SUPPLEMENTAL MATERIALS

By Zhiyi Chi

In this note we collect proofs of the theoretical results stated in the manuscript. For convenience of reference, all the equations and statements that appear in the main text will be indexed as in there.

We will denote by Bern(p) the Bernoulli distribution with mean p, Bin(n, p) the binomial distribution for n iid trials with success rate p, Unif(a, b) the uniform distribution on (a, b), and Exp(c) the exponential distribution with mean c. The distribution of a random variable X is denoted by $\mathcal{L}(X)$. The total variation distance between two distributions μ and ν on $\mathbb{Z}$ is denoted $d_{\mathrm{TV}}(\mu, \nu) = \sum_k |\mu_k - \nu_k|$.

Given unadjusted marginal p-values $p_1, \ldots, p_n$, one for a different null hypothesis, let $p_{n:1} \le \cdots \le p_{n:n}$ be their order statistics. Set $p_{n:0} = 0$ and $p_{n:n+1} = 1$. Given target FDR control level $\alpha \in (0, 1)$, the BH procedure rejects all hypotheses with p-values $\le p_{n:R_n}$, where

$$R_n = \max\left\{j \ge 0 : p_{n:j} \le \frac{\alpha j}{n}\right\}. \tag{S-1}$$

As a result, $R = R_n$. The number of false rejections and power are

$$V_n = \#\left\{j \le n : p_j \le p_{n:R_n},\ \text{the } j\text{th null is true}\right\}, \quad \text{power}_n = \frac{R_n - V_n}{(n - N) \vee 1},$$

respectively.

The p-values are assumed to be sampled from a random effects model as follows. Let the population fraction of false nulls among all the nulls be a fixed $\pi \in (0, 1)$. Then (p_1, θ_1), (p_2, θ_2), … are iid, such that

$$\theta_j := \mathbf{1}\left\{\text{the } j\text{th null is false}\right\} \sim \text{Bern}(\pi)$$
$$\Pr\{p_j \le u \,|\, \theta_j = 0\} = u, \quad \Pr\{p_j \le u \,|\, \theta_j = 1\} = G(u), \quad u \in [0, 1].$$

Let $\eta_1, \eta_2, \ldots \overset{\text{iid}}{\sim} \text{Unif}(0, 1)$ be independent of $p_1, p_2, \ldots$. Then under the random effects model, $\theta_1, \theta_2, \ldots$ can be represented by

$$\theta_j = \mathbf{1}\left\{\eta_j > \rho(p_j)\right\}, \quad \text{where} \ \ \rho(x) = \frac{1 - \pi}{1 - \pi + \pi F'(x)}. \tag{S-2}$$

The following representation will be used repeatedly. Under the random effects model,

$$\xi_1 = F(p_1),\ \xi_2 = F(p_2), \ldots \overset{\text{iid}}{\sim} \text{Unif}(0, 1), \tag{S-3}$$

and $p_j = F^*(\xi_j)$, where $F^*(t) = \inf\{x : t \le F(x)\}$ (cf. [5], pp5–8). Given $n \ge 1$, let $\xi_{n:1} < \ldots < \xi_{n:n}$ be the order statistics of $\xi_1, \ldots, \xi_n$ and $\xi_{n:0} = 0$

and $\xi_{n:n+1} = 1$. Then $p_{n:j} = F^*(\xi_{n:j})$. Because $F^*(u) \le v$ if and only if $u \le F(v)$ for any $u, v \in (0,1)$, therefore,

$$R_n = \max\{j \le n : \xi_{n:j} \le z_{n:j}\}, \tag{S-4}$$

where henceforth we denote

$$z_{n:j} = F(\alpha j/n).$$

Henceforth denote the uniform empirical distribution and its inverse by

$$\mathbb{U}_n(x) = \frac{1}{n}\sum_{j=1}^n \mathbf{1}\{\xi_j \le x\}, \quad \mathbb{U}_n^*(t) = \inf\{x : t \le \mathbb{U}_n(x)\}.$$

We will also use Poisson representations for $\xi_1, \xi_2, \ldots$. For this reason, $\gamma_1, \gamma_2, \ldots$ will always denote a sequence of random variables $\stackrel{\text{iid}}{\sim}$ Exp(1) independent of $p_1, p_2, \ldots$.

Henceforth for $x > 1$, denote $\log_2 x = \log\log x$,

$$\mathbb{L}(x) = \sqrt{x\log_2 x}, \quad \mathbb{D}(x) = \mathbb{L}(x)/x = \sqrt{\log_2 x/x}.$$

1. Proof of Theorem 2.1.

Recall the theorem is stated as follows.

THEOREM 2.1. *Suppose $\alpha \in (0, \alpha_*)$. Let $c = \alpha F'(0)$. Then, as $n \to \infty$,*

$$R_n \xrightarrow{\text{d}} \tau := \max\{j : S_j < 0\},$$

the last time of excursion into $(-\infty, 0)$ by the random walk $S_0 = 0$, $S_j = S_{j-1} + \gamma_j - c$, $j \ge 1$, with $\gamma_1, \gamma_2, \ldots$ iid $\sim$ Exp(1) with density e^{-x}, $x > 0$. The distribution of τ is

$$\Pr\{\tau = k\} = \frac{k^k}{k!}(1-c)c^k e^{-kc}, \quad k = 0, 1, \ldots. \tag{2.5}$$

Consequently, the power of the BH procedure is of order $O_p(1/n)$. Furthermore,

$$pFDR \to \beta_* \tag{2.6}$$

and

$$\sum_{k=1}^{\infty} d_{\text{TV}}\left(\mathcal{L}(V_n \,|\, R_n = k),\ \text{Bin}(k, \beta_*)\right)\Pr\{R_n = k\} \to 0, \tag{2.7}$$

where $d_{\text{TV}}(\mu, \nu) := \sum_k |\mu_k - \nu_k|$ denotes the total variation distance of two distributions μ and ν on $\mathbb{Z}$, and $\mathcal{L}(V_n \,|\, R_n = k)$ the conditribution distribution of V_n.

1.1. *Proof of Eq. (2.5).* From the conditions of Theorem 2.1, $z_{n:j} = F(\alpha j/n) < j/n$. Recall $\xi_j = F(p_j)$. In order to prove (2.5), we need the following two lemmas.

LEMMA S-1. *There is $M = M(\alpha/\alpha_*) > 0$, such that $\sum_n \Pr\{R_n \ge M\log n\} < \infty$ and hence a.s., for $n \gg 1$, $R_n \le M\log n$.*

LEMMA S-2. *For $n \ge 1$, define*

$$\zeta_{n:j} = (\xi_{n:j} - \xi_{n:j-1})(\gamma_1 + \cdots + \gamma_{n+1}), \quad j = 1, \ldots, n.$$

Then $\zeta_{n:1}, \ldots, \zeta_{n:n+1} \overset{\text{iid}}{\sim} \mathrm{Exp}(1)$, and $\xi_{n:j} = \sum_{i=1}^{j} \zeta_{n:i} / \sum_{i=1}^{n+1} \zeta_{n:i}$.

PROOF OF THEOREM 2.1. Fix $k \ge 1$. Given $M > 0$ as in Lemma S-1, let $a_n = \lfloor M\log n\rfloor$. From the definition of R_n and Lemma S-2,

$$\{R_n < k\} = \bigcap_{j=k}^{n} \{\xi_{n:j} > z_{n:j}\} = \underbrace{\bigcap_{j=k}^{a_n} \left\{\frac{\zeta_{n:1} + \cdots + \zeta_{n:j}}{\gamma_1 + \cdots + \gamma_n} > z_{n:j}\right\}}_{A_n} \bigcap B_n,$$

where $B_n = \{R_n \le a_n\}$. Then by Lemma S-1, $|\Pr\{R_n \le k\} - \Pr\{A_n\}| \le \Pr\{B_n^c\} \to 0$.

Let

$$H_n = \{\gamma_1 + \cdots + \gamma_n < n - 2\mathbb{L}(n)\}$$
$$\Gamma_{n:j} = \left\{\gamma_1 + \cdots + \gamma_j > \alpha F'(0)j - n^{-1/3}\right\}$$

for $n \ge 1$, $1 \le j \le n$. Let $C > 0$ such that $|F(\alpha t) - \alpha F'(0)t| \le Ct^2$ for $0 \le t \ll 1$. Then for $n \gg 1$ and $j \le a_n$, we have

$$z_{n:j} > \alpha F'(0)j/n - C(j/n)^2,$$

and hence

$$\begin{aligned}
&\Pr\{A_n \cap H_n^c\} \\
&\le \Pr\left\{\bigcap_{j=k}^{a_n}\left\{\zeta_{n:1} + \cdots + \zeta_{n:j} \ge [1 - 2\mathbb{D}(n)]\left(\alpha F'(0)j - \frac{Cj^2}{n}\right)\right\}\right\} \\
&\le \Pr\left\{\bigcap_{j=k}^{a_n} \Gamma_{n:j}\right\},
\end{aligned}$$

with the second inequality due to Lemma S-2. Let

$$E_{n:j} = \{\gamma_1 + \cdots + \gamma_j > \alpha F'(0)j\} \subset \Gamma_{n:j}$$
$$x_n = \sum_{j=k}^{a_n} \Pr\left\{\alpha F'(0)j - n^{-1/3} < \gamma_1 + \cdots + \gamma_j \le \alpha F'(0)j\right\}.$$

Then $E_{n:j} \subset \Gamma_{n:j}$ and

$$0 \le \Pr\left\{\bigcap_{j=k}^{a_n} \Gamma_{n:j}\right\} - \Pr\left\{\bigcap_{j=k}^{a_n} E_{n:j}\right\} \le x_n.$$

Since the density of $\gamma_1 + \cdots + \gamma_j$ is bounded by 1, $x_n \le a_n n^{-1/3} \to 0$. Following the proof for Lemma S-1, it can be seen that

$$\sum_{n=1}^{\infty} \sum_{j=a_n+1}^{n} \Pr\{E_{n:j}^c\} < \infty.$$

By the Law of Iterated Logarithm (LIL), $\Pr\{H_n\} \to 0$ as $n \to \infty$. Combining these results,

$$\begin{aligned}
\varlimsup_{n\to\infty} \Pr\{R_n < k\} &= \varlimsup_{n\to\infty} \Pr\{A_n\} = \varlimsup_{n\to\infty} \Pr\{A_n \cap H_n^c\} \\
&\le \varlimsup_{n\to\infty} \Pr\left\{\bigcap_{j=k}^{a_n} E_{n:j}\right\} = \varlimsup_{n\to\infty} \Pr\left\{\bigcap_{j=k}^{\infty} E_{n:j}\right\} \\
&= \Pr\left\{\gamma_1 + \cdots + \gamma_j > \alpha F'(0) j, \text{ for all } j \ge k\right\} \\
&= \Pr\{\tau < k\}.
\end{aligned}$$

Similarly, $\varliminf_n \Pr\{R_n < k\} \ge \Pr\{\tau < k\}$ and hence $\Pr\{R_n < k\} \to \Pr\{\tau < k\}$. □

