o(1/p)$. Thus, using the same arguments as those for proving Lemma 3.4, part (1b) we obtain that with probability at least $1 - o(p^{-1})$,

$$Sep(t) = \tilde{M}_p(t)/\sqrt{\tilde{V}_p(t)} \leq L_p n_p^{-1/2}. \tag{A.64}$$

Using similar arguments as in case (b), we prove that the desired claim in the lemma continue to hold in case (c). This completes the proof of the lemma. □

A.14. Proof of Lemma 3.1. The following lemma is proved in Section A.15.

LEMMA A.13. *As $p \to \infty$, there is a constant $C > 0$ such that with probability at least $1 - o(1/p^3)$, for all $0 < t < \sqrt{2\log(p)}$,*

$$\frac{\sqrt{p}|\widetilde{F}_p(t) - \widetilde{F}(t)|}{\sqrt{\widetilde{F}(t)(1 - \widetilde{F}(t))}} \leq \begin{cases} CK_p^3(\log(p))^{1/2}, & \text{if } \frac{p}{2} > p\widetilde{F}(t) \geq \log^{5/4}(p), \\ CK_p^3\big(\log(p)\big)^{7/4}, & \text{if } p\widetilde{F}(t) < \log^{5/4}(p). \end{cases}$$

We now prove Lemma 3.1. Put an evenly spaced grid on $[0, \sqrt{2\log p}]$ by $t_k = (\sqrt{2\log p}/p^2)k$, $0 \leq k \leq p^2$. Denote by $V(t) = \sqrt{p}(\widetilde{F}_p(t) - \widetilde{F}(t))(\widetilde{F}(t)(1 - \widetilde{F}(t)))^{-1/2}$. For each $0 \leq i \leq p^2 - 1$, we claim that

$$\sup_{\{t_i \leq t \leq t_{i+1}\}} |V(t)| \leq \max\{|V(t_i)|, |V(t_{i+1})|\} + L_p/p. \tag{A.65}$$

In fact, as both $\widetilde{F}_p(t)$ and $\widetilde{F}(t)$ are monotone functions, we have

$$\frac{\widetilde{F}_p(t_{i+1}) - \widetilde{F}(t_i)}{\sqrt{\widetilde{F}(t_i)}} \leq \frac{\widetilde{F}_p(t) - \widetilde{F}(t)}{\sqrt{\widetilde{F}(t)}} \leq \frac{\widetilde{F}_p(t_i) - \widetilde{F}(t_{i+1})}{\sqrt{\widetilde{F}(t_{i+1})}}.$$

Let $h_i = \frac{\widetilde{F}(t_{i+1})}{\widetilde{F}(t_i)}$. Since $\widetilde{F}(t) \leq \frac{1}{2}$, $\sup_{\{t_i \leq t \leq t_{i+1}\}}\big\{|V(t)|\big\}$ does not exceed

$$2\Big(\max\{\sqrt{\frac{1}{h_i}}|V(t_i)|, \sqrt{h_i}|V(t_{i+1})|\} + \frac{\sqrt{p}|\widetilde{F}(t_i) - \widetilde{F}(t_{i+1})|}{\sqrt{\widetilde{F}(t_i)}} + \frac{\sqrt{p}|\widetilde{F}(t_i) - \widetilde{F}(t_{i+1})|}{\sqrt{\widetilde{F}(t_{i+1})}}\Big). \tag{A.66}$$

Since the derivative of $(-\widetilde{F}(t))$ is the density of a location normal mixture, and is therefore bounded from above. Moreover, for $0 < t < \sqrt{2\log p}$ and sufficiently large p, $\widetilde{F}(t) \geq \widetilde{F}(\sqrt{2\log p}) \geq 2(1 - K_p\epsilon_p)\bar{\Phi}(\sqrt{2\log p}) \geq p^{-1}L_p$. Using Taylor expansion,

$$\frac{\sqrt{p}|\widetilde{F}(t_i) - \widetilde{F}(t_{i+1})|}{\sqrt{\widetilde{F}(t_i)}} + \frac{\sqrt{p}|\widetilde{F}(t_i) - \widetilde{F}(t_{i+1})|}{\sqrt{\widetilde{F}(t_{i+1})}} \leq \frac{L_p}{\sqrt{p^3\widetilde{F}(t_i)}} + \frac{L_p}{\sqrt{p^3\widetilde{F}(t_{i+1})}} \leq L_p/p. \tag{A.67}$$

Similarly, we can show $|h_i - 1| \leq L_p/p$. Inserting this and (A.67) into (A.66) gives (A.65).

Combining (A.65) with Lemma A.13, the claim follows from

$$\sup_{\{0 \leq t \leq \sqrt{2\log(p)}\}} \Big[\frac{\sqrt{p}|\widetilde{F}_p(t) - \widetilde{F}(t)|}{\sqrt{\widetilde{F}(t)(1 - \widetilde{F}(t))}}\Big] = \sup_{\{0 \leq t \leq \sqrt{2\log(p)}\}} |V(t)| \leq C \sup_{\{0 \leq i \leq p^2\}} |V(t_i)| + \frac{L_p}{p},$$

where $C > 0$ is some constant. $\square$

A.15. Proof of Lemma A.13.

The following lemma is proved in Section A.16.

Lemma A.14. *There are partitions $\{1, 2, \ldots, p\} = R_1' \cup R_2' \ldots \cup R_{N_1}' = R_1'' \cup R_2'' \ldots \cup R_{N_2}''$ such that $N_1 \leq K_p$, $N_2 \leq K_p^2$, and that for any fixed $1 \leq j \leq N_1$ and $1 \leq k \leq N_2$, the collection of random variables $\{\tilde{Z}(i) - \tilde{\mu}(i), i \in R_j'\}$ are independent of each other, and the same are $\{\tilde{\mu}(i), i \in R_k''\}$.*

We now show Lemma A.13. The key idea is to combine Lemma 1.1 with the well-known Bennett's inequality (e.g., [3]). The Bennett's inequality only applies to sum of independent random variables. To apply it in the current setting, note that by Lemma A.14, we can partition $\{1, 2, \ldots, p\}$ into N different subsets $R_1, \ldots, R_N$, where $N \leq K_p^3$, such that the collection of random variables $\{\tilde{Z}(i) : i \in R_k\}$ are independent, for each $1 \leq k \leq N$. In light of this, we write $\widetilde{F}_p(t) = \frac{1}{p}\sum_{k=1}^N S_p^{(k)}(t)$, where $S_p^{(k)}(t) = \sum_{i \in R_k} 1\{|\tilde{Z}(i)| \geq t\}$ is the sum of independent random variables, to which the Bennet's inequality can be applied directly.

In detail, let $S^{(k)}(t) = E[S_p^{(k)}(t)]$ and $s_k = |R_k|$, $1 \leq k \leq N$, and $S(t) = \sum_{k=1}^N S^{(k)}(t)$. Since we are only interested in the region of t such that $\widetilde{F}(t) \leq 1/2$, it follows easily that

$$\frac{\sqrt{p}|\widetilde{F}_p(t) - \widetilde{F}(t)|}{\sqrt{\widetilde{F}(t)(1-\widetilde{F}(t))}} \lesssim \frac{\sqrt{2}|S_p(t) - S(t)|}{\sqrt{S(t)}} \leq \sum_{k=1}^N \frac{\sqrt{2}|S_p^{(k)}(t) - S^{(k)}(t)|}{\sqrt{S(t)}}. \tag{A.68}$$

For each $1 \leq k \leq N$, using Bennet's inequality [3, Page 851] yields

$$P\big(\big|S_p^{(k)} - S^{(k)}(t)\big| \geq \lambda\big) \leq 2\exp\bigg(-\frac{\lambda^2}{2s_k\sigma_k^2}\psi\big(\frac{\lambda}{s_k\sigma_k^2}\big)\bigg), \tag{A.69}$$

where ψ is as in [3, Page 851] and $s_k\sigma_k^2 = \mathrm{Var}(S_p^{(k)}(t))$. First, note that $x\psi(x)$ is monotonely increasing in $x \in (0, \infty)$. Second, by definitions and basic property of Bernoulli random variables, $s_k\sigma_k^2 \leq S^{(k)}(t) \leq S(t)$. Inserting these into (A.69) gives

$$P\bigg(\big[S_p^{(\ell)} - S^{(\ell)}(t)\big] \geq \lambda\bigg) \leq \exp\bigg(-\frac{\lambda^2}{2S(t)}\psi\big(\frac{\lambda}{S(t)}\big)\bigg).$$

