SUPPLEMENT TO "HYPOTHESIS TESTING FOR HIGH-DIMENSIONAL SPARSE BINARY REGRESSION"

NOTATIONS

We begin by briefly summarizing notation. We recall the definition of our chosen prior π for the sake of completeness. We choose π to be uniform over all k sparse subsets of $\mathbb{R}^p$ with signal strength either A or $-A$. Let $M(k,p)$ be the collection of all subsets of $\{1,\ldots,p\}$ of size k. For each $m \in M(k,p)$, let $\xi^m = (\xi_j)_{j\in m}$ be a sequence of independent Rademacher random variables taking values in $\{+1,-1\}$ with equal probability. Given $A > 0$ for testing (2.3), a realization from the prior distribution π on $\mathbb{R}^p$ can be expressed as $\beta_{\xi,m} = \sum_{j\in m} A\xi_j e_j$, where $(e_j)_{j=1}^p$ is the canonical basis of $\mathbb{R}^p$ and m is uniformly chosen from $M(k,p)$. In the following we will define m_1, m_2 to be two independent draws at random from $M(k,p)$ and $\xi_1 = (\xi_1^j)_{j\in m}, \xi_2 = (\xi_2^j)_{j\in m}$ the corresponding draws of a sequence of Radamacher random variables. Further we denote by m_3 and m_4 the set valued random variables $m_3 := \{j \in m_1 \cap m_2 : \xi_1^j = \xi_2^j\}$ and $m_4 := \{j \in m_1 \cap m_2 : \xi_1^j = -\xi_2^j\}$. Also ϕ, Φ and $\overline{\Phi}$ denote the standard normal pdf, cdf and survival functions respectively. We let Hypergeometric(N,m,n) denote the hypergeometric distribution counting the number of red balls in n draws from an urn containing m red balls out of N. Also throughout C will denote generic positive constants whenever necessary.

PRELIMINARY LEMMAS

We will use the following results many times and hence present them as useful lemmas.

The first result compares the hypergeometric distribution with a related binomial distribution, which is in general simpler to work with.

LEMMA A.1. *If $W \sim$ Hypergeometric(N,m,n) and $Y \sim \mathrm{Bin}(n, \frac{m}{N-m})$ then W is stochastically smaller than Y, i.e., $\mathbb{P}(Y \geq t) \geq \mathbb{P}(X \geq t)$ for all $t \in \mathbb{R}$. Moreover this implies that for any non-decreasing function g one has $\mathbb{E}(g(W)) \leq \mathbb{E}(g(Y))$.*

PROOF. The proof can be found in Arias-Castro, Candès and Plan (2011) and follows by noting that if the balls are picked one by one without replacement, then at each stage, the probability of selecting a red ball is smaller than $m/(N-m)$. □

The next result presents an inequality about the tail probability of a binomial random variable (Carter and Pollard, 2004)

Lemma A.2. *Let $X \sim \mathrm{Bin}(n, \frac{1}{2})$ with $n \geq 28$. Define*

$$\gamma(\epsilon) = \frac{(1+\epsilon)\log(1+\epsilon) + (1-\epsilon)\log(1-\epsilon) - \epsilon^2}{2\epsilon^4} = \sum_{l=0}^{\infty} \frac{\epsilon^2 l}{(2l+3)(2l+4)},$$

an increasing function. Suppose $\frac{n}{2} < k' \leq n-1$. Define $\epsilon = (2K-N)/N$, where $K = k'-1$ and $N = n-1$. Then there exists a λ_n such that $\frac{1}{12n+1} < \lambda_n < \frac{1}{12n}$ and a constant C such that

$$\mathbb{P}(X \geq k') = \overline{\Phi}(\epsilon\sqrt{N})e^{A_n(\epsilon)}$$

where

$$A_n(\epsilon) = -N\epsilon^4\gamma(\epsilon) - \frac{1}{2}\log(1-\epsilon^2) - \lambda_{n-k} + r_{k'}$$

and

$$-C\log N \leq N r_{k'} \leq C$$

for all ϵ corresponding to the range $\frac{n}{2} < k' \leq n-1$.

The next lemma shows that any random draw of a subset of size k from $\{1,\ldots,p\}$ can have at most one element in each block. The proof of the lemma is similar to the proof of Lemma A.8 of Hall and Jin (2010) and is omitted.

Lemma A.3. *Let $t_1 < t_2 < \ldots < t_k$ be k distinct indices randomly sampled from $\{1,\ldots,p\}$ without replacement. Then for any $1 \leq Q \leq p$ we have $\mathbb{P}(\min_{1\leq i\leq k-1}|t_{i+1}-t_i| \leq Q) \leq Qk(k+1)/p$.*

The next Lemma is tailored towards controlling the contribution of the i^{th} row in the expression for $\mathbb{E}_0(L_\pi^2)$.

Lemma A.4. *Suppose for the i^{th} row of $\mathbf{X}$ one has $|S_i| \leq Q$ and that the elements of $\mathbf{X}$ are bounded by M in absolute value. Then for any $\beta, \beta' \sim \pi$,*

$$\theta(\mathbf{x}_i^t\boldsymbol{\beta})\theta(\mathbf{x}_i^t\boldsymbol{\beta}') + \theta(-\mathbf{x}_i^t\boldsymbol{\beta})\theta(-\mathbf{x}_i^t\boldsymbol{\beta}') \leq \theta^2(QMA) + \theta^2(-QMA).$$

where θ is the distribution function of a symmetric random variable, i.e., θ satisfies Equation 2.2.

PROOF. We begin by noting that for any i,

$$\theta(\mathbf{x}_i^t\boldsymbol{\beta})\theta(\mathbf{x}_i^t\boldsymbol{\beta}^{'})+\theta(-\mathbf{x}_i^t\boldsymbol{\beta})\theta(-\mathbf{x}_i^t\boldsymbol{\beta}^{'}) \leq \sup_{s_1,s_2\in[-MQ,MQ]} \theta(s_1A)\theta(s_2A)+\theta(-s_1A)\theta(-s_2A).$$

Hence by symmetry of the above supremum in s_1, s_2 and using the fact that $\theta(z) + \theta(-z) = 1$ for all w, we have that

$$\theta(\mathbf{x}_i^t\boldsymbol{\beta})\theta(\mathbf{x}_i^t\boldsymbol{\beta}^{'}) + \theta(-\mathbf{x}_i^t\boldsymbol{\beta})\theta(-\mathbf{x}_i^t\boldsymbol{\beta}^{'}) \leq \max_{s\in[0,MQ]} (\theta(sA))^2 + (1-\theta(sA))^2.$$

Now noting that $(1-w)^2 + w^2$ is an increasing function of w for $w \geq \frac{1}{2}$ and using the fact that $\theta(sA) \geq \frac{1}{2}$ for $s \geq 0$, we have the desired result. □

PROOF OF MAIN RESULTS

PROOF OF THEOREM 3.1 . We will produce one prior $\pi_0 \sim \pi$ for which the theorem holds. Hence, for any other $\pi^* \sim \pi$, since one also has $\pi^* \sim \pi_0$ we have the result holding by a similar proof. We begin by noting that

$$\theta(\mathbf{x}_i^t\boldsymbol{\beta})\theta(\mathbf{x}_i^t\boldsymbol{\beta}^{'}) + \theta(-\mathbf{x}_i^t\boldsymbol{\beta})\theta(-\mathbf{x}_i^t\boldsymbol{\beta}^{'}) \leq 1 \text{ for all } i, \boldsymbol{\beta}, \boldsymbol{\beta}^{'} \tag{A.1}$$

The proof of (A.1) follows from noting that for any two real numbers w_1, w_2, one has by symmetry $\theta(w_1)\theta(w_2)+\theta(-w_1)\theta(-w_2) \leq \sup_{w\in\mathbb{R}}[2\theta^2(w)-2\theta(w)+1]$. Since θ is a distribution function of a symmetric random variable as posed by equation (2.2), it is easy to show that $2\theta^2(w) - 2\theta(w) + 1$ is an increasing function of w. Hence we have that the supremum equals 1 and thus proving (A.1). Now, recall that it suffices to bound from below the second moment $\mathbb{E}_0(L_\pi^2)$ where by Fubini's Theorem

$$\begin{aligned}
\mathbb{E}_0(L_\pi^2) &= 2^n \iint \prod_{i=1}^n \Big[\theta(\mathbf{x}_i^t\boldsymbol{\beta})\theta(\mathbf{x}_i^t\boldsymbol{\beta}^{'}) + \theta(-\mathbf{x}_i^t\boldsymbol{\beta})\theta(-\mathbf{x}_i^t\boldsymbol{\beta}^{'})\Big] d\pi(\boldsymbol{\beta})d\pi(\boldsymbol{\beta}^{'}) \\
&\leq \iint 2^{n-\sum\limits_{i=1}^n \mathcal{I}(\min\{|m_1\cap S_i|,|m_2\cap S_i|\}=0)} d\pi(\boldsymbol{\beta})d\pi(\boldsymbol{\beta}^{'}) \\
&= \iint 2^{\sum\limits_{i=1}^n \mathcal{I}(\min\{|m_1\cap S_i|,|m_2\cap S_i|\}>0)} d\pi(\boldsymbol{\beta})d\pi(\boldsymbol{\beta}^{'}).
\end{aligned} \tag{A.2}$$

The inequality in the second to last line above follows from noting that, when i is such that one of $S_i \cap m_1$ or $S_i \cap m_2$ is empty, then the integrand $\theta(\mathbf{x}_i^t\boldsymbol{\beta})\theta(\mathbf{x}_i^t\boldsymbol{\beta}^{'}) + \theta(-\mathbf{x}_i^t\boldsymbol{\beta})\theta(-\mathbf{x}_i^t\boldsymbol{\beta}^{'}) = \frac{1}{2}$, whereas for any other i, the integrand is less than or equal to 1 by (A.1). Applying Lemma A.3 we obtain that when $\alpha > \frac{1}{2}$, *i.e.*, $k = p^{1-\alpha} \ll \sqrt{p}$, it makes negligible difference by restricting π to $R_p = \{\{t_1,\ldots,t_k\}, \min_{1\le i\le k-1}|t_{i+1}-t_i| > \sigma_p\}$ where by assumption σ_p is such that $\sigma_p \ll p^\epsilon$ for all $\epsilon > 0$. If we denote this restricted prior by π_0, then we have $\pi_0 \sim \pi$ and $R_{\pi_0} = R_p$. Now by elementary combinatorics,

$$|R^N_{m_1}(\sigma_p)| \lesssim \binom{k}{N}(2\sigma_p)^N \binom{p-N}{k-N} \le \binom{k}{N}(2\sigma_p)^N \binom{p}{k-N}.$$

Also by direct calculation,

$$\frac{\binom{k}{N}\binom{p}{k-N}}{\binom{p}{k}} = \frac{1}{N!}\left(\frac{k!}{(k-N)!}\right)^2 \frac{(p-k)!}{(p-k+N)!} \lesssim \frac{1}{N!}\left(\frac{k^2}{p}\right)^N.$$

Hence from (A.2) and assumption of the Theorem we have that

$$\begin{aligned}
\mathbb{E}_0(L_\pi^2) &\le \binom{p}{k}^2 \sum_{m_1\in R_{\pi_0}} \sum_{N=0}^{k} \sum_{m_2 \in R^N_{m_1}(\sigma_p)} 2^{\sum_{i=1}^{n} \mathcal{I}(\min\{|m_1\cap S_i|,|m_2\cap S_i|\}>0)}(1+o(1)) \\
&\le \binom{p}{k}^2 \sum_{m_1\in R_{\pi_0}} \sum_{N=0}^{k} \sum_{m_2 \in R^N_{m_1}(\sigma_p)} 2^{N\delta_p}(1+o(1)) \\
&\lesssim \binom{p}{k} \sum_{m_1\in R_{\pi_0}} \sum_{N=0}^{\infty} \frac{2\frac{k^2}{p}\sigma_p 2^{\delta_p}{}^N}{N!}(1+o(1)) \\
&= \binom{p}{k} \sum_{m_1\in R_{\pi_0}} e^{2\frac{k^2}{p}\sigma_p 2^{\delta_p}}(1+o(1)) \\
&= e^{2\frac{k^2}{p}\sigma_p 2^{\delta_p}}(1+o(1))
\end{aligned}$$

Since σ_p is a poly-logarithmic factor of p and $k = p^{1-\alpha}$ with $\alpha > \frac{1}{2}$, we have that $\delta_p \ll \log(p)$ implies that $\mathbb{E}_0(L_\pi^2) = 1 + o(1)$. Hence all tests are asymptotically powerless as required. $\square$

Proof of Theorem 3.2. The proof relies on verifying the assumptions and conditions of Theorem 3.1. To begin with we produce a prior that is equivalent to π as follows. Let π_0 be the restriction of π to $R_p = \{\{t_1,\ldots,t_k\}, \min_{1\le i\le k-1}|t_{i+1}-t_i| > \sigma_p\}$ and let $\pi_{0,1}$ be the restriction of π_0 to $(\bigcup_{i\notin\Omega} S_i)^c$ where $\sigma_p \ge 2l^*$ is such that $\sigma_p \ll p^\epsilon$ for all $\epsilon > 0$. We note that such a σ_p can be found since we have by assumption $l^* \ll p^\epsilon$ for all $\epsilon > 0$. Since $k = p^{1-\alpha}$ with $\alpha > \frac{1}{2}$, by Lemma A.3 and the fact $|\bigcup_{i\notin\Omega} S_i| \ll p$

we have that $\pi_{0,1} \sim \pi_0 \sim \pi$. Since any draw from $\pi_{0,1}$ does not intersect with S_i with $i \notin \Omega$, we have that

$$\sum_{i=1}^{n} \mathcal{I}(\min\{|m_1 \cap S_i|, |m_2 \cap S_i|\} > 0) = \sum_{i \in \Omega} \mathcal{I}(\min\{|m_1 \cap S_i|, |m_2 \cap S_i|\} > 0).$$

Let m_1 and m_2 be two independent draws from $\pi_{0,1}$ with $m_2 \in \tilde{R}_p^N(2\sigma_p)$. We have that there must exist exactly N blocks $T_{j_1}, \ldots, T_{j_N}$ which have elements from m_1 and m_2 σ_p-mututallyclose. In the rest of the $M - N$ blocks there is either no element of m_1 or no element of m_2. Hence the total number of rows corresponding to $\mathcal{I}(\min\{|m_1 \cap S_i|, |m_2 \cap S_i|\} > 0)$ equals $\sum_{l=1}^{N} c_{j_l} \leq Nc^*$. Hence we have

$$\sum_{i=1}^{n} \mathcal{I}(\min\{|m_1 \cap S_i|, |m_2 \cap S_i|\} > 0) \leq Nc^*$$

for the prior $\pi_{0,1} \sim \pi$ and all m_1, m_2 drawn from $\pi_{0,1}$ with $m_2 \in \tilde{R}_p^N(2\sigma_p)$. So by Theorem 3.1, we have that if $c^* \ll \log(p)$ then all tests are asymptotically powerless. □

Proof of Theorem 3.3 . The proof follows by arguments similar to that of Theorem 3.2 and hence is omitted. □

Proof of Theorem 5.2. Since for each $t > 0$, $W_p(t)$ is a normalized mean of i.i.d random variables, by the union bound and Chebyshev's Inequality,

$$\begin{aligned}\mathbb{P}(\mathsf{T}_{\mathrm{HC}} > \log(p)) &\leq \sum_{t \in [1, \sqrt{3\log(p)}] \cap \mathbb{N}} \mathbb{P}(W_p(t) > \log(p)) \\ &\leq 2\sqrt{3\log(p)} \frac{1}{(\log(p))^2} = o(1)\end{aligned}$$

□

Proof of Theorem 6.3. The proof of this theorem follows techniques similar to the proof of Theorem 6.5. However, this can be proved from much simpler combinatorial arguments and hence we provide the proof for the sake of interest. We divide the proof of the theorem into three paragraphs, namely, two-sided alternatives, one-sided alternative for sparse regime and one-sided alternative for dense regime, which correspond to the three parts of the theorem.

Proof of Part(1): Two-Sided Alternatives.

We do the proof for logistic regression for the sake of clarity and note that the proof for general binary regression is exactly same, because the proof only uses the fact $\theta(x) + \theta(-x) = 1$ for the logistic distribution function which is symmetric. Using Remark 6.1, the proof also holds for problem 6.2. Although the following proof is carried out in the usual way of analyzing the second moment of the likelihood ratio as in the proof of Theorem 6.6, here we provide a more direct combinatorial proof.