Proof of Lemma S-1. By definition, for $i \le n$,

$$\{R_n \ge i\} = \bigcup_{j=i}^{n} \{\xi_{n:j} \le z_{n:j}\}.$$

Let $q_{n:j} = \Pr\{\xi_{n:j} \le z_{n:j}\}$ and $Z_1, \ldots, Z_n \overset{\text{iid}}{\sim} \text{Bern}(z_{n:j})$. Then $q_{n:j} = \Pr\left\{\sum_{i=1}^{n} Z_i \ge j\right\}$. Since $j - z_{n:j} n > 0$, by Hoeffding's inequality (cf. Appendix),

$$q_{n:j} \le \exp\left\{-2n(j/n - z_{n:j})^2\right\}.$$

Fix $\epsilon, \delta > 0$, such that $a := \alpha/\alpha_* + \delta \in (0,1)$, $b := \inf_{t\in(0,\epsilon)} F(t)/t > 0$, and for $t \in (0, \epsilon)$, $F(t) < (a/\alpha)t$. By assumption, $d := \inf\{t - F(\alpha t) : t \ge \epsilon\} > 0$. Therefore, for $j > \epsilon n$, $q_{n:j} \le \exp\{-2nd^2\}$, yielding

$$\begin{aligned}
\sum_{\epsilon n \le i \le n} \Pr\{R_n \ge i\} &\le \sum_{\epsilon n \le i \le n} \sum_{j=i}^{n} \Pr\{\xi_{n:j} \ge z_{n:j}\} \\
&\le \sum_{\epsilon n \le i \le n} n e^{-2d^2 n} \to 0, \quad n \to \infty. \qquad \text{(S-1)}
\end{aligned}$$

For $j \le \epsilon n$, $c = F(\alpha j/n) \in [b(\alpha j/n),\ aj/n]$, so $\alpha b j \le cn$ and $j - cn \ge (1/a-1)cn > 0$. Let $\varphi(t) = (1+t)\log(1+t) - t$. By Bennett's inequality (cf. Appendix),

$$q_{n:j} \le \exp\left\{-\frac{cn}{1-c}\,\varphi\left(\frac{j-cn}{cn}\right)\right\} \le e^{-\rho j}, \text{ with } \rho = \alpha b \varphi(1/a-1) > 0.$$

Therefore, for any $i \le n$,

$$\begin{aligned}\Pr\{R_n \ge i\} &= \Pr\left\{\bigcup_{j=i}^{n}\{\xi_{n:j} \le z_{n:j}\}\right\} \\ &\le \sum_{j=i}^{\lfloor \epsilon n\rfloor} \Pr\{\xi_{n:j} \le z_{n:j}\} + \sum_{j=\lfloor \epsilon n\rfloor+1}^{n} \Pr\{\xi_{n:j} \le z_{n:j}\} \\ &\le (1-e^{-\rho})^{-1} e^{-\rho i} + n e^{-2nd^2}.\end{aligned}$$

Then $\sum_n \Pr\{R_n \ge 2\log n/\rho\} < \infty$. By Borel-Cantelli lemma,

$$R_n < 2\log n/\rho, \quad \text{a.s.}$$

The proof is complete by letting $M = 1/\rho$. □

Proof of Lemma S-2. The joint density of $X_j = \xi_{n:j}$, $1 \le j \le n$ and $Y = \sum_{j=1}^{n+1} \gamma_j$ is $G(x_1,\ldots,x_n,y) = \mathbf{1}\{0 < x_1 < \ldots < x_n < 1\}\, y^n e^{-y}$. Because $(\zeta_{n:1},\ldots,\zeta_{n:n},\zeta_{n:n+1}) = \phi(X_1,\ldots,X_n,Y)$, where

$$\phi(x_1,\ldots,x_n,y) = (x_1 y,\ (x_2-x_1)y,\ \ldots,\ (x_n - x_{n-1})y,\ (1-x_n)y),$$

$\zeta_{n:1},\ldots,\zeta_{n:n+1}$ have joint density $e^{-(z_1+\cdots+z_{n+1})}$ and therefore are $\stackrel{\text{iid}}{\sim}$ Exp(1). □

1.2. *The other statements of Theorem 2.1.* We only prove (2.7). Eq. (2.6) follows immediately. Let $\theta_{n:1}, \ldots, \theta_{n:n}$ be the re-ordered $\theta_1,\ldots,\theta_n$ corresponding to $p_{n:j}$. Then by (S-2),

$$V_n = \sum_{j=1}^{R_n} (1-\theta_{n:j}) \stackrel{\text{d}}{=} \sum_{j=1}^{R_n} \mathbf{1}\{\eta_j < \rho(p_{n:j})\}.$$

Given $k \ge 1$, almost surely, as $n \to \infty$, $p_{n:s} \to 0$ and hence $\rho(p_{n:i}) \to \beta_*$ for all $i \le k$. Then

$$\begin{aligned}&(\mathbf{1}\{U_1 < \rho(p_{n:1})\},\ldots,\mathbf{1}\{U_k < \rho(p_{n:k})\}) \\ &\xrightarrow{\text{a.s.}} (\mathbf{1}\{U_1 < \beta_*\},\ldots,\mathbf{1}\{U_k < \beta_*\})\end{aligned}$$

and hence by η's and R_n being independent,

$$d_{\text{TV}}(\mathcal{L}(V_n \,|\, R_n = k),\ \text{Bern}(k,\beta_*)) \to 0.$$

Since $R_n \xrightarrow{\text{d}} \tau$, by dominant convergence, (2.7) is proved.

2. Proof of Theorem 2.3. We shall prove Theorem 2.3 before Theorem 2.2, since some of the ideas for the proof of the former can be applied to the latter. Recall

THEOREM 2.3. *Suppose F is twice differentiable at 0 and $F'(0)u > F(u)$ for $u > 0$. Then $\alpha_* = 1/F'(0)$. Suppose $I = \left\{i > 1 : F^{(i)}(0) \neq 0\right\} \neq \emptyset$. Let $\ell = \min I$. If $\alpha = \alpha_*$, then*

$$\lim_{n\to\infty} \frac{\log R_n}{\log n} = \nu_0 := \frac{2\ell - 2}{2\ell - 1} \quad a.s. \tag{2.10}$$

The upper bound of R_n can be strengthened to

$$\varlimsup_{n\to\infty} \frac{R_n}{n^{\nu_0}(\log n)^{1-\nu_0}} \le \left\{ \frac{\ell! \sqrt{2(1+\nu_0)}\, F'(0)^\ell}{|F^{(\ell)}(0)|} \right\}^{2(1-\nu_0)} \qquad a.s. \tag{2.11}$$

Furthermore, a.s., $V_n/R_n \to \beta_*$ and true discoveries $R_n - V_n \sim (1-\beta_*)R_n$.

2.1. *Upper limit.* The proof of Eq. (2.11) relies on Lemma S-2 and the following two Lemmas.

LEMMA S-1. *Let $m_n \in \mathbb{N}$ such that $m_n \to \infty$ as $n \to \infty$. Given $c > 1$ and $\delta > 0$,*

$$\Pr\left\{ \bigcup_{j=m_n}^{n} \left(\gamma_1 + \cdots + \gamma_j \le j - \sqrt{2c(\delta+1) j \log j} \right) \right\}$$
$$\le m_n^{-\delta}/\delta, \quad \textit{for all } n \gg 1.$$

LEMMA S-2. *If $\alpha = 1/F'(0)$, then there is constant $M > 0$, such that*

$$\varlimsup_{n\to\infty} \frac{R_n}{n^{\nu_0}(\log n)^{1-\nu_0}} \le M \quad a.s.$$

PROOF OF THE UPPER LIMIT EQ. (2.11). Denote

$$A_0 := \left\{ \ell! \sqrt{4 - \frac{2}{2\ell-1}} \frac{F'(0)^\ell}{|F^{(\ell)}(0)|} \right\}^{\frac{2}{2\ell-1}}$$

and $Q_n = n^{\nu_0}(\log n)^{1-\nu_0}$. It suffices to show that for any $A > A_0$,

$$\Pr\{\xi_{n:j} > z_{n:j}, \text{ for all } j \ge AQ_n \text{ and } n \gg 1\} = 1. \tag{S-1}$$

Denote $a = (4\ell-3)/(\ell-1) > 2$. Given $\epsilon > 0$, let $c \in (1, 1+\frac{\epsilon}{a})$ and $\delta = \frac{a+\epsilon}{2c} - 1$. Then $\nu_0\delta > 1$. Let $\zeta_{n:1}, \ldots, \zeta_{n:n}$ be defined as in Lemma S-2 and $m_n = \lceil An^{\nu_0} \rceil$. By Lemma S-1,

$$\Pr\left\{ \bigcup_{j=\lceil An^{\nu_0}\rceil}^{n} \left\{ \zeta_{n:1} + \cdots + \zeta_{n:j} < j - \sqrt{(a+\epsilon) j \log j} \right\} \right\} \le \frac{A^{-\delta} n^{-\nu_0\delta}}{\delta}$$

have a finite sum over n. Therefore, by Borel-Cantelli lemma, a.s.,

$$\zeta_{n:1} + \ldots + \zeta_{n:j} \ge j - \sqrt{(a+\epsilon) j \log j}, \quad \text{for all } j \ge A n^{\nu_0} \text{ and } n \gg 1. \tag{S-2}$$

Let $E_i = \{R_i < M i^{\nu_0} (\log i)^{1-\nu_0}\}$. By Lemma S-2, for $M \gg 0$,

$$\lim_{n\to\infty} \Pr\left\{ \bigcap_{i \ge n} E_i \right\} = 1.$$

On $\bigcap_{i\ge n} E_i$, if $\xi_{n:j} \le z_{n:j}$, then $j \le R_n = o(n)$, so by Taylor's expansion,

$$\xi_{n:j} = \frac{\zeta_{n:1} + \cdots + \zeta_{n:j}}{\gamma_1 + \cdots + \gamma_{n+1}} \le z_{n:j} \le \frac{j}{n} - \frac{(1-\epsilon)}{\ell!} \frac{|F^{(\ell)}(0)|}{F'(0)^\ell} \left(\frac{j}{n}\right)^\ell.$$

On the other hand, by the LIL, almost surely, $\gamma_1 + \cdots + \gamma_n \le n + \sqrt{2+\epsilon}\,\mathbb{L}(n)$ for $n \gg 1$. Together with (S-2), the inequalities lead to

$$\begin{aligned} &1 - \sqrt{\frac{(a+\epsilon)\log j}{j}} \\ &\le \left[1 - \frac{(1-\epsilon)}{\ell!} \frac{|F^{(\ell)}(0)|}{F'(0)^\ell} \left(\frac{j}{n}\right)^{\ell-1}\right] \left(1 + \sqrt{2+\epsilon}\,\mathbb{D}(n)\right). \end{aligned} \tag{S-3}$$

Assume that for infinitely many n, there exists $j \ge A Q_n$ such that $\xi_{n:j} \le F(\alpha j/n)$. Then by $\sqrt{\log n/n} = o(j/n)$ and (S-3), for $n \gg 1$,

$$\begin{aligned} &\sqrt{\frac{(a+\epsilon)\log j}{j}} \ge \frac{(1-\epsilon)^2}{\ell!} \frac{|F^{(\ell)}(0)|}{F'(0)^\ell} \left(\frac{j}{n}\right)^{\ell-1} \\ \Longrightarrow \quad &\frac{j^{\ell-1/2}}{\sqrt{\log j}} \le \frac{\ell! \sqrt{a+\epsilon}\, F'(0)^\ell n^{\ell-1}}{(1-\epsilon)^2 |F^{(\ell)}(0)|}. \end{aligned}$$

On the other hand, there are $r_n \downarrow 0$ such that

$$\frac{j^{\ell-1/2}}{\sqrt{\log j}} \ge \frac{(1-r_n) A^{\ell-1/2} n^{\ell-1}}{\sqrt{(2\ell-2)/(2\ell-1)}}$$

Combine the above two inequalities to get

$$A^{\ell-1/2} \le \ell! \sqrt{\frac{2\ell-2}{2\ell-1}} \frac{\sqrt{a+\epsilon}\, F'(0)^\ell}{|F^{(\ell)}(0)|} \frac{1}{(1-\epsilon)^2 (1-r_n)}.$$

Let $n\to\infty$ and then $\epsilon \to 0$. Then $A^{\ell-1/2} \le A_0^{\ell-1/2}$ and hence $A \le A_0$. The contradiction shows that (S-1) holds almost surely. This then finishes the proof. □