Let $\lambda = C\sqrt{(\log p)S(t)}$ if $S(t) \geq \frac{1}{2}(\log p)^{5/4}$ and $\lambda = C(\log p)^{3/2}$ if $S(t) < \frac{1}{2}(\log p)^{5/4}$, where $C > 0$ is a constant. By elementary calculus and the property of ψ,

$$P\bigg(\big[S_p^{(\ell)} - S^{(\ell)}(t)\big] \geq \lambda\bigg) \leq \begin{cases} \exp\big(-\frac{C^2 \log p}{2}\big), & S(t) \geq \frac{1}{2}(\log p)^{5/4} \\ \exp\big(-\frac{C \log p}{2}\big), & S(t) < \frac{1}{2}(\log p)^{5/4}. \end{cases}$$

Inserting this into (A.68) and noting that $p\widetilde{F}(t) \geq (\log p)^{-1/2}$ give the claim. $\square$

A.16. Proof of Lemma A.14. Recall that $\tilde{Z} - \tilde{\mu} \sim N(0, \Omega)$, the first claim follows directly from Lemma 1.1. For the second claim, introduce a graph $\mathcal{G} = (V, E)$ where $V = \{1, 2, \ldots, p\}$, and nodes i and j are connected if and only if $S_i \cap S_j = \emptyset$, where $S_i = \{1 \leq k \leq p : \Omega(i,k) \neq 0\}$, $1 \leq i \leq p$. Since Ω is K_p-sparse, $\mathcal{G}$ is K_p^2-sparse. Also, $\tilde{\mu}(i)$ and $\tilde{\mu}(j)$ are independent if and only if nodes i and j are disconnected. Applying Lemma 1.1 to $\mathcal{G}$ gives the claim. $\square$

A.17. Proof of Lemma 3.3. Recall that $n_p = p^\theta$, $\hat{Z} = \hat{\Omega} Z$, and $\tilde{Z} = \Omega Z$. A direct result of Lemma 3.2 is that there is a term $0 < \eta_p \leq C K_p^3 (\log p) p^{-\theta/2}$ such that with probability at least $1 - o(1/p)$,

$$|1\{|\hat{Z}(j)| \geq t\} - 1\{|\tilde{Z}(j)| \geq t\}| \leq 1\{t - \eta_p \leq |\tilde{Z}(j)| < t + \eta_p\}, \ \forall t > 0 \text{ and } 1 \leq j \leq p.$$

Let $G_p(t) = \widetilde{F}_p(t - \eta_p) - \widetilde{F}_p(t + \eta_p)$ and $G(t) = \widetilde{F}(t - \eta_p) - \widetilde{F}(t + \eta_p)$. By the above inequality, it is seen that with probability at least $1 - o(1/p)$,

$$|\bar{F}_p(t) - \widetilde{F}_p(t)| \leq G_p(t). \tag{A.70}$$

We now analyze $G_p(t)$. By definitions and the triangle inequality,

$$G_p(t) \leq G(t) + |\widetilde{F}_p(t - \eta_p) - \widetilde{F}(t - \eta_p)| + |\widetilde{F}_p(t + \eta_p) - \widetilde{F}(t + \eta_p)|. \tag{A.71}$$

A key fact is that there is a universal constant $C > 0$ such that

$$|\widetilde{F}'(t)| \leq C(K_p \tau_p + t) \widetilde{F}(t). \tag{A.72}$$

To see the point, we write $\widetilde{F}(t) = \frac{1}{p} \sum_{i=1}^p E[\bar{\Psi}_{\sqrt{n_p}\tilde{\mu}(i)}(t)]$ and $\widetilde{F}'(t) = -\frac{1}{p} \sum_{i=1}^p E[\phi(t - \sqrt{n_p}\tilde{\mu}(i)) + \phi(t + \sqrt{n_p}\tilde{\mu}(i))]$, where ϕ is the density function of $N(0,1)$. Note that there is a constant $C > 0$ such that $\phi(x) \leq C|x|\bar{\Phi}(x)$, and that $|t \pm \sqrt{n_p}\tilde{\mu}(i)| \leq t + K_p \tau_p$ for all $1 \leq i \leq p$, the desired claim follows.

Now, first, write $G(t) = \widetilde{F}(t - \eta_p) - \widetilde{F}(t + \eta_p) = 2\eta_p \widetilde{F}'(\xi)$ for some number ξ with $|\xi - t| < \eta_p$. Using (A.72), $|\widetilde{F}'(\xi)| \leq C K_p \tau_p \widetilde{F}(\xi) \sim C K_p \tau_p \widetilde{F}(t)$. It follows

$$G(t) \leq C K_p \tau_p \widetilde{F}(t) \eta_p. \tag{A.73}$$

Second, by Lemma 3.1 and monotonicity, with probability at least $1 - o(1/p)$, $|\widetilde{F}_p(t \pm \eta_p) - \widetilde{F}(t \pm \eta_p)| \leq C K_p^3 (\log p)^{7/4} p^{-1/2} (\widetilde{F}(t \pm \eta_p))^{1/2}$, where by (A.72), $\widetilde{F}(t \pm \eta_p) \asymp \widetilde{F}(t)$. It follows that with probability at least $1 - o(1/p)$,

$$|\widetilde{F}_p(t \pm \eta_p) - \widetilde{F}(t \pm \eta_p)| \leq C K_p^3 (\log p)^2 p^{-1/2} (\widetilde{F}(t))^{1/2}. \tag{A.74}$$

Recall that $\eta_p \leq K_p^3(\log p)p^{-\theta/2}$. Inserting (A.73)-(A.74) into (A.71) gives

$$G_p(t) \leq CK_p^4(\log p)^{3/2}p^{-\theta/2}\widetilde{F}(t) + CK_p^3(\log p)^2 p^{-1/2}(\widetilde{F}(t))^{1/2}. \tag{A.75}$$

Combining (A.75) with (A.70) gives

(A.76)
$$\frac{\sqrt{p}|\bar{F}_p(t) - \widetilde{F}_p(t)|}{\sqrt{\widetilde{F}(t)(1-\widetilde{F}(t))}} \leq \frac{\sqrt{p}|G_p(t)|}{\sqrt{\widetilde{F}(t)(1-\widetilde{F}(t))}} \leq C\big(K_p^4(\log p)^{3/2}(p^{1-\theta}\widetilde{F}(t))^{1/2} + K_p^3(\log p)^2\big),$$

and the claim follows. □

A.18. Proof of Theorem 3.1. We consider the case when $p\widetilde{F}(t) < K_p^6(\log(p))^5$ and when $p\widetilde{F}(t) \geq K_p^6(\log(p))^5$ separately.

In the first case, it is sufficient to show that $|HC(t, \bar{F}_p)| \leq L_p$ and $|HC(t, \widetilde{F}_p)| \leq L_p$. By Lemmas 3.3 and 3.1, with probability at least $1 - o(1/p)$, $p|\bar{F}_p(t) - \widetilde{F}(t)| \leq L_p$. By Lemma A.4, $\widetilde{F}(t) \geq (1 - K_p\epsilon_p)\bar{\Psi}(t)$ and thus, $p\bar{\Psi}(t) \leq L_p$. Since $HC(t, \bar{F}_p)$ is defined in a way such that $\bar{F}_p(t) \geq 1/p$, it is easy to see that $HC(t, \bar{F}_p) \leq p|\bar{F}_p(t) - \bar{\Psi}(t)| \leq p|\bar{F}_p(t) - \widetilde{F}(t)| + p\widetilde{F}(t) + p\bar{\Psi}(t) \leq L_p$. Similarly, we can prove that $HC(t, \widetilde{F}_p) \leq L_p$. The claim follows easily.

In the second case, let $h(t) = (\widetilde{F}(t)(1-\widetilde{F}(t)))/(\bar{F}_p(t)(1-\bar{F}_p(t))$ and write for short $g(t) = \sqrt{p}\big(\bar{F}_p(t) - \widetilde{F}_p(t)\big)\big(\widetilde{F}(t)(1-\widetilde{F}(t)\big)^{-1/2}$. By definitions, we can write

$$HC(t, \bar{F}_p) - HC(t, \widetilde{F}) = g(t)\sqrt{h(t)} + HC(t, \widetilde{F})(\sqrt{h(t)} - 1). \tag{A.77}$$

We first prove $|h(t) - 1| \leq o(1)$. To see this, note that (A.76) and Lemma A.13 ensure that with probability at least $1 - o(1/p)$,

$$\begin{aligned} |\bar{F}_p(t)/\widetilde{F}(t) - 1| &\leq |(\bar{F}_p(t) - \widetilde{F}_p(t))/\widetilde{F}(t)| + |\widetilde{F}_p(t)/\widetilde{F}(t) - 1| \\ &\leq CK_p^4(\log p)^{3/2}p^{-\theta/2} + CK_p^3(\log p)^2(p\widetilde{F}(t))^{-1/2}. \end{aligned} \tag{A.78}$$

