Supplemental Materials

for “Positive false discovery proportions: intrinsic bounds and adaptive control” by Z. Chi and Z. Tan

S1 Notations

In this Supplemental Materials, we prove the main theoretical results in the article, i.e., Theorems 4.2, 4.3, 5.1 and 5.2. The proofs for the first two theorems are similar to those for the other two. Since the latter ones are more of interest to applications, they will be demonstrated in detail. The proof of Theorems 4.2 and 4.3 will only be outlined afterwards.

The following notations will be used. If $X_n$ and $Y_n$ are random variables, then $X_n \geq_p Y_n$, $X_n \leq_p Y_n$, and $X_n \sim_p Y_n$ denote $P(X_n \geq Y_n) \rightarrow 1$, $P(X_n \leq Y_n) \rightarrow 1$, and $X_n/Y_n \xrightarrow{p} 1$, respectively. The notation $X_n = o_p(Y_n)$ means “$|X_n| \leq_p \epsilon |Y_n|$ for any $\epsilon > 0$”, whereas $X_n = O_p(Y_n)$ means “for any $\epsilon > 0$, there are $M > 0$ and $n_0 > 0$, such that $P(|X_n| \geq M|Y_n|) < \epsilon$ for all $n \geq n_0$”.

Finally, when necessary, to make explicit the dependence on the number $n$ of tested hypotheses, we use a superscript to index a random variable. For example, denote by $R^{(n)}$ the number of rejections when there are $n$ null hypotheses.

It will be easier to work with continuous time to prove the theorems. For procedure (5.1), given $p$-values $\xi_1^{(n)}, \ldots, \xi_n^{(n)}$, the $R^{(n)}_{\tau_n}$ smallest ones are rejected, where

$$\tau_n = \sup_{t \in [0,1]} \left\{ qbin\left( \Gamma_*(t); R^{(n)}_t, \frac{n(t \lor \xi_n^{(n)})}{R_t^{(n)} \lor k_n} \land 1 \right) \leq \alpha(R^{(n)}_t \lor 1) \right\}. \quad (S1.1)$$

Henceforth denote

$$q_n(t; z) := qbin\left( z; R^{(n)}_t, \frac{nt}{R_t^{(n)}} \land 1 \right).$$
For brevity, write $\tau = \tau_n$, $R_t = R_t^{(n)}$, and $V_t = V_t^{(n)}$. By the notations in Sections 4 and 5, $R = R_\tau$ and $V = V_\tau$. The same relationship holds for the BH procedure (4.4), except that

$$\tau = \tau_n = \sup \{ t \in [0,1] : nt \leq R_t \} \quad (S1.2)$$

### S2 Subcritical case with increasingly sparse false nulls

Define $\eta_t = \eta^{(n)}_t = nt/R_t$, $\theta_t = \theta^{(n)}_t = t/F_n(t)$, and $\rho = 1/\alpha - 1 > 0$. Then $q_n(z;t) = q_{\text{bin}}(z; R_t, \eta_t \wedge 1)$ and by $u_n = \alpha F_n(u_n)$,

$$\theta_{u_n} = \alpha, \quad (\pi_n + \rho)u_n = \pi_n G_n(u_n). \quad (S2.1)$$

Suppose the subcritical condition (5.4) is satisfied. Then we have the following lemmas.

**Lemma S2.1** For procedure (5.1) and the BH procedure (4.4), $\tau/u_n \xrightarrow{P} 1$ and

$$(a) \quad \frac{G_n(\tau)}{G_n(u_n)} \xrightarrow{P} 1, \quad (b) \quad \theta_\tau \xrightarrow{P} \alpha, \quad (c) \quad \eta_\tau \xrightarrow{P} \alpha, \quad (d) \quad \frac{R_\tau}{(\log n)^4} \xrightarrow{P} \infty.$$  

**Lemma S2.2** For both procedures, $\theta_\tau R_\tau - n\tau = \alpha R_{u_n} - nu_n + o_p(\sqrt{nu_n})$.

**Lemma S2.3** For procedure (5.1), $\alpha R_\tau = q_n(\tau; \Gamma(\tau)) + O_p(1)$, and for the BH procedure (4.4), $\alpha R_\tau = n\tau$.