Proof. Proof of Lemma S-1 By $\log E[e^{\theta(1-\gamma_1)}] = \theta - \log(1+\theta) = \frac{1}{2}\theta^2 - \frac{1}{3}\theta^3 + \cdots$, fix $\theta_0 > 0$, such that $\Lambda(\theta) < \frac{1}{2}c\theta^2$ for $\theta \in (0, \theta_0)$. Let $g_j(\theta) = \frac{1}{2}c\theta^2 - \theta\sqrt{2c(1+\delta)\log j/j}$. Then by Chernoff's inequality, for $j \ge 1$ and $\theta \in (0, \theta_0)$,

$$\lambda_j := \Pr\left\{\gamma_1 + \cdots + \gamma_j - j \le -\sqrt{2c(1+\delta)j\log j}\right\} \le \exp\{jg_j(\theta)\}.$$

For $n \gg 1$ and $j = m_n, \ldots, n$, $\hat\theta = \sqrt{2(1+\delta)c^{-1}\log j\,/\,j} < \theta_0$, $g_j(\hat\theta) = -(1+\delta)\log j/j$, and so $\lambda_j \le j^{-(1+\delta)}$. The proof is then complete by the inequality for union of events. □

Proof of Lemma S-2. Following Lemma S-1, we estimate

$$q_{n:j} = \Pr\{\xi_{n:j} \le z_{n:j}\}.$$

Given $\epsilon \in (0,1)$, again, (S-1) holds, so almost surely, $R_n < \epsilon n$ for all $n \gg 1$. On the other hand, by $\alpha = 1/F'(0)$, if $\epsilon \ll 1$, then $0 < \frac{1}{2}\kappa t^\ell < t - F(\alpha t) < 2\kappa t^\ell$ for all $t \in [0, \epsilon]$, where $\kappa = \frac{|F^{(\ell)}(0)|}{\ell! F'(0)}$. It follows that there is a constant $\delta > 0$, such that

$$\frac{n z_{n:j}}{1 - z_{n:j}}\,\varphi\left(\frac{j - z_{n:j}n}{z_{n:j}n}\right) \ge \frac{\delta j^{2\ell-1}}{n^{2\ell-2}}, \quad \text{for all } j \le \epsilon n \text{ and } n \gg 1.$$

Therefore, if $M > (2/\delta)^{1-\nu_0}$ and $a_n = M n^{\nu_0}(\log n)^{1-\nu_0}$, then by Bennett's inequality,

$$\Pr\{a_n \le R_n \le \epsilon n\} \le n \exp\left\{-\frac{\delta a_n^{2\ell-1}}{n^{2\ell-2}}\right\} = n^{-(\delta M^{2\ell-1}-1)},$$

which have a finite sum (cf. Appendix). The proof is then complete by Borel-Cantelli Lemma. □

2.2. *Lower limit.* The proof of the lower bound is based on two lemmas. Given $\eta > 1$, denote $M_j = M_j(\eta) = \lfloor \eta^j \rfloor$. Given $j \le m < n$, let $R(n, m, j) = \#\{i \le n : \xi_i \le \xi_{m:j}\}$. Clearly $\xi_{m:j} = \xi_{n:R(n,m,j)}$.

Lemma S-3. *Fix $0 < \epsilon \ll 1$ and $0 < a < b < 1$. Then almost surely, for $m \gg 1$, there exists $j \in [am, bm] \cap \mathbb{N}$ such that*

$$\xi_{M_m:M_j} < [\,1 - \sqrt{2-\epsilon}\,\mathbb{D}(\eta^j)\,]\eta^{j-m}.$$

Lemma S-4. *Given $0 < a < b < 1$, almost surely, for all $m \gg 1$, $j \in [am, bm] \cap \mathbb{N}$, and $M_{m+1} \le n \le M_{m+2}$, we have*

$$\left|\frac{1}{n}R(n, M_m, M_j) - \eta^{j-m}\right| \le C_\eta\,\mathbb{D}(\eta^j)\eta^{j-m},$$

where $C_\eta = 11\eta(\eta+1)\sqrt{\eta - 1}$.

Proof of the lower limit for Eq. (2.10). Assume the above two lemmas for now. To show $\underline{\lim}_n \frac{\log R_n}{\log n} \ge \nu_0$, it suffices to show that given $\nu \in (0, \nu_0)$, almost surely, for $n \gg 1$, $R_n > n^\nu$.

Fix $\lambda \in (\nu, \nu_0)$, $\eta \in (1, 2)$, and $\epsilon \in (0, 1-\eta^{\nu-1})$, such that $2\ell\lambda - \nu < 2\ell - 2$ and $C_\eta < \frac{1}{2}$. By Lemma S-3, a.s., for each m large enough, there exists $J \in [\nu m, \lambda m] \cap \mathbb{N}$, such that

$$\xi_{M_m:M_J} < [\,1 - \mathbb{D}(\eta^J)\,]\eta^{J-m}. \tag{S-4}$$

For $n \gg 1$, if $M_{m+1} \le n \le M_{m+2}$, then by Lemma S-4,

$$\left|\frac{1}{n} R(n, M_m, M_J) - \eta^{J-m}\right| \le \frac{1}{2}\mathbb{D}(\eta^J)\eta^{J-m}. \tag{S-5}$$

By $J \ge \nu m$ and $n \ge \eta^{m+1}$, (S-5) implies $R(n, M_m, M_J) \ge n\eta^{J-m}(1-\epsilon) \ge n^\nu$. On the other hand, by $J \le \lambda m$ and $n \le \eta^{m+2}$, (S-5) also implies

$$R(n, M_m, M_J) \le n\eta^{J-m}(1+\epsilon) \le \eta^{2(1-\lambda)} n^\lambda (1+\epsilon). \tag{S-6}$$

It only remains to show that almost surely, for $n \gg 1$,

$$\xi_{n:R(n,M_m,M_J)} < z_{n:R(n,M_m,M_J)}.$$

If this is shown, then $R_n \ge R(n, M_m, M_J) \ge n^\nu$ and the proof is complete. Assume that with a positive probability, $\xi_{n:R(n,M_m,M_J)} \ge z_{n:R(n,M_m,M_J)}$ for infinitely many n. Since $\xi_{n:R(n,M_m,M_J)} = \xi_{M_m:M_J}$ for $M_{m+1} \le n \le M_{m+2}$, by (S-4) – (S-6), almost surely, for $n \gg 1$,

$$\begin{aligned}
&[\,1 - \mathbb{D}(\eta^J)\,]\eta^{J-m} \ge z_{n:R(n,M_m,M_J)} \\
&\ge \frac{R(n, M_m, M_J)}{n} - \frac{2|F^{(\ell)}(0)|}{\ell! F'(0)^\ell}\left(\frac{R(n, M_m, M_j)}{n}\right)^\ell \\
&\ge [\,1 - \mathbb{D}(\eta^J)/2\,]\eta^{J-m} - \frac{2|F^{(\ell)}(0)|}{\ell! F'(0)^\ell}\left[\eta^{2(1-\lambda)}(1+\epsilon)\right]^\ell n^{\ell(\lambda-1)} \\
&\ge [\,1 - \mathbb{D}(\eta^J)/2\,]\eta^{J-m} - K\eta^{\ell(\lambda-1)m},
\end{aligned}$$

where $K > 0$ is a constant independent of m. Thus by $\nu m \le J \le \lambda m$, $\eta^{-m}\mathbb{L}(\eta^{\nu m}) \le 2K\eta^{\ell(\lambda-1)m}$. Since the inequality holds for infinitely many m, it follows that $\nu/2 - 1 \le \ell(\lambda - 1)$, giving $2\ell - 2 \le 2\ell\lambda - \nu < 2\ell - 2$. The contradiction finishes the proof. □

Our proof of Lemma S-3 shall be based on the Poisson representation in Lemma S-2. We need several inequalities for Gamma distributions (cf. [5], section 11.9).

Lemma S-5. *Given $c, r > 1$ and ϵ, $\nu \in (0, 1)$, we have*

$$\Pr\left\{\bigcup_{j=\lfloor n^\nu \rfloor}^{n} \left\{\gamma_1 + \cdots + \gamma_j \ge j + \sqrt{2cr}\,\mathbb{L}(j)\right\}\right\} \le \epsilon(\log n)^{1-r}, \tag{S-7}$$

all $n \gg 1$.

PROOF. Given $\eta \in (1, c)$, let

$$a_n = \max\{i : \eta^i \le n^\nu\}, \quad b_n = \min\{i : \eta^i > n\}.$$

Then $b_n/a_n \to 1/\nu$. Denote the left hand side of (S-7) by λ_n. Let $Z_j = \gamma_j - 1$. Then

$$\lambda_n \le \Pr\left\{ \bigcup_{s=a_n}^{b_n} \left\{ \max_{\eta^{s-1} \le j \le \eta^s} (Z_1 + \cdots + Z_j) \ge \sqrt{2cr}\mathbb{L}(\eta^{s-1}) \right\} \right\}$$
$$\le \sum_{s=a_n}^{b_n} \underbrace{\Pr\left\{ \max_{1 \le j \le \eta^s} (Z_1 + \cdots + Z_j) \ge \sqrt{2cr}\mathbb{L}(\eta^{s-1}) \right\}}_{q_s}.$$

Since $EZ_j = 0$, by martingale inequality (cf. [1], p256, Example)

$$q_s \le (Ee^{\theta Z_1})^{\eta^s} \exp\left\{-\theta\sqrt{2cr}\mathbb{L}(\eta^{s-1})\right\} = e^{\eta^s f_s(\theta)}, \quad \text{any } \theta \in (0,1). \tag{S-8}$$

where $f_s(\theta) = \Lambda(\theta) - \theta\sqrt{2cr}\mathbb{D}(\eta^{s-1})/\eta$ with $\Lambda(\theta) = \log(Ee^{\theta Z_1}) = -\log(1-\theta) - \theta$. By $\Lambda(0) = \Lambda'(0) = 0$, and $\Lambda''(0) = 1$, given $C \in (1, c/\eta)$, there is $0 < \theta_C \ll 1$, such that for $\theta \in (0, \theta_C)$, $\Lambda(\theta) \le C\theta^2/2$ and so $f_s(\theta) \le g_s(\theta) := \frac{1}{2}C\theta^2 - \theta\sqrt{2cr}\mathbb{D}(\eta^{s-1})/\eta$. Let $\theta_s = C^{-1}\sqrt{2cr}\mathbb{D}(\eta^{s-1})/\eta$. For $n \gg 1$, because $\theta_s < \theta_C$ for all $s = a_n, a_n + 1, \ldots, b_n$,

$$q_s \le e^{\eta^s f_s(\theta_s)} \le e^{\eta^s g_s(\theta_s)} = \exp\left\{-\frac{cr\log(s - 1 + \log\eta)}{C\eta}\right\} \le s^{-r}.$$

Let $\delta = r - 1$. Then

$$\lambda_n \le \sum_{s=a_n}^{b_n} s^{-r} \le a_n^{-\delta}/\delta \sim \frac{(\log\eta)^\delta}{\delta\nu^\delta}(\log n)^{-\delta}.$$

To finish the proof, choose η close to 1 so that $(\log\eta)^\delta/(\delta\nu^\delta) < \epsilon$. □

We need the following lemma on moderate deviation.