By the assumption of $p\widetilde{F}(t) \geq K_p^6(\log p)^5$, the right hand side of (A.78) tends to 0. Thus, with probability at least $1 - o(1/p)$, $0 \leq \bar{F}_p(t), \widetilde{F}(t) < 2/3$ for all $t > \bar{\Psi}^{-1}(1/2)$ and $p\tilde{F}(t) \geq K_p^6(\text{log}p)^5$. Note that for all $x, y \in (0, 2/3)$, $|[x(1-x)]/[y(1-y)] - 1| \leq C|x/y - 1|$. It follows from (A.78) and definitions that

$$|h(t) - 1| \leq C|\bar{F}_p(t)/\widetilde{F}(t) - 1| \leq L_p(p^{-\theta/2} + (p\widetilde{F}(t))^{-1/2}), \tag{A.79}$$

where the right hand side tends to 0 since $p\widetilde{F}(t) \geq K_p^6(\log p)^5$. At the same time, since $\widetilde{F}(t) \geq (1 - K_p\epsilon_p)\bar{\Psi}(t)$, we have $|\widetilde{F}(t) - \bar{\Psi}(t)| \leq \widetilde{F}(t) + \bar{\Psi}(t) \lesssim 2\widetilde{F}(t)$. It follows from $1 - \widetilde{F}(t) \geq 1 - \bar{\Psi}(t) - K_p\epsilon_p \geq 1/2 - K_p\epsilon_p$ that

$$|HC(t, \widetilde{F})| = \sqrt{p}|\widetilde{F}(t) - \bar{\Psi}(t)|(\widetilde{F}(t)(1 - \widetilde{F}(t))^{-1/2} \leq C(p\widetilde{F}(t))^{1/2}. \tag{A.80}$$

Combining (A.79) and (A.80) gives

$$HC(t, \widetilde{F})|\sqrt{h(t)} - 1| \leq L_p[(p^{1-\theta}\widetilde{F}(t))^{1/2} + 1]. \tag{A.81}$$

At the same time, a direct use of Lemma 3.3 also gives that with probability at least $1 - o(1/p)$,

$$g(t) \leq L_p[(p^{1-\theta}\widetilde{F}(t))^{1/2} + 1]. \tag{A.82}$$

Inserting (A.81) and (A.82) into (A.77) and recalling $|h(t) - 1| \to 0$ gives the claim. $\square$

A.19. Proof of Theorem 3.2. Write for short $\widehat{W}_p(t) = p^{-1/2}HC(t, \bar{F}_p)$ and $W(t) = p^{-1/2}HC(t, \widetilde{F}_p)$.

First consider the case of $\theta \geq \frac{1}{2}$. By triangle inequality, Theorem 3.1, and Lemma 2.3 we have

(A.83)

$$\begin{aligned}\sup_{\bar{\Psi}^{-1}(\frac{1}{2})<t<s_p^*} |\widehat{W}_p(t) - W_0(t)| &\leq \sup_{\bar{\Psi}^{-1}(\frac{1}{2})<t<s_p^*} |\widehat{W}_p(t) - W(t)| + \sup_{t>\bar{\Psi}^{-1}(\frac{1}{2})} |W(t) - W_0(t)| \\ &\leq \sup_{\bar{\Psi}^{-1}(\frac{1}{2})<t<s_p^*} L_p(p^{-\beta} + p^{-\theta/2}\sqrt{\widetilde{F}(t)} + p^{-1/2}) \leq L_p(p^{-\theta/2} + p^{-\beta}).\end{aligned}$$

This result is parallel to Lemma 2.3. When $r < \beta$, similar to (A.48) we can obtain that for all u satisfying $|u| \leq c_4/\tau_p$,

(A.84)

$$\widehat{W}_p(t_p^{**} + u) - \widehat{W}_p(t_p^{**}) \leq L_p(p^{-\frac{\theta}{2}} + p^{-\beta}) + [L_p p^{-c_0(\beta,r,a)} - \frac{c_7u^2}{1+u^2}] \sup_{\{t\geq 0\}} \widetilde{W}_0(t),$$

for some constant $c_7 > 0$, where t_p^{**} is as in (A.48). It is easy to check that $\sup_{\{t\geq 0\}} \widetilde{W}_0(t) = L_p p^{-\delta(\beta,r)} > p^{-\frac{1-\theta}{2}} > p^{-\theta/2}$, $c_0(\beta, r, a, \theta) < \beta$, and $p^{-c_0(\beta,r,a)} \sup_{t\geq 0} \widetilde{W}_0(t) \geq p^{-\beta}$. Thus, for any $u > L_p p^{-c_2(\beta,r,a)}$ with $c_2(\beta, r, a) < \min\{\frac{\theta - 2\delta(\beta,r)}{4}, \frac{c_0(\beta,r,a)}{2}\}$, it holds that $\widehat{W}_p(t_p^{**}+u) - \widehat{W}_p(t_p^{**}) = -L_p p^{-2c_2(\beta,r,a)}(1+o(1)) < 0$ for all $|u| \leq c_4/\tau_p$. Again, using similar arguments as in Theorem

2.1, we can prove that $\widehat{W}_p(t) - \widehat{W}_p(t_p^{**}) < 0$ for all $|t - t_p^{**}| > c_4/\tau_p$. Thus, we have proved that

$$|t_p^{HC} - t_p^{**}| = |T_{HC}(\bar{F}_p) - t_p^{**}| \leq L_p p^{-c_2(\beta,r,a)}.$$

This together with (A.49) completes the proof of the Theorem when $r < \beta$.

Now we consider the case where $r \geq \beta$. If $t > \tau_p$ or $t < \sqrt{2\beta \log p} - \Delta_1$ with $\Delta_1 = d_0 \log\log p/\sqrt{\log p}$, by Lemma 2.4 and (A.83), it holds $\widehat{W}_p(t) = W_0(t) + (\widehat{W}_p(t) - W_0(t)) \lesssim \frac{1}{\sqrt{2}} p^{-\beta/2} + L_p p^{-\beta} + L_p p^{-\theta/2}$. Recall that $\beta < 1 - \theta \leq \theta$. Thus $\widehat{W}_p(t) \lesssim \frac{1}{\sqrt{2}} p^{-\beta/2}(1 + o(1))$. If $\sqrt{2\beta \log p} - \Delta_1 < t < \tau_p$, using similar argument we obtain that $\widehat{W}_p(t) = W_0(t) + (\widehat{W}_p(t) - W_0(t)) \gtrsim p^{-\beta/2}(1 - o(1))$. Thus,

$$t_p^{HC} \in (\sqrt{2\beta \log p} - \Delta_1, \tau_p)$$

and the claim in the theorem follows.

Next we consider the case where $\theta < \frac{1}{2}$. By Theorem 3.1 and Lemma 2.4 and noting that $1 - \theta > \beta > \frac{1-\theta}{2}$, for any $t, t + u \in [s_p(\theta), s_p^*]$ we have

$$\widehat{W}_p(t+u) - \widehat{W}_p(t) = (\widehat{W}_p(t+u) - W_0(t+u)) - (\widehat{W}_p(t) - W_0(t))$$
$$+ (W_0(t+u) - W_0(t)) \leq L_p p^{-\theta/2}\sqrt{\widetilde{F}(t)} + L_p p^{-\beta} + (W_0(t+u) - W_0(t)).$$

Since $p^{-\theta}\widetilde{F}(t) \leq p^{-1+\theta}$ and $\beta > (1-\theta)/2$, it follows that

$$\widehat{W}_p(t+u) - \widehat{W}_p(t) \leq L_p p^{-(1-\theta)/2} + L_p p^{-\beta} + (W_0(t+u) - W_0(t)). \tag{A.85}$$

So the stochastic behavior of $W_0(t)$ in the range $t \in [s_p(\theta), s_p^*]$ determines the stochastic behavior of $\widehat{W}_p(t+u) - \widehat{W}_p(t)$. By direct calculations, we obtain that if (β, r, θ) falls in either of the six sub-regions as follows