For logistic regression, we have

$$L_\pi = 2^p \int \prod_{j=1}^{p} \frac{e^{\beta_j y_j}}{1+e^{\beta_j}} d\pi(\boldsymbol{\beta}) = 2^p \cdot \frac{1}{2^k} \frac{1}{\binom{p}{k}} \sum_{m,\xi} \prod_{j=1}^{p} \frac{e^{\beta_j y_j}}{1+e^{\beta_j}}.$$

Take any instance of (m, ξ), say, $m = \{j_1, \dots, j_k\} \subseteq \{1, \dots, p\}$ and $\xi = \{\sigma_1, \dots, \sigma_k\}$, $\sigma_l \in \{-1, 1\}$, $l = 1, \dots, k$. Then

$$\prod_{j=1}^{p} \frac{e^{\beta_j y_j}}{1+e^{\beta_j}} = \left(\frac{1}{2}\right)^{p-k} \prod_{j \in m} \frac{e^{\beta_j y_j}}{1+e^{\beta_j}} \ .$$

Hence,

$$\begin{aligned} L_\pi &= \frac{1}{\binom{p}{k}} \sum_{m,\xi} \prod_{j\in m} \frac{e^{\beta_j y_j}}{1+e^{\beta_j}} \\ &= \frac{1}{\binom{p}{k}} \sum_{\{i_1,\dots,i_k\}\subseteq\{1,\dots,p\}} \sum_{r=0}^{k} \sum_{\{j_1,\dots,j_r\}\subseteq\{i_1,\dots,i_k\}} \frac{e^{Ay_{j_1}} \cdots e^{Ay_{j_r}} e^{A(1-y_{j_{r+1}})} \cdots e^{A(1-y_{j_k})}}{(1+e^A)^k} \end{aligned}$$

where $\{j_{r+1}, \dots, j_k\} = \{i_1, \dots, i_k\} \cap \{j_1, \dots, j_r\}^c$. Now we claim that for any subset $\{i_1, \dots, i_k\} \subseteq \{1, \dots, p\}$,

$$\sum_{r=0}^{k} \sum_{\{j_1,\dots,j_r\}\subseteq\{i_1,\dots,i_k\}} \frac{e^{Ay_{j_1}} \cdots e^{Ay_{j_r}} e^{A(1-y_{j_{r+1}})} \cdots e^{A(1-y_{j_k})}}{(1+e^A)^k} = 1$$

for any sample $(y_1, \dots, y_p)$. To see this, given a sample $(y_1, \dots, y_p)$ and a subset $\{i_1, \dots, i_k\} \subseteq \{1, \dots, p\}$, the number of times the summand equals $\frac{e^{Al}}{(1+e^A)^k}$ is $\binom{k}{l}$ for any $l = 0, 1, \dots, k$ (because any y_j is either 0 or 1)and this exhausts the sum. Hence the total equals

$$\sum_{r=0}^{k} \sum_{\{j_1,\dots,j_r\}\subseteq\{i_1,\dots,i_k\}} \frac{e^{Ay_{j_1}} \cdots e^{Ay_{j_r}} e^{A(1-y_{j_{r+1}})} \cdots e^{A(1-y_{j_k})}}{(1+e^A)^k} = \sum_{l=0}^{k} \frac{\binom{k}{l} e^{Al}}{(1+e^A)^k} = 1$$

as claimed. Hence $L_\pi = 1$ for any sample. Hence by noting that for any test T, $\mathrm{Risk}_\pi(T) \geq 1 - \frac{1}{2}\mathbb{E}_0|L_\pi - 1| \geq 1$ we have that all tests are powerless.

Proof of Part(2a): One-Sided Alternatives, Dense Regime.

We divide our proof into that of lower bound and upper bound.
Proof of Lower Bound. We will do the proof for general binary regression i.e. $\mathbb{E}(Y_j) = \theta(\beta_j)$, $j = 1, \ldots, p$ where θ is any distribution function of a symmetric random variable, *i.e.*, $\theta(x) + \theta(-x) = 1$ for all x and $\theta \in BC^1(0)$. Hence, by Remark 6.1, the proof for lower bound in problem 6.2 follows. Note that one can express $\mathbb{E}_0(L_\pi^2)$ as follows:

$$
\begin{aligned}
\mathbb{E}_0(L_\pi^2) &= \mathbb{E}_{m_1,m_2,\xi_1,\xi_2}[\{4\theta^2(A) - 4\theta(A) + 2\}^{|m_1\cap m_2|\frac{|\xi_1+\xi_2|}{2}}\{4\theta(A)\theta(-A)\}^{|m_1\cap m_2|\frac{|\xi_1-\xi_2|}{2}}] \\
&= \frac{1}{2}\mathbb{E}_{m_1,m_2}[\{4\theta^2(A) - 4\theta(A) + 2\}^{|m_1\cap m_2|} + \{4\theta(A)\theta(-A)\}^{|m_1\cap m_2|}] \\
&\leq \mathbb{E}_{m_1,m_2}[\{4\theta^2(A) - 4\theta(A) + 2\}^{|m_1\cap m_2|}].
\end{aligned}
$$

The last line is true because $4\theta^2(A) - 4\theta(A) + 2 \geq \max\{1, 4\theta(A)\theta(-A)\}$. Now we note that $|m_1 \cap m_2| \sim \mathrm{Hypergeometric}(p, k, k)$ which is stochastically smaller than $\mathrm{Bin}(k, \frac{k}{p-k})$ by Lemma A.1. Since $4\theta^2(A) - 4\theta(A) + 2 \geq \max\{1, 4\theta(A)\theta(-A)\}$ one has that for $Z \sim \mathrm{Bin}(k, \frac{k}{p-k})$,

$$
\begin{aligned}
\mathbb{E}_0(L_\pi^2) &\leq \mathbb{E}_{m_1,m_2}[\{4\theta^2(A) - 4\theta(A) + 2\}^{|m_1\cap m_2|}] \leq \mathbb{E}_Z[\{4\theta^2(A) - 4\theta(A) + 2\}^Z] \\
&= \left[\frac{p-2k}{p-k} + \frac{k}{p-k}(4\theta^2(A) - 4\theta(A) + 2)\right]^k = \left[1 + \frac{\frac{k^2}{p-k}(2\theta(A)-1)^2}{k}\right]^k \\
&= \left[1 + \frac{\frac{k^2}{p-k}(2A\theta'(0) + O(A^2))^2}{k}\right]^k = 1 + o(1)
\end{aligned}
$$

since $p^{1-2\alpha}A \to 0$
Proof of Upper Bound. The proof is similar to the proof of upper bound in Theorem 6.5 in the main text and is based on comparing second moment and variance of the test statistic under the alternative. Hence we skip the details of the proof.

Proof of Part(2b): One-Sided Alternatives, Sparse Regime.

We give the proof for logistic regression and note that the proof for general binary regression is exactly same because the proof uses only the fact $\theta(x) + \theta(-x) = 1$ for the logistic distribution function which is symmetric.

Using Remark 6.1, the proof also holds for problem 6.2. Although the following proof can be proved in the usual way of analyzing the second moment of the likelihood ratio, here we provide a more combinatorial proof without using Lemma A.1.

Note that we have by Fubini's Theorem,

$$\begin{aligned}
\mathbb{E}_0(L_\pi^2) &= 2^p.\frac{1}{4}.\frac{1}{\binom{p}{k}^2}\sum_{(m_1,\xi_1),(m_2,\xi_2)}\left(\frac{1}{2}\right)^{|m_1\Delta m_2|}\left\{\frac{1+e^{2A}}{(1+e^A)^2}\right\}^{|m_1\cap m_2|\frac{\xi_1+\xi_2}{2}}\left\{\frac{2e^A}{(1+e^A)^2}\right\}^{|m_1\cap m_2|\frac{|\xi_1-\xi_2|}{2}} \\
&= 2^p.\frac{1}{4}.\frac{1}{\binom{p}{k}^2}\sum_{r=0}^{k}\sum_{(m_1,\xi_1),(m_2,\xi_2):|m_1\cap m_2|=r}\left(\frac{1}{2}\right)^{p-r}\left\{\frac{1+e^{2A}}{(1+e^A)^2}\right\}^{r\frac{\xi_1+\xi_2}{2}}\left\{\frac{2e^A}{(1+e^A)^2}\right\}^{r\frac{|\xi_1-\xi_2|}{2}}
\end{aligned}$$

where $(m_1,\xi_1),(m_2,\xi_2)$ are i.i.d.

First consider $r=0$. Then $m_1\cap m_2=\Phi$. The number of such tuples (m_1,m_2) is $\binom{p}{k}\binom{p-k}{k}$. For each such $\binom{p}{k}\binom{p-k}{k}$ combinations of $(m_1,\xi_1),(m_2,\xi_2)$ the summand above equals $(\frac{1}{2})^p$. Hence total $=\frac{\binom{p}{k}\binom{p-k}{k}}{\binom{p}{k}^2}=1+o(1)$ by Stirling's Theorem since $k<<p$.

Now consider any $k>r\geq 1$. Then one has that the number of tuples for which $|m_1\cap m_2|=r$ and $\xi_1=\xi_2$ equals $2\binom{p}{r}\binom{p-r}{k-r}\binom{p-k}{k-r}$ and the number of tuples for which $|m_1\cap m_2|=r$ and $\xi_1=-\xi_2$ also equals $2\binom{p}{r}\binom{p-r}{k-r}\binom{p-k}{k-r}$. Hence the total sum can be bounded by $2^r\frac{1}{4}\frac{1}{\binom{p}{k}^2}2\binom{p}{r}\binom{p-r}{k-r}\binom{p-k}{k-r}\{[\frac{1+e^{2A}}{(1+e^A)^2}]^r+[\frac{2e^A}{(1+e^A)^2}]^r\}\leq 2^r\frac{1}{\binom{p}{k}^2}\binom{p}{r}\binom{p-r}{k-r}\binom{p-k}{k-r}$ because $[\frac{1+e^{2A}}{(1+e^A)^2}]^r+[\frac{2e^A}{(1+e^A)^2}]^r\leq 2$. Hence,

$$\begin{aligned}
\mathbb{E}_0(L_\pi^2) &\leq \frac{\binom{p}{k}\binom{p-k}{k}}{\binom{p}{k}^2}+\sum_{r=1}^{k}\frac{2^r\binom{p}{r}\binom{p-r}{k-r}\binom{p-k}{k-r}}{\binom{p}{k}^2} \\
&= \frac{\binom{p}{k}\binom{p-k}{k}}{\binom{p}{k}^2}+\frac{2^k}{\binom{p}{k}}+\sum_{r=1}^{k-1}2^r\frac{(p-k)\cdots(p-2k+r+1)}{p\cdots(p-k+1)}\frac{k!k!}{r!(k-r)!(k-r)!} \\
&\leq \frac{\binom{p-k}{k}}{\binom{p}{k}}+\frac{2^k}{\binom{p}{k}}+\sum_{r=1}^{k-1}2^r\frac{(p-k+1)^{k-r}}{(p-k+1)^k\frac{[k\cdots(k-r+1)]^2}{r!}} \\
&\leq \frac{\binom{p-k}{k}}{\binom{p}{k}}+\frac{2^k}{\binom{p}{k}}+\sum_{r=1}^{k-1}2^r\frac{1}{(p-k+1)^r}\frac{k^{2r}}{r!}\leq \frac{\binom{p-k}{k}}{\binom{p}{k}}+\frac{2^k}{\binom{p}{k}}+\sum_{r=1}^{k-1}\left(\frac{2k^2}{(p-k+1)}\right)^r\frac{1}{r^re^{-r}} \\
&\leq \frac{\binom{p-k}{k}}{\binom{p}{k}}+\frac{2^k}{\binom{p}{k}}+\sum_{r=1}^{k-1}\left(\frac{2ek^2}{(p-k+1)}\right)^r=\frac{\binom{p-k}{k}}{\binom{p}{k}}+\frac{2^k}{\binom{p}{k}}+\frac{1-(\frac{2ek^2}{(p-k+1)})^{k-1}}{1-\frac{2ek^2}{(p-k+1)}}-1
\end{aligned}$$

The last step holds because $k^2 \ll p$ since $\alpha > 1/2$. For $r = k$ we have the factor $\binom{p}{r}\binom{p-r}{k-r}\binom{p-k}{k-r}$ replaced by $\binom{p}{k}$. Now since $\frac{2^k}{\binom{p}{k}} \leq (\frac{2k}{p-k+1})^k = o(1)$ and $\frac{1-(\frac{2ek^2}{(p-k+1)})^{k-1}}{1-\frac{2ek^2}{(p-k+1)}} = 1 + o(1)$ we have that $\mathbb{E}_0(L_\pi^2) \leq 1 + o(1)$.

□

Proof of Theorem 6.5. We first present the proof of the lower bound. We will estimate the second moment of the likelihood ratio as follows.

$$\begin{aligned}
\mathbb{E}_0(L_\pi^2) &= 2^{-2k}\binom{p}{k}^{-2} \sum_{m_1,m_2,\xi_1,\xi_2} \left(\frac{1+4\Delta^2}{1-4\Delta^2}\right)^{r|m_3|} (1-4\Delta^2)^{r|m_1\cap m_2|} \\
&= \mathbb{E}_{|m_3|,|m_1\cap m_2|}\left[\left(\frac{1+4\Delta^2}{1-4\Delta^2}\right)^{r|m_3|} (1-4\Delta^2)^{r|m_1\cap m_2|}\right]
\end{aligned}$$

where $m_3 = \{j \in m_1 \cap m_2 : \xi_1^j = \xi_2^j\}$. Now given $|m_1 \cap m_2|$, $|m_3| \sim \text{Bin}(|m_1 \cap m_2|, \frac{1}{2})$. Hence

$$\begin{aligned}
&\mathbb{E}_0(L_\pi^2) \qquad \text{(A.3)} \\
&= \mathbb{E}_{|m_1\cap m_2|}\left[\left(\frac{1}{2}+\frac{1}{2}\left(\frac{1+4\Delta^2}{1-4\Delta^2}\right)^r\right)^{|m_1\cap m_2|} (1-4\Delta^2)^{r|m_1\cap m_2|}\right] \\
&= \mathbb{E}_{|m_1\cap m_2|}\left[\left(\frac{1}{2}\right)^{|m_1\cap m_2|} \left((1+4\Delta^2)^r + (1-4\Delta^2)^r\right)^{|m_1\cap m_2|}\right] \\
&= \mathbb{E}_Z\left[\left(\frac{1}{2}\right)^Z (a^r+b^r)^Z\right] = \mathbb{E}_Z\left[2^{(r-1)Z}(a_1^r+b_1^r)^Z\right]
\end{aligned}$$

where $Z \sim \text{Hypergeometric}(p, k, k)$ and $a = (1 + 4\Delta^2)^r$, $b = (1 - 4\Delta^2)^r$ and $(a_1, b_1) = (a/2, b/2)$. Thus $a_1 + b_1 = 1$ and hence $(a_1^r + b_1^r)2^{r-1} \geq 1$. Now since $Z \sim \text{Hypergeometric}(p, k, k)$, Z is stochastically smaller than W where $W \sim \text{Bin}(k, \frac{k}{p-k})$. Hence

$$\begin{aligned}
\mathbb{E}_0(L_\pi^2) &= \mathbb{E}_Z\left[2^{(r-1)Z}(a_1^r+b_1^r)^Z\right] \\
&\leq \mathbb{E}_W\left[2^{(r-1)W}(a_1^r+b_1^r)^W\right] \\
&= \left[1+\frac{\frac{k^2}{p-k}(2^{r-1}(a_1^r+b_1^r)-1)}{k}\right]^k
\end{aligned}$$

We complete our proof by showing that $\frac{k^2}{p-k}(2^{r-1}(a_1^r+b_1^r)-1) \to 0$ when $\Delta \ll \sqrt{\frac{p^{\frac{1}{2}}}{kr}}$ and hence rendering all tests asymptotically powerless. To this end, note that by Taylor series expansion up to 4^{th} order around 0 and analyzing the remainder, we have

$$\begin{aligned}\frac{k^2}{p-k}(2^{r-1}(a_1^r+b_1^r)-1) &= \frac{k^2}{p-k}\left(192\frac{\Delta^4}{4!}r(r-1)+O(\Delta^4 r^2)\right)\\ &= O\left(\frac{k^2r^2\Delta^4}{p}\right) \to 0\end{aligned}$$

where the last line holds since $\Delta \ll \sqrt{\frac{p^{\frac{1}{2}}}{kr}}$. This completes the proof of the lower bound for problem 6.2. The proof of lower bound in 2.3 follows by noting that $\theta(A) = \frac{1}{2}+\Delta$ and the fact that $\theta \in BC^1(0)$.