LEMMA S-6. *Let $X_1, X_2, \ldots$ be iid with $EX_1 = 0$, $\mathrm{Var}(X_1) = \sigma^2 > 0$, and $E|X_1|^q < \infty$ for any $q > 0$. Fix $c > 0$. Suppose $t_j \le c\sqrt{\log j}$ and $t_j \to \infty$, then*

$$\Pr\{X_1 + \cdots + X_j \ge \sigma t_j\sqrt{j}\} = (1 + o(1))\frac{e^{-t_j^2/2}}{\sqrt{2\pi}\,t_j}, \quad \text{as } j \to \infty.$$

PROOF. By Lemma A.3 in [3], $\Pr\{X_1 + \cdots + X_j \ge \sigma t_j\sqrt{j}\} = (1 + o(1))\bar\Phi(t_j)$, where $\bar\Phi(x) = \Pr\{N(0,1) \ge x\} \sim x^{-1}e^{-x^2/2}/\sqrt{2\pi}$, as $x \to \infty$. □

Let $X_i = 1 - \gamma_i$. Then $EX_i = 0$ and $\mathrm{Var}(X_i) = 1$. Then from Lemma S-6,

$$\Pr\{\gamma_1 + \cdots + \gamma_j \le j - t_j\sqrt{j}\} = (1+o(1))\frac{e^{-t_j^2/2}}{\sqrt{2\pi}\, t_j}\,, \tag{S-9}$$

Lemma S-7. *Let $\epsilon \in (0,1)$. If L is large enough, then for any $a, b, \delta \in (0,1)$ with $a < b$, almost surely, for $m \gg 1$ and $N_m := \left\lfloor L^{\delta m} \right\rfloor$, there exists $j \in [a\delta m, b\delta m] \cap \mathbb{N}$ such that*

$$\zeta_{N_m:1} + \cdots + \zeta_{N_m:\lfloor L^j \rfloor} \le L^j - \sqrt{2-\epsilon}\,\mathbb{L}(L^j)$$

Proof. For convenience, for $x, y \in \mathbb{R}_+$, denote by $\bigcup_{j=x}^{y}$, $\prod_{j=x}^{y}$, etc. operations over $j \in [x,y] \cap \mathbb{N}$. For $m \ge 1$, let

$$E_m = \bigcup_{j=L^{a\delta m}}^{L^m} \left\{\zeta_{N_m:1} + \ldots + \zeta_{N_m:j} \ge j + \sqrt{4+\epsilon}\,\mathbb{L}(j)\right\}.$$

Then by Lemma S-5, $\Pr\{E_m\} = O(m^{-\rho})$ for some $\rho > 1$. Choose $L \gg 1$ such that

$$\sqrt{2-\tfrac{\epsilon}{2}}\,\mathbb{L}(L^{m+1} - L^m) - \sqrt{4+\epsilon}\,\mathbb{L}(L^m) \ge \sqrt{2-\epsilon}\,\mathbb{L}(L^{m+1}).$$

For $m \ge 1$ and $j \le m$, let

$$\begin{aligned}
A_m &= \bigcap_{i=a\delta m}^{b\delta m} \left\{\zeta_{N_m:1} + \ldots + \zeta_{N_m:\lfloor L^i \rfloor} > L^i - \sqrt{2-\epsilon}\,\mathbb{L}(L^i)\right\} \\
\Delta_j &= L^{j+1} - L^j - \sqrt{2-\tfrac{\epsilon}{2}}\,\mathbb{L}(L^{j+1} - L^j) \\
G_{m:j} &= \left\{\zeta_{N_m:\lfloor L^j \rfloor+1} + \cdots + \zeta_{N_m:\lfloor L^{j+1} \rfloor} > \Delta_j\right\}, \\
G_m &= \bigcap_{i=a\delta m}^{b\delta m-1} G_{m:i}.
\end{aligned}$$

Then by the selection of L, $A_m \setminus E_m \subset G_m$. For fixed m, $G_{m:j}$, $j \ge 1$, are independent of each other. Therefore,

$$\begin{aligned}
\Pr\{A_m \setminus E_m\} &\le \prod_{j=a\delta m}^{b\delta m-1} \Pr\{G_{m:j}\} \\
&\le \prod_{j=a\delta m}^{b\delta m-1} \left(1 - \Pr\left\{\gamma_1 + \ldots + \gamma_{\lfloor L^{j+1} \rfloor - \lfloor L^j \rfloor} \le \Delta_j\right\}\right).
\end{aligned}$$

By (S-9), given $\eta \in (0, \frac{1}{4}\epsilon)$, there is a constant $C = C(L)$ such that for all $j \gg 1$,

$$\Pr\left\{\gamma_1 + \ldots + \gamma_{\lfloor L^{j+1} \rfloor - \lfloor L^j \rfloor} \le \Delta_j\right\} \ge C j^{-(1-\eta)}.$$

Then for some constant $C_2 = C_2(a, b, \delta, \eta, L)$,

$$\Pr\{A_m \setminus E_m\} \le \exp\left\{-C \sum_{j=a\delta m}^{b\delta m-1} j^{-1+\eta}\right\} \le \exp\left\{-C_2 m^{\eta}\right\}.$$

As a result, $\sum_m \Pr\{A_m \setminus E_m\} < \infty$. Together with $\Pr\{E_m\} = O(m^{-\rho})$, this yields $\sum \Pr\{A_m\} < \infty$. The lemma then follows according to Borel-Cantelli lemma. □

Proof of Lemma S-3. Given $0 < \epsilon_1 \ll \epsilon$, by Lemma S-7, there is $L \gg 1$, such that for any $\delta \in (0, 1)$, a.s., for $m \gg 1$, there exists $j \in [a\delta m, b\delta m] \cap \mathbb{N}$ with $\zeta_{N_m:1} + \cdots + \zeta_{N_m:\lfloor L^j \rfloor} \le L^j - \sqrt{2 - \epsilon_1}\, \mathbb{L}(L^j)$, where $N_j = \left\lfloor L^{\delta j} \right\rfloor$. In particular, let $\delta = \log_L \eta$. Then $N_j \equiv M_j$ and there is $j \in [am, bm] \cap \mathbb{N}$, such that $\zeta_{M_m:1} + \cdots + \zeta_{M_m:M_j} \le \eta^j - \sqrt{2 - \epsilon_1}\, \mathbb{L}(\eta^j)$. On the other hand, by the LIL, when $m \gg 1$, $|\gamma_1 + \cdots + \gamma_{M_m+1} - \eta^m| \le \sqrt{2 + \epsilon_1} \mathbb{L}(\eta^m)$. By the two inequalities and Lemma S-2,

$$\begin{aligned}\xi_{M_m:M_j} &\le \frac{\eta^j - \sqrt{2 - \epsilon_1}\, \mathbb{L}(\eta^j)}{\eta^m - \sqrt{2 + \epsilon_1}\, \mathbb{L}(\eta^m)} \\ &= [\,1 - \sqrt{2 - \epsilon_1}\, \mathbb{D}(\eta^j)\,]\eta^{j-m}\,(1 + o(1)).\end{aligned}$$

Thus, if $m \gg 1$, then for the above j, $\xi_{M_m:M_j} \le [1 - \sqrt{2 - \epsilon}\, \mathbb{D}(\eta^j)\,]\eta^{j-m}$. □

The proof of Lemma S-4 is based on two preliminary results.

Lemma S-8. *For any $0 < a < b < 1$, $\epsilon \in (0, 1)$, and $\eta > 1$, almost surely, for all $m \gg 1$ and $j \in [am, bm] \cap \mathbb{N}$, we have*

$$\left|\xi_{\lfloor \eta^m \rfloor : \lfloor \eta^j \rfloor} - \eta^{j-m}\right| \le 3\, \mathbb{D}(\eta^j)\eta^{j-m}.$$

Proof. By a result of Csörgő and Révész (cf. [5], Theorem 1, p 616), for $u_n = 9 \log_2 n/n$,

$$\varlimsup_{n\to\infty} \frac{n}{\mathbb{L}(n)} \sup_{t \in [u_n, 1-u_n]} \left|\frac{\mathbb{U}_n^*(t) - t}{\sqrt{t(1-t)}}\right| \le 2\sqrt{2}, \quad \text{a.s.}$$

For $j \in [am, bm] \cap \mathbb{N}$, let $t_j = \lfloor \eta^j \rfloor / \lfloor \eta^m \rfloor$. Then $\xi_{\lfloor \eta^m \rfloor : \lfloor \eta^j \rfloor} = \mathbb{U}^*(t_j)$. For $m \gg 1$, $t_j \in [u_{M_m}, 1 - u_{M_m}]$ for all $j \in [am, bm] \cap \mathbb{N}$ and $\log_2 \eta^m / \log_2 \eta^j \to 1$ uniformly. Therefore, by the above inequality, given $\epsilon < 3 - 2\sqrt{2}$, for $m \gg 1$,

$$\begin{aligned}\left|\xi_{\lfloor \eta^m \rfloor : \lfloor \eta^j \rfloor} - \eta^{j-m}\right| &\le (2\sqrt{2} + \epsilon)\mathbb{L}(\eta^m)\sqrt{\eta^{j-m}(1 - \eta^{j-m})}/\eta^m \\ &< 3\mathbb{D}(\eta^j)\eta^{j-m}.\end{aligned}$$

□

The proof of the next elementary result is omitted for brevity.

LEMMA S-9. *There is* $\theta_0 > 0$*, such that* $\Lambda(\theta, x) \le x\theta^2$ *for all* $\theta \in [-\theta_0, \theta_0]$ *and* $x \in [0, 1]$*, where* $\Lambda(\theta, x) = \log(1 - x + xe^\theta) - x\theta$.

LEMMA S-10. *For* $n \ge 1$*, let* $a_n \in (0, \frac{1}{2})$ *and* $f_n(x) > 0$ *with*

$$\sup_{x\in[a_n,1/2]} \frac{f_n(x)}{nx} \to 0.$$

Then for all $n \gg 1$ *and* $x \in [a_n, \frac{1}{2}]$*, we have*

$$\Pr\left\{\max_{1\le j\le n} j\,|\mathbb{U}_j(x) - x| \ge f_n(x)\right\} \le 2\exp\left\{-\frac{f_n^2(x)}{4nx}\right\}.$$

PROOF. For $x \in (0, 1)$, $Z_j := \mathbf{1}\{\xi_j \le x\} - x$ are iid with $EZ_1 = 0$ and $\Pr\{Z_j = 1-x\} = x = 1-\Pr\{Z_j = -x\}$. Since $j(\mathbb{U}_j(x)-x) = Z_1+\cdots+Z_j$, by martingale inequality (cf. [1], p256, Example 9), for all $\theta > 0$,

$$\begin{aligned}\lambda_n(x) := \Pr\left\{\max_{1\le j\le n} j(\mathbb{U}_j(x) - x) \ge f_n(x)\right\}\\ \le \exp\{n\Lambda(\theta, x) - f_n(x)\theta\}.\end{aligned}$$

By Lemma S-9, there is $\theta_0 > 0$, such that $\Lambda(\theta, x) \le x\theta^2$ for all $\theta \in [0, \theta_0]$ and $x \in [0, 1]$. Then $\lambda_n(x) \le \exp\{nx\theta^2 - f_n(x)\theta\}$. By the assumption, if $n \gg 1$, then $\frac{f_n(x)}{2nx} \in [0, \theta_0]$ for all $x \in [a_n, 1/2]$ and hence $\lambda_n(x) \le \exp\left\{-\frac{f_n^2(x)}{4nx}\right\}$. The inequality

$$\Pr\left\{\max_{1\le j\le n} j(x - \mathbb{U}_j(x)) \ge f_n(x)\right\} \le \exp\left\{-\frac{f_n^2(x)}{4nx}\right\},$$

can be similarly proved. Then the proof is complete. □

PROOF OF LEMMA S-4. Let

$$E_m = \left\{\frac{1}{2}\,\eta^{j-m} < \xi_{M_m:M_j} < 2\eta^{j-m},\ am \le j \le bm\right\}.$$