- $1/3 < \theta \leq 1/2$, $(1-\theta)/2 < \beta < 1 - \theta$, $r > \max\{\rho_\theta^*(\beta), \frac{1-2\theta}{4}\}$,
- $\frac{1}{4} < \theta \leq \frac{1}{3}$, $(1-\theta)/2 < \beta \leq 1 - 2\theta$, $r > \max\{\frac{1-2\theta}{4}, \rho_\theta^*(\beta)\}$, $|r - \sqrt{1-2\theta}| \geq \sqrt{1 - 2\theta - \beta}$
- $\frac{1}{4} < \theta \leq \frac{1}{3}$, $1 - 2\theta < \beta \leq 1 - \theta$, $r > \max\{\frac{1-2\theta}{4}, \rho_\theta^*(\beta)\}$
- $0 < \theta \leq \frac{1}{4}$, $(1-\theta)/2 < \beta \leq 3(1-2\theta)/4$, $r > \max\{\frac{\beta}{3}, \rho_\theta^*(\beta)\}$, $|r - \sqrt{1-2\theta}| \geq \sqrt{1-2\theta-\beta}$
- $0 < \theta \leq \frac{1}{4}$, $3(1-2\theta)/4 < \beta \leq 1 - 2\theta$, $r > \max\{\frac{1-2\theta}{4}, \rho_\theta^*(\beta)\}$, $|r - \sqrt{1-2\theta}| \geq \sqrt{1-2\theta-\beta}$
- $0 < \theta \leq \frac{1}{4}$, $1 - 2\theta < \beta < 1 - \theta$, $r > \max\{\frac{1-2\theta}{4}, \rho_\theta^*(\beta)\}$,

then $t_p^{**} \in (s_p(\theta), s_p^*)$ and the maximum of $W_0(t)$ is achieved in $(s_p(\theta), s_p^*)$. So it reduces to the $\theta > 1/2$ case. Note that the six regions above can be summarized into Condition (a)-(b) in Theorem 1.3. By (A.85) and using similar proof as that for $\theta \geq \frac{1}{2}$ we finish the proof of Theorem 3.2. □

A.20. Proof of Lemma 3.4. Introduce $u_p(t) = u_p(t, \epsilon_p, \tau_p, \Omega) = \sum_{j=1}^p E\big[\tilde{\mu}(j)^2 \cdot 1\{|\tilde{Z}(j)| \geq t\}\big]$. The following lemma is proved in APPENDIX B.

LEMMA A.15. *For any $t > 0$, there are universal constants $C_1 > 0$ and $C_2 > 0$ such that for sufficiently large p, $C_1 \min\{t, \frac{1}{K_p\sqrt{2\log p}}\}\sqrt{n_p} \leq \frac{m_p(t,\epsilon_p,\tau_p,\Omega)}{u_p(t,\epsilon_p,\tau_p,\Omega)} \leq C_2(1+t)\sqrt{n_p}$ and $m_p(t, \epsilon_p, \tau_p, \Omega) \leq C_2(1+t)K_p^2\tau_p^2 n_p^{-1/2} p\widetilde{F}(t)$, where $\widetilde{F}(t)$ is defined in Lemma A.2.*

The following lemma is proved in Section A.21.

LEMMA A.16. *There is a constant $C > 0$ such that with probability at least $1 - o(1/p)$, for all $0 \leq t \leq \sqrt{2\log(p)}$,*

$$\sqrt{n_p}|M_p(t, \tilde{Z}, \mu) - m_p(t, \epsilon_p, \tau_p, \Omega)| \leq CK_p^5(\log p)^{13/4}\sqrt{p\widetilde{F}(t)}, \tag{A.86}$$

$$|V_p(t, \tilde{Z}, \mu) - v_p(t, \epsilon_p, \tau_p, \Omega)| \leq CK_p^4(\log p)^{3/2}\sqrt{p\widetilde{F}(t)}. \tag{A.87}$$

Write for short $\tilde{V}_p(t) = V_p(t, \tilde{Z}, \Omega)$, $\tilde{M}_p(t) = M_p(t, \tilde{Z}, \Omega, \mu)$, $m_p(t) = m_p(t, \epsilon_p, \tau_p, \Omega)$, $v_p(t) = v_p(t, \epsilon_p, \tau_p, \Omega)$, $\widetilde{Sep}(t) = \widetilde{Sep}(t, \epsilon_p, \tau_p, \Omega)$, $Sep(t) = Sep(t, \tilde{Z}, \mu, \Omega)$, $\widetilde{F}(t) = \widetilde{F}(t, \epsilon_p, \tau_p, \Omega)$ and $\widetilde{F}_p(t) = \widetilde{F}_p(t, \tilde{Z}, \mu, \Omega)$. We consider the two cases 1) $t > \tau_p + \tilde{s}_p$ or $p\widetilde{F}(t) \leq K_p^8(\log p)^4$, and 2) $t \leq \tau_p + \tilde{s}_p$ and $p\widetilde{F}(t) > K_p^8(\log p)^4$, separately, where $\tilde{s}_p$ is defined in Lemma A.4.

Consider the first case. It suffices to show (1a) $p^{(\theta-1)/2}\widetilde{Sep}(t) \leq L_p p^{-1/2} + L_p p^{-\max\{4\beta-2r, 3\beta+r\}/4}$ and (1b) $p^{(\theta-1)/2}Sep(t) \leq L_p p^{-\max\{4\beta-2r,3\beta+r\}/4} + L_p p^{-1/2}$. Claim (1a) can be proved using the same arguments as in Lemma 2.1, so we only need to prove (1b).

Consider (1b). Let η be a $p \times 1$ vector such that $\eta(j) = 1\{(\Omega\hat{\mu}_t^{\tilde{Z}})(j) \neq 0\}$, $1 \leq j \leq p$. Also, for any $p \times 1$ vectors x and y, let $x \circ y$ be the $p \times 1$ vector such that $(x \circ y)(j) = x(j)y(j)$, $1 \leq j \leq p$. By definitions, it is seen that $\tilde{M}_p(t) = (\hat{\mu}_t^{\tilde{Z}})'\Omega\mu = (\hat{\mu}_t^{\tilde{Z}})'\Omega(\mu \circ \eta)$. Using Cauchy-Schwartz inequality, $|\tilde{M}_p(t)| \leq \big((\hat{\mu}_t^{\tilde{Z}})'\Omega\hat{\mu}_t^{\tilde{Z}}\big)^{1/2}\big((\mu \circ \eta)'\Omega(\mu \circ \eta)\big)^{1/2}$. Recalling that $\tilde{V}_p(t) = (\hat{\mu}_t^{\tilde{Z}})'\Omega\hat{\mu}_t^{\tilde{Z}}$, it follows that

$$|Sep(t)| = 2|\tilde{M}_p(t)|(V_p(t))^{-1/2} \leq 2\big((\mu \circ \eta)'\Omega(\mu \circ \eta)\big)^{1/2}.$$

Since the largest eigenvalue of Ω is no greater than K_p, the last term above $\leq 2K_p^{1/2}\|\mu\circ\eta\|$ and so $|Sep(t)| \leq 2K_p^{1/2}\|\mu\circ\eta\|$. At the same time, by Lemma 3.1, with probability at least $1 - o(1/p)$, $p\widetilde{F}_p(t) \leq p|\widetilde{F}_p(t) - \widetilde{F}(t)| + p\widetilde{F}(t) \leq L_p(p\widetilde{F}(t))^{1/2} + p\widetilde{F}(t) \leq L_p p^{1-\max\{4\beta-2r,3\beta+r\}/2}$ if $t \geq \tau_p + \tilde{s}_p$. Similarly, we can show that $p\widetilde{F}_p(t) \leq L_p$ if $p\widetilde{F}(t) \leq K_p^8(\log p)^4$. Thus, in case (1b) we have $p\widetilde{F}_p(t) \leq L_p p^{1-\max\{4\beta-2r,3\beta+r\}/2} + L_p$. By definitions, this implies that $\hat{\mu}_t^{\tilde{Z}}$ has no more than $L_p p^{1-\max\{4\beta-2r,3\beta+r\}/2} + L_p$ non-zero coordinates. Since Ω is K_p-sparse, η also has no more than $L_p p^{1-\max\{4\beta-2r,3\beta+r\}/2} + L_p$ nonzero coordinates. Therefore, $\|\mu \circ \eta\| \leq L_p p^{\frac{1-\theta}{2}-\max\{4\beta-2r,3\beta+r\}/4} + L_p p^{-\theta/2}$, and (1b) follows from the assumption that $K_p \leq L_p$.

Consider the second case. Denote $h(t) = v_p(t)/\tilde{V}_p(t)$. The key is to show

$$|h(t) - 1| \leq L_p(p\widetilde{F}(t))^{-1/2}. \tag{A.88}$$

Towards this end, we write $|h(t) - 1| = I \cdot II \cdot h(t) \cdot (p\widetilde{F}(t))^{-1/2}$, where $I = |\tilde{V}_p(t) - v_p(t)|(p\widetilde{F}(t))^{-1/2}$, and $II = (p\widetilde{F}(t))/v_p(t)$. First, by Lemma A.16, $I \leq L_p$ with probability at least $1 - o(1/p)$. Second, by Lemma A.4, $II \leq C$ with some constant $C > 0$ whose value depends on whether $r < \beta$ and $t \leq \tau_p + \tilde{s}_p$ or $r \geq \beta$. Last, by Lemma A.16 and (A.87), with probability at least $1-o(1/p)$, $\tilde{V}_p(t)/v_p(t) \geq 1 - CK^4(\log p)^{3/2}\frac{(p\widetilde{F}(t))^{1/2}}{v_p(t)} \geq 1-o(1)$, where we note that $p\widetilde{F}(t) \geq K_p^8(\log p)^4$ and $CK_p^4(\log(p))^{3/2}(p\widetilde{F}(t))^{1/2}(v_p(t))^{-1} \lesssim K_p^4(\log(p))^{3/2}(p\widetilde{F}(t))^{-1/2} = o(1)$. As a result, with probability at least $1 - o(1/p)$, $h(t) = \frac{\tilde{V}_p(t)}{v_p(t)} \lesssim 1$. Combining these gives (A.88).