Let $R_{t-} = R_{t-}^{(n)}$ and $V_{t-} = V_{t-}^{(n)}$ be the numbers of rejected nulls and rejected true nulls, respectively, whose $p$-values are strictly less than $t$.

**Lemma S2.4** Given $t \in (0,1)$ and $k > 0$, for procedure (5.1), conditioning on $\tau = t$ and $R_{t-} = k$, $V_{t-} \sim \text{Bin}(k, t/F_n(t))$. The statement holds as well for the BH procedure (4.4).
Recall that if \( p_n \in (0, 1) \) satisfies \( np_n(1 - p_n) \to \infty \), then by Lindeberg’s CLT, for \( S_n \sim \text{Bin}(n, p_n) \) and \( z \in (0, 1) \),

\[
\frac{S_n - np_n}{\sqrt{np_n(1 - p_n)}} \overset{d}{\to} N(0, 1), \quad \frac{\text{qbin}(z; n, p_n) - np_n}{\sqrt{np_n(1 - p_n)}} \to \Phi^*(z). \tag{S2.2}
\]

**Proof of Theorem 5.1.** Assume the Lemmas are true for now. We show (a)–(d) in sequel.

(a) By Lemma S2.1, \( \eta_{\tau} \overset{P}{\to} \alpha \) and

\[
n\tau(1 - \eta_{\tau}) = R_{\tau}\eta_{\tau}(1 - \eta_{\tau}) \sim \text{p}(1 - \alpha)R_{\tau} \overset{P}{\to} \infty.
\]

Then by Lemma S2.3,

\[
P(V_{\tau} \leq \alpha R_{\tau}) = P(V_{\tau} \leq q_{n}(\tau; \Gamma(\tau)) + O_p(1) + o(1)
= P\left(\frac{V_{\tau} - n\tau}{\sqrt{n\tau(1 - \eta_{\tau})}} \leq \frac{q_{n}(\tau; \Gamma(\tau)) - n\tau}{\sqrt{n\tau(1 - \eta_{\tau})}} + o_p(1)\right) + o(1).
\]

Since \( q_{n}(\tau; \Gamma(\tau)) = \text{qbin}(\Gamma(\tau); R_{\tau}, \eta_{\tau}) \) and \( n\tau = R_{\tau}\eta_{\tau} \), Lindeberg’s CLT yields

\[
\frac{q_{n}(\tau; \Gamma(\tau)) - n\tau}{\sqrt{n\tau(1 - \eta_{\tau})}} \sim \text{p}\Phi^*(\Gamma(\tau)) = \sqrt{\frac{1 - \tau}{1 - \alpha}}\Phi^*(1 - \gamma). \tag{S2.3}
\]

Write \( (V_{\tau} - n\tau)/\sqrt{n\tau(1 - \eta_{\tau})} = Z_1 Z + Z_2 \), where

\[
Z_1 = \frac{V_{\tau} - \theta_\tau R_{\tau}}{\sqrt{\theta_\tau(1 - \theta_\tau)R_{\tau}}}, \quad Z = \sqrt{\frac{\theta_\tau(1 - \theta_\tau)R_{\tau}}{n\tau(1 - \eta_{\tau})}}, \quad Z_2 = \frac{\theta_\tau R_{\tau} - n\tau}{\sqrt{n\tau(1 - \eta_{\tau})}}.
\]

By Lemma S2.4, conditioning on \( \tau = t \) and \( R_{\tau-} = k \), \( V_{\tau-} \sim \text{Bin}(k, \theta_\tau) \).

Since \( R_{\tau} - R_{\tau-}, V_{\tau} - V_{\tau-} \in \{0, 1\} \), and \( \theta_\tau R_{\tau} \to \infty \), it follows that \( Z_1 \overset{d}{\to} N(0, 1) \). By Lemma S2.1, \( Z \overset{P}{\to} 1 \). By Lemma S2.2, \( Z_2 = Z_2' + o_p(1) \), where \( Z_2' = (\alpha R_{\tau n} - n\omega_n)/\sqrt{n\tau(1 - \eta_{\tau})} \). From \( F_n(u_n) = u_n/\alpha \to 0 \), \( R_{\tau n} \sim
Bin\((n, F_n(u_n))\), \(\tau/u_n \xrightarrow{p} 1\), and \(\eta_{\tau} \xrightarrow{p} \alpha\), it follows that