Now we prove the upper bound. Recall T_{GLRT} from (5.1). Once again we will provide proof for problem 6.2. The proof of lower bound in problem 2.3 follows by noting that $\theta(A) = \frac{1}{2}+\Delta$ and the fact that $\theta \in BC^1(0)$.

We will show that if $t_p \to \infty$ at a sufficiently slow rate, the test is asymptotically powerful. It suffices to show $\sup_{\nu\in\Theta_k^A}\mathbb{P}_\nu(\frac{\mathsf{T}_{\text{GLRT}}-p}{\sqrt{2p}} \le t_p) \to 0$. We will show that $\sup_{\nu\in\Theta_k^A}\frac{\mathbb{E}_\nu(\frac{\mathsf{T}_{\text{GLRT}}-p}{\sqrt{2p}})}{t_p} \to \infty$ and $\frac{\text{Var}_\nu(\frac{\mathsf{T}_{\text{GLRT}}-p}{\sqrt{2p}})}{(\mathbb{E}_\nu(\frac{\mathsf{T}_{\text{GLRT}}-p}{\sqrt{2p}}))^2} \to 0$ when $\frac{A^2kr}{\sqrt{p}} \to \infty$.

Fix $\nu^* \in \Xi_k^\Delta$. Under the measure $\mathbb{P}_{\nu^*}$, exactly k of the Z_j's are distributed as i.i.d $\text{Bin}(r, \frac{1}{2}+\Delta)$ and the rest of the $p-k$ Z_j's are distributed as i.i.d $\text{Bin}(r,\frac{1}{2})$. Let $O = \{j : \beta_j^* \neq 0\}$. Hence we have, for $j \in O$,

$$\mathbb{E}_{\nu^*}\left[(Z_j-\frac{r}{2})^2\right] = r\left(\frac{1}{4}-\Delta^2\right)+r^2\Delta^2\,. \tag{A.4}$$

For $j \in O^c$, $\mathbb{E}_{\nu^*}[(Z_j-\frac{r}{2})^2] = \frac{r}{4}$. Hence,

$$\begin{aligned}\frac{\mathbb{E}_{\nu^*}(\frac{\mathsf{T}_{\text{GLRT}}-p}{\sqrt{2p}})}{t_p} &= \frac{\frac{\frac{4}{r}[kr(\frac{1}{2}-\Delta^2)+kr^2\Delta^2+(p-k)\frac{r}{4}]-p}{\sqrt{2p}}}{t_p}\\ &= \frac{\frac{p+4kr\Delta^2-\frac{k\Delta^2}{4}-p}{\sqrt{2p}}}{t_p} \gtrsim \frac{kr\Delta^2}{t_p\sqrt{p}} \approx \frac{krA^2}{\sqrt{p}t_p}\,.\end{aligned} \tag{A.5}$$

Since $\frac{krA^2}{\sqrt{p}} \to \infty$ and t_p can be chosen to grow to ∞ at a sufficiently slow rate, (A.5) goes to infinity.

Now we compute the variance. For $j \in O$,

$$\mathbb{E}_{\nu^*}\left(Z_j - \frac{r}{2}\right)^4 = r\left(\frac{1}{4} - \Delta^2\right)\left[3r\left(\frac{1}{4} - \Delta^2\right) + 6\Delta\left(\frac{1}{2} + \Delta\right) - 8r\Delta^2 + 6r\Delta^2\right] + r^4\Delta^4 \,.$$

Using the above and (A.4), a straightforward calculation yields that

$$\sum_{j \in O} \mathrm{Var}_{\nu^*}\left(\frac{4(Z_j - \frac{r}{2})^2}{r}\right) = 16k\left(\frac{1}{4} - \Delta^2\right)\left[2\left(\frac{1}{4} - \Delta^2\right) + 4r\Delta^2 + 6\Delta\left(\frac{1}{2} + \Delta\right) - 8\Delta^2\right] \,.$$

Also, by another direct calculation

$$\sum_{j \in O^c} \mathrm{Var}_{\nu^*}\left(\frac{4(Z_j - \frac{r}{2})^2}{r}\right) = 2(p-k)(1 - \frac{1}{r})$$

Combining the above two,

$$\begin{aligned}\mathrm{Var}_{\nu^*}\left(\frac{\mathsf{T}_{\mathrm{GLRT}} - p}{\sqrt{2p}}\right) &= \left[16k\left(\frac{1}{4} - \Delta^2\right)\left\{2\left(\frac{1}{4} - \Delta^2\right) + 4r\Delta^2 + 6\Delta\left(\frac{1}{2} + \Delta\right) - 8\Delta^2\right\}\right]/2p \\ &\quad + 2(p-k)(1 - \frac{1}{r})/2p \\ &\le \frac{4p + 32kr\Delta^2}{2p} = \frac{2p + 16kr\Delta^2}{p} \,.\end{aligned}$$

Also $\left(\mathbb{E}_{\nu^*}(\frac{\mathsf{T}_{\mathrm{GLRT}} - p}{\sqrt{2p}})\right)^2 \ge \frac{kr\Delta^2}{4p}$. Hence,

$$\frac{\mathrm{Var}_{\nu^*}(\frac{\mathsf{T}_{\mathrm{GLRT}} - p}{\sqrt{2p}})}{(\mathbb{E}_{\nu^*}(\frac{\mathsf{T}_{\mathrm{GLRT}} - p}{\sqrt{2p}}))^2} \le 4\frac{2p + 16kr\Delta^2}{kr\Delta^2} \to 0$$

since $k^2r^2\Delta^4 \gg p$.

Now note that if ν^* had k_1 elements which are greater than or equal to A and k_2 elements less than equal to $-A$, then a similar calculation yields $\mathrm{Var}_{\nu^*}(\frac{\mathsf{T}_{\mathrm{GLRT}} - p}{\sqrt{2p}}) \le \frac{2p + 16kr\Delta^2}{p}$ and $(\mathbb{E}_{\nu^*}(\frac{\mathsf{T}_{\mathrm{GLRT}} - p}{\sqrt{2p}}))^2 \ge \frac{kr\Delta^2}{4p}$ where $k = k_1 + k_2$ equals the number of nonzero coefficients in β^*. Hence we have $\max_{\nu \in \Xi_k^\Delta}[\mathbb{P}_\nu(\mathsf{T}_{\mathrm{GLRT}} \le t_p)] \to 0$ when $\alpha \le \frac{1}{2}$ and $\frac{\Delta^2 kr}{\sqrt{p}} \to \infty$. This proves the GLRT is asymptotically powerful.

□

PROOF OF THEOREM 6.6. We will provide an argument for problem 6.2. The proof for problem 2.3 follows from noting that $\theta(A) = \frac{1}{2} + \Delta$.

We will estimate the second moment of the likelihood ratio similar to before. Following the same line of arguments as in proof of Theorem 6.5, we note that

$$\begin{aligned}\mathbb{E}_0(L_\pi^2) &= \mathbb{E}_Z\left[2^{(r-1)Z}\{a_1^r + b_1^r\}^Z\right] \text{ where } Z \sim \text{Hypergeometric}(p,k,k)\\ &\leq \left[1 + \frac{\frac{k^2}{p-k}(2^{r-1}(a_1^r + b_1^r) - 1)}{k}\right]^k .\end{aligned}$$

Now $\alpha > \frac{1}{2}$ implies that $\frac{k^2}{p-k} \to 0$. Also the quantity $\frac{k^2}{p-k}(2^{r-1}(a_1^r + b_1^r) - 1) = O(k^2 2^r/p)$. Hence if $r \ll \frac{\log(p)}{\log(2)}$, we have that $\mathbb{E}_0(L_\pi^2) \to 1$ and thus all tests are asymptotically powerless.

□

Proof of Theorem 6.8. We will provide proof for the lower bound in problem 2.3 where $\theta \in BC^2(0)$. Using Remark 6.1, the proof also holds for problem 6.2. Since directly bounding $\mathbb{E}_0(L_\pi^2)$ yields trivial bounds we invoke a truncation trick which breaks down the analysis into parts related to extreme tails and non-extreme tails of the Z-statistics. In particular, define the interval

$$H_p = \left(\frac{r}{2} - \sqrt{2\log(p)}\sqrt{\frac{r}{4}}, \frac{r}{2} + \sqrt{2\log(p)}\sqrt{\frac{r}{4}}\right) . \tag{A.6}$$

and put

$$D = \{Z_l \in H_p,\, l = 1, \ldots, p\}, \quad Z_l = \sum_{s=1}^{r} y_{(l-1)r+s},\ l = 1, \ldots, p. \tag{A.7}$$

By Hölder's inequality it can be shown that for proving a lower bound it suffices to prove,

$$\mathbb{E}_0(L_\pi \mathcal{I}_{D^c}) = o(1), \qquad \mathbb{E}_0(L_\pi^2 \mathcal{I}_D) = 1 + o(1) . \tag{A.8}$$

We first prove the first inequality of (A.8). Since $\mathcal{I}_{D^c} \leq \sum_{l=0}^{p-1} \mathcal{I}_{(Z_{l+1} \in H_p^c)}$ and

$$L_\pi = 2^n \int \prod_{j=1}^{p} \left\{\frac{\theta(\beta_j)}{\theta(-\beta_j)}\right\}^{Z_j} \{\theta(-\beta_j)\}^r d\pi(\boldsymbol{\beta})$$

we have

$$L_\pi \mathcal{I}_{D^c} \leq 2^n \int \sum_{l=1}^{p} \prod_{j=1}^{p} \left\{\frac{\theta(\beta_j)}{\theta(-\beta_j)}\right\}^{Z_j} \{\theta(-\beta_j)\}^r \mathcal{I}(Z_l \in H_p^c) d\pi(\boldsymbol{\beta})$$

Hence

$$
\begin{aligned}
&\mathbb{E}_0(L_\pi \mathcal{I}_{D^c})\\
&\leq 2^n \int \sum_{l=1}^{p} \mathbb{E}_0\left[\prod_{j=1}^{p}\left\{\frac{\theta(\beta_j)}{\theta(-\beta_j)}\right\}^{Z_j}\{\theta(-\beta_j)\}^r \mathcal{I}(Z_l \in H_p^c)\right] d\pi(\boldsymbol{\beta})\\
&= 2^n \int \sum_{l=1}^{p} \mathbb{E}_0\left[\prod_{j\neq l}^{p}\left\{\frac{\theta(\beta_j)}{\theta(-\beta_j)}\right\}^{Z_j}\{\theta(-\beta_j)\}^r\right] \mathbb{E}_0\left[\left\{\frac{\theta(\beta_l)}{\theta(-\beta_l)}\right\}^{Z_l}\{\theta(-\beta_l)\}^r \mathcal{I}(Z_l \in H_p^c)\right] d\pi(\boldsymbol{\beta})\\
&= 2^n \int \sum_{l=1}^{p} \left[\prod_{j\neq l}^{p}\left(\frac{1}{2}\right)^r\left(1+\frac{\theta(\beta_j)}{\theta(-\beta_j)}\right)^r\{\theta(-\beta_j)\}^r\right]\\
&\qquad \times \mathbb{E}_0\left[\left\{\frac{\theta(\beta_l)}{\theta(-\beta_l)}\right\}^{Z_l}\{\theta(-\beta_l)\}^r \mathcal{I}(Z_l \in H_p^c)\right] d\pi(\boldsymbol{\beta})\\
&= 2^n \int \sum_{l=1}^{p} \left(\frac{1}{2}\right)^{(p-1)r} \mathbb{E}_0\left[\left\{\frac{\theta(\beta_l)}{\theta(-\beta_l)}\right\}^{Z_l}\{\theta(-\beta_l)\}^r \mathcal{I}(Z_l \in H_p^c)\right] d\pi(\boldsymbol{\beta})\\
&= \int \sum_{l=1}^{p} 2^r \mathbb{E}_0\left[\left\{\frac{\theta(\beta_l)}{\theta(-\beta_l)}\right\}^{Z_l}\{\theta(-\beta_l)\}^r \mathcal{I}(Z_l \in H_p^c)\right] d\pi(\boldsymbol{\beta})
\end{aligned}
$$

Letting $m_1^1 = \{j \in m_1 : \xi_1^j = +1\}$ and $m_1^{-1} = \{j \in m_1 : \xi_1^j = -1\}$, we have

$$
\begin{aligned}
&\mathbb{E}_0(L_\pi \mathcal{I}_{D^c})\\
&\leq \binom{p}{k}^{-1} 2^{-k} 2^r \sum_{m_1,\xi_1}\Big[\sum_{j\in m_1^1} \mathbb{E}_0\Big(\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{Z_j}\Big\{\theta(-A)\Big\}^r \mathcal{I}(Z_j \in H_p^c)\Big)\\
&\quad + \sum_{j\in m_1^{-1}} \mathbb{E}_0\Big(\Big\{\frac{\theta(-A)}{\theta(A)}\Big\}^{Z_j}\{\theta(A)\}^r \mathcal{I}(Z_j \in H_p^c)\Big)\\
&\quad + \sum_{j\in m_1^c} \mathbb{E}_0\Big(\Big\{\frac{\theta(0)}{\theta(-0)}\Big\}^{Z_j}\{\theta(-0)\}^r \mathcal{I}(Z_j \in H_p^c)\Big)\Big]\\
&= \binom{p}{k}^{-1} 2^{-k} 2^r \sum_{m_1,\xi_1}\Big[\sum_{j\in m_1^1} \mathbb{E}_0\Big(\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{Z_j}\{\theta(-A)\}^r \mathcal{I}(Z_j \in H_p^c)\Big)\\
&\quad + \sum_{j\in m_1^{-1}} \mathbb{E}_0\Big(\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{r-Z_j}\{\theta(-A)\}^r \mathcal{I}(Z_j \in H_p^c)\Big) + \sum_{j\in m_1^c}\Big(\frac{1}{2}\Big)^r \mathbb{P}_0(Z_j \in H_p^c)\Big]\\
&= k\{2\theta(-A)\}^r \mathbb{E}_0\Big(\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{Z_1}\mathcal{I}(Z_1 \in H_p^c)\Big) + (p-k)\mathbb{P}_0(Z_1 \in H_p^c) \qquad \text{(A.9)}
\end{aligned}
$$

where we have used the fact that $r - Z_l \overset{d}{=} Z_l$ and that the set D in (A.7) is symmetric in Z_l and $r - Z_l$.