Next, write

$$|Sep(t) - \widetilde{Sep}(t)| = |\frac{\tilde{M}_p(t)}{\sqrt{\tilde{V}_p(t)}} - \frac{m_p(t)}{\sqrt{v_p(t)}}| \leq III + IV, \tag{A.89}$$

where $III = |\tilde{M}_p(t) - m_p(t)|\sqrt{h(t)}/\sqrt{v_p(t)}$ and $IV = m_p(t)|\sqrt{h(t)} - 1|/\sqrt{v_p(t)}$. Recall that $h(t) \lesssim 1 + L_p$ and that $Cp\widetilde{F}(t) \leq v_p(t)$. It follows from Lemma A.16 that with probability at least $1-o(1/p)$, $III \lesssim |\tilde{M}_p(t) - m_p(t)|(p\bar{F}_p(t))^{-1/2} \leq L_p n_p^{-1/2}$. At the same time, note that $IV \leq |h(t)-1| m_p(t)(v_p(t))^{-1/2}$. On one hand, by Lemmas A.4 and A.15, $m_p(t) \leq L_p n_p^{1/2} u_p(t) \leq L_p K_p^2 n_p^{-1/2} p\widetilde{F}(t)$. On the other hand, since $v_p(t) \geq Cp\widetilde{F}(t)$, by (A.88), we have $IV \leq L_p n_p^{-1/2}$ with probability at leats $1 - o(1/p)$. Combining these with (A.89) gives the claim.

By going through the proof above we see that if further $\Omega \in \widetilde{M}_p^*(a, b, K_p)$, then the two cases at the very beginning can be reduced to 1) $p\widetilde{F}(t) \leq K_p^8(\log p)^4$, and 2) $p\widetilde{F}(t) > K_p^8(\log p)^4$, and the claim $|Sep(t) - \widetilde{Sep}(t)| \leq L_p n_p^{-1/2}$ can be proved using the same arguments. Thus, Lemma 3.4 is proved.

□

A.21. Proof of Lemma A.16. Write for short $\tilde{M}_p(t) = M_p(t, \tilde{Z}, \mu, \Omega)$, $\tilde{V}_p(t) = V_p(t, \tilde{Z}, \Omega)$, $m_p(t) = m_p(t, \epsilon_p, \tau_p, \Omega)$, and $v_p(t) = E[V_p(t, \tilde{Z}, \Omega)]$. The following lemma is proved in Section A.22.

LEMMA A.17. *For any* $t \in (0, \sqrt{2\log p}]$,

$$P\Big(\sqrt{n_p}|\tilde{M}_p(t) - m_p(t)| \geq K_p^3 \lambda\Big) \leq K_p^3 \exp\Big(- \frac{\lambda^2 c_2}{2K_p\sqrt{2\log(p) n_p}\, m_p(t)} \psi\big(\frac{\lambda c_2}{\sqrt{n_p} m_p(t)}\big)\Big),$$

$$P\Big(|\tilde{V}_p(t) - v_p(t)| \geq K_p^3 \lambda\Big) \leq K_p^3 \exp\left(- \frac{\lambda^2}{4K_p p\widetilde{F}(t)} \psi\Big(\frac{\lambda}{2K_p p\widetilde{F}(t)}\Big)\right),$$

where ψ is as in Bennett's lemma [3, Page 851].

Since the proofs are very similar, we only show the first one. The goal is to show that with probability $1 - o(1/p^3)$, $|\tilde{M}_p(t) - m_p(t)| \leq CK_p^5(\log(p))^{13/4}(p\widetilde{F}(t))^{1/2}$ for any $0 \leq t \leq \sqrt{2\log(p)}$. Once this is shown, we lay out an evenly spaced grid on $[0, \sqrt{2\log(p)}]$ with an inter-distance of $1/p$, and the claim follows by similar argument as in the proof of Lemma 3.1.

Since Lemma A.15 ensures that $m_p(t) \leq CK_p^2(\log p)^{3/2} p n_p^{-1/2}\widetilde{F}(t)$, by the monotonicity of $x\psi(x)$ and Lemma A.17,

$$P\Big(\sqrt{n_p}|\tilde{M}_p(t) - m_p(t)| \geq K_p^3 \lambda\Big) \leq K_p^3 \exp\Big(- \frac{\lambda^2 c_2}{2CK_p^3(\log p)^2 p\widetilde{F}(t)} \psi\big(\frac{\lambda c_2}{CK_p^2(\log p)^{3/2} p\widetilde{F}(t)}\big)\Big). \tag{A.90}$$

We now show the desired claim for the case $p\widetilde{F}(t) \geq (\log(p))^{5/4}$ and the case $p\widetilde{F}(t) < (\log(p))^{5/4}$ separately.

Consider the first case. Let $\lambda = CK_p^2(\log p)^{3/2}\sqrt{p\widetilde{F}(t)}$. Direct calculations show that $\lambda/[K_p^2(\log p)^{3/2} p\widetilde{F}(t)] \leq C(p\widetilde{F}(t))^{-1/2}$ and $\lambda^2/[K_p^3(\log p)^2 p\widetilde{F}(t)] \geq C^2\log(p)K_p$. By (A.90) and noting that $\lim_{x\to 0+}\psi(x) = 1$,

$$P\left(\sqrt{n_p}|\tilde{M}_p(t) - m_p(t)| \geq CK_p^5(\log p)^{3/2}\sqrt{p\widetilde{F}(t)\}}\right) \leq K_p^3 \exp\Big(- \frac{C^2 K_p(\log p)}{2}\Big) \leq o(1/p^3).$$

Consider the second case. Let $\lambda = CK_p^2(\log p)^3$. It is seen that $\lambda/[K_p^2(\log p)^{3/2} p\widetilde{F}(t)] \geq C(\log(p))^{3/2}/(p\widetilde{F}(t))$. Using Lemma A.17 where we note that $\psi(x) \sim \frac{\log(x)}{x}$

when $x \to \infty$ [3, Page 852],

$$P\left(\sqrt{n_p}|\tilde{M}_p(t) - m_p(t)| \geq CK_p^5(\log p)^3\right) \leq K_p^3 \exp\Big(- \frac{C(\log p)}{2}\Big) \leq o(1/p^3).$$

This together with $p\widetilde{F}(t) \gtrsim p(1 - K_p\epsilon_p)\bar{\Psi}(t) \gtrsim (\log p)^{-1/2}$ yields the desired claim. □

A.22. Proof of Lemma A.17. Since the proofs are similar, we only show the first one. By Lemma 1.1, we can partition $\{1, \cdots, p\}$ into $N = N_1N_2 \leq K_p^3$ sets $R_1, \cdots, R_N$ such that for any fixed index $1 \leq k \leq N$, the collection of bivariate random variables $\{(\tilde{\mu}(j), \tilde{Z}(j)) : j \in R_k\}$ are independent of each other. Recall that $\tilde{M}_p(t) = \sum_{j=1}^p \tilde{\mu}(j)\text{sgn}(\tilde{Z}(j))1\{|\tilde{Z}(j)| \geq t\}$ and $m_p(t) = E[\tilde{M}_p(t)]$. The partition allows us to write $\tilde{M}_p(t) - m_p(t) = \sum_{k=1}^N[\tilde{M}_p^{(k)}(t) - m_p^{(k)}(t)]$, where $M_p^{(k)}(t) = \sum_{j \in R_k} \tilde{\mu}(j)\text{sgn}(\tilde{Z}(j))1\{|\tilde{Z}(j)| \geq t\}$ and $m_p^{(k)}(t) = E[M_p^{(k)}(t)]$, $1 \leq k \leq N$. It follows that for any $\lambda > 0$,

$$P(\sqrt{n_p}|\tilde{M}_p(t) - m_p(t)| \geq N\lambda) \leq \sum_{k=1}^N P(\sqrt{n_p}|\tilde{M}_p^{(k)}(t) - m_p^{(k)}(t)| \geq \lambda). \tag{A.91}$$