\[
Z_2' = \frac{\alpha (R_n - nF_n(u_n))}{\sqrt{nF_n(u_n)(1 - F_n(u_n))}} \sqrt{\frac{u_n(1 - F_n(u_n))}{\alpha \tau (1 - \eta_{\tau})}} \xrightarrow{d} \sqrt{\frac{\alpha}{1 - \alpha}} N(0, 1)
\]

and hence \(Z_2 \xrightarrow{d} \sqrt{\alpha/(1 - \alpha)} N(0, 1)\).

Let \(U_1, U_2\) be iid \(\sim N(0, 1)\). We next show

\[
(Z_1, Z_2) \xrightarrow{d} (U_1, \sqrt{\alpha/(1 - \alpha)} U_2).
\]

Let \(f(x, y) = E(e^{ixZ_1 + iyZ_2})\). Then by Lemma S2.4 and CLT, for any \(a_n \to \infty\) and \(t_n \in (0, 1)\), as long as \(a_n \theta_t(1 - \theta_t) \to \infty\),

\[
\lim_{n \to \infty} E(e^{ixZ_1 | R\tau = a_n, \tau = t_n}) = e^{-x^2/2}.
\]

Since \(Z_2\) is a deterministic function of \(\tau\) and \(R\tau\), by \(R\tau \theta_t(1 - \theta_t) \xrightarrow{p} \infty\) and dominated convergence,

\[
E(e^{ixZ_1 + iyZ_2}) = E(E(e^{ixZ_1 + iyZ_2 | R\tau, \tau})) \sim e^{-x^2/2}E(e^{iyZ_1}) \to \exp \left\{ -\frac{x^2}{2} - \frac{\alpha y^2}{2(1 - \alpha)} \right\}.
\]

Combining all the above results, it follows that

\[
\frac{V_{\tau} - n\tau}{\sqrt{n\tau(1 - \eta_{\tau})}} \xrightarrow{d} \frac{U}{\sqrt{1 - \alpha}} \quad \text{with} \quad U \sim N(0, 1),
\]

which, together with (S2.3) and \(\tau \xrightarrow{p} 0\), implies

\[
P(V_{\tau} \leq \alpha R_{\tau}) \sim P(U \leq \sqrt{1 - \alpha} \Phi^*(\Gamma(\tau))) \to 1 - \gamma.
\]

This completes the proof of part (a).

(b) This directly follows from Lemma S2.1(d).
(c) For the BH procedure (4.4), from (S1.2), \( 0 \leq \alpha R - n\tau \leq \alpha \). By Lemma S2.2,

\[
P(V_{\tau} \leq \alpha R) = P(V_{\tau} \leq n\tau + O(1)) = P(Z_1 \leq Z_2 + o(1))
\]

with

\[
Z_1 = (V_{\tau} - \theta_{\tau}R_{\tau})/\sqrt{\theta_{\tau}(1 - \theta_{\tau})R_{\tau}},
\]

\[
Z_2 = (n\tau - \theta_{\tau}R_{\tau})/\sqrt{\theta_{\tau}(1 - \theta_{\tau})R_{\tau}}.
\]

Following the argument for part (a), \((Z_1, Z_2) \xrightarrow{d} (U_1, \sqrt{\alpha/(1 - \alpha)}U_2)\), where \(U_1\) and \(U_2\) are iid \(\sim N(0, 1)\). As a result, \(P(V_{\tau} \leq \alpha R) \rightarrow 1/2\).