Now by Lemma A.2 we have that $(p-k)\mathbb{P}(Z_1 \in H_p^c) = o(1)$ since $r \gg \log(p)$. To see this we put $n = r$ and $k' = \frac{r}{2} + \sqrt{2\log(p)}\sqrt{\frac{r}{4}}$ in Lemma A.2 to obtain $\epsilon = \frac{2\sqrt{\frac{r}{4}}\sqrt{2\log(p)}-1}{r-1} = o(1)$ since $r \gg \log(p)$ and also $\epsilon\sqrt{r} \to \infty$. This implies

$$\begin{aligned}
\mathbb{P}\Big(Z_l > \frac{r}{2} + \sqrt{2\log(p)}\sqrt{\frac{r}{4}}\Big) &= \overline{\Phi}(\epsilon\sqrt{n})e^{-(r-1)((1+\epsilon)\log(1+\epsilon)+(1-\epsilon)\log(1-\epsilon)-\epsilon^2)} \\
&\le \frac{e^{-\frac{4.\frac{1}{4}.2\log(p)}{2}}}{\epsilon\sqrt{r}} e^{r\epsilon^2 - r(1+\epsilon)\log(1+\epsilon)} \text{ using Lemma ??} \\
&\le \frac{e^{-\log(p)}}{\epsilon\sqrt{r}} e^{r\epsilon^2 - r\epsilon} \text{ since } \log(1+\epsilon) \ge \frac{\epsilon}{(1+\epsilon)} \\
&\ll \frac{1}{p\epsilon\sqrt{r}}. \qquad \text{(A.10)}
\end{aligned}$$

Hence $(p-k)\mathbb{P}(Z_1 \in H_p^c) = o(1)$ as needed. Next we need to control $k\{2\theta(-A)\}^r \mathbb{E}_0\{\frac{\theta(A)}{\theta(-A)}\}^{Z_1}\mathcal{I}(Z_1 \in H_p^c)$. To this end note that

$$\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{Z_1} = e^{Z_1\log(\frac{\theta(A)}{\theta(-A)})} = e^{(2\frac{\theta'(0)}{\theta(0)}A + o(A^2))Z_1} = e^{(4\theta'(0)A+o(A^2))Z_1}.$$

Hence by Hölder's Inequality for any $f > 1$ and complementary $g > 1$ such that $\frac{1}{f} + \frac{1}{g} = 1$, one has

$$\begin{aligned}
k\{2\theta(-A)\}^r\mathbb{E}_0\Big(\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{Z_1}\mathcal{I}(Z_1 \in H_p^c)\Big) &\le \{k^f\{2\theta(-A)\}^{rf}\mathbb{E}_0[e^{4\theta'(0)AfZ_1}\mathcal{I}(Z_1 \in H_p^c)]\}^{1/f} \\
&\quad \times \{\mathbb{E}_0[e^{g\epsilon Z_1}]\}^{1/g} \qquad \text{(A.11)}
\end{aligned}$$

where $\epsilon = o(A^2)$. Our next task is hence to control $k^f\{2\theta(-A)\}^{rf}\mathbb{E}_0[e^{4\theta'(0)AfZ_1}\mathcal{I}(Z_1 \in H_p^c)]$ for an appropriately chosen $f > 1$ and then subsequently bound $\{\mathbb{E}_0[e^{g\epsilon Z_1}]\}^{1/g}$ for the corresponding $g > 1$. We first analyze $\mathbb{E}_0[e^{4\theta'(0)AfZ_1}\mathcal{I}(Z_1 \in H_p^c)]$ for arbitrary $f > 1$ and we will make the choice of the pair (f, g) clear later:

$$\begin{aligned}
\mathbb{E}_0[e^{4\theta'(0)AfZ_1}\mathcal{I}(Z_1 \in H_p^c)] &= \mathbb{E}_0\left[e^{4\theta'(0)AfZ_1}\mathcal{I}\left(Z_1 > \frac{r}{2} + \sqrt{2\log p}\sqrt{\frac{r}{4}}\right)\right] \\
&\quad + \mathbb{E}_0\left[e^{4\theta'(0)AfZ_1}\mathcal{I}\left(Z_1 < \frac{r}{2} - \sqrt{2\log p}\sqrt{\frac{r}{4}}\right)\right] := I_1 + I_2.
\end{aligned}$$

We will analyze I_1 in detail; the analysis of I_2 is very similar and is omitted. Since

$$I_1 = e^{4\theta^{'}(0)f\frac{Ar}{2}}\mathbb{E}_0\left[e^{4\theta^{'}(0)f\frac{A^*}{2}\frac{Z_1-\frac{r}{2}}{\sqrt{\frac{r}{4}}}}\mathcal{I}\left(Z_1 > \frac{r}{2}+\sqrt{2\log p}\sqrt{\frac{r}{4}}\right)\right]$$

where $A^* = A\sqrt{r}$, we will first control $\mathbb{E}_0\left[e^{4\theta^{'}(0)f\frac{A^*}{2}\frac{Z_1-\frac{r}{2}}{\sqrt{\frac{r}{4}}}}\mathcal{I}\left(Z_1 > \frac{r}{2}+\sqrt{2\log p}\sqrt{\frac{r}{4}}\right)\right] =$ I_1' (say). Denoting $\frac{Z_1-\frac{r}{2}}{\sqrt{\frac{r}{4}}}$ by W_r, by the Komlos-Major-Tusnady strong embedding theorem (Komlós, Major and Tusnády, 1975), there exists a version of standard Brownian Motion B_r on the same probability space as W_r such that

$$\mathbb{P}(|W_r - B_r| \geq C\log(r) + s) \leq Ke^{-\lambda s} \tag{A.12}$$

where C, K, λ do not depend on r. For notational conveneience we will take $C = 1$ w.l.o.g. Let $x > 0$ which we will choose appropriately later. Hence

$$\begin{aligned}
I_1' &= \mathbb{E}_0\left[e^{4\theta^{'}(0)f\frac{A^*}{2\sqrt{r}}W_r}\mathcal{I}\left(Z_1 > \frac{r}{2}+\sqrt{2\log p}\sqrt{\frac{r}{4}}\right)\right] \\
&= \mathbb{E}_0\Big[e^{4\theta^{'}(0)f\frac{A}{2}(W_r)}\mathcal{I}(W_r > \sqrt{2\log p}\sqrt{r}) \\
&\qquad \times \{\mathcal{I}(|W_r - B_r| \leq \log(r)+x) + \mathcal{I}(|W_r - B_r| > \log(r)+x)\}\Big] \\
&:= I_{11} + I_{12}.
\end{aligned}$$

Hence we will need to control both $k^f\{2\theta(-A)\}^{rf}e^{4\theta^{'}(0)f\frac{Ar}{2}}I_{11}$ and $k^f\{2\theta(-A)\}^{rf}e^{4\theta^{'}(0)f\frac{Ar}{2}}I_{12}$. Now

$$\begin{aligned}
&I_{11} \\
&= \mathbb{E}_0[e^{4\theta^{'}(0)f\frac{A}{2}(W_r)}\mathcal{I}(W_r > \sqrt{2\log p}\sqrt{r})\mathcal{I}(|W_r - B_r| \leq \log(r)+x)] \\
&\leq e^{4\theta^{'}(0)f\frac{A}{2}(\log(r)+x)}\mathbb{E}_0[e^{4\theta^{'}(0)f\frac{A}{2}B_r}\mathcal{I}(B_r > \sqrt{2\log p}\sqrt{r} - (\log(r)+x))] \\
&= e^{4\theta^{'}(0)f\frac{A}{2}(\log(r)+x)}\mathbb{E}_0[e^{4\theta^{'}(0)f\frac{A^*}{2}\frac{B_r}{\sqrt{r}}}\mathcal{I}(\frac{B_r}{\sqrt{r}} > \sqrt{2\log p} - \frac{(\log(r)+x)}{\sqrt{r}})] \\
&= e^{4\theta^{'}(0)f\frac{A}{2}(\log(r)+x)}\int_{T_p}^{\infty}\frac{e^{4\theta^{'}(0)f\frac{A^*}{2}v-\frac{v^2}{2}}}{\sqrt{2\pi}}dv \text{ where } T_p = \sqrt{2\log p} - \frac{(\log(r)+x)}{\sqrt{r}} \\
&= e^{4\theta^{'}(0)f\frac{A}{2}(\log(r)+x)+2\theta^{'}(0)^2f^2(A^*)^2}\overline{\Phi}(T_p - 2\theta^{'}(0)fA^*)
\end{aligned}$$

$$\leq Ce^{\{4\theta^{'}(0)f\frac{A}{2}(\log(r)+x)+2\theta^{'}(0)^2f^2(A^*)^2-\frac{T_p^2-4\theta^{'}(0)^2(A^*)^2f^2+4\theta^{'}(0)A^*fT_p}{2}\}} \text{ if } T_p-2\theta^{'}(0)fA^*>1$$
$$= Ce^{\{-\log(p)(1-4\theta^{'}(0)\sqrt{t}f)-\frac{(\log r+x)^2}{2r}+\frac{\sqrt{2\log p}(\log r+x)}{2\sqrt{r}}\}} \tag{A.13}$$

Since I_{11} is multiplied outside by $\{2\theta(-A)\}^{rf}e^{4\theta^{'}(0)f\frac{Ar}{2}}$ we bound that coefficient as follows:

$$\begin{aligned}\{2\theta(-A)\}^{rf}e^{4\theta^{'}(0)f\frac{Ar}{2}} &= (2\theta(-A)e^{2\theta^{'}(0)A})^{rf}\\ &= e^{rf\log(2\theta(-A)e^{2\theta^{'}(0)A})} = e^{rf(\log 2+\log\theta(_A)+2\theta^{'}(0)A)}\\ &= e^{rf\{\log 2+2\theta^{'}(0)A+\log\theta(0)+\frac{\theta^{'}(0)}{\theta(0)}A(-1)-\frac{1}{2!}\frac{\theta^{''}(0)\theta(0)(-1)-\theta^{'}(0)^2(-1)}{\theta(0)^2}A^2+o(A^2)\}}\\ &= e^{rf\{\log 2+2\theta^{'}(0)A-\log 2-2\theta^{'}(0)A-2\theta^{'}(0)^2A^2+o(A^2)\}} \text{ since } \theta^{''}(0)=0\\ &= e^{\{-f4\theta^{'}(0)^2t\log(p)+rf\epsilon'\}} \text{ where } \epsilon'=o(A^2)\end{aligned} \tag{A.14}$$

Finally collecting the terms from (A.13) and (A.14), we bound $k^f\{2\theta(-A)\}^{rf}e^{4\theta^{'}(0)f\frac{Ar}{2}}I_{11}$ as follows:

$$k^f\{2\theta(-A)\}^{rf}e^{4\theta^{'}(0)f\frac{Ar}{2}}I_{11} \leq Ce^{-\log(p)\{f(1-\alpha-(1-2\theta^{'}(0)\sqrt{t})^2)+(f-1)\}-\frac{(\log r+x)^2}{2r}+\frac{\sqrt{2\log p}(\log r+x)}{2\sqrt{r}}+rf\epsilon'}. \tag{A.15}$$

Now since $t<\rho^*_{\text{binary}}(\alpha)$, $1-\alpha-(1-2\theta^{'}(0)\sqrt{t})^2<0$. Hence we can choose $f>1$ sufficiently close to 1 such that $f(1-\alpha-(1-2\theta^{'}(0)\sqrt{t})^2)+(f-1)<0$. We note that since $r\gg\log(p)$, there exists a sequence $a_{r,p}\to\infty$ such that $r\gg a_{r,p}\log(p)$. If we chose $x=a_{r,p}\log(p)$ then $T_p-2\theta^{'}(0)A^*f>1$ as required for the conclusions to hold since $4\theta^{'}(0)^2t<1$ and $r\gg a_{r,p}\log(p)$. Also again since $r\gg a_{r,p}\log(p)$ we have $-\log(p)\{f(1-\alpha-(1-2\theta^{'}(0)\sqrt{t})^2)+(f-1)\}-\frac{(\log r+x)^2}{2r}+\frac{\sqrt{2\log p}(\log r+x)}{2\sqrt{r}}+rf\epsilon'\leq-\delta\log(p)$ for some $\delta>0$ for sufficiently large r,p. Hence for such x, we have $k^f\{2\theta(-A)\}^{rf}e^{4\theta^{'}(0)f\frac{Ar}{2}}I_{11}\to 0$. In order to bound I_{12} from above we repeatedly apply the Cauchy-Schwarz Inequality and use the fact that $\cosh(s)=1+s^2/2+o(s^2)$ for small s as follows:

$$\begin{aligned}I_{12} &= \mathbb{E}_0[e^{4\theta^{'}(0)f\frac{A}{2}W_r}\mathcal{I}(W_r>\sqrt{2\log p}\sqrt{r})\mathcal{I}(|W_r-B_r|>(\log r+x))]\\ &\leq \left\{\mathbb{E}_0[e^{4\theta^{'}(0)fAW_r}\mathcal{I}(W_r>\sqrt{2\log p}\sqrt{r})]\mathbb{P}_0(|W_r-B_r|>(\log r+x))\right\}^{\frac{1}{2}}\\ &\leq \left\{\mathbb{E}_0[e^{8\theta^{'}(0)fAW_r}]\mathbb{P}_0(W_r>\sqrt{2\log p}\sqrt{r})(\mathbb{P}_0(|W_r-B_r|>(\log r+x)))^2\right\}^{\frac{1}{4}}\\ &= \left\{(\cosh(8\theta^{'}(0)fA))^r\mathbb{P}_0(W_r>\sqrt{2\log p}\sqrt{r})(\mathbb{P}_0(|W_r-B_r|>(\log r+x)))^2\right\}^{\frac{1}{4}}\end{aligned}$$

$$= \left\{ e^{r\log(1+32\theta'(0)^2f^2A^2+o(A^2))}\mathbb{P}_0(W_r > \sqrt{2\log p}\sqrt{r})(\mathbb{P}_0(|W_r - B_r| > (\log r + x)))^2 \right\}^{\frac{1}{4}}$$
$$\leq \left\{ e^{r(32\theta'(0)^2f^2A^2+o(A^2))}\mathbb{P}_0(W_r > \sqrt{2\log p}\sqrt{r})(\mathbb{P}_0(|W_r - B_r| > (\log r + x)))^2 \right\}^{\frac{1}{4}}$$
$$\leq C \left\{ e^{\{64\theta'(0)^2f^2t\log(p)-\log(p)+\frac{(\log(p))^2}{r}-\frac{\log\log(p)}{2}-2\lambda x+r\epsilon''\}} \right\}^{\frac{1}{4}} \tag{A.16}$$

where $\epsilon'' = o(A^2)$, the second last line uses the fact that $\log(1+x) \leq x$ for $x \geq 0$ and the last line follows from (A.10) and (A.12) for some constant $C > 0$. Recall from (A.14) that $k^f\{2\theta(-A)\}^{rf}e^{4\theta'(0)f\frac{Ar}{2}} = e^{\{-4\theta'(0)^2ft\log(p)+rf\epsilon'+(1-\alpha)\log(p)\}}$ where $\epsilon' = o(A^2)$. Hence by combining terms from (A.16) and (A.14), we obtain that for a constant K depending on f, t and $\theta'(0)$,

$$k^f\{2\theta(-A)\}^{rf}e^{4\theta'(0)f\frac{Ar}{2}}I_{12} \leq Ce^{K\log(p)-2\lambda x}.$$

Now since $x = a_{r,p}\log(p)$ for some $a_{r,p} \to \infty$ such that $r \gg a_{r,p}\log(p)$, it follows that

$$k^f\{2\theta(-A)\}^{rf}e^{4\theta'(0)f\frac{Ar}{2}}I_{12} = o(1) \tag{A.17}$$

as required.

Next considering the g-factor from (A.11) we have

$$\begin{aligned}
\{\mathbb{E}_0[e^{g\epsilon Z_1}]\}^{1/g} &= e^{\frac{r}{g}\log(\frac{1+e^{g\epsilon}}{2})} \text{ where } \epsilon = o(A^2) \\
&= e^{\frac{r}{g}\log(1+\frac{e^{g\epsilon}-1}{2})} = e^{\frac{r}{g}\log(1+\frac{g\epsilon+g^2\epsilon^2+o(g^2\epsilon^2)}{2})} \\
&\leq e^{\frac{r}{g}\log(1+\frac{(2g+2g^2)\epsilon}{2})} \\
&= e^{r\epsilon O(1)} = e^{o(1)} \to 1
\end{aligned} \tag{A.18}$$

Hence collecting terms from (A.15),(A.17) and (A.18) in (A.11) we finish proving $\mathbb{E}_0(L_\pi\mathcal{I}_{D^c}) = o(1)$ which is the first inequality of (A.8).