Fix $1 \leq k \leq N$, using Bennett's inequality [3, Page 851],

$$P(\sqrt{n_p}|\tilde{M}_p^{(k)}(t) - m_p^{(k)}(t)| \geq \lambda) \leq \exp\left(- \frac{\lambda^2}{2|R_k|\sigma_k^2}\psi\Big(\frac{\lambda K_p\sqrt{2\log p}}{|R_k|\sigma_k^2}\Big)\right),$$

where ψ is as in [3, Page 851], and $|R_k|\sigma_k^2$ is the variance of $\sqrt{n_p}\tilde{M}_p^{(k)}(t)$. Using Lemma A.15, $|R_k|\sigma_k^2 \leq n_pu_p(t) \leq c_2^{-1}K_p\sqrt{2\log(p)n_p}m_p(t)$. By the monotonicity of the function $x\psi(x)$ [3, Page 851], it follows that

$$P\Big(\sqrt{n_p}|\tilde{M}_p^{(k)}(t) - m_p^{(k)}(t)| \geq \lambda\Big) \leq \exp\Big(- \frac{\lambda^2 c_2}{2K_p\sqrt{2\log(p)n_p}m_p(t)}\psi\big(\frac{\lambda c_2}{\sqrt{n_p}m_p(t)}\big)\Big).$$

Inserting this into (A.91), the claim follows by recalling $N \leq K_p^3$. □

A.23. Proof of Lemma 3.5. Write for short $\hat{M}_p(t) = M_p(t, \hat{Z}, \mu, \Omega)$, $\tilde{M}_p(t) = M_p(t, \tilde{Z}, \mu, \Omega)$, $\hat{V}_p(t) = V_p(t, \hat{Z}, \Omega)$, and $\tilde{V}_p(t) = V_p(t, \tilde{Z}, \Omega)$, $m_p(t) = m_p(t, \epsilon_p, \tau_p, \Omega)$, and $v_p(t) = v_p(t, \epsilon_p, \tau_p, \Omega)$. We discuss the case 1) $t > \tau_p + \tilde{s}_p$ or $p\widetilde{F}(t) \leq K_p^{10}(\log p)^7$ and the case 2) $t \leq \tau_p + \tilde{s}_p$ and $p\widetilde{F}(t) > K_p^{10}(\log p)^7$ separately.

Consider the first case. First, in the proof of Lemma 3.4, we have shown that $Sep(t, \tilde{Z}, \mu, \Omega) \leq L_pp^{\frac{1-\theta}{2} - \frac{1}{4}\max\{4\beta - 2r, 3\beta + r\}} + L_pp^{-\theta/2}$. Second, by similar argument as in the proof Lemma 3.4 part (1b), and using Lemma 3.3, we can prove that $Sep(t, \hat{Z}, \mu, \hat{\Omega}) \le L_p p^{\frac{1-\theta}{2} - \frac{1}{4}\max\{4\beta-2r, 3\beta+r\}} + L_p p^{-\theta/2}$. Combining these gives the claim.

Consider the second case. The key is that with probability at least $1 - o(1/p)$,

(A.92)
$$\max\{\sqrt{n_p}|\hat{M}_p(t) - \tilde{M}_p(t)|, |\hat{V}_p(t) - \tilde{V}_p(t)|\} \le L_p \cdot [p^{1-\theta/2}\widetilde{F}(t) + (p\widetilde{F}(t))^{1/2}],$$
(A.93)
$$\max\{|(\tilde{V}_p(t) - v_p(t))/v_p(t)|, |\sqrt{n_p}(\tilde{M}_p(t) - m_p(t))/v_p(t)|\} = o(1).$$

for all $t \le \tau_p + \tilde{s}_p$. To see (A.92), note that $|\hat{M}_p(t) - \tilde{M}_p(t)| = |(\hat{\mu}_t^{\hat{Z}} - \hat{\mu}_t^{\tilde{Z}})'\Omega\mu| \le \|\hat{\mu}_t^{\hat{Z}} - \hat{\mu}_t^{\tilde{Z}}\|_1 \cdot \|\Omega\mu\|_\infty$, where by the K_p-sparsity of Ω, $\|\Omega\mu\|_\infty \le K_p\tau_p n_p^{-1/2}$, and so $|\hat{M}_p(t) - \tilde{M}_p(t)| \le K_p\tau_p n_p^{-1/2}\|\hat{\mu}_t^{\hat{Z}} - \hat{\mu}_t^{\tilde{Z}}\|_1$. Similarly, $\|\Omega(\hat{\mu}_t^{\hat{Z}} + \hat{\mu}_t^{\tilde{Z}})\|_\infty \le \|\Omega\|_1\|\hat{\mu}_t^{\hat{Z}} + \hat{\mu}_t^{\tilde{Z}}\|_\infty \le 2K_p$, and so $|\hat{V}_p(t) - \tilde{V}_p(t)| \le |(\hat{\mu}_t^{\hat{Z}} - \hat{\mu}_t^{\tilde{Z}})'\Omega(\hat{\mu}_t^{\hat{Z}} + \hat{\mu}_t^{\tilde{Z}})| \le 2K_p\|\hat{\mu}_t^{\hat{Z}} - \hat{\mu}_t^{\tilde{Z}}\|_1$. By similar argument as in the proof of Lemma 3.3, it is seen that with probability at least $1 - o(1/p)$, $\|\hat{\mu}_t^{\hat{Z}} - \hat{\mu}_t^{\tilde{Z}}\|_1 \le pG_p(t)$, where $G_p(t)$ is defined therein. It is shown in Lemma 3.3 that $G_p(t) \le CK_p^4(\log p)^{3/2}p^{-\theta/2}\widetilde{F}(t) + CK_p^3(\log p)^2 p^{-1/2}(\widetilde{F}(t))^{1/2}$ with probability at least $1 - o(1/p)$. Combining these gives (A.92).

To see (A.93), note that by Lemma A.16, with probability at least $1 - o(1/p)$,

$$|\tilde{V}_p(t) - v_p(t)| \le CK_p^4((\log(p))^{3/2}(p\widetilde{F}(t))^{1/2}. \tag{A.94}$$

Recall that by Lemma A.4, $v_p(t) \ge Cp\widetilde{F}(t)$ with some constant $C > 0$ whose value depends on whether $r < \beta$ or $r \ge \beta$. Combining this with the fact that $p\widetilde{F}(t) \ge K_p^{10}(\log p)^7$ for all $t \le \tau_p + \tilde{s}_p$, it is seen that $CK_p^4((\log(p))^{3/2}(p\widetilde{F}(t))^{1/2} = o(p\widetilde{F}(t)) = o(v_p(t))$. Inserting this into (A.94) gives that $|(\tilde{V}_p(t) - v_p(t))/v_p(t)| = o(1)$ with probability at least $1 - o(1/p)$. By similar argument, $|\sqrt{n_p}(\tilde{M}_p(t) - m_p(t))/v_p(t)| = o(1)$ with probability at least $1 - o(1/p)$. Combining these gives (A.93).

We now proceed to show the lemma in the second case. Let $h(t) = \hat{V}_p(t)/\tilde{V}_p(t)$. Write

(A.95)
$$\sqrt{n_p}|Sep(t, \hat{Z}, \mu, \hat{\Omega}) - Sep(t, \tilde{Z}, \mu, \Omega)| \le \sqrt{1/h(t)} \cdot I + |\sqrt{1/h(t)} - 1| \cdot II,$$

where $I = \sqrt{n_p}|\hat{M}_p(t) - \tilde{M}_p(t)|(\tilde{V}_p(t))^{-1/2}$ and $II = \sqrt{n_p}\tilde{M}_p(t)(\tilde{V}_p(t))^{-1/2}$. Recall that by Lemmas A.4 and A.15, $\sqrt{n_p}m_p(t) \le K_p^2(\log p)^{3/2}p\widetilde{F}(t) \lesssim$

$K_p^2(\log p)^{3/2}v_p(t)$. Using Lemma A.4 and (A.92)-(A.93),

$$|h(t)-1| = \frac{p\widetilde{F}(t)}{v_p(t)}\frac{v_p(t)}{\tilde{V}_p(t)}|\tilde{V}_p(t)-\hat{V}_p(t)|(p\widetilde{F}(t))^{-1} \leq L_p[p^{-\theta/2}+(\log p)^{5/2}(p\widetilde{F}(t))^{-1/2}],$$

$$I \leq L_p(p\widetilde{F}(t)/v_p(t))^{1/2}(p\widetilde{F}(t))^{-1/2}[p^{1-\theta/2}\widetilde{F}(t)+(p\widetilde{F}(t))^{1/2}] \leq L_p\cdot[p^{-\theta/2}(p\widetilde{F}(t))^{1/2}+1],$$

and

$$II \lesssim (p\widetilde{F}(t)/v_p(t))(p\widetilde{F}(t))^{-1}\sqrt{n_p}[|\tilde{M}_p(t)-m_p(t)|+m_p(t)] \leq L_p.$$

Recall that $p\widetilde{F}(t) \geq K_p^{10}(\log p)^7$. This together with the inequality above for $h(t)$ ensures that $|h(t)-1| \leq o(1)$. Inserting these into (A.95) gives

$$|Sep(t,\hat{Z},\mu,\hat{\Omega}) - Sep(t,\tilde{Z},\mu,\Omega)| \leq L_p\cdot n_p^{-1/2}[p^{-\theta/2}(p\widetilde{F}(t))^{1/2}+1],$$

and the claim follows.