Denote $D_m = M_{m+2} - M_m$. For each $j \in [am, bm] \cap \mathbb{N}$, let

$$\begin{aligned}\Gamma_{m:j} = \bigcup_{n=M_{m+1}}^{M_{m+2}} \Big\{&|R(n, M_m, M_j) - M_j - (n - M_m)\xi_{M_m:M_j}|\\ &\ge 4\sqrt{D_m \xi_{M_m:M_j} \log_2 M_m}\Big\}.\end{aligned}$$

Because $R(n, M_m, M_j) - M_j = \#\{i = M_m + 1, \ldots, n : \ \xi_i \le \xi_{M_m:M_j}\}$ and $\xi_i \overset{\text{iid}}{\sim} \text{Unif}(0,1)$ for $i > M_m$ and are independent of $\xi_{M_m:M_j}$, by Lemma S-10,

$$\begin{aligned}
&\Pr\{\Gamma_{m:j} \cap E_m\} \\
&\le \sup_{\eta^{j-m}/2 \le x \le 2\eta^{j-m}} \Pr\left\{ \max_{1 \le n \le D_m} n\, |\mathbb{U}_n(x) - x| \ge 4\sqrt{D_m\, x \log_2 M_m} \right\} \\
&\le 2e^{-4 \log_2 M_m} = 2(\log M_m)^4 \sim 2(m \log \eta)^{-4},
\end{aligned}$$

yielding $\sum_m \sum_{am \le j \le bm} \Pr\{\Gamma_{m:j} \cap E_m\} < \infty$. By Lemma S-8, $\Pr\{\bigcap_{m=n}^\infty E_m\} \to 1$ as $n \to \infty$. Then by Borel-Cantelli lemma, almost surely, if $m \gg 1$, then for all $j \in [am, bm] \cap \mathbb{N}$ and $M_{m+1} \le n \le M_{m+2}$, we have $\xi_{M_m:M_j} < 2\eta^{j-m}$, $\log_2 \eta^m \sim \log_2 \eta^j$, and

$$\begin{aligned}
\left| \frac{R(n,\, M_m,\, M_j) - M_j}{n - M_m} - \xi_{M_m:M_j} \right| &\le \frac{4\sqrt{D_m \xi_{M_m:M_j} \log_2 \eta^m}}{n - M_m} \\
&\le \frac{8\eta\, \mathbb{L}(\eta^j)}{\sqrt{\eta - 1}\, \eta^m}.
\end{aligned}$$

By Lemma S-8, $\left|\xi_{M_m:M_j} - M_j/M_m\right| \le 3\mathbb{L}(\eta^j)/\eta^m$. Combining the results,

$$\left| \frac{R(n,\, M_m,\, M_j) - M_j}{n - M_m} - \frac{M_j}{M_m} \right| \le \frac{11\eta\, \mathbb{L}(\eta^j)}{\sqrt{\eta - 1}\, \eta^m}.$$

Then

$$\begin{aligned}
\left| \frac{R(n, M_m, M_j)}{n} - \frac{M_j}{M_m} \right| &= \frac{n - M_m}{n} \left| \frac{R(n, M_m, M_j) - M_j}{n - M_m} - \frac{M_j}{M_m} \right| \\
&\le \frac{11\eta(\eta^2 - 1)\, \mathbb{L}(\eta^j)}{\sqrt{\eta - 1}\, \eta^m}.
\end{aligned}$$

By $\mathbb{L}(\eta^j)/\eta^m = \mathbb{D}(\eta^j)\, \eta^{j-m}$, the proof is complete. □

2.3. *The last statement of Theorem 2.3.* By (2.10), it is apparent that $R_n \to \infty$ a.s. To show $V_n/R_n \to \beta_*$ a.s., apply the representation (S-2). For each $n \ge 1$, sorting $p_1, \ldots, p_n$ in increasing order rearranges $\eta_1, \ldots, \eta_n$ as $\eta_{n:1}, \ldots, \eta_{n:n}$. Then

$$\frac{V_n}{R_n} = \frac{1}{R_n} \sum_{j=1}^{R_n} \mathbf{1}\{\eta_{n:j} < \rho(p_{n:j})\}.$$

Almost surely, $p_{n:R_n} \to 0$ and hence $\rho(p_{n:i}) \to \beta_*$ uniformly for $i = 1, \ldots, R_n$. Then from the Strong Law of Large Numbers (SLLN), $V_n/R_n \to \beta_*$.

2.4. *Bandwidth for kernel estimation of $F'(0)$.* Because the behavior of the BH procedure relies on $F'(0)$, it is desirable to estimate the latter. Given an appropriate kernel function $K(t) \ge 0$, $F'(0)$ can be estimated by

$$\hat{F}_n'(0) = \frac{2}{nb_n} \sum_{j=1}^n K\left(\frac{\epsilon_j p_j}{b_n}\right) \tag{S-10}$$

where $\epsilon_1, \epsilon_2, \ldots$ are iid with $\Pr\{\epsilon_1 = 1\} = \Pr\{\epsilon_1 = -1\} = 1/2$ and $nb_n \to \infty$. We have

PROPOSITION S-1. *Suppose for (S-10), the kernel function K is smooth and nonnegative with bounded support, such that $K(x) = K(-x)$ and $\int K = 1$. Let*

$$\mu = \frac{F^{(\ell)}(0)}{(\ell-1)!} \int K(x)|x|^{\ell-1}\,dx, \quad \kappa = 2\int K^2.$$

Then the minimum MSE of $\hat{F}_n'(0)$ is

$$\textit{MIN MSE} = (1+o(1))\left(\frac{\kappa}{\nu_0}\right)^{\nu_0}\left(\frac{\mu^2}{1-\nu_0}\right)^{1-\nu_0} n^{-\nu_0}$$

with the corresponding optimal bandwidth

$$b_n = \left[\left(\frac{1}{\nu_0} - 1\right)\frac{\kappa}{\mu^2}\right]^{1-\nu_0} \frac{1}{n^{1-\nu_0}}.$$

It is interesting that for the kernel estimator (S-10), the minimum MSE is of the same order as $n^{-\nu_0}$ and the optimal bandwidth is of the same order as n^{ν_0-1}. Whether there is any connection between the kernel estimation and criticality remains to be seen.

PROOF. Because the Rademacher process $\epsilon_1, \epsilon_2, \ldots$ and $p_1, p_2, \ldots$ are independent, $\epsilon_1 p_1$, $\epsilon_2 p_2$, $\ldots$ are iid with density $f(x) = \frac{1}{2}F'(|x|)$. So the bias of $\hat{F}_n'(0)$ is

$$\begin{aligned}
E[\hat{F}_n'(0)] - F'(0) &= \frac{2}{b_n}\int K\left(\frac{x}{b_n}\right) f(x)\,dx - F'(0) \\
&= \frac{1}{b_n}\int K\left(\frac{x}{b_n}\right) F'(|x|)\,dx - F'(0) \\
&= \int K(x)\,[\,F'(b_n|x|) - F'(0)\,]\;dx \\
&= \frac{b_n^{\ell-1}F^{(\ell)}(0)}{(\ell-1)!}\int K(x)|x|^{\ell-1} + o(x^{\ell-1}) \\
&= \mu b_n^{\ell-1}(1+o(1)).
\end{aligned}$$

On the other hand, the variance of $\hat{F}_n'(0)$ is

$$\frac{2(1+o(1))}{nb_n^2}\int K^2\left(\frac{x}{b_n}\right) f(x)\,dx = \frac{\kappa}{nb_n}(1+o(1)).$$

Therefore, the MSE with bandwidth b_n is $(\mu^2 b_n^{2(\ell-1)} + \frac{\kappa}{nb_n})(1+o(1))$. Minimizing over b_n then finishes the proof. □

3. Proof of Theorem 2.2.

THEOREM 2.2. *Suppose $\alpha \in (\alpha_*, 1)$ and $\Delta = 1 - \alpha F'(u_*) > 0$. Let $q_* = 1 - p_*$. Then*

$$\overline{\lim_n} \pm \frac{R_n - np_*}{\sqrt{n\log_2 n}} = \frac{\sqrt{2p_*q_*}}{\Delta}, \quad a.s. \tag{2.8}$$

Furthermore, R_n/n is asymptotically proportional to the power:

$$\mathrm{power}_n = \frac{R_n}{n}\left(\frac{1-\alpha}{\pi} + \alpha\right) + o_p(1) \to G(u_*), \quad a.s. \tag{2.9}$$ □

We first prove (2.8). By (2.9), $R_n/n \xrightarrow{\text{a.s.}} p_*$. Then by SLLN, $V_n/R_n \xrightarrow{\text{a.s.}} (1-\pi)\alpha$. Therefore,

$$\begin{aligned}
\mathrm{power}_n &= \frac{R_n - V_n}{n - N_n} = \frac{R_n}{n}\frac{1 - V_n/R_n}{1 - N_n/n} \\
&= \frac{R_n}{n}\frac{1-(1-\pi)\alpha}{\pi} + o_p(1) \\
&\xrightarrow{\text{a.s.}} \frac{p_*[1-(1-\pi)\alpha]}{\pi} = G(\alpha p_*).
\end{aligned}$$

On the other hand,

$$G(\alpha p_*) = \alpha\left[\frac{1/\alpha - 1}{\pi} + 1\right] p_* \sim \alpha\left[\frac{F'(0)-1}{\pi} + 1\right] p_*$$

Combining the above two formulas then proves (2.8).

Most of the effort will be devoted to $\underline{\lim}_n (R_n - np_*)/\mathbb{L}(n) = -\sqrt{2p_*q_*}/\Delta$, which is equivalent to $\overline{\lim}_n -(R_n - np_*)/\mathbb{L}(n) = \sqrt{2p_*q_*}/\Delta$. Recall that $\Delta = 1 - \alpha F'(\alpha p_*)$. Then $\overline{\lim}_n (R_n - np_*)/\mathbb{L}(n) = \sqrt{2p_*q_*}/\Delta$ will be shown by a "time reversal" argument.

3.1. *Lower bound for $\underline{\lim}_n (R_n - np_*)/\mathbb{L}(n)$.* In this section we show $\underline{\lim}_n (R_n - np_*)/\mathbb{L}(n) \ge -\sqrt{2p_*q_*}/\Delta$, a.s. In the following, denote

$$t(n,a) = \lfloor np_* - a\,\mathbb{L}(n)\rfloor, \quad n \ge 1,\ a > 0.$$

Recall the definition of $R(n,m,j)$ given just before Lemma S-4. We need two lemmas.