(d) From (5.4), \(n\pi_n \rightarrow \infty\). Since \(n - N_0 \sim \text{Bin}(n, \pi_n)\), then by the weak law of large numbers (WLLN), \((n - N_0)/(n\pi_n) \xrightarrow{P} 1\). By Lemmas S2.1, S2.2, and S2.4,

\[
\frac{R_{\tau}}{nF_n(u_n)} \xrightarrow{P} 1, \quad \frac{R_{\tau} - V_{\tau}}{R_{\tau}} \xrightarrow{P} 1 - \alpha.
\]

Therefore, by (S2.1) and \(\pi_n \rightarrow 0\),

\[
\frac{\psi_n}{G_n(u_n)} = \frac{R_{\tau} - V_{\tau}}{G_n(u_n)(n - N_0)} \sim_p \frac{(1 - \alpha)nF_n(u_n)}{G_n(u_n)n\pi_n} = \frac{\rho u_n}{\pi_nG_n(u_n)} \xrightarrow{P} 1.
\]

The proof for the BH procedure (4.4) is similar and hence is omitted. □

To show the lemmas, the following representation of the p-values \(\xi_{1}^{(n)}, \ldots, \xi_{n}^{(n)}\) will be used. Let \(\zeta_k^{(n)} = F_n(\xi_k^{(n)})\). Then \(\zeta_1^{(n)}, \ldots, \zeta_n^{(n)}\) are iid \(\sim U(0, 1)\). Let

\[
W_t := W_{t}^{(n)} = \# \left\{ k \geq 1 : \zeta_k^{(n)} \leq t \right\},
\]

so that \(R_t = W_{F_n(t)}\). Recall the following result (Shorack & Wellner 1986, p 600). Let \(b_n = \sqrt{2\log \log n}, c_n = 2\log \log n + \log \sqrt{\log \log n} - \log \sqrt{4\pi}, \) and

S-5
\[ Z_t = \frac{(W_t - nt)}{\sqrt{nt(1-t)}}. \] Then for any \( x \in (-\infty, \infty) \), as \( n \to \infty \)

\[ P \left( b_n \sup_{t \in [0,1]} |Z_t| \geq c_n + x \right) \to e^{-4e^x}. \quad (S2.5) \]

From (5.4), it can be seen that as \( n \to \infty \), \( nu_n/(\log n)^4 \to \infty \) and

\[ \frac{\sqrt{n}}{(\log n)^2} \frac{\pi_n^2 [\lambda G_n(u_n) - G_n(\lambda u_n)]}{\sqrt{u_n}} \to \begin{cases} 
\infty, & \text{if } \lambda > 1 \\
-\infty, & \text{if } 0 < \lambda < 1 
\end{cases} \quad (S2.6) \]

Indeed, because \( G_n \) is strictly concave, if \( \lambda > 1 \), (5.4) implies that

\[ \frac{\sqrt{n u_n}}{(\log n)^2} \pi_n \left( \lambda - \frac{G_n(\lambda u_n)}{G_n(u_n)} \right) \to \infty. \]

Then the first limit in (S2.6) follows by (S2.1). The second limit similarly holds.

**Proof of Lemma S2.1** We will only show the Lemma for procedure (5.1).

The proof for the BH procedure (4.4) is similar.

The main part of the proof is devoted to \( \tau /u_n \overset{p}{\to} 1 \). Denote

\[ d_n = \sqrt{n \log \log n}, \quad f_n(t) := \pi_n G_n(t) - (\pi_n + \rho)t. \]

Because \( G_n \) is strictly concave, so is \( f_n \). By (S2.1), \( f_n(u_n) = 0 \). Also, \( f_n(t) \gtrless (\less) 0 \) for \( t \less (\gtr) u_n \). Given \( \lambda > 1 \), let \( v_n = u_n/\lambda \). On the one hand, by (S2.2), \( q_n(v_n; \Gamma(v_n)) \leq nv_n + \sqrt{n v_n} A_n \), with \( A_n \to \Phi^*(\Gamma(v_n)) \). On the other, by (S2.5), for large \( n \), \( R_{vn} = W_{F_n(v_n)} \geq p n F_n(v_n) - 2d_n \sqrt{F_n(v_n)} \). Therefore, by \( F_n(v_n) \leq \alpha v_n \) and \( F_n(v_n) \leq F_n(u_n) = u_n/\alpha \),

\[ R_{vn} - \frac{1}{\alpha} q_n(v_n; \Gamma(v_n)) \]

\[ \geq p n (F_n(v_n) - (1 + \rho)v_n) - \sqrt{n F_n(v_n) / \alpha A_n - 2d_n \sqrt{F_n(v_n)}} \]

\[ \geq p n f_n(v_n) - 3d_n \sqrt{u_n / \alpha}. \]
By \( \pi_n + \rho = \pi_n G_n(u_n)/u_n \),
\[
n f_n(v_n) = n \pi_n \left( G_n(v_n) - \frac{\pi_n G_n(u_n)}{u_n} \right) \geq n \pi_n \left( \frac{G_n(u_n/\lambda) - G_n(u_n)}{\lambda} \right).
\]