Next we prove the second inequality in (A.8). Since definition of D does not depend on $\boldsymbol{\beta}$, it follows that

$$\begin{aligned}
L_\pi^2\mathcal{I}_D &= (L_\pi\mathcal{I}_D)^2 \\
&= 2^{2n}\iint\prod_{j=1}^{p}\left\{\frac{\theta(\beta_j)\theta(\beta_j')}{\theta(-\beta_j)\theta(-\beta_j')}\right\}^{Z_j}\left\{\theta(-\beta_j)\theta(-\beta_j')\right\}^r\mathcal{I}(Z_j \in H_p)d\pi(\boldsymbol{\beta})d\pi(\boldsymbol{\beta})'.
\end{aligned}$$

Hence by Fubini's Theorem and independence of the Z_j's,

$$\mathbb{E}_0(L_\pi^2 \mathcal{I}_D) = 2^{2n} \iint \prod_{j=1}^{p} \mathbb{E}_0\Big[\left\{\frac{\theta(\beta_j)\theta(\beta'_j)}{\theta(-\beta_j)\theta(-\beta'_j)}\right\}^{Z_j} \left\{\theta(-\beta_j)\theta(-\beta'_j)\right\}^r \mathcal{I}(Z_j \in H_p)\Big] d\pi(\boldsymbol{\beta}) d\pi(\boldsymbol{\beta}')$$

$$= 2^{2n} 2^{-2k} \binom{p}{k}^{-2} \sum_{m_1, m_2, \xi_1, \xi_2} \prod_{j=1}^{p} \mathbb{E}_0\Big[\left\{\frac{\theta(\beta_j)\theta(\beta'_j)}{\theta(-\beta_j)\theta(-\beta'_j)}\right\}^{Z_j} \left\{\theta(-\beta_j)\theta(-\beta'_j)\right\}^r \mathcal{I}(Z_j \in H_p)\Big] \tag{A.19}$$

For any two i.i.d draws (m_1, ξ_1) and $(m_2, , \xi_2)$, set for $j = 1, \dots, p$

$$T_j = \mathbb{E}_0\Big[\left\{\frac{\theta(\beta_j)\theta(\beta'_j)}{\theta(-\beta_j)\theta(-\beta'_j)}\right\}^{Z_j} \left\{\theta(-\beta_j)\theta(-\beta'_j)\right\}^r \mathcal{I}(Z_j \in H_p)\Big] .$$

We divide into the following cases. For each $j \in \{1, \dots, p\}$,

1. $j \in m_1^c \cap m_2^c$: $T_j = \frac{\mathbb{P}_0(Z_j \in H_p)}{2^r}$.
2. $j \in m_1 \cap m_2^c \cap \{l : \xi_1^l = 1\}$: $T_j = \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{Z_j} \mathcal{I}(Z_j \in H_p)\Big](\frac{\theta(-A)}{2})^r$.
3. $j \in m_1 \cap m_2^c \cap \{l : \xi_1^l = -1\}$:

$$T_j = \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{r-Z_j} \mathcal{I}(Z_j \in H_p)\Big](\frac{\theta(-A)}{2})^r$$
$$= \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{Z_j} \mathcal{I}(Z_j \in H_p)\Big](\frac{\theta(-A)}{2})^r$$

   since $r - Z_j \stackrel{d}{=} Z_j$ and the definition of the set D is also symmetric in Z_j and $r - Z_j$ for all Z_j.
4. $j \in m_1^c \cap m_2 \cap \{l : \xi_2^l = 1\}$: $T_j = \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{Z_j} \mathcal{I}(Z_j \in H_p)\Big](\frac{\theta(-A)}{2})^r$.
5. $j \in m_1^c \cap m_2 \cap \{l : \xi_2^l = -1\}$: $T_j = \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{Z_j} \mathcal{I}(Z_j \in H_p)\Big](\frac{\theta(-A)}{2})^r$ by the symmetry argument made in case (3) above.
6. $j \in m_3 \cap \{l : \xi_1^l = \xi_2^l = 1\}$: $T_j = \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_j} \mathcal{I}(Z_j \in H_p)\Big](\theta(-A))^{2r}$.
7. $j \in m_3 \cap \{j : \xi_1^l = \xi_2^l = -1\}$: $T_j = \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_j} \mathcal{I}(Z_j \in H_p)\Big](\theta(-A))^{2r}$ again by the symmetry argument.

8. $j \in m_4$: $T_j = \{\theta(A)\theta(-A)\}^r \mathbb{P}_0(Z_j \in H_p)$.

Grouping the terms in (A.19) by the above cases and collecting terms,

$\mathbb{E}_0(L_\pi^2 I(D))$

$$= \frac{2^{2n-2k}}{\binom{p}{k}^2} \sum_{\substack{m_1,m_2\\ \xi_1,\xi_2}} \Big[\prod_{j\in m_1^c\cap m_2^c} \frac{\mathbb{P}_0(Z_j\in H_p)}{2^r} \prod_{j\in m_1^c\Delta m_2^c} \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{Z_j}\mathcal{I}(Z_j\in H_p)\Big]\Big(\frac{\theta(-A)}{2}\Big)^r$$

$$\times \prod_{j\in m_3} \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_j}\mathcal{I}(Z_j\in H_p)\Big](\theta(-A))^{2r} \prod_{j\in m_4}\{\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_j\in H_p)\Big]$$

$$= \frac{2^{2n-2k}}{\binom{p}{k}^2} \sum_{\substack{m_1,m_2\\ \xi_1,\xi_2}} \Big[\prod_{j\in m_1^c\cap m_2^c} \frac{\mathbb{P}_0(Z_1\in H_p)}{2^r} \prod_{j\in m_1^c\Delta m_2^c} \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{Z_1}\mathcal{I}(Z_1\in H_p)\Big](\frac{\theta(-A)}{2})^r$$

$$\times \prod_{j\in m_3} \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](\theta(-A))^{2r} \prod_{j\in m_4}\{\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)\Big]$$

$$= \frac{2^{2n-2k}}{\binom{p}{k}^2} \sum_{\substack{m_1,m_2\\ \xi_1,\xi_2}} \Big[\Big(\frac{\mathbb{P}_0(Z_1\in H_p)}{2^r}\Big)^{|m_1^c\cap m_2^c|}\Big(\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{Z_1}\mathcal{I}(Z_1\in H_p)\Big](\frac{\theta(-A)}{2})^r\Big)^{|m_1\Delta m_2|}$$

$$\times\Big(\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](\theta(-A))^{2r}\Big)^{|m_3|}\Big(\{\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)\Big)^{|m_4|}\Big]$$

$$= \frac{1}{2^{2k}\binom{p}{k}^2} \sum_{\substack{m_1,m_2\\ \xi_1,\xi_2}} \Big[\Big(\mathbb{P}_0(Z_1\in H_p)\Big)^{|m_1^c\cap m_2^c|}\Big(\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^r\Big)^{|m_1\Delta m_2|}$$

$$\times\Big(\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^{2r}\Big)^{|m_3|}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)\Big)^{|m_4|}\Big]$$

$$\leq \frac{1}{2^{2k}\binom{p}{k}^2} \sum_{\substack{m_1,m_2\\ \xi_1,\xi_2}} \Big[\Big(\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^{2r}\Big)^{|m_3|}$$

$$\times\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)\Big)^{|m_4|}\Big]$$

$$= \frac{1}{2^{2k}\binom{p}{k}^2} \sum_{\substack{m_1,m_2\\ \xi_1,\xi_2}} \Big[\Big\{\frac{\Big(\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^{2r}\Big)}{\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)\Big)}\Big\}^{|m_3|}$$

$$\times\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)\Big)^{|m_1\cap m_2|}\Big]$$

$$= \frac{1}{\binom{p}{k}^2}\sum_{m_1,m_2}\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)$$

$$+\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^{2r}\Big)^{|m_1\cap m_2|}\Big]$$

$$= \mathbb{E}_W\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)+\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^{2r}\Big)\Big]^W$$

where $W \sim \text{Hypergeometric}(p,k,k)$. Now we observe that by Lemma A.1,

$$\mathbb{E}_W\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)+\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^{2r}\Big)\Big]^W$$
$$\le \mathbb{E}_U\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)+\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^{2r}\Big)\Big]^U$$

where $U \sim \text{Bin}(k, \frac{k}{p-k})$, provided the following holds:

$$\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)+\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^{2r}\Big)\Big]\ge 1\ . \tag{A.20}$$

Hence under the inequality (A.20) we have

$$\begin{aligned}
&\mathbb{E}_0(L_\pi^2 I(D))\\
&\le \mathbb{E}_U\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)+\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^{2r}\Big)\Big]^U\\
&=\Big\{1+\frac{k}{p-k}\Big(\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)\\
&\quad+\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^{2r}\Big)\Big]-1\Big)\Big\}
\end{aligned}$$

Hence in order to prove the second inequality of (A.8) it suffices to verify the inequality (A.20) and prove

$$\frac{k^2}{p}\Big(\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)+\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^{2r}\Big)\Big]-1\Big)=o(1)\ . \tag{A.21}$$

We first verify (A.21). We note that

$$\begin{aligned}
&\frac{k^2}{p}\Big(\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p)+\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^{2r}\Big)\Big]-1\Big)\\
&:=E_1+E_2+E_3\ .
\end{aligned}\tag{A.22}$$

where

$$E_1=\frac{k^2}{2p}\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z_1\in H_p),$$
$$E_2=\frac{k^2}{2p}\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1\in H_p)\Big](2\theta(-A))^{2r}$$

and

$$E_3 = \frac{k^2}{p}.$$

Since $\alpha > \frac{1}{2}$, trivially $E_3 = o(1)$. Hence it suffices to prove that $E_1 = o(1)$ and $E_2 = o(1)$. To this end, first note that

$$\begin{aligned} E_1 &= \frac{k^2}{2p}\{4\theta(A)\theta(-A)\}^r \mathbb{P}_0(Z_1 \in H_p) \le \frac{k^2}{p}\{4\theta(A)\theta(-A)\}^r \\ &= \frac{k^2}{p} e^{r\log(4\theta(A)\theta(-A))} = \frac{k^2}{p} e^{r(2\theta'(0)A - 2\theta'(0)^2A^2 - 2\theta'(0)A - 2\theta'(0)^2A^2 + o(A^2))} \\ &= \frac{k^2}{p} e^{r(-4\theta'(0)^2A^2 + o(A^2))} = o(1) \end{aligned} \tag{A.23}$$

as required. Next we control E_2 as follows:

$$\begin{aligned} E_2 &= \frac{k^2}{2p}\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1 \in H_p)\Big](2\theta(-A))^{2r} \\ &\le \frac{k^2}{p}\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1 \in H_p)\Big](2\theta(-A))^{2r} \\ &= \frac{k^2}{p}\mathbb{E}_0\Big[e^{2Z_1\log\{\frac{\theta(A)}{\theta(-A)}\}}\mathcal{I}(Z_1 \in H_p)\Big](2\theta(-A))^{2r} \\ &= \mathbb{E}_0\Big[e^{2Z_1(4\theta'(0)A+\epsilon)}\mathcal{I}(Z_1 \in H_p)\Big](2\theta(-A))^{2r} \text{ where } \epsilon = o(A^2) \\ &\le \Big\{\mathbb{E}_0\Big[e^{8\theta'(0)AZ_1}\mathcal{I}(Z_1 \in H_p)(2\theta(-A))^{2r}\frac{k^2}{p}\Big]^f\Big\}^{1/f}\Big\{\mathbb{E}_0\Big[e^{2g\epsilon Z_1}\Big]\Big\}^{1/g}. \end{aligned} \tag{A.24}$$

where the last line is by Hölder's Inequality for any $f > 1$ and complementary $g > 1$ such that $\frac{1}{f} + \frac{1}{g} = 1$. Our next task is hence to control $\mathbb{E}_0\Big[e^{8\theta'(0)AZ_1}\mathcal{I}(Z_1 \in H_p)(2\theta(-A))^{2r}\frac{k^2}{p}\Big]^f$ for an appropriately chosen $f > 1$ and then subsequently bound $\Big\{\mathbb{E}_0\Big[e^{2g\epsilon Z_1}\Big]\Big\}^{1/g}$ for the corresponding $g > 1$. We first analyze $\mathbb{E}_0\Big[e^{8\theta'(0)AZ_1}\mathcal{I}(Z_1 \in H_p)(2\theta(-A))^{2r}\frac{k^2}{p}\Big]^f$ for arbitrary $f > 1$ and we will make the choice of the pair (f, g) clear later. To that end, we have

$$\begin{aligned} &\mathbb{E}_0\Big[e^{8\theta'(0)AfZ_1}\mathcal{I}(Z_1 \in H_p)(2\theta(-A))^{2rf}\Big] \\ &= \mathbb{E}_0\left[e^{8\theta'(0)AfZ_1}(2\theta(-A))^{2rf}\left\{\mathcal{I}\left(Z_1 \le \frac{r}{2} + \sqrt{2\log p}\sqrt{\frac{r}{4}}\right) + \mathcal{I}\left(Z_1 \ge \frac{r}{2} - \sqrt{2\log p}\sqrt{\frac{r}{4}}\right)\right\}\right] \end{aligned}$$

$$:= I_1 + I_2 - I_3 \tag{A.25}$$

where $I_1 = \mathbb{E}_0\Big[e^{8\theta^{'}(0)AfZ_1}(2\theta(-A))^{2rf}\{\mathcal{I}(Z_1 \leq \frac{r}{2} + \sqrt{2\log p}\sqrt{\frac{r}{4}})\}\Big]$ and $I_2 = \mathbb{E}_0\Big[e^{8\theta^{'}(0)AfZ_1}(2\theta(-A))^{2rf}\{\mathcal{I}(Z_1 \geq \frac{r}{2} - \sqrt{2\log p}\sqrt{\frac{r}{4}})\}\Big]$ and I_3 is the remainder. We will analyze I_1 in detail; the analysis of I_2 is very similar and is omitted. The proof of $I_3 = o(1)$ is easier and can be also done following similar techniques and is hence also omitted. Recalling the definition of $W_r := \frac{Z_1 - \frac{r}{2}}{\sqrt{\frac{r}{4}}}$, we have

$$\mathbb{E}_0\left[e^{8\theta^{'}(0)AfZ_1}\mathcal{I}\left(Z_1 \leq \frac{r}{2} + \sqrt{2\log p}\sqrt{\frac{r}{4}}\right)\right] = e^{4\theta^{'}(0)fAr}\mathbb{E}_0\left[e^{4\theta^{'}(0)fAW_r}\mathcal{I}\left(\frac{W_r}{\sqrt{r}} \leq \sqrt{2\log p}\right)\right]$$

Arguing similarly as in proof of the first inequality of (A.8), it can be shown that it suffices to analyze $e^{4\theta^{'}(0)fAr}\mathbb{E}_0\Big[e^{4\theta^{'}(0)fAB_r}\{\mathcal{I}(\frac{B_r}{\sqrt{r}} \leq \sqrt{2\log p})\}\Big]$ where B_r is the version of Brownian Motion on the same probability space as W_r satisfying (A.12). Of course in the proof of the first inequality of (A.8) we went through complete details in choosing an appropriate $x > 0$ which calibrates the degree of approximation between W_r and B_r. However we note that the same choice of x as before goes through and the essence of the proof boils down to controlling $e^{4\theta^{'}(0)fAr}\mathbb{E}_0\Big[e^{4\theta^{'}(0)fAB_r}\{\mathcal{I}(\frac{B_r}{\sqrt{r}} \leq \sqrt{2\log p})\}\Big]$. Now

$$\begin{aligned}
&\mathbb{E}_0\left[e^{4\theta^{'}(0)fAB_r}\mathcal{I}\left(\frac{B_r}{\sqrt{r}} \leq \sqrt{2\log p}\right)\right] \\
&= \mathbb{E}_0\left[e^{4\theta^{'}(0)fA^*\frac{B_r}{\sqrt{r}}}\mathcal{I}\left(\frac{B_r}{\sqrt{r}} \leq \sqrt{2\log p}\right)\right] \\
&= \int_{-\infty}^{\sqrt{2\log p}} e^{4\theta^{'}(0)fA^*v}\frac{e^{-\frac{v^2}{2}}}{\sqrt{2\pi}}dv \\
&= \int_{-\infty}^{\sqrt{2\log p}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(v^2-8\theta^{'}(0)fA^*v+16\theta^{'}(0)^2f^2(A^*)^2)}e^{8\theta^{'}(0)^2f^2(A^*)^2}dv \\
&= \Phi(\sqrt{2\log p} - 4\theta^{'}(0)fA^*)e^{8\theta^{'}(0)^2f^2(A^*)^2}
\end{aligned}$$

Considering the expression for I_1 in (A.25), we have the following:

$$e^{4\theta^{'}(0)fAr}(2\theta(-A))^{2rf} = e^{-4\theta^{'}(0)^2(A^*)^2f+rf\epsilon'} \text{ where } \epsilon' = o(A^2)$$

since $\theta^{''}(0) = 0$. Hence we have, as in the proof of the first inequality of (A.8),

$$I_1 \lesssim e^{(1-2\alpha)f\log(p)+8\theta^{'}(0)^2(A^*)^2f^2-4\theta^{'}(0)(A^*)^2f+rf\epsilon'}\Phi(\sqrt{2\log p} - 4\theta^{'}(0)fA^*)$$

$$= e^{\{(1-2\alpha)f+16\theta'(0)^2f^2t-8\theta'(0)ft\}\log(p)+rf\epsilon'}\Phi(\sqrt{2\log p}-4\theta'(0)fA^*)$$

Now the behavior of the bounds on $\Phi(s)$ is different depending on whether $s \geq 0$ or $s < 0$ and we have $\Phi(s) \leq 1$ when $s \geq 0$ and $\Phi(s) < \phi(s)$ if $s < 0$. But $\sqrt{2\log p} - 4\theta'(0)fA^* \lesseqgtr 0$ accordingly as $t \gtreqless \frac{1}{16\theta'(0)^2f^2}$. Hence we divide our analysis into two parts according to the range of t.