Lemma S-1. *If $a > \sqrt{2p_*q_*}/\Delta$, then there are $\delta = \delta(a) > 0$ and $c = c(a) > 1$, such that given $\eta \in (1, c)$, almost surely, $\xi_{M_m:T_m} \le z_{M_m:T_m} - \delta\mathbb{D}(\eta^m)$ for all $m \gg 1$, where $M_m = \lfloor \eta^m \rfloor$ and $T_m = t(M_m, a)$. Recall $z_{n:j} = F(\alpha j/n)$ for $1 \le j \le n$.*

Lemma S-2. *Given $\eta \in (1, 2)$, define M_m and T_m as in Lemma S-1. Then almost surely, for $m \gg 1$,*

$$|R(n, M_m, T_m)/n - T_m/M_m| \le 66\sqrt{\eta - 1}\mathbb{D}(\eta^m).$$

Proof of lower bound. Given $a > 2p_*q_*/\Delta^2$, fix $\delta = \delta(a)$, $c = c(a)$ as in Lemma S-1. Let $\eta \in (1, c)$ such that $66\sqrt{\eta - 1} < \delta$. Clearly $T_m/M_m \to p_*$. By Lemma S-1, almost surely, for $m \gg 1$, $\xi_{M_m:T_m} \le z_{M_m:T_m} - \delta\mathbb{D}(\eta^m)$. On the other hand, by Taylor expansion, for all $M_{m+1} \le n \le M_{m+2}$, $z_{n:R(n,M_m,T_m)} \ge z_{M_m:T_m} - |R(n, M_m, T_m)/n - T_m/M_m|$. Then by Lemma S-2, $z_{n:R(n,M_m,T_m)} > z_{M_m:T_m} - \delta\mathbb{D}(\eta^m)$, yielding

$$\xi_{n:R(n,M_m,T_m)} = \xi_{M_m:T_m} \le z_{M_m:T_m} - \delta\mathbb{D}(\eta^m) < z_{n:R(n,M_m,T_m)}$$

Thus $R_n \ge R(n, M_m, T_m)$. Because by Lemma S-2,

$$\begin{aligned} R(n, M_m, T_m) - np_* &\ge \left(T_m/M_m - 66\sqrt{\eta - 1}\mathbb{D}(\eta^m) - p_*\right) n \\ &\ge -\left(a + 66\sqrt{\eta - 1}\right) \mathbb{D}(\eta^m) n \\ &\ge -\eta \left(a + 66\sqrt{\eta - 1}\right) \mathbb{L}(n), \end{aligned}$$

we get $\underline{\lim}_n (R_n - np_*)/\mathbb{L}(n) \ge -\eta(a + 66\sqrt{\eta - 1})$. Since $\eta > 1$ and $a > \sqrt{2p_*q_*}/\Delta$ are arbitrary, this then finishes the proof. □

Following Lemma S-10, it can be shown that if $c_n \to \infty$ and $a \in (0, 1)$, then for $d_n > 0$ and $\delta > 0$, for all $n \gg 1$ and $c_n d_n \le j \le an$,

$$\Pr\{\pm(j - n\mathbb{U}_n(j/n)) \ge d_n\} \le \exp\left\{-\frac{d_n^2}{2(1+\delta)j(1 - j/n)}\right\}. \tag{S-1}$$

Proof of Lemma S-1. Indeed, it suffices to choose $\delta = \delta(a) = \frac{1}{2}(a\Delta - \sqrt{2p_*q_*})$ and $c = c(a) = \frac{1}{6}a\Delta/\sqrt{2p_*q_*} + \frac{5}{6} > 1$. To see this, define

$$\begin{aligned} r_n &= \left|\xi_{n:t(n,a)} + \mathbb{U}_n\left(\frac{t(n,a)}{n}\right) - \frac{2t(n,a)}{n}\right|, \\ \Gamma_n &= \left\{r_k \le \frac{2\log k}{k^{3/4}},\ \text{all } k \ge n\right\}, \\ E_n &= \left\{\xi_{n:t(n,a)} > z_{n:t(n,a)} - \delta\mathbb{D}(n)\right\}. \end{aligned}$$

Fix $0 < \epsilon < \frac{1}{6}(\Delta - \sqrt{2p_* q_*}/a)$. Because $t(n,a) - np_* \sim -a\mathbb{L}(n)$, for $n \gg 1$, on $E_n \cap \Gamma_n$, by Taylor expansion of $z_{n:t(n,a)} = F(\alpha t(n,a)/n)$ around αp_*,

$$\begin{aligned} &\frac{2t(n,a)}{n} - \mathbb{U}_n\left(\frac{t(n,a)}{n}\right) \\ &> F(\alpha p_*) + (\alpha F'(\alpha p_*) + \epsilon)\left(\frac{t(n,a)}{n} - p_*\right) - r_n - \delta\mathbb{D}(n) \\ &> p_* + (\alpha F'(\alpha p_*) + 2\epsilon)\left(\frac{t(n,a)}{n} - p_*\right) - \delta\mathbb{D}(n)\,, \\ \Longrightarrow \quad & t(n,a) - n\mathbb{U}_n\left(\frac{t(n,a)}{n}\right) \\ &\ge (np_* - t(n,a))(\Delta - 2\epsilon) - \delta\mathbb{L}(n) \ge c\sqrt{2p_* q_*}\mathbb{L}(n). \end{aligned}$$

Then by (S-1), given $\eta \in (1, c)$, for $n \gg 1$,

$$\Pr\{E_n \cap \Gamma_n\} \le \exp\left\{-c^2 \log_2 n/\eta\right\} < (\log n)^{-c}.$$

Consequently $\sum_{m=1}^{\infty} \Pr\{E_{M_m} \cap \Gamma_{M_m}\} < \infty$. By Borel-Cantelli lemma, almost surely, for $m \gg 1$, either $E^c_{M_m}$ or $\Gamma^c_{M_m}$ occurs. However, by the Bahadur-Kiefer representation, almost surely, for all $n \gg 1$, Γ_n occurs. As a result, for $m \gg 1$, $E^c_{M_m}$ occurs, i.e., $\xi_{M_m:T_m} \le z_{M_m:T_m} - \delta\mathbb{D}(\eta^m)$. □

Because the proof of Lemma S-2 is very similar to that for Lemma S-4, with M_j being replaced by T_m, the detail of the proof is omitted.

3.2. *Upper bound for* $\underline{\lim}_n (R_n - np_*)/\mathbb{L}(n)$. In this section we show $\underline{\lim}_n (R_n - np_*)/\mathbb{L}(n) \le -\sqrt{2p_* q_*}/\Delta$. Together with the lower bound, this implies $\overline{\lim}_n -(R_n - np_*)\mathbb{L}(n) = \sqrt{2p_* q_*}/\Delta$. We need two lemmas.

Lemma S-3. *If* $a \in (0, \sqrt{2p_* q_*}/\Delta)$, *then there are* $\delta = \delta(a) > 0$ *and* $L = L(a) \gg 1$, *such that* $\Pr\{E_{M_m}$ *i.o.*$\} = 1$, *where*

$$E_n = \{\xi_{n:t(n,a)} - z_{n:t(n,a)} > \delta\mathbb{D}(n)\}$$

and $M_m = \lfloor L^m \rfloor$.

Lemma S-4. *Let* E_n *and* M_m *be defined as in Lemma S-3 and* $T_m = t(M_m, a)$. *Then there is a constant* $\kappa > 0$, *such that*

$$\sum_{m=1}^{\infty} \Pr\{R_{M_m} > T_m + \kappa m,\ E_{M_m}\} < \infty.$$

Proof of upper bound. Let

$$\Gamma_m = \{R_{M_m} > T_m + \kappa m\}\,.$$

By Lemma S-4 and Borel-Cantelli lemma, a.s., $\Gamma^c_m \cup E^c_{M_m}$ occurs for all $m \gg 1$. By Lemma S-3, for infinitely many m, E_{M_m} occurs, and hence $R_{M_m} \le T_m + \kappa m$. Therefore, $\underline{\lim}_n (R_n - np_*)/\mathbb{L}(n) \le -a$. Because $a < \sqrt{2p_* q_*}/\Delta$ is arbitrary, the upper bound is proved. □

The proof of Lemma S-3 is based on the following result.

Lemma S-5. *Let $a \in (0, \sqrt{2p_* q_*}/\Delta)$. There are $\delta = \delta(a) > 0$ and $L = L(a) \gg 1$, such that a.s., $t_m - \mathbb{U}_{M_m}(t_m) \geq \Delta(p_* - t_m) + \delta\mathbb{D}(M_m)$ for infinitely many m, where $M_m = \lfloor L^m \rfloor$, $T_m = t(M_m, a)$, and $t_m = T_m/M_m$*

Proof. Let $D_m = D_m(L) = M_{m+1} - M_m$. Fix $\delta > 0$, $b \in (0, a)$, and $L > 0$, such that

$$
\begin{gathered}
a\Delta + 2\delta < \sqrt{2p_* q_*}, \\
b + b\Delta + \delta > a + a\Delta, \\
b\sqrt{L/(L-1)} < a, \quad \text{and} \\
(b + b\Delta + 2\delta)\mathbb{L}(D_m) > (a + a\Delta + \delta)\mathbb{L}(M_{m+1}) + \sqrt{3p_* q_*}\mathbb{L}(M_m), \\
\text{for all } m \gg 1.
\end{gathered}
$$

It is not hard to see such δ, b, and L indeed exist. Denote $x_n = t(n, b)/n$ and let

$$
\lambda_n = \Pr\left\{ n x_n - \sum_{j=1}^{n} \mathbf{1}\left\{\xi_j \leq x_n\right\} \geq (b\Delta + 2\delta)\mathbb{L}(n) \right\}.
$$

Given $\eta > 0$ with $2p_* q_* > (1 + \eta)(b\Delta + 2\delta)^2$, since $x_n - \mathbf{1}\left\{\xi_j \leq x_n\right\}$ are iid with mean 0 and variance $x_n(1 - x_n)$, by Lemma S-6, it can be seen that

$$
\lambda_n \geq \exp\left\{ -\frac{(1+\eta)(b\Delta + 2\delta)^2 \log_2 n}{2x_n(1 - x_n)} \right\}, \quad \text{for all } n \gg 1.
$$

Since $x_n \to p_*$, there is $\rho \in (0, 1)$, such that $\lambda_n > (\log n)^{-\rho}$ for $n \gg 1$. Let

$$
\Gamma_m = \left\{ D_m x_{D_m} - \sum_{j=M_m+1}^{M_{m+1}} \mathbf{1}\left\{\xi_j \leq x_{D_m}\right\} \geq (b\Delta + 2\delta)\mathbb{L}(D_m) \right\}.
$$

Then Γ_m are independent of each other and $\Pr\{\Gamma_m\} \geq (\log D_m)^{-\rho} > (m \log L)^{-\rho}$ for all $m \gg 1$. Therefore, $\sum_m \Pr\{\Gamma_m\} = \infty$, and by Borel-Cantelli lemma, $\Pr\{\Gamma_m \text{ i.o.}\} = 1$.

Because $b\sqrt{L/(L-1)} < a$, for $m \gg 1$, $x_{D_m} > t_{m+1}$ and hence on Γ_m,

$$
D_m x_{D_m} - \sum_{j=M_m+1}^{M_{m+1}} \mathbf{1}\left\{\xi_j \leq t_{m+1}\right\} \geq (b\Delta + 2\delta)\mathbb{L}(D_m). \tag{S-2}
$$

Let $G_m = \left\{ p_* M_m - \sum_{j=1}^{M_m} \mathbf{1}\left\{\xi_j \leq p_*\right\} \geq -\sqrt{3p_* q_*}\mathbb{L}(M_m) \right\}$. By the LIL, as $n \to \infty$, $\Pr\{\bigcap_{m \geq n} G_m\} \to 1$. Then by $t_{m+1} < p_*$, almost surely, for all $m \gg 1$,

$$
p_* M_m - \sum_{j=1}^{M_m} \mathbf{1}\left\{\xi_j \leq t_{m+1}\right\} \geq -\sqrt{3p_* q_*}\mathbb{L}(M_m). \tag{S-3}
$$

Add (S-2) and (S-3). Observe

$$D_m x_{D_m} + p_* M_m \le p_* M_{m+1} - b\mathbb{L}(D_m) \\ \le T_{m+1} + a\mathbb{L}(M_{m+1}) - b\mathbb{L}(D_m) + 1.$$

Then by the selection of δ, b, and L, for $m \gg 1$, on $\Gamma_m \cap G_m$,

$$\begin{aligned} &T_{m+1} - \sum_{j=1}^{M_{m+1}} \mathbf{1}\{\xi_j \le t_{m+1}\} \\ &\ge (b + b\Delta + 2\delta)\mathbb{L}(D_m) - a\mathbb{L}(M_{m+1}) - \sqrt{3p_*q_*}\mathbb{L}(M_m) \\ &\ge (a\Delta + \delta)\mathbb{L}(M_{m+1}) \\ &> \Delta(p_* M_{m+1} - T_{m+1}) + \delta\mathbb{L}(M_{m+1}). \end{aligned}$$

Dividing both sides by M_{m+1} yields

$$t_{m+1} - \mathbb{U}_{M_{m+1}}(t_{m+1}) \ge \Delta(p_* - t_{m+1}) + \delta\mathbb{D}(M_{m+1}).$$

Since $\Pr\{\Gamma_m \cap G_m \text{ i.o.}\} = 1$, the proof is then complete. □

Proof of Lemma S-3. Let $\delta = \delta(a)$ and $L = L(a)$ as in Lemma S-5. Fix $\epsilon \in (0, \frac{\delta}{2a})$. Because $t_m \to p_*$ and $t_m < p_*$, by Taylor expansion of F around αp_*, $z_{M_m:T_m} < p_* + (\alpha F'(\alpha p_*) - \epsilon)(t_m - p_*) = p_* - (1-\Delta-\epsilon)(p_* - t_m)$. By Lemma S-5, it follows that that almost surely, for infinitely many m,

$$\begin{aligned} 2t_m - \mathbb{U}_{M_m}(t_m) &\ge t_m + \Delta(p_* - t_m) + \delta\mathbb{D}(M_m) \\ &= p_* - (1-\Delta)(p_* - t_m) + \delta\mathbb{D}(M_m) \\ &\ge z_{M_m:T_m} + (\delta/2)\mathbb{D}(M_m). \end{aligned}$$

On the other hand, by the Bahadur-Kiefer representation, $\xi_{M_m:T_m} = 2t_m - \mathbb{U}_{M_m}(t_m) + o(\mathbb{D}(M_m))$. Combining the inequalities and redefining δ as $\delta/3$, the proof is complete. □

Lemma S-4 is based on the following result. The argument in [2] for in-probability convergence can be modified to show the lemma. Therefore we omit the proof for brevity.