Similarly to Lemma 3.4, we see that if in addition $\Omega \in \widetilde{M}_p^*(a,b,K_p)$, then the term $L_p p^{\frac{1-\theta}{2}-\frac{1}{4}\max\{4\beta-2r,3\beta+r\}}$ in the upper bound of the claim can be removed using the same proof as above. This concludes the proof of the lemma. □

APPENDIX B: PROOFS FOR SECONDARY LEMMAS

B.1. Proof of Lemma A.3. Note that $P(|X|\geq t,|Y|\geq t) = P(X\geq t, Y\geq t)+P(-X\geq t, Y\geq t)+P(X\geq t,-Y\geq t)+P(-X\geq t,-Y\geq t) \equiv I_1+I_2+I_3+I_4$. Consider I_3. Define $\tilde{Y} = 2\tau - Y$. Then $(X,\tilde{Y})$ has joint normal distribution with mean $(0,\tau)$ and correlation $-\rho$. Since $\tau \geq 0$, it is seen that $I_3 = P(X\geq t,\tilde{Y}\geq t+2\tau) \leq P(X\geq t,\tilde{Y}\geq t)$. Similarly, we can obtain that $I_4 \leq P(\tilde{X}\geq t,\tilde{Y}\geq t)$ with $\tilde{X} = -X$ and $\tilde{Y} = 2\tau - Y$. So we only need to bound I_1 and I_2.

Since the proofs are similar, we only show the case $\rho \geq 0$. Write $P(X\geq t|Y\geq t) = P(X\geq t, Y\geq t)/P(Y\geq t)$. First, by elementary calculus,

$$P(X\geq t, Y\geq t) \leq \begin{cases} C\exp(-\frac{t^2}{2}), & (t-\tau)\leq \rho t, \\ C\exp(-\frac{t^2-2\rho t(t-\tau)+(t-\tau)^2}{2(1-\rho^2)}), & (t-\tau)\geq \rho t. \end{cases}$$

Second, note that when $0\leq t\leq\tau$, $P(Y\geq t)\geq 1/2$, and that when $t\geq\tau$, $P(Y\geq t) = \bar{\Phi}(t-\tau) \geq C[1+(t-\tau)]^{-1}\phi(t-\tau)$ (e.g., by Mills' ratio [4]), where we note that $[1+(t-\tau)]^{-1}\geq(1+t)^{-1}$. Combining these with elementary algebra,

$$P(X\geq t|Y\geq t) \leq \begin{cases} C\exp(-t^2/2), & 0\leq t\leq\tau, \\ C(1+t)\exp(-\frac{t^2-(t-\tau)^2}{2}), & \tau < t < \frac{\tau}{1-\rho}, \\ C(1+t)\exp(-\frac{((1-\rho)t+\rho\tau)^2}{2(1-\rho^2)}), & t > \frac{1}{1-\rho}\tau. \end{cases}$$

Since $0 \le \rho \le a$, the claim follows by basic algebra. □

B.2. Proof of Lemma A.7. Define $f(t) = (1+t)\exp\big(-\frac{\tilde{a}t^2}{2}\big)\widetilde{W}_0(t)$ with $\tilde{a} = (1-a)/(1+a)$. We separate the cases of $q \le r$ and $q > r$. In the first case, $\bar{\Psi}_{\tau_p}(t_p(q)) = \bar{\Phi}(t_p(q) - \tau_p) + \bar{\Phi}(t_p(q) + \tau_p) \sim C$, where $C > 0$ is some constant. At the same time, by Mills' ratio we have $\bar{\Phi}(t_p(q)) \sim p^{-q}/t_p(q)$ as $p \to \infty$. Thus $\bar{\Psi}(t_p(q)) \sim 2p^{-q}/t_p(q)$. Combining these and noting that $\exp(-\frac{\tilde{a}}{2}t_p(q)^2) = p^{-\tilde{a}q}$, $t_p(q) = L_p$, and $\epsilon_p = p^{-\beta}$, we obtain that as $p \to \infty$,

$$f(t_p(q)) \sim L_p p^{-\tilde{a}q-\beta}/\sqrt{p^{-q} + p^{-\beta}}, \text{ if } q \le r. \tag{B.1}$$

In the second case, by Mills' ratio we have $\bar{\Psi}_{\tau_p}(t_p(q)) \sim L_p p^{-(\sqrt{q}-\sqrt{r})^2}$. Thus, similarly we have

$$f(t_p(q)) \sim L_p p^{-\tilde{a}q-\beta-(\sqrt{q}-\sqrt{r})^2}/\sqrt{p^{-q} + p^{-\beta-(\sqrt{q}-\sqrt{r})^2}}, \text{ if } q > r. \tag{B.2}$$

Define $\tilde{\delta}(q; r, \beta) = \beta + (\sqrt{q} - \sqrt{r})^2$ if $q \le r$, and $\tilde{\delta}(q; r, \beta) = \beta$ if $q > r$. Then combining (B.1) with (B.2) yields

$$f(t_p(q)) \sim L_p p^{-\tilde{a}q-\tilde{\delta}(q;r,\beta)}/\sqrt{p^{-q} + p^{-\tilde{\delta}(q;r,\beta)}}.$$

Since $r < \beta$, direct calculation shows that

$$\sup_{0<q<1} f(t_p(q)) \sim L_p p^{-\tilde{c}_0(\beta,r,a)-\delta(\beta,r)} = L_p p^{-\tilde{c}_0(\beta,r,a)} \sup_{0<q<1} \widetilde{W}_0(t_p(q)), \text{ as } p \to \infty,$$

where $\tilde{c}_0(\beta, r, a)$ is defined in (2.12).

B.3. Proof of Lemma A.10. Write $h(t) = \bar{\Phi}(t)/\phi(t)$ for short. For positive functions $f(t)$ and $g(t)$ defined over $(0, \infty)$, we say that $f(t) \asymp g(t)$ if there are constants $C_2 > C_1 > 0$ such that $C_1 \le f(t)/g(t) \le C_2$ for all $t > 0$. The following claims can be proved by elementary calculus and Mills' ratio [4] so we omit the proof. (a) $h(t) \asymp C\min\{1, 1/t\}$, (b) $h'(t)/h(t) = t - 1/h(t)$ and $(t^{-1} - t^{-3}) < h(t) < (t^{-1} - t^{-3} + 6t^{-5})$, and (c) $h'(-t)/h(-t) \le -C\max\{1, t\}$ for all $t > 0$.

To show the lemma, it suffices to show that $m_2'(t) < 0$ for all $t > 0$. Write

$$m_2(t) = \frac{1}{h(t)}\frac{\bar{\Phi}(t-\tau_p) + \bar{\Phi}(t+\tau_p)}{\phi(t-\tau_p) + \phi(t+\tau_p)} \equiv \frac{1}{h(t)}\frac{h(t-\tau_p)\phi(t-\tau_p) + h(t+\tau_p)\phi(t+\tau_p)}{\phi(t-\tau_p) + \phi(t+\tau_p)}.$$

We show this for the case of $t \ge \tau_p$ and the case of $t < \tau_p$ separately.

Consider the first case. By direct calculations, it is seen

(B.3)
$$m_2(t) = \frac{1}{1+e^{-2\tau_p t}}[h(t-\tau_p)/h(t)] + \frac{e^{-2\tau_p t}}{1+e^{-2\tau_p t}}[h(t+\tau_p)/h(t)] \equiv m_{2a}(t) + m_{2b}(t).$$

Write for short $\xi(t) = h'(t-\tau_p)/h(t-\tau_p) - h'(t)/h(t)$. By (a)-(b) and direct calculations,

$$|m_{2b}'(t)| \le C\tau_p e^{-t\tau_p}, \qquad m_{2a}'(t) = \xi(t)[h(t-\tau_p)/h(t)] + O(\tau_p t e^{-\tau_p t}),$$

where we note $h(t-\tau_p)/h(t) \ge C$. Note that the claim follows trivially if $t \le \tau_p + 3$. Therefore, to show the claim, it is sufficient to show $\xi(t) \le -C\tau_p^{-1}\min\{1, (\tau_p/t)^2\}$ for all $t > \tau_p + 3$. Toward this end, note that by basic algebra and (b),

$$\xi(t) = -\tau_p - \frac{1}{h(t-\tau_p)} + \frac{1}{h(t)} \le -\tau_p - \frac{(t-\tau_p)}{(1-(t-\tau_p)^{-2} + 6(t-\tau_p)^{-4})} + \frac{t}{1-t^{-2}}.$$

By basic algebra, we have that for sufficiently large τ_p and $t > \tau_p + 3$,

$$\xi(t) \le -(t-\tau_p)^{-1}\Big[\frac{1-6(t-\tau_p)^{-2}}{1-(t-\tau_p)^{-2}+6(t-\tau_p)^{-4}}\Big] + 1/t + 2t^{-3}.$$

The claim now follows from elementary calculus.