When $t \leq \frac{1}{16\theta'(0)^2f^2}$, *i.e.*, $\sqrt{2\log p} - 4\theta'(0)fA^* \geq 0$ we have

$$I_1 \lesssim e^{\{(1-2\alpha)f+16\theta'(0)^2f^2t-8\theta'(0)ft\}\log(p)+rf\epsilon'}$$

Now the coefficient of $\log(p)$ in the above exponent is

$$\begin{aligned} f\left[(1-2\alpha)+8\theta'(0)^2t\,(2f-1)\right] &= 2f\left[\left(\frac{1}{2}-\alpha\right)+4\theta'(0)^2t\,(2f-1)\right] \\ &= 8f\theta'(0)^2\left[\frac{\frac{1}{2}-\alpha}{4\theta'(0)^2}+t\,(2f-1)\right] \end{aligned}$$

For $\alpha \leq \frac{3}{4}$, since $t < \rho^*_{\text{binary}}(\alpha) = \frac{\alpha-\frac{1}{2}}{4\theta'(0)^2}$, we have there exists $\delta_1(\alpha,t) > 0$ such that $\frac{\frac{1}{2}-\alpha}{4\theta'(0)^2} + t(2f-1) < 0$ whenever $f = 1+\delta$ with $\delta \leq \delta_1(\alpha,t)$. For $\alpha > \frac{3}{4}$, since $t \leq \frac{1}{16\theta'(0)^2f^2}$, $\frac{\alpha-\frac{1}{2}}{4\theta'(0)^2}$ is monotone increasing in α and $\rho^*_{\text{binary}}(\frac{3}{4}) = \frac{\frac{3}{4}-\frac{1}{2}}{4\theta'(0)^2} = \frac{1}{16\theta'(0)^2f^2}$, we have that there exists $\delta_2(\alpha,t) > 0$ such that $\frac{\frac{1}{2}-\alpha}{4\theta'(0)^2} + t(2f-1) < 0$ whenever $f = 1+\delta$ with $\delta \leq \delta_2(\alpha,t)$.

When $t > \frac{1}{16\theta'(0)^2f^2}$ we have

$$\begin{aligned} I_1 &\lesssim e^{\{(1-2\alpha)f+16\theta'(0)^2f^2t-8\theta'(0)ft\}\log(p)+rf\epsilon'}\phi(\sqrt{2\log p}-4\theta'(0)fA^*) \\ &= e^{f\log(p)(1-2\alpha-8\theta'(0)^2t-1+8\theta'(0)\sqrt{t})+\log(p)(f-1)+rf\epsilon'} \\ &= e^{f\log(p)(1-2\alpha-8\theta'(0)^2t-1+8\theta'(0)\sqrt{t})+\log(p)(f-1)+rf\epsilon'} \\ &= e^{f\log(p)\{2(1-\alpha)-2(1-2\theta'(0)\sqrt{t})^2\}+(f-1)\log(p)+rf\epsilon'} \end{aligned}$$

Since $t < \rho^*_{\text{binary}}(\alpha)$, $2(1-\alpha) - 2(1-2\theta'(0)\sqrt{t})^2 < 0$ and hence there exists $\delta_3(\alpha,t) > 0$ such that $f\{\{2(1-\alpha)-2(1-2\theta'(0)\sqrt{t})^2\}+(f-1) < 0$ whenever $f = 1+\delta$ with $\delta \leq \delta_3(\alpha,t)$.

Hence choosing $f = 1+\delta$ with $\delta = \min\{\delta_1(\alpha,t),\delta_2(\alpha,t),\delta_3(\alpha,t)\}$ yields $I_1 = o(1)$ as required. Controlling the corresponding g-factor in (A.24) is similar to that in (A.11) and can be done along the lines of deriving (A.18).

Next we prove (A.20). We note that it suffices to prove that $\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1 \in H_p)\Big](2\theta(-A))^{2r} \to \infty$. As before

$$\begin{aligned}
&\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_1}\mathcal{I}(Z_1 \in H_p)\Big](2\theta(-A))^{2r}\\
&\quad= \mathbb{E}_0\Big[e^{2Z_1\log\{\frac{\theta(A)}{\theta(-A)}\}}\mathcal{I}(Z_1 \in H_p)\Big](2\theta(-A))^{2r}\\
&\quad= \mathbb{E}_0\Big[e^{(4\theta'(0)+\epsilon)A2Z_1}\mathcal{I}(Z_1 \in H_p)\Big](2\theta(-A))^{2r} \text{ where } \epsilon A = o(A^2)\\
&\quad= e^{(4\theta'(0)+\epsilon)Ar}(2\theta(-A))^{2r}\mathbb{E}_0\Big[e^{(4\theta'(0)+\epsilon)AW_r}\mathcal{I}(Z_1 \in H_p)\Big].
\end{aligned}$$

Now,

$$\begin{aligned}
&\mathbb{E}_0\Big[e^{(4\theta'(0)+\epsilon)AW_r}\mathcal{I}(Z_1 \in H_p)\Big]\\
&\geq \mathbb{E}_0\Big[e^{(4\theta'(0)+\epsilon)AW_r}\mathcal{I}(|W_r - B_r| \leq (\log r + x))\mathcal{I}(-\sqrt{2\log p}\sqrt{r} \leq W_r \leq \sqrt{2\log p}\sqrt{r})\\
&\geq e^{-(4\theta'(0)+\epsilon)(\log r+x)A}\mathbb{E}_0\Big[e^{(4\theta'(0)+\epsilon)AB_r}\mathcal{I}(|W_r - B_r| \leq (\log r + x))\\
&\qquad \times \mathcal{I}(-\sqrt{2\log p}\sqrt{r} + (\log r + x) \leq B_r \leq \sqrt{2\log p}\sqrt{r} - (\log r + x))\Big]\\
&= e^{-(4\theta'(0)+\epsilon)(\log r+x)A}\mathbb{E}_0\Big[e^{(4\theta'(0)+\epsilon)A^*\frac{B_r}{\sqrt{r}}}\\
&\quad \mathcal{I}(-\sqrt{2\log p}\sqrt{r} + \frac{(\log r + x)}{\sqrt{r}} \leq \frac{B_r}{\sqrt{r}} \leq \sqrt{2\log p} - \frac{(\log r + x)}{\sqrt{r}})\Big]\\
&- \quad e^{-(4\theta'(0)+\epsilon)(\log r+x)A}\mathbb{E}_0\Big[e^{(4\theta'(0)+\epsilon)A^*\frac{B_r}{\sqrt{r}}}\\
&\quad \mathcal{I}(-\sqrt{2\log p}\sqrt{r} + \frac{(\log r + x)}{\sqrt{r}} \leq \frac{B_r}{\sqrt{r}} \leq \sqrt{2\log p} - \frac{(\log r + x)}{\sqrt{r}})\\
&\qquad \times \mathcal{I}(|W_r - B_r| > (\log r + x))\Big]\\
&:= S_1 - S_2\,.
\end{aligned} \tag{A.26}$$

Hence it is enough to prove that $e^{(4\theta'(0)+\epsilon)Ar}(2\theta(-A))^{2r}S_1 \to \infty$ and $e^{(4\theta'(0)+\epsilon)Ar}(2\theta(-A))^{2r}S_2 = O(1)$.

Now

$$S_1 := e^{-(4\theta'(0)+\epsilon)(\log r+x)A}\mathbb{E}_0\Big[e^{(4\theta'(0)+\epsilon)A^*\frac{B_r}{\sqrt{r}}}$$

$$\mathcal{I}(-\sqrt{2\log p}\sqrt{r}+\frac{(\log r+x)}{\sqrt{r}}\leq\frac{B_r}{\sqrt{r}}\leq\sqrt{2\log p}-\frac{(\log r+x)}{\sqrt{r}})\Big]$$
$$=e^{-(4\theta^{'}(0)+\epsilon)(\log r+x)A}e^{\frac{1}{2}(4\theta^{'}(0)+\epsilon)^2(A^*)^2}\Phi^{(\sqrt{2\log p}-\frac{(\log r+x)}{\sqrt{r}}-(4\theta^{'}(0)+\epsilon)A^*)}. \tag{A.27}$$

Also

$$e^{(4\theta^{'}(0)+\epsilon)Ar}(2\theta(-A))^{2r}=e^{-4\theta^{'}(0)^2A^2r+ro(A^2)+\epsilon Ar}=e^{-4\theta^{'}(0)^2A^2r+ro(A^2)}. \tag{A.28}$$

Hence by (A.27) and (A.28) we have

$$e^{(4\theta^{'}(0)+\epsilon)Ar}(2\theta(-A))^{2r}S_1=e^{\{\frac{1}{2}(4\theta^{'}(0)+\epsilon)^2(A^*)^2-(4\theta^{'}(0)+\epsilon)(\log r+x)A-4\theta^{'}(0)^2A^2r+ro(A^2)\}}$$
$$\times\Phi\Big(\sqrt{2\log p}-\frac{(\log r+x)}{\sqrt{r}}-(4\theta^{'}(0)+\epsilon)A^*\Big). \tag{A.29}$$

The behavior of the above quantity depends on $\Phi(\eta)$ where $\eta=\sqrt{2\log p}-\frac{(\log r+x)}{\sqrt{r}}-(4\theta^{'}(0)+\epsilon)A^*$. Hence we divide our study in the following cases.

First suppose $t\leq\frac{1}{16\theta^{'}(0)^2}$. If $\epsilon=-\delta<0$, then $\eta\geq\frac{\sqrt{2\log p}\delta}{4\theta^{'}(0)}-\frac{(\log r+x)}{\sqrt{r}}$. Hence $\Phi(\eta)\geq\frac{1}{2}+o(1)$. Hence from (A.29) we have

$$\begin{aligned}
&e^{(4\theta^{'}(0)+\epsilon)Ar}(2\theta(-A))^{2r}S_1\\
&\geq(\frac{1}{2}+o(1))e^{\{\frac{1}{2}(4\theta^{'}(0)+\epsilon)^2(A^*)^2-(4\theta^{'}(0)+\epsilon)(\log r+x)Ar-4\theta^{'}(0)^2A^2r+ro(A^2)\}}\\
&=(\frac{1}{2}+o(1))e^{4\theta^{'}(0)(A^*)^2+\kappa-(4\theta^{'}(0)+\epsilon)(\log r+x)A+ro(A^2)}\ \text{where } |\kappa|\ll\log(p)\\
&=(\frac{1}{2}+o(1))e^{8t\theta^{'}(0)\log(p)+\kappa-(4\theta^{'}(0)+\epsilon)(\log r+x)A+ro(A^2)}.
\end{aligned} \tag{A.30}$$

Now $(4\theta^{'}(0)+\epsilon)(\log r+x)A<\frac{5\theta^{'}(0)(\log r+x)\sqrt{2\log p}}{\sqrt{r}}\ll\log(p)$ if $x=a_{r,p}\log(p)$ is such that $a_{r,p}\to\infty$ ensuring both $r\gg a_{r,p}\log(p)$ and $\frac{a_{r,p}\log(p)\sqrt{2\log p}}{\sqrt{r}}\ll\log(p)$. Thus $e^{(4\theta^{'}(0)+\epsilon)Ar}(2\theta(-A))^{2r}S_1\geq e^{c\log(p)}$ for some $c>0$ and hence diverges.

If $\epsilon>0$, then $\eta\geq-\frac{\sqrt{2\log p}\epsilon}{4\theta^{'}(0)}-\frac{(\log r+x)}{\sqrt{r}}$ and hence $-\eta\leq\frac{\sqrt{2\log p}\epsilon}{4\theta^{'}(0)}+\frac{(\log r+x)}{\sqrt{r}}\ll\tau$ for some divergent $\tau\ll\sqrt{\log(p)}$. Hence, by Lemma **??**,

$$\Phi(\eta)=\overline{\Phi}(-\eta)\geq(1-\frac{1}{\tau^2})\frac{\phi(\tau)}{\tau}$$

$$\geq \frac{\phi(\tau)}{2\tau} \text{ for sufficiently large } \; r, p$$
$$= \frac{e^{-\tau^2/2}}{\tau\sqrt{2\pi}}.$$

Hence similar to the calculations in deriving (A.30) we have

$$\begin{aligned} e^{(4\theta'(0)+\epsilon)Ar}(2\theta(-A))^{2r}S_1 &\geq \frac{e^{-\tau^2/2}}{\tau\sqrt{2\pi}} e^{8t\theta'(0)\log(p)+\kappa-(4\theta'(0)+\epsilon)(\log r+x)A+ro(A^2)} \\ &\geq \frac{e^{-\tau^2/2}}{\tau\sqrt{2\pi}} e^{c\log(p)} \text{ for some } c > 0 \\ &= \frac{e^{c\log(p)-\tau^2/2}}{\tau\sqrt{2\pi}} \\ &\geq \frac{e^{c'\log(p)}}{\sqrt{\log(p)}} \text{ for some } c' > 0 \text{ since } \tau \ll \sqrt{\log(p)} \\ &\to \infty. \end{aligned} \tag{A.31}$$

Now suppose $\frac{1}{16\theta'(0)^2} < t < \rho^*_{\text{binary}}(\alpha)$. If $\epsilon = -\delta < 0$, then $\eta < \frac{\sqrt{2\log p}\delta}{4\theta'(0)} - \frac{(\log r+x)}{\sqrt{r}}$. If $\eta \in (-2, \frac{\sqrt{2\log p}\delta}{4\theta'(0)} - \frac{(\log r+x)}{\sqrt{r}})$ then since $\Phi(\eta) \geq \Phi(-2)$ we have by the same argument as in (A.30) that $e^{(4\theta'(0)+\epsilon)Ar}(2\theta(-A))^{2r}S_1 \to \infty$. Now suppose $\eta \leq -1$. Then once again using the fact that $\Phi(\eta) = \overline{\Phi}(-\eta)$ and Lemma **??** we have that

$$\begin{aligned} &e^{(4\theta'(0)+\epsilon)Ar}(2\theta(-A))^{2r}S_1 \\ &= e^{\{\frac{1}{2}(4\theta'(0)+\epsilon)^2(A^*)^2-(4\theta'(0)+\epsilon)(\log r+x)A-4\theta'(0)^2A^2r+ro(A^2)\}}\overline{\Phi}(-\eta) \\ &\geq (1-\frac{1}{\eta^2})\frac{\phi(\eta)}{-\eta}e^{\{\frac{1}{2}(4\theta'(0)+\epsilon)^2(A^*)^2-(4\theta'(0)+\epsilon)(\log r+x)A-4\theta'(0)^2A^2r+ro(A^2)\}} \\ &= \frac{(1-\frac{1}{\eta^2})}{-\eta}e^{\{\log(p)(1-2(1-2\theta'(0)\sqrt{t})^2)+\kappa'\}} \end{aligned} \tag{A.32}$$

where $|\kappa'| \ll \log(p)$. Now

$$\begin{aligned} \inf_{\frac{1}{16\theta'(0)^2}<t<\rho^*_{\text{binary}}(\alpha)}\{1-2(1-2\theta'(0)\sqrt{t})^2\} &\geq \inf_{t<\rho^*_{\text{binary}}(\alpha)}\{1-2(1-2\theta'(0)\sqrt{t})^2\} \\ &= 1-2(1-2\theta'(0)\sqrt{\rho^*_{\text{binary}}(\alpha)})^2 \end{aligned}$$

$$= 1 - 2(1 - 2\theta^{'}(0)\sqrt{\frac{(1-\sqrt{1-\alpha})^2}{4\theta^{'}(0)^2}})^2$$
$$= 1 - 2(1-(1-\sqrt{1-\alpha}))^2 = 2\alpha - 1 > 0$$

since $\alpha > \frac{1}{2}$. Hence from (A.32) we have that

$$e^{(4\theta^{'}(0)+\epsilon)Ar}(2\theta(-A))^{2r}S_1 \geq \frac{3}{-4\eta}e^{c'' \log(p)} \text{ for some } c'' > 0$$
$$\to \infty$$

since $|\eta| \lesssim \log(p)$. This completes the proof of $e^{(4\theta^{'}(0)+\epsilon)Ar}(2\theta(-A))^{2r}S_1 \to \infty$.