Lemma S-6. *Given $r > 0$, a.s., for $n \gg 1$, $R_n \le n(p_* + r)$.*

Proof of Lemma S-4. Given $\kappa > 0$, let $\Gamma_m = \{R_{M_m} \ge T_m + \kappa m\} \cap E_{M_m}$. Then

$$\Gamma_m \subset \{\xi_{M_m:T_m} > z_{M_m:T_m} \text{ and there is } j > T_m + \kappa m, \text{ s.t. } \xi_{M_m:j} \le z_{M_m:j}\}.$$

Let $\epsilon \in (0, \Delta)$ and $c = \alpha F'(\alpha p_*) + \epsilon < 1$. Fix $r > 0$ and let $I = [\alpha p_* - 2r, \alpha p_* + 2r]$, such that $|F(x) - F(y)| < c|x - y|$ for $x, y \in I$. Then $\Gamma_m \subset G_m \cup N_m$, where

$$G_m = \Big\{\exists j \in (\kappa m,\, rM_m) \cap \mathbb{N}, \quad \text{s.t.} \ \ \xi_{M_m:T_m+j} - \xi_{M_m:T_m} \le cj/M_m\Big\}$$
$$N_m = \Big\{R_{M_m} > (p_* + r/2)M_m\Big\}$$
$$\qquad \bigcup \Big\{\xi_{M_m:T_m+j} \notin I, \ \text{for some } 0 \le j \le rM_m\Big\}.$$

By Lemma S-6 and SLLN, $\Pr\{\bigcup_{m=n} N_m\} \to 0$ as $n \to \infty$. On the other hand, observe that $(\xi_{M_m:T_m+j} - \xi_{M_m:T_m},\ 0 < j \le n - T_m)$ has the same distribution as $(\xi_{M_m:j},\ 0 < j \le n - T_m)$. Therefore,

$$\Pr\{G_m\} \le \lambda_m := \Pr\left\{\bigcup_{j > \kappa m} \{\xi_{M_m:j} < cj/M_m\}\right\}.$$

Following the proof of Lemma S-1, it is seen that if $\kappa > 0$ is large enough, then $\sum_m \lambda_m < \infty$, and hence by Borel-Cantelli lemma, $\Pr\{\bigcup_{m=n} G_m\} \to 0$ as $n \to \infty$. Then almost surely, for $m \gg 1$, $\Gamma_m^c = \{R_{M_m} < T_m + \kappa m\} \cup E_{M_m}^c$ occurs. By Lemma S-3, the proof is thus complete. ☐

3.3. *Proof for* $\overline{\lim}_n (R_n - np_*)/\mathbb{L}(n)$. The upper limit of $(R_n - np_*)/\mathbb{L}(n)$ can be deduced from an appropriate lower limit based on "time reversal". The idea is that R_n has a bounded difference with the first index j that satisfies $\xi_{n:j} > z_{n:j} \sim p_*$. Let τ_n be this first index. Then for $\tilde{\xi}_k = 1 - \xi_k$ and $\tilde{\phi}(x) = 1 - F(\alpha(1 - x))$, $\tilde{R}_n = n - \tau_n$ is the last index satisfying $\tilde{\xi}_{n:j} < \tilde{\phi}(j/n) \sim \tilde{p_*}$, with $\tilde{p_*} = 1 - p_*$ a fixed point of $\phi(x)$. Since we can handle the lower limit of $\tilde{R}_n$, the upper limit for R_n can be obtained.

To complete the proof, we shall remove the condition that p_* be the largest fixed point of $F(\alpha x)$ and that F be a distribution function. Note that by Lemma S-6, R_n can be defined as $\max\{j \ge 0 : \xi_{n:j} \le z_{n:j}$ and $|j/n - p_*| < r\}$ for any $r > 0$. This makes it possible to generalize the result to any fixed point of ϕ.

Lemma S-7. *Suppose $\phi \in C([0,1])$ is increasing and has a fixed point $p_* \in (0,1)$ such that $\Delta = 1 - \phi'(p_*) > 0$. Denote $q_* = 1 - p_*$. Let $c \in (\phi'(p_*), 1)$ and $r \in (0, q_*)$ such that*

$$\phi(y) - \phi(x) < c(y - x), \quad \text{if} \ \ p_* - r < x < y < p_* + r;$$
$$c(q_* + r) < q_* - r.$$

Define $R_n = \max\{j \ge 0 :\ \xi_{n:j} \le \phi(j/n),$ and $|j/n - p_*| < r\}$, then $\underline{\lim}_n (R_n - np_*)/\mathbb{L}(n) = \sqrt{2p_*q_*}/\Delta$.

Lemma S-8. *Under the same conditions as Lemma S-7, define*

$$\tau_n = \begin{cases} \min\{j \ge 0 : \xi_{n:j+1} \ge \phi(j/n) \text{ and } |j/n - p_*| < r\}, & \text{if such } j \text{ exist;} \\ \infty, & \text{otherwise.} \end{cases}$$

Then $R_n - \tau_n \ge 0$ is stochastically bounded and there is a constant $\kappa > 0$, such that almost surely, for $n \gg 1$, $R_n \le \tau_n + \kappa \log n$.

Proof of $\overline{\lim}_n (R_n - np_*)/\mathbb{L}(n) = \sqrt{2p_*q_*}/\Delta$. Let $\tilde{\xi}_j = 1 - \xi_j$. Then for $n \ge 1$ and $j = 0, \ldots, n+1$, $\tilde{\xi}_{n:j} = 1 - \xi_{n:n-j+1}$. Let $\phi(x) = F(\alpha x)$ and $\tilde{\phi}(x) = 1 - \phi(1-x)$. Then q_* is a fixed point of $\tilde{\phi}$ such that $\tilde{\phi}'(q_*) = \phi'(p_*) = \alpha F'(\alpha p_*) < 1$. Moreover,

$$n - \tau_n = \max\{j : \tilde{\xi}_{n:j} \le \tilde{\phi}(j/n) \text{ and } |j/n - q_*| < r\}.$$

Therefore, by Lemma S-7,

$$\underline{\lim}_{n\to\infty} \frac{np_* - \tau_n}{\mathbb{L}(n)} = \underline{\lim}_{n\to\infty} \frac{n - \tau_n - nq_*}{\mathbb{L}(n)} = -\sqrt{2p_*q_*}/\Delta, \quad \text{a.s.}$$

and hence $\overline{\lim}_n (\tau_n - np_*)/\mathbb{L}(n) = \sqrt{2p_*q_*}/\Delta$. Then by Lemma S-8, almost surely, $R_n - \tau_n = O(\log n) = o(\mathbb{L}(n))$, which gives $\overline{\lim}_n (R_n - np_*)/\mathbb{L}(n) = \sqrt{2p_*q_*}/\Delta$. □

Lemma S-7 follows exactly the same argument as in last two subsections. We therefore omit its proof.

Proof of Lemma S-8. By definition, $\xi_{n:\tau_n} < \phi(\tau_n/n)$ and so $\tau_n \le R_n$. Let $I = (p_* - r, p_* + r)$ and

$$E_n = \left\{\tau_j/j,\ R_j/j, \text{ and } \xi_{j:\tau_j+1} \in I \text{ for all } j \ge n\right\}.$$

Given $k \ge 1$,

$$\begin{aligned}
&\Pr\{R_n - \tau_n = k,\ E_n\} \\
=& \sum_{s:\ s/n,\,(k+s)/n \in I} \Pr\{R_n = k+s,\ \tau_n = s,\ E_n\} \\
\le& \sum_{s:\ s/n,\,(k+s)/n \in I} \Pr\{\xi_{n:k+s} - \xi_{n:s+1} \le \phi((k+s)/n) - \phi(s/n), \\
&\qquad \tau_n = s,\ \xi_{n:\tau_n+1} \in I\} \\
\le& \sum_{s:\ s/n \in I} \Pr\{\xi_{n:k+s} - \xi_{n:s+1} \le ck/n,\ \tau_n = s,\ \xi_{n:\tau_n+1} \in I\}.
\end{aligned}$$

Let $\xi'_1, \xi'_2, \ldots \overset{\text{iid}}{\sim} \text{Unif}(0,1)$ and independent of $\xi_1, \xi_2, \ldots$. Because τ_n is a stopping time with respect to $\xi_{n:1}, \ldots, \xi_{n:n}$, conditioning on $\tau_n = s$,

$$\xi_{n:j+s+1} - \xi_{n:s+1}, \quad j \ge 0$$

have the same joint distribution as

$$(1-\xi_{n:s+1})\xi'_{n-s-1:j}, \quad j \ge 0.$$

Therefore,

$$\begin{aligned}
&\Pr\{R_n - \tau_n = k,\, E_n\} \\
&\le \sum_{s:\ s/n\in I} \Pr\{\xi'_{n-s-1:k-1}(1-\xi_{n:s+1}) \le ck/n,\ \tau_n = s,\ \xi_{n:s+1}\in I\} \\
&\le \sum_{s:\ s/n\in I} \Pr\{\xi'_{n-s-1:k-1}(q_*-r) \le c(q_*+r)k/(n-s-1)\}\Pr\{\tau_n = s\} \\
&\le \sup_{j\ge(q_*-r)n} \Pr\{\xi_{j:k-1} < c'k/j\},
\end{aligned}$$

where $c' = c(q_*+r)/(q_*-r) < 1$. From Lemma S-6, $\Pr\{E_n\} \to 1$. The lemma then follows from almost the identical argument for the the case $\alpha < \alpha_*$. □

4. Proofs for Procedure M.

4.1. *Proof of Proposition 4.1.* The result involves the following convergences:

$$\text{(S-1)} \qquad \begin{aligned} p_{n:L_n(t)} &\to u^- := \inf\{x\in[0,t] : t-x \le \alpha(F(t)-F(x))\}, \\ p_{n:L_n(t)} &\to u^+ := \sup\{x\in[t,1] : x-t \le \alpha(F(x)-F(t))\}. \end{aligned}$$

Proposition 4.1. $L_n(t) = R_n^o(T_n^-(t)) + 1$ and $U_n(t) = R_n(T_n^+(t))$, where

$$\text{(S-2)} \qquad \begin{aligned} T_n^-(t) &= \inf\left\{x \le t : \frac{t-x}{\alpha} \le \frac{[R_n^o(t)-R_n^o(x)]\vee 1}{n}\right\}, \\ T_n^+(t) &= \sup\left\{x \ge t : \frac{x-t}{\alpha} \le \frac{[R_n(x)-R_n(t)]\vee 1}{n}\right\}. \end{aligned}$$

Therefore, Step 2 of Procedure M is the same as rejecting all nulls with p-values in $[T_n^-(t_{i_k}), T_n^+(t_{i_k})]$, $k = 0,\dots,m$. Furthermore, $T_n^\pm(t) \xrightarrow{\text{a.s.}} u_t^\pm$ and (S-1) holds.