Then by (S2.6), \( n f_n(v_n) - 3d_n \sqrt{u_n/\alpha} \xrightarrow{P} \infty \), yielding
\[
\alpha R_{v_n} - q_n(v_n; \Gamma(v_n)) \xrightarrow{P} \infty
\]
and hence \( P(\tau > v_n) \to 1 \).

Now let \( w_n = \lambda u_n \). Then for all \( t \geq w_n \), \( F_n(t) < t/\alpha \). Similar to the above argument, the probability that
\[
R_t - \frac{1}{\alpha} q_n(t; \Gamma(t)) \leq f_n(t) + 3d_n \sqrt{F_n(t)} \leq f_n(t) + 3d_n \sqrt{t/\alpha}, \quad \text{all } t \geq w_n
\]
tends to 1. Because \( f_n(t) \) is concave, it is upper bounded by \( f_n(w_n)t/w_n \).

Note \( f_n(w_n) < 0 \). Therefore, the probability that
\[
R_t - \frac{1}{\alpha} q_n(t; \Gamma(t)) \leq f_n(w_n)t/w_n + 3d_n \sqrt{t/\alpha}
\]
\[
\leq \sqrt{t/w_n} \left( f_n(w_n) + 3d_n \sqrt{w_n/\alpha} \right)
\]
tends to 1. Similar to the above argument, \( x_n \to -\infty \) by (S2.6). Then \( P(\tau \leq w_n) \to 1 \). Together with \( P(\tau > v_n) \to 1 \) and \( \lambda > 1 \) being arbitrary, \( \tau/u_n \xrightarrow{P} 1 \). Now we can show parts (a)–(d) in sequel.

(a) Because \( G_n \) is concave,
\[
G_n(u_n/\lambda) > G_n(u_n)/\lambda, \quad G_n(\lambda u_n) < \lambda G_n(u_n).
\]

Then from \( \tau/u_n \xrightarrow{P} 1 \), \( P(G_n(u_n)/\lambda \leq G_n(\tau) < \lambda G_n(u_n)) \to 1 \) and hence \( G_n(\tau)/G_n(u_n) \xrightarrow{P} 1 \).
(b) Since \( F_n(t) = (1 - \pi_n) t + \pi_n G_n(t) \), by the above results,
\[
F_n(\tau)/F_n(u_n) \overset{P}{\to} 1.
\]
By \( F_n(u_n) = u_n/\alpha \), \( \theta_\tau = \tau/F_n(\tau) \overset{P}{\to} \alpha \).

(c) Given \( \lambda > 1 \), define \( v_n \) and \( w_n \) as above. By \( P(R_\tau > R_{v_n}) \to 1 \) and \( P(\tau < w_n) \to 1 \), \( P(\eta_\tau < n w_n/R_{v_n}) \to 1 \). On the other hand, since \( n v_n \to \infty \), by the CLT in (S2.2), \( R_{v_n} = W_{F_n(v_n)} \sim n F_n(v_n) \) and \( w_n/F_n(v_n) = \lambda^2 v_n/F_n(v_n) \leq \alpha \lambda^2 \). Then \( P(\eta_\tau < \alpha \lambda^2) \to 1 \). Similarly, \( P(\eta_\tau > \alpha/\lambda^2) \to 1 \). Hence \( \eta_\tau \overset{P}{\to} \alpha \).

(d) This follows from \( R_\tau \geq R_{v_n} \sim n F_n(v_n) \geq n v_n \) and \( n v_n/(\log n)^4 \to \infty \).

\[ \square \]

By the weak version of Hungarian construction (Shorack & Wellner 1986, p.494), for each \( n \), there exist a Brownian bridge \( B^{(n)}_t = Z_t - t Z_1 \) and a stochastic process \( r^{(n)}_t \) defined on the same probability space as \( \zeta_1^{(n)}, \ldots, \zeta_n^{(n)} \), where \( Z_t \) is a standard Brownian motion, such that \( \sup_{t \in [0,1]} |r^{(n)}_t| = O_p(1) \) and \( W_t = nt + \sqrt{n} B^{(n)}_t + r^{(n)}_t (\log n)^2 \).