Next we prove $e^{(4\theta^{'}(0)+\epsilon)Ar}(2\theta(-A))^{2r}S_2 = O(1)$. To this end note, that by the Cauchy-Schwarz Inequality,

$$\begin{aligned} &S_2 \\ &\leq e^{-(4\theta^{'}(0)-\epsilon)(\log r + x)A}(\mathbb{E}_0[e^{(8\theta^{'}(0)-2\epsilon)A^* V}]\mathbb{P}_0(|W_r - B_r| > (\log r + x)))^{1/2} \text{ where } V \sim \mathrm{N}(0,1) \\ &\leq (e^{-(4\theta^{'}(0)-\epsilon)\frac{(\log r + x)}{\sqrt{r}}\sqrt{2t\log(p)}+(8\theta^{'}(0)-2\epsilon)^2 t\log(p) - \lambda x})^{1/2} \text{ by Equation (A.12)} \end{aligned} \tag{A.33}$$

Hence from (A.33) and (A.28) we have that $e^{(4\theta^{'}(0)+\epsilon)Ar}(2\theta(-A))^{2r}S_2 \to 0$ since $x = a_{r,p}\log(p)$ where $a_{r,p}$ was chosen to diverge at a slow enough rate. This completes the verification of (A.20) and hence proves the theorem. □

Proof of Theorem 6.10. We will provide proof for the lower bound in problem 2.3 where $\theta \in BC^2(0)$. Using Remark 6.1, the proof also holds for problem 6.2. To analyze the power of the Higher Criticism test, we need to define the following quantities. Let $\frac{1}{2} + \delta = \theta(A)$. Also define S_1 to be a generic $\mathrm{Bin}(r, \frac{1}{2}+\delta)$ random variable and let $\mathbb{B}_1, \overline{\mathbb{B}}_1$ respectively denote the distribution function and survival function of S_1. Then

$$\mathbb{B}_1(t) = \mathbb{P}(\frac{|S_1 - \frac{r}{2}|}{\sqrt{\frac{r}{4}}} \leq t),\ \overline{\mathbb{B}}_1(t) = 1 - \mathbb{B}_1(t)\ .$$

The proof of the rest of the theorem relies on the following lemma.

Lemma A.5. *Let $r \gg \log(p)$ and $t > \rho^*_{\text{logistic}}(\alpha)$. Then there exists $s \in [1, \sqrt{3\log(p)}]$ such that*

1. $\frac{k}{\sqrt{p}}\frac{\overline{\mathbb{B}}_1(s) - \overline{\mathbb{B}}(s)}{\sqrt{\overline{\mathbb{B}}(s)(1-\overline{\mathbb{B}}(s))}} \gg \log(p)$

2. $\frac{(p-k)\overline{\mathbb{B}}(s)(1-\overline{\mathbb{B}}(s))+k\overline{\mathbb{B}}_1(s)(1-\overline{\mathbb{B}}_1(s))}{k^2(\overline{\mathbb{B}}_1(s)-\overline{\mathbb{B}}(s))^2} \to 0$

Now we return to the proof of the main result. For any $z \in [1, \sqrt{3\log(p)}] \cap \mathbb{N}$, $\mathsf{T}_{\mathrm{HC}} \geq W_p(z)$ where $W_p(z) = \sqrt{p}\frac{\overline{\mathbb{F}}_p(z)-\overline{\mathbb{B}}(z)}{\sqrt{\overline{\mathbb{B}}(z)(1-\overline{\mathbb{B}}(z))}}$. Hence by Chebysev's inequality it suffices to prove that there exists $s \in [1, \sqrt{3\log(p)}]$ such that uniformly in $\beta \in \Theta_k^A$, $\frac{\mathbb{E}_\beta(W_p(s))}{\sqrt{2\log\log(p)}} \to \infty$ and $\frac{\mathrm{Var}_{\boldsymbol{\beta}}(W_p(s))}{(\mathbb{E}_{\boldsymbol{\beta}}(W_p(s)))^2} \to 0$ when $t > \rho^*_{\mathrm{logistic}}(\alpha)$. Fix $\boldsymbol{\beta}^* \in \Theta_k^A$; thus $\boldsymbol{\beta}^*$ has 0 in $p-k$ locations, A in k_1 locations (say) and $-A$ in $k-k_1 = k_2$ locations. Now note that by symmetry $\mathbb{P}(|\mathrm{Bin}(r, \frac{1}{2}+\delta) - \frac{r}{2}| > t) = \mathbb{P}(|\mathrm{Bin}(r, \frac{1}{2}-\delta) - \frac{r}{2}| > t)$ for all $t > 0$. Hence it is easy to show that irrespective of k_1, k_2, $\mathbb{E}_{\boldsymbol{\beta}^*}(W_p(s)) = \frac{k}{\sqrt{p}}\frac{\overline{\mathbb{B}}_1(s)-\overline{\mathbb{B}}(s)}{\sqrt{\overline{\mathbb{B}}(s)(1-\overline{\mathbb{B}}(s))}}$ and $\mathrm{Var}_{\boldsymbol{\beta}^*}(W_p(s)) = \frac{p-k}{p} + \frac{k}{p}\frac{\overline{\mathbb{B}}_1(s)(1-\overline{\mathbb{B}}_1(s))}{\overline{\mathbb{B}}(s)(1-\overline{\mathbb{B}}(s))}$. Hence to show $\frac{\mathbb{E}_\beta(W_p(s))}{\log(p)} \to \infty$ it suffices to show $\frac{k}{\sqrt{p}}\frac{\overline{\mathbb{B}}_1(s)-\overline{\mathbb{B}}(s)}{\sqrt{\overline{\mathbb{B}}(s)(1-\overline{\mathbb{B}}(s))}} \gg \sqrt{\log(p)}$ which is true by item 1 of Lemma A.5. Similarly to show that $\frac{\mathrm{Var}_{\boldsymbol{\beta}}(W_p(s))}{(\mathbb{E}_{\boldsymbol{\beta}}(W_p(s)))^2} \to 0$ it suffices to show that $\frac{(p-k)\overline{\mathbb{B}}(s)(1-\overline{\mathbb{B}}(s))+k\overline{\mathbb{B}}_1(s)(1-\overline{\mathbb{B}}_1(s))}{k^2(\overline{\mathbb{B}}_1(s)-\overline{\mathbb{B}}(s))^2} \to 0$ which is also true by item 2 of Lemma A.5. This completes the proof. □

Proof of Lemma A.5. By inspecting the expressions, it suffices to prove $\frac{k}{\sqrt{p}}\frac{\overline{\mathbb{B}}_1(s)-\overline{\mathbb{B}}(s)}{\sqrt{\overline{\mathbb{B}}(s)(1-\overline{\mathbb{B}}(s))}} \to \infty$ as some positive power of p. Put $s = \lfloor 2\sqrt{2q\log(p)} \rfloor$ where $q = \min\{4t\theta'(0)^2, \frac{1}{4}\}$. By the choice of q, $s \in [1, \sqrt{3\log(p)}] \cap \mathbb{N}$. Now by the Berry-Esseen approximation and Mill's Ratio,

$$\frac{k}{\sqrt{p}}\frac{\overline{\mathbb{B}}_1(s)-\overline{\mathbb{B}}(s)}{\sqrt{\overline{\mathbb{B}}(s)(1-\overline{\mathbb{B}}(s))}} \approx p^{\frac{1}{2}-\alpha}\frac{\overline{\Phi}\left(s-\frac{r\delta}{\sqrt{\frac{r}{4}}}\right)}{\sqrt{\overline{\Phi}(s)}} \approx p^{\frac{1}{2}-\alpha}\frac{\overline{\Phi}\left(s-\frac{r\theta^{'}(0)A}{\sqrt{\frac{r}{4}}}\right)}{\sqrt{\overline{\Phi}(s)}}$$

$$= p^{\frac{1}{2}-\alpha}\frac{\overline{\Phi}\left(\frac{\sqrt{2q\log(p)}\sqrt{r}-\sqrt{2t\log(p)}\theta^{'}(0)\sqrt{r}}{\sqrt{\frac{r}{4}}}\right)}{\sqrt{\overline{\Phi}\left(\frac{\sqrt{2q\log(p)}\sqrt{r}}{\sqrt{\frac{r}{4}}}\right)}}$$

$$= p^{\frac{1}{2}-\alpha}\frac{\overline{\Phi}\left(\sqrt{8q\log(p)}-\sqrt{8t\log(p)}\theta^{'}(0)\right)}{\sqrt{\overline{\Phi}\left(\frac{\sqrt{2q\log(p)}}{\sqrt{\frac{1}{4}}}\right)}}$$

$$\approx e^{\frac{1}{2}-\alpha-\frac{8}{2}\log(p)(\sqrt{q}-\sqrt{t}\theta^{'}(0))^2+\frac{8}{4}q\log(p)}$$

$$= p^{\frac{1}{2}-\alpha+2q-4(\sqrt{q}-\sqrt{t}\theta^{'}(0))^2} .$$

The exponent of p above is given by

$$\frac{1}{2}-\alpha+2q-4(\sqrt{q}-\sqrt{t}\theta^{'}(0))^2 =: f(q) \text{ say} .$$

The function $f(q)$ is maximized at $q = 4t\theta^{'}(0)^2$ for $t \leq \frac{1}{16\theta^{'}(0)^2}$. The maximum value is $(\frac{1}{2}-\alpha)+4t\theta^{'}(0)^2 > 0$ since $t > \rho^*_{\text{binary}}(\alpha)$. For $t > \frac{1}{16\theta^{'}(0)^2}$ if we put $q = \frac{1}{4}$, then $f(q) = (1-\alpha)-(1-2\sqrt{t}\theta^{'}(0))^2 > 0$ since $t > \max\{\rho^*_{\text{binary}}(\alpha), \frac{1}{16\theta^{'}(0)^2}\}$. Hence taking $s = \sqrt{2q\log(p)}$ where $q = \min\{4t\theta^{'}(0)^2, \frac{1}{4}\}$ proves the lemma. □

Proof of Proposition 6.11. Set $V_{(j)} = |Z-\frac{r}{2}|_{(j)}$ so that

$$\sup_{t\in[V_{(p-j)},V_{(p-j+1)})}\sqrt{p}\frac{\overline{\mathbb{F}}_p(t)-\overline{\mathbb{B}}(t)}{\sqrt{\overline{\mathbb{B}}(t)(1-\overline{\mathbb{B}}(t))}}$$

$$= \sup_{t\in[V_{(p-j)},V_{(p-j+1)})}\sqrt{p}\frac{\frac{j}{p}-\overline{\mathbb{B}}(t)}{\sqrt{\overline{\mathbb{B}}(t)(1-\overline{\mathbb{B}}(t))}}$$

$$= \sqrt{p}\frac{\frac{j}{p}-\inf_{t\in[V_{(p-j)},V_{(p-j+1)})}\overline{\mathbb{B}}(t)}{\sqrt{\inf_{t\in[V_{(p-j)},V_{(p-j+1)})}\overline{\mathbb{B}}(t)(1-\inf_{t\in[V_{(p-j)},V_{(p-j+1)})}\overline{\mathbb{B}}(t))}} .$$

Now $\overline{\mathbb{B}}(t)$ is a decreasing function of t and thus $\inf_{t\in[V_{(p-j)},V_{(p-j+1)})}\overline{\mathbb{B}}(t)) \geq \overline{\mathbb{B}}(V_{(p-j+1)})) = q_{(j)}$. Therefore we obtain that

$$\begin{aligned}\sup_{t\in[V_{(p-j)},V_{(p-j+1)})} &\sqrt{p}\frac{\overline{\mathbb{F}}_p(t)-\overline{\mathbb{B}}(t)}{\sqrt{\overline{\mathbb{B}}(t)(1-\overline{\mathbb{B}}(t))}}\\ &= \sqrt{p}\frac{\frac{j}{p}-\inf_{t\in[V_{(p-j)},V_{(p-j+1)})}\overline{\mathbb{B}}(t)}{\sqrt{\inf_{t\in[V_{(p-j)},V_{(p-j+1)})}\overline{\mathbb{B}}(t)(1-\inf_{t\in[V_{(p-j)},V_{(p-j+1)})}\overline{\mathbb{B}}(t))}}\\ &\leq \sqrt{p}\frac{\frac{j}{p}-q_{(j)}}{\sqrt{q_{(j)}(1-q_{(j)})}}\end{aligned}$$

since $\frac{c-x}{\sqrt{x(1-x)}}$ is a decreasing function of $x\in[0,1]$ for $c\in[0,1]$ and the proof is done. □

Proof of Theorem 6.12. We will provide proof for the lower bound in problem 2.3 where $\theta \in BC^2(0)$. Using Remark 6.1, the proof also holds for problem 6.2. As in proof of Theorem 6.8, we denote by $B^{'}_{jr}$ the version of Brownian Motion approximating $W^{'}_{jr}$ where $W^{'}_{jr} = W_{jr}\sqrt{r}$ and we can choose $B^{'}_{jr}$ independent for $j = 1,\ldots,p$. Let $B_{jr} = \frac{B^{'}_{jr}}{\sqrt{r}}$. For any $t_p > 0$,

$$\begin{aligned}\mathbb{P}(\max_{1\leq j\leq p}|W_{jr}|\leq t_p) &= \mathbb{P}(\max_{1\leq j\leq p}|W_{jr}-B_{jr}+B_{jr}|\leq t_p)\\ &\geq \mathbb{P}(\max_{1\leq j\leq p}|W_{jr}-B_{jr}|+\max_{1\leq j\leq p}|B_{jr}|\leq t_p)\\ &\geq \mathbb{P}\Big(\max_{1\leq j\leq p}|B_{jr}|\leq t_p-\frac{\log(r)+x}{\sqrt{r}}\Big)+o(1)\end{aligned}$$

for some $x > 0$. By a similar token we can show that $\mathbb{P}(\max_{1\leq j\leq p}|W_{jr}| \leq t_p) \leq \mathbb{P}(\max_{1\leq j\leq p}|B_{jr}| \leq t_p + \frac{\log(r)+x}{\sqrt{r}}) + o(1)$ for the same x above. Now by Lemma 11 of Arias-Castro, Candès and Plan (2011) we have that

$$\mathbb{P}(\max_{1\leq j\leq p}|B_{jr}|\leq \kappa_p+\frac{s}{\sqrt{2\log(p)}})\to e^{-e^{-s}}$$

as $p\to\infty$ where $\kappa_p = \sqrt{2\log(p)}-\frac{\log\log(p)+4\pi-4}{2\sqrt{2\log(p)}}$. Hence if $r \gg (\log(r))^2\log(p)$ then $\frac{\log(r)+x}{\sqrt{r}} = \frac{o(1)}{\sqrt{2\log(p)}}$ for appropriately chosen x. Therefore, by following the arguments of Lemma 11, Lemma 12 and proof of Theorem 5 of Arias-Castro, Candès and Plan (2011) we have the result when $r \gg (\log(r))^2\log(p)$

if we can choose x appropriately. We choose it to be the same as our choice in the proof of Theorem 6.8. To be precise, since $r \gg \log(p)$, there exists a sequence $a_{r,p} \to \infty$ such that $r \gg a_{r,p} \log(p)$. Take $x = a_{r,p}$. We skip the rest of the details. □

Proof of Theorem 7.1. We divide the proof into proofs of lower bound and upper bound respectively.