Proof. For (S-2), we only show $L_n(t) = R_n^o(T_n^-(t)) + 1$, i.e. $r_n^-(t) = J := R_n^o(t) - R_n^o(T_n^-(t))$. First, if $J > 0$, then by $T_n^-(t) \le p_{n:R_n^o(T_n^-(t))+1} = p_{n:R_n^o(t)-J+1}$ and the right continuity of R_n^o,

$$\begin{aligned}
\frac{1}{\alpha}(t - p_{n:R_n^o(t)-J+1}) &\le \frac{1}{\alpha}(t - T_n^-(t)) \\
&\le \frac{[R_n^o(t) - R_n^o(T_n^-(t))]\vee 1}{n} = \frac{\alpha J}{n}.
\end{aligned}$$

If $J = 0$, then $t - p_{n:R_n^o(t)-J+1} = t - p_{n:R_n^o(t)+1} \le 0 = J/n$. In either case, by definition, $r_n^-(t) \ge J$. On the other hand, for $j > J$, $p_{n:R_n^o(t)-j+1} \le p_{n:R_n^o(t)-J} = p_{n:R_n^o(T_n^-(t))} < T_n^-(t)$, then by the definition of $T_n^-(t)$ and $j \ge 1$,

$$\frac{1}{\alpha}(t - p_{n:R_n^o(t)-j+1}) > \frac{[R_n^o(t) - R_n^o(p_{n:R_n^o(t)-j+1})] \vee 1}{n} = \frac{\alpha j}{n}$$

and hence $r_n^-(t) < j$, implying $r_n^-(t) \le J$.

To show (S-1), notice $p_{n:L_n(t)-1} < T_n^-(t) \le p_{n:L_n(t)}$. Following the argument in [2], $T_n^-(t) \to u^-(t)$. Because $p_{n:L_n(t)} - p_{n:L_n(t)-1} \to 0$, the first convergence in (S-1) is proved. The second convergence in (S-1) can be proved likewise. □

4.2. *Proof of Theorem 4.1.* Let $p_*^+(t) = F(u_t^+) - F(t)$, $p_*^-(t) = F(t) - F(u_t^-)$. The theorem is stated as follows.

THEOREM 4.1. *Suppose $t_i - \alpha F(t_i)$ are different from each other.*

1) *If $p_*^\pm(t_i) = 0$ and $F'(t_i) < 1/\alpha$ for all i, then a.s., for $n \gg 1$, $R_n = 0$. Furthermore,*

$$\begin{aligned}&(r_n^+(0), r_n^-(t_1), r_n^+(t_1), \ldots, r_n^-(t_{M-1}), r_n^+(t_{M-1}), r_n^-(1))\\ &\xrightarrow{\mathrm{d}} (\tau_0, \tilde\tau_1, \tau_1, \ldots, \tilde\tau_{M-1}, \tau_{M-1}, \tilde\tau_M), \quad \textit{as } n \to \infty,\end{aligned}$$

where the τ's and $\tilde\tau$'s are independent, each τ_k and $\tilde\tau_k$ following the distribution of the last excursion time into $(-\infty, 0)$ of the random walk $S_0 = 0$, $S_k = S_{k-1} + \gamma_k - \alpha F'(t_k)$, with $\gamma_1, \gamma_2, \ldots$ iid $\sim$ Exp(1).

2) *If $p_*^+(t_i) + p_*^-(t_i) > 0$ for some $i = 0, \ldots, M$, then a.s.,*

$$\varlimsup_{n\to\infty} FDR = \varlimsup_{n\to\infty} pFDR \le (1-\pi)\alpha, \quad \lim_{n\to\infty} \mathrm{power}_n = \left(\frac{1-\alpha}{\pi} + \alpha\right)\Pi$$

where

$$\Pi = \max_{\substack{S \subset \{t_0,\ldots,t_M\}:\ [u_s^-, u_s^+] \\ \text{are disjoint for } s \in S}} \left\{ \sum_{s\in S} [p_*^-(s) + p_*^+(s)] \right\}.$$

3) *Suppose for each i, $I_i := \left\{k > 1 : F^{(k)}(t_i) \ne 0\right\} \ne \emptyset$. Let $\ell_i = \min I_i$. If $p_*^\pm(t_i) = 0$ for all i but $F'(t_i) = 1/\alpha$ for at least one of them, then*

$$\lim_{n\to\infty} FDR = \lim_{n\to\infty} pFDR = (1-\pi)\alpha, \quad \frac{\log R_n}{\log n} \xrightarrow{\text{a.s.}} \frac{2\ell - 2}{2\ell - 1},$$

where $\ell = \max\{\ell_i : F'(t_i) = 1/\alpha\}$. Additionally, a.s., for $n \gg 1$, the set of rejected p-values consists exactly of those in $[T_n^-(t_i), T_n^+(t_i)]$ with $F'(t_i) = 1/\alpha$.

We need some notations for the proof. From Proposition 4.1, Step 2 of Procedure M is the same as rejecting all hypotheses whose p-values are in $p_j \in [T_n^-(t_{i_k}), T_n^+(t_{i_k})]$, $k = 0, \ldots, m$. Given $t \in [0, 1]$, for the forward and backward BH-type procedures with reference point t, $r_n^+(t) = R_n(T_n^+(t)) - R_n(t)$ and $r_n^-(t) = R_n^o(t) - R_n^o(T_n^-(t))$, and the numbers of false rejections are $V_n(T_n^+(t)) - V_n(t)$ and $V_n^o(t) - V_n^o(T_n^-(t))$, respectively, where

$$V_n(x) = \#\{1 \le j \le n : p_j \le x, \theta_j = 0\},$$
$$V_n^o(x) = \#\{1 \le j \le n : p_j < x, \theta_j = 0\}.$$

Similar to the BH procedure (S-1),

$$\text{(S-3)} \qquad \left.\begin{aligned} E\left[\frac{V_n(T_n^+(t)) - V_n(t)}{r_n^+(t) \vee 1}\right] &\le (1-\pi)\alpha, \\ E\left[\frac{V_n^o(t) - V_n^o(T_n^-(t))}{r_n^-(t) \vee 1}\right] &\le (1-\pi)\alpha. \end{aligned}\right\}$$

Most part of Theorem 4.1 follows from the results on the BH procedure. We will only give proof to statements that are specific to Procedure M.

Proof of Part 1). Following Lemma S-1, a.s., for $n \gg 1$, $r_n \pm (t_i) < (\log n)^2$ for all i. Then $R_n = 0$ by Condition 1) in Step 2.

The convergence of the joint distribution of $r_n^\pm(t_i)$ can be proved as follows. First, for each i, following the proof of Theorem 2.1, each $r_n^\pm(t_i)$ converges weakly to the corresponding excursion time of random walk. Second, given $\epsilon > 0$ such that $[t_i - \epsilon, t_i + \epsilon]$ are disjoint from each other, for $n \gg 1$, the distribution of $r_n^+(t_i)$ (resp. $r_n^-(t_i)$) only depends on the p-values in $[t_i, t_i + \epsilon]$ (resp. $[t_i - \epsilon, t_i)$). Conditioning on the numbers of p-values in these intervals, $r_n^\pm(t_i)$ are independent of each other. Since the number of p-values in each interval tends to ∞, it then follows that the conditional distribution of $r_n^\pm(t_i)$ converges to its unconditional distribution. The convergence of the joint distribution then follows. □

Proof of Part 2). For any t, if $p_*^+(t) + p_*^-(t) > 0$, then by SLLN and Proposition 4.1

$$\frac{r_n(t)}{n} \xrightarrow{\text{a.s.}} F(u^+(t)) - F(u^-(t)),$$
$$\frac{V_n(T_n^+(t)) - V_n^o(T_n^-(t))}{n} \xrightarrow{\text{a.s.}} (1-\pi)(u^+(t) - u^-(t)),$$

and

$$\frac{V_n(T_n^+(t)) - V_n^o(T_n^-(t))}{r_n(t)} \xrightarrow{\text{a.s.}} (1-\pi)\frac{u^+(t) - u^-(t)}{F(u^+(t)) - F(u^-(t))} \le (1-\pi)\alpha.$$

Since R_n (resp. V_n) is the sum of a subset of $r_n(t_i)$ (resp. $V_n(T_n^+(t_i)) - V_n^o(T_n^-(t_i))$), with $p_*^+(t_i) + p_*^-(t_i) > 0$, it then follows that $\overline{\lim}\, V_n/R_n \le (1-\pi)\alpha$. The convergence of power can be shown following Theorem 2.2 as well. □

PROOF OF PART 3). In this case, $F'(t_i) \le 1/\alpha$ for all i. For $n \gg 1$, $A_i = [T_n^-(t_i), T_n^+(t_i)]$ are disjoint from each other. Therefore, Step 1 of Procedure M includes all A_i with $r_n(t_i) > (\log n)^2$. By Theorems 2.2 and 2.3, $r_n(t_i) > (\log n)^2$ if and only if $F'(t_j) = 1/\alpha$. The rest of the statement can be shown following the proof for Theorem 2.3. □

References.

[1] CHOW, Y. S. AND TEICHER, H. (1997). *Probability theory: independence, interchangeability, martingales*, 3 ed. Springer Texts in Statistics. Springer-Verlag, New York.

[2] GENOVESE, C. AND WASSERMAN, L. (2002). Operating characteristics and extensions of the false discovery rate procedure. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64**, 3, 499–517.

[3] JIANG, T. F. (2004). The asymptotic distributions of the largest entries of sample correlation matrices. *Ann. Appl. Probab.* **14**, 2, 865–880.

[4] POLLARD, D. (1984). *Convergence of stochastic processes*. Springer Series in Statistics. Springer-Verlag, New York.

[5] SHORACK, G. R. AND WELLNER, J. A. (1986). *Empirical processes with applications to statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York.

Appendix.

Kiefer's result on Bahadur representation. (cf. [5], section 15.1):

$$\overline{\lim_{n\to\infty}} \frac{\sup_{t\in[0,1]} |\mathbb{U}_n(t) + \mathbb{U}_n^*(t) - 2t|}{n^{-3/4}\sqrt{\log n \log_2 n}} = 1, \quad \text{a.s.}$$

Hoeffding's inequality. (cf. [4], p. 191): If $X_1, \ldots, X_n$ are iid with $a \le X_1 \le b$ and $EX_1 = \mu$, then for all $c > 0$,

$$\Pr\left\{\sum_{k=1}^n X_k - n\mu \ge c\right\} \le \exp\left\{-\frac{2c^2}{n(b-a)^2}\right\}.$$

Bennett's inequality. (cf. [5], p. 440): Let

$$\varphi(t) = (1+t)\log(1+t) - t = t^2 \sum_{k=0}^\infty \frac{(-1)^k t^k}{(k+1)(k+2)},$$

which is nonnegative and strictly increasing for $t \ge 0$. If $X \sim \text{Bin}(n,p) - np$ and $p \in [0, 1/2]$, then

$$\Pr\{X \ge x\} \le \exp\left\{-\frac{np}{1-p}\varphi\left(\frac{x}{np}\right)\right\}, \quad \text{all } x > 0.$$