Consider the second case. Rewrite

$$m_2(t) = \frac{1}{[1+e^{-2\tau_p t}]h(t)} h(t-\tau_p) + \frac{e^{-2\tau_p t}}{1+e^{-2\tau_p t}} \frac{h(t+\tau_p)}{h(t)} \equiv m_{2c}(t)h(t-\tau_p) + m_{2d}(t),$$

and so

$$m_2'(t) = m_{2c}'(t)h(t-\tau_p) + m_{2c}(t)h'(t-\tau_p) + m_{2d}'(t).$$

Similarly, by (a)-(c),

$$|m_{2d}'(t)| \le C\tau_p^{-1},\ m_{2c}'(t) \le C,\ m_{2c}(t)h'(t-\tau_p) \le -C\max\{1,t\}\cdot\max\{1,(\tau_p-t)\}h(t-\tau_p).$$

Combining these gives

$$m_2'(t) \le C[-\max\{1,t\}\cdot\max\{1,(\tau_p-t)\} + C]h(t-\tau_p) + C.$$

Since $h(t-\tau_p) \ge C$, it is seen that $m_2'(t) < 0$ for sufficiently large τ_p and the claim follows.

The second claim $m_2(t) > 1$ follows directly from the first claim and $\lim_{t\to\infty} m_2(t) = 1$, which can be obtained immediately by (B.3). □

B.4. Proof of Lemma A.15. Recall that $\bar{\Phi} = 1 - \Phi$ is the survival function of $N(0,1)$. The following lemma is proved below.

LEMMA B.1. *For any $t > 0$ and $u > 0$, there are universal constants $C_1 > 0$ and $C_2 \geq 1$ such that* $C_1 \min\{t, \frac{1}{u}\} \leq \frac{1}{u} \cdot \frac{\bar{\Phi}(t-u)-\bar{\Phi}(t+u)}{\bar{\Phi}(t-u)+\bar{\Phi}(t+u)} \leq C_2(1+t)$.

We now show Lemma A.15. Let $\tilde{\mu} = \Omega\mu$ for short. First, by definitions, $\sqrt{n_p} m_p(t) = \sum_{j=1}^p E[\sqrt{n_p}\tilde{\mu}(j)\mathrm{sgn}(\tilde{z}(j))1\{|\tilde{Z}(j)| \geq t\}] = \sum_{j=1}^p E[\sqrt{n_p}\tilde{\mu}(j)(\bar{\Phi}(t- \sqrt{n_p}\tilde{\mu}(j)) - \bar{\Phi}(t+\sqrt{n_p}\tilde{\mu}(j)))]$. Noting that for any fixed $t > 0$, $u[\bar{\Phi}(t-u) - \bar{\Phi}(t+u)]$ is a symmetric function,

$$\sqrt{n_p} m_p(t) = \sum_{j=1}^p E[|\sqrt{n_p}\tilde{\mu}(j)|(\bar{\Phi}(t - |\sqrt{n_p}\tilde{\mu}(j)|) - \bar{\Phi}(t+|\sqrt{n_p}\tilde{\mu}(j)|))]. \tag{B.4}$$

Similarly, we have

$$n_p u_p(t) = \sum_{j=1}^p E[n_p \tilde{\mu}^2(j)(\bar{\Phi}(t - |\sqrt{n_p}\tilde{\mu}(j)|) + \bar{\Phi}(t+|\sqrt{n_p}\tilde{\mu}(j)|))]. \tag{B.5}$$

Since that Ω is K_p-sparse and that $|\sqrt{n_p}\mu(j)| \leq \tau_p \leq \sqrt{2\log(p)}$, $|\sqrt{n_p}\tilde{\mu}(j)| = \sum_{k=1}^p |\Omega(j,k)| \cdot |\sqrt{n_p}\mu(k)| \leq K_p\sqrt{2\log(p)}$. Comparing (B.4) and (B.5), the first claim follows by Lemma B.1. The second claim follows easily from the first claim and that $|\sqrt{n_p}\tilde{\mu}(j)| \leq K_p \tau_p$. □

B.5. Proof of Lemma B.1. Consider the first inequality first. Let $\phi(\cdot)$ be the density of $N(0,1)$. For any real number v, write

$$\frac{\bar{\Phi}(t-v)}{\phi(t-v)} = \frac{\int_0^\infty \phi(x+(t-v))dx}{\phi(t+v)} = \int_0^\infty e^{-(t-v)x} e^{-x^2/2} dx,$$

where the right hand side is strictly monotone in v. Therefore, $\bar{\Phi}(t-u)/\phi(t-u) \geq \bar{\Phi}(t+u)/\phi(t+u)$ or equivalently, $\bar{\Phi}(t+u)/\bar{\Phi}(t-u) \leq \phi(t+u)/\phi(t-u)$. Combining this with basic algebra,

$$\frac{1}{u}\left[\frac{\bar{\Phi}(t-u)-\bar{\Phi}(t+u)}{\bar{\Phi}(t-u)+\bar{\Phi}(t+u)}\right] \geq \frac{1}{u}\left[\frac{\phi(t-u)-\phi(t+u)}{\phi(t-u)+\phi(t+u)}\right] = \frac{t}{ut}\left[\frac{e^{tu}-e^{-tu}}{e^{tu}+e^{-tu}}\right]. \tag{B.6}$$

When $0 < ut \leq 1$, the right hand side $\geq t \cdot \inf_{0<x<1}\{\frac{1}{x}\frac{e^x-e^{-x}}{e^x+e^{-x}}\}$. When $ut \geq 1$, by the monotonicity of the function $(e^x - e^{-x})/(e^x + e^{-x})$, the right hand side $\geq (1/u)\cdot[(e^{tu}-e^{-tu})/(e^{tu}+e^{-tu})] \geq (1/u)\cdot[(e-e^{-1})/(e+e^{-1})]$. Letting $C_1 = \min\{\inf_{0<x<1}\{\frac{1}{x}\frac{e^x-e^{-x}}{e^x+e^{-x}}\}, (e-e^{-1})/(e+e^{-1})\}$ gives the claim.

Consider the second inequality. When $u > 1$, the claim follows trivially, so we consider the case $0 < u \leq 1$ only. By Taylor expansion, there is a constant $c_3 \geq 1$ such that

$$\frac{1}{u}\frac{\bar{\Phi}(t-u)-\bar{\Phi}(t+u)}{\bar{\Phi}(t-u)+\bar{\Phi}(t+u)} \leq \frac{2u\max_{\{t-u<s<t+u\}}\{\phi(s)\}}{\bar{\Phi}(t-u)} \leq c_3\frac{\phi(t-u)}{\bar{\Phi}(t-u)}, \tag{B.7}$$

where in the second inequality we have used $t > 0$ and $u < 1$. At the same time, By Mills' ratio [4], there is a constant $c_4 > 0$ such that $\bar{\Phi}(t) \leq c_4 \cdot (t\phi(t))$. Therefore, $\phi(t-u)/\bar{\Phi}(t-u) \leq c_4(1+|t-u|) \leq 2c_4(1+t)$. Insert this into (B.7). The claim follows by letting $C_2 = \max\{1, 2c_3c_4\}$. □

REFERENCES

[1] Donoho, D. and Jin, J. (2009). Feature selection by higher criticism thresholding achieves the optimal phase diagram. *Phil. Trans. R. Soc. A* 367 4449-4470.

[2] Fan, Y., Jin, J. and Yao, Z. (2013). Optimal classification in sparse Gaussian graphic model. *arXiv:1212.5332*.

[3] Shorack, G. R. and Wellner, J. A. (1986). *Empirical processes with applications to statistics*. Wiley, New York.

[4] Wasserman, L. (2006). *All of nonparametric statistics*. Springer, New York.

Y. Fan
Department of Data Sciences and Operations
Marshall School of Business
University of Southern California
Los Angeles, CA 90089
USA
E-mail: fanyingy@marshall.usc.edu

J. Jin
Department of Statistics
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
USA
E-mail: jiashun@stat.cmu.edu

Z. Yao
Section de Mathématiques
Ecole Polytechnique Fédérale de Lausanne
EPFL Station 8, 1015 Lausanne
Switzerland
E-mail: zhigang.yao@epfl.ch