Part 1 : Proof of Lower Bound. For the purpose of brevity assume that $\mathbf{X} = [\mathbf{X}_1^t : \mathbf{X}_2^t]^t$ where $\mathbf{X}_1$ is an $n^* \times p$ matrix whose rows comprises exactly of the rows in Ω^* and $\mathbf{X}_2$ is an $n_* \times p$ matrix whose rows consists of the rows of $\mathbf{X}$ with more than one non-zero element in its support. Note that, this can always be achieved by a permutation of the rows of $\mathbf{X}$ and hence this does change the validity of the theorem. Let

$$f(\mathbf{X}_1, \boldsymbol{\beta}, \boldsymbol{\beta}^{'}) = \prod_{i \in \Omega^*} \left[\theta(\mathbf{x}_i^t \boldsymbol{\beta})\theta(\mathbf{x}_i^t \boldsymbol{\beta}^{'}) + \theta(-\mathbf{x}_i^t \boldsymbol{\beta})\theta(-\mathbf{x}_i^t \boldsymbol{\beta}^{'})\right]$$

and

$$f(\mathbf{X}_2, \boldsymbol{\beta}, \boldsymbol{\beta}^{'}) = \prod_{i \notin \Omega^*} \left[\theta(\mathbf{x}_i^t \boldsymbol{\beta})\theta(\mathbf{x}_i^t \boldsymbol{\beta}^{'}) + \theta(-\mathbf{x}_i^t \boldsymbol{\beta})\theta(-\mathbf{x}_i^t \boldsymbol{\beta}^{'})\right].$$

Note that by Lemma A.4, we have that $f(\mathbf{X}_1, \boldsymbol{\beta}, \boldsymbol{\beta}^{'}) \le \left[\theta^2(QA) + \theta^2(-QA)\right]^{n_*}$ for any realizations $\boldsymbol{\beta}, \boldsymbol{\beta}'$ from π:

$$\begin{aligned} \mathbb{E}_0(L_\pi^2) &= 2^{n^* + n_*} \iint f(\mathbf{X}_1, \boldsymbol{\beta}, \boldsymbol{\beta}^{'}) f(\mathbf{X}_2, \boldsymbol{\beta}, \boldsymbol{\beta}^{'}) d\pi(\boldsymbol{\beta}) d\pi(\boldsymbol{\beta}^{'}) \qquad \text{(A.34)} \\ &\le 2^{n_*} \left[\theta^2(QA) + \theta^2(-QA)\right]^{n_*} 2^{n^*} \iint f(\mathbf{X}_1, \boldsymbol{\beta}, \boldsymbol{\beta}^{'}) d\pi(\boldsymbol{\beta}) d\pi(\boldsymbol{\beta}^{'}). \end{aligned}$$

Now, using $\theta(A) = \frac{1}{2} + \Delta$ we have that $A^2 \ll \frac{\sqrt{p}}{kr^*}$ implies $\Delta^2 \ll \frac{\sqrt{p}}{kr^*}$ since $\theta \in BC^2(0)$. Following the exact arguments as in the proof of lower bound Theorem 6.5, one has

$$\begin{aligned} 2^{n^*} \iint f(\mathbf{X}_1, \boldsymbol{\beta}, \boldsymbol{\beta}^{'}) d\pi(\boldsymbol{\beta}) d\pi(\boldsymbol{\beta}^{'}) &= \mathbb{E}_0 \left[\left(\frac{1+4\Delta^2}{1-4\Delta^2}\right)^{\sum_{j \in m_3} r_j} (1-4\Delta^2)^{\sum_{j \in m_1 \cap m_2} r_j}\right] \\ &= \mathbb{E}_0 \left[\prod_{\in m_1 \cap m_2} \left(\frac{1+4\Delta^2}{1-4\Delta^2}\right)^{r_j \mathcal{I}(j \in m_3)} (1-4\Delta^2)^{r_j}\right] \\ &= \mathbb{E}_0 \left[\prod_{j \in m_1 \cap m_2} \frac{1}{2}\left((1+4\Delta^2)^{r_j} + (1-4\Delta^2)^{r_j}\right)\right] \\ &\le \mathbb{E}_0 \left[\left(\frac{1}{2}\right)^{|m_1 \cap m_2|} \left((1+4\Delta^2)^{r^*} + (1-4\Delta^2)^{r^*}\right)^{|m_1 \cap m_2|}\right] \end{aligned}$$

where the second to the last line follows since given $j \in m_1 \cap m_2$, $\mathcal{I}(j \in m_3) \sim \text{Bernoulli}(\frac{1}{2})$, independent for all j and the last line follows from noting that for any $\lambda \in (0,1)$, $(1+\lambda)^x + (1-\lambda)^x$ is an increasing function of $x \geq 1$. Hence following the same argument as in Theorem 6.5 after equation A.3, we have that there exists a constant $C > 0$ such that

$$2^{n^*} \iint f(\mathbf{X}_1, \boldsymbol{\beta}, \boldsymbol{\beta}^{'}) d\pi(\boldsymbol{\beta}) d\pi(\boldsymbol{\beta}^{'}) \leq \left(1 + C\frac{\frac{k^2 r^{*2} \Delta^4}{p}}{k}\right)^k \to 0 \quad \text{(A.35)}$$

since $\Delta \ll \sqrt{\frac{\sqrt{p}}{kr^*}}$. Hence, by A.34 and A.35 we have $\mathbb{E}_0(L_\pi^2) = 1 + o(1)$ if $2^{n_*}\left[\theta^2(QA) + \theta^2(-QA)\right]^{n_*} = 1 + o(1)$.

However, since $\theta \in BC^2(0)$, there exists a constant $C_1 > 0$ such that $\theta^2(QA) + \theta^2(-QA) \leq \frac{1}{2}(1 + C_1 Q^2 A^2)$. Hence

$$2^{n_*}\left[\theta^2(QA) + \theta^2(-QA)\right]^{n_*} = O\left(\left(1 + C_1\frac{Q^2 A^2 n_*}{n_*}\right)^{n_*}\right)$$

Now by assumption, $A^2 \ll \frac{\sqrt{p}}{kr^*}$ and $\frac{Q^2 n_*}{r^*} \ll p^{\frac{1}{2}-\alpha} = \frac{k}{\sqrt{p}}$, one has that $C_1 Q^2 A^2 n_* \to 0$ and hence $2^{n_*}\left[\theta^2(QA) + \theta^2(-QA)\right]^{n_*} = 1 + o(1)$ as required. This completes the proof of the lower bound. □

Part 2 : Proof of Upper Bound. We begin by noting that when $n_* = 0$, then the proof follows along the same lines as the power analysis argument of GLRT in Theorem 6.5 by using the fact $r_j \geq r_* \gg \log(p)$ for all $j = 1, \ldots, p$. The proof then immediately follows by noting that the definition of the GLRT does not depend on n_* and solely depends on the observations corresponding to indices in Ω^*, *i.e*, on $(y_i, \mathbf{x}_i^t)_{i \in \Omega^*}$ and does not even consider the data corresponding to $\mathbf{X}_2$.

PROOF OF THEOREM 7.2. The proof follows from Theorem 3.2 and is omitted. □

PROOF OF THEOREM 7.4. We divide the proof into the proof of lower bound and upper bound respectively.

Part 1 : Proof of Lower Bound. Define the intervals for $j = 1, \ldots, p$:

$$H_{p,j} = \left(\frac{r_j}{2} - \sqrt{2\log(p)}\sqrt{\frac{r_j}{4}}, \frac{r_j}{2} + \sqrt{2\log(p)}\sqrt{\frac{r_j}{4}}\right). \quad \text{(A.36)}$$

and put

$$D = \{Z_j \in H_{p,j},\ j = 1, \ldots, p\},\ Z_j = \sum_{i \in \Omega_j} y_i,\ l = 1, \ldots, p. \quad \text{(A.37)}$$

By Hölder's inequality it can be shown that for proving a lower bound it suffices to prove,

$$\mathbb{E}_0(L_\pi \mathcal{I}_{D^c}) = o(1), \qquad \mathbb{E}_0(L_\pi^2 \mathcal{I}_D) = 1 + o(1)\ . \quad \text{(A.38)}$$

We first prove the first equality of (A.38). Since $\{y_i, i \notin \Omega^*\}$ is independent of $\{Z_j, j = 1, \ldots, p\}$, we have by a calculation similar to proof of Theorem 6.8,

$$\begin{aligned}
\mathbb{E}_0(L_\pi \mathcal{I}_{D^c}) \le & \binom{p}{k}^{-1} 2^{-k} \sum_{m_1, \xi_1} \Big[\sum_{j \in m_1^1} \mathbb{E}_0 \left(2^{r_j} \{ \frac{\theta(A)}{\theta(-A)} \}^{Z_j} \{\theta(-A)\}^r \mathcal{I}(Z_j \in H_{p,j}^c) \right) \\
& + \sum_{j \in m_1^{-1}} \mathbb{E}_0 \left(2^{r_j} \{ \frac{\theta(A)}{\theta(-A)} \}^{r - Z_j} \{\theta(-A)\}^r \mathcal{I}(Z_j \in H_{p,j}^c) \right) \\
& + \sum_{j \in m_1^c} 2^{r_j} (\frac{1}{2})^{r_j} \mathbb{P}_0(Z_j \in H_{p,j}^c) \Big] \\
= & \binom{p}{k}^{-1} \sum_{m_1} \Big[\sum_{j \in m_1} (2\theta(-A))^{r_j} \mathbb{E}_0 \Big(\left(\frac{\theta(A)}{\theta(-A)} \right)^{Z_j} \mathcal{I}(Z_j \in H_{p,j}^c) \Big) + \sum_{j \in m_1^c} \mathbb{P}_0(Z_j \in H_{p,j}^c) \Big].
\end{aligned} \quad \text{(A.39)}$$

Now note that, by the same argument as proof of A.10, we have by an application of Lemma A.2,

$$\sum_{j \in m_1^c} \mathbb{P}_0(Z_j \in H_{p,j}^c) \le 2 \sum_{j \in m_1^c} \frac{e^{-\log(p)}}{\epsilon_j \sqrt{r_j}} e^{r_j \epsilon_j^2 - r_j \epsilon_j}$$

where $\epsilon_j = \frac{2\sqrt{\frac{r_j}{4}}\sqrt{2\log(p)} - 1}{r_j - 1}$. Hence there exists a constant $C > 0$ which does not depend on j such that $\epsilon_j \le C\sqrt{\frac{2\log p}{r_j}}$. Therefore, $r_j \epsilon_j^2 - r_j \epsilon_j \le C\sqrt{2\log p}(C\sqrt{2\log p} - r_j) \le C\sqrt{2\log p}(C\sqrt{2\log p} - r_*)$. Also,there exists a constant $c > 0$, not depending on j such that $\epsilon_j \sqrt{r_j} \ge c\sqrt{2\log p}$. Thus

$$\sum_{j \in m_1^c} \mathbb{P}_0(Z_j \in H_{p,j}^c) \le 2 \frac{C\sqrt{2\log p}(C\sqrt{2\log p} - r_*)(p - k)}{pc\sqrt{2\log p}} \to 0 \quad \text{(A.40)}$$

since $r_* \gg \log(p)$.

Next we control the term $\sum_{j\in m_1}(2\theta(-A))^{r_j}\mathbb{E}_0\Big(\Big(\frac{\theta(A)}{\theta(-A)}\Big)^{Z_j}\mathcal{I}(Z_j \in H^c_{p,j})\Big)$. To this end note that, by a proof similar to that of controlling A.11, one has using $r_* \gg \log(p)$ and $t < \rho^*_{\text{binary}}(\alpha)$ that there exists a sequence of real numbers $\lambda_p = o(1)$ which does not depend j such that $k(2\theta(-A))^{r_j}\mathbb{E}_0\Big(\Big(\frac{\theta(A)}{\theta(-A)}\Big)^{Z_j}\mathcal{I}(Z_j \in H^c_{p,j})\Big) \le \lambda_p$. In particular, this sequence can be taken to be polynomially p, as the proof of Theorem 6.8 suggests. This implies that

$$\begin{aligned}&\sum_{j\in m_1}(2\theta(-A))^{r_j}\mathbb{E}_0\Big(\Big(\frac{\theta(A)}{\theta(-A)}\Big)^{Z_j}\mathcal{I}(Z_j \in H^c_{p,j})\Big)\\&=\frac{1}{k}\sum_{j\in m_1}k(2\theta(-A))^{r_j}\mathbb{E}_0\Big(\Big(\frac{\theta(A)}{\theta(-A)}\Big)^{Z_j}\mathcal{I}(Z_j \in H^c_{p,j})\Big) \le \lambda_p.\end{aligned} \tag{A.41}$$

Hence, by A.39, we have using A.40 and A.41 that

$$\mathbb{E}_0(L_\pi\,\mathcal{I}_{D^c}) = o(1)$$

as required. This completes the proof of the first equality of (A.38). Next we prove the second claim of (A.38). Arguing similarly as in analysis of equation A.19 in proof of Theorem 6.8 and using Lemma A.4 we have that

$$\begin{aligned}&\mathbb{E}_0(L^2_\pi\mathcal{I}_D)\\&\le 2^{n_*}\left[\theta^2(QA)+\theta^2(-QA)\right]^{n_*}\\&\times\mathbb{E}_{m_1\cap m_2}\Big\{\prod_{j=1}^{p}\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^{r_j}\mathbb{P}_0(Z_j\in H_{p,j})\\&\quad+\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_j}\mathcal{I}(Z_j\in H_{p,j})\Big](2\theta(-A))^{2r_j}\Big)\Big]^{\mathcal{I}(j\in m_1\cap m_2)}\Big\}.\end{aligned}$$

As in proof of claim A.20 in Theorem 6.8, we have that for any $j = 1,\ldots,p$,

$$\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^{r_j}\mathbb{P}_0(Z_j\in H_{p,j})+\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_j}\mathcal{I}(Z_j\in H_{p,j})\Big](2\theta(-A))^{2r_j}\Big)\Big]\ge 1$$

since $r_* \gg \log(p)$ and $t < \rho^*_{\text{binary}}(\alpha)$. Let

$$j^* = \underset{j}{\text{argmax}}\Big\{\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^{r_j}\mathbb{P}_0(Z_j\in H_{p,j})+\mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z_j}\mathcal{I}(Z_j\in H_{p,j})\Big](2\theta(-A))^{2r_j}\Big)\Big\}$$

and let $r = r_{j^*},\ Z = Z_{j^*}, H_p = H_{p,j^*}$. Hence, one has with $U \sim Bin(k, \frac{k}{p-k})$ and $\varphi_{n,p} = 2^{n_*}\left[\theta^2(QA) + \theta^2(-QA)\right]^{n_*}$, that

$$
\begin{aligned}
&\mathbb{E}_0(L_\pi^2 \mathcal{I}_D) \\
&\leq \varphi_{n,p}\mathbb{E}_{m_1\cap m_2}\Big\{\prod_{j=1}^{p}\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z\in H_p) \\
&\qquad + \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z}\mathcal{I}(Z\in H_p)\Big](2\theta(-A))^{2r}\Big)\Big]^{\mathcal{I}(j\in m_1\cap m_2)}\Big\} \\
&= \varphi_{n,p}\mathbb{E}_{m_1\cap m_2}\Big\{\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z\in H_p) \\
&\qquad + \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z}\mathcal{I}(Z\in H_p)\Big](2\theta(-A))^{2r}\Big)\Big]^{|m_1\cap m_2|}\Big\} \\
&\leq \varphi_{n,p}\mathbb{E}_U\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z\in H_p) + \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z}\mathcal{I}(Z\in H_p)\Big](2\theta(-A))^{2r}\Big)\Big]^U \\
&= \varphi_{n,p}\Big\{1 + \frac{k}{p-k}\Big(\Big[\frac{1}{2}\Big(\{4\theta(A)\theta(-A)\}^r\mathbb{P}_0(Z\in H_p) \\
&\qquad + \mathbb{E}_0\Big[\Big\{\frac{\theta(A)}{\theta(-A)}\Big\}^{2Z}\mathcal{I}(Z\in H_p)\Big](2\theta(-A))^{2r}\Big)\Big] - 1\Big)\Big\}
\end{aligned}
\tag{A.42}
$$

where the second to the last line follows from Lemma A.1. Using the fact that $r = r_{j^*} \geq r_* \gg \log(p)$, one has by similar argument as in the proof of A.21 in Theorem 6.8 that

$$\mathbb{E}_0(L_\pi^2\mathcal{I}_D) = \varphi_{n,p}(1 + o(1))$$

when $t < \rho^*_{\text{binary}}(\alpha)$. Hence the verification of the second claim in (A.38) will be complete if we prove that $\varphi_{n,p} = 1 + o(1)$. Now, since $\theta \in BC^2(0)$, there exists a constant $C_1 > 0$ such that $\theta^2(QA) + \theta^2(-QA) \leq \frac{1}{2}(1 + C_1Q^2A^2)$. Hence

$$\varphi_{n,p} = 2^{n_*}\left[\theta^2(QA) + \theta^2(-QA)\right]^{n_*} = O\left(\left(1 + C_1\frac{Q^2A^2n_*}{n_*}\right)^{n_*}\right)$$

Now by assumption, $A^2 = \frac{2t\log(p)}{r^*}$ with $t < \rho^*_{\text{binary}}(\alpha)$ and $\frac{Q^2n_*}{r^*} \ll \log(p)$, one has that $C_1Q^2A^2n_* \to 0$ and hence $2^{n_*}\left[\theta^2(QA) + \theta^2(-QA)\right]^{n_*} = 1 + o(1)$. This justifies second equality of (A.38) and hence completes the proof of Theorem 7.4.

Part 2 : Proof of Upper Bound. We begin by noting that when $n_* = 0$, then the proof follows by very similar way as the power analysis argument of the Higher Criticism test in Theorem 6.10 by using the fact $r_j \geq r_* \gg \log(p)$ for all $j = 1, \ldots, p$ and hence we omit the details. The proof then immediately follows by noting that the definition of the Higher Criticism test does not depend on n_* and solely depends on the observations corresponding to indices in Ω^*, *i.e*, on $(y_i, \mathbf{x}_i^t)_{i \in \Omega^*}$ and does not even consider the data corresponding to $\mathbf{X}_2$. □

REFERENCES

Arias-Castro, E., Candès, E. J. and Plan, Y. (2011). Global testing under sparse alternatives: ANOVA, multiple comparisons and the higher criticism. *The Annals of Statistics* **39** 2533–2556.

Carter, A. and Pollard, D. (2004). Tusnádys inequality revisited. *The Annals of Statistics* **32** 2731–2741.

Hall, P. and Jin, J. (2010). Innovated higher criticism for detecting sparse signals in correlated noise. *The Annals of Statistics* 1686–1732.

Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent rv'-s, and the sample df. I. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **32** 111–131.