**Proof of Lemma S2.2** By \( R_t = W_{F_n(t)} \) and the Hungarian construction,
\[
\theta_t R_t - nt = \sqrt{n} \theta_t B^{(n)}_{F_n(t)} + (\log n)^2 \theta_t r^{(n)}_{F_n(t)}.
\]
Note \( \theta_{u_n} = \alpha \) and by Lemma S2.1, \( \theta_\tau = O_p(1) \). Since \( (\log n)^2/\sqrt{u_n} \to 0 \) and \( r^{(n)}_t \) is bounded, in order to show Lemma S2.2 for \( \tau \), it is enough to show that
\[
\frac{1}{\sqrt{u_n}} \left[ \theta_\tau B^{(n)}_{F_n(\tau)} - \alpha B^{(n)}_{F_n(u_n)} \right] \overset{P}{\to} 0.
\]
Write the left hand side as $I_1 + \alpha I_2$, where $I_1 = (\theta_\tau - \alpha)B^{(n)}_{F_n(\tau)}/\sqrt{u_n}$ and $I_2 = (B^{(n)}_{F_n(\tau)} - B^{(n)}_{F_n(u_n)})/\sqrt{u_n}$. Given $\lambda > 1$, since $u_n \to 0$, $P(\tau < \lambda u_n) \to 1$ and $F_n(\lambda u_n) < \lambda F_n(u_n) = \lambda u_n/\alpha$, it is seen that $|B^{(n)}_{F_n(\tau)}|$ asymptotically is dominated by $\sup_{t \leq \lambda u_n/\alpha} |B^{(n)}_t|$, hence stochastically dominated by $\sup_{t \leq \lambda u_n/\alpha} |Z_t| \leq (\lambda u_n/\alpha)|Z_1|$. Then $B^{(n)}_{F_n(\tau)}/\sqrt{u_n} = O_P(1)$. By $\theta_\tau \xrightarrow{P} \alpha$, $I_1 \xrightarrow{P} 0$.

Similarly, letting $D_n = \lambda u_n - u_n/\lambda$, $B^{(n)}_{F_n(\tau)} - B^{(n)}_{F_n(u_n)}$ asymptotically is stochastically dominated by $\sup_{t \in [0,D_n]} |Z_t| + D_n|Z_1|$. Therefore, $I_2$ asymptotically is stochastically dominated by $\sqrt{\lambda - 1}/\lambda \sup_{t \in [0,1]} |Z_t| + o_p(1)$. Because $\lambda$ is arbitrary, $I_2 \xrightarrow{P} 0$. \hfill \square

Recall that $\text{qbin}(z; n, p)$ is increasing and left-continuous in $z$ and $p$ respectively; for $z, p \in (0,1)$,

$$\lim_{x \downarrow z} \text{qbin}(x; n, p), \lim_{x \uparrow p} \text{qbin}(z; n, p) \in \{\text{qbin}(z; n, p), \text{qbin}(z; n, p) + 1\};$$

and $\text{qbin}(z; n, p) \leq \text{qbin}(z; n - 1, p) + 1$.

**Proof of Lemma S2.3** By the definition of $\tau$, when $\tau > 0$ and $R_\tau > 0$, for all $t > \tau$, $q_n(t; \Gamma(t)) > \alpha R_\tau$. If $t - \tau > 0$ is small enough, $R_t = R_\tau$. Since $\Gamma(t)$ is decreasing in $t$, $\text{qbin}(\Gamma(t); R_\tau, nt/R_\tau) \geq \alpha R_\tau$. Letting $t \downarrow \tau$ then yields $q_n(\tau, \Gamma(\tau)) \geq \alpha R_\tau - 1$.

On the other hand, there is a sequence $t_j \uparrow \tau$, such that $q_n(t_j; \Gamma(t_j)) \leq \alpha R_{t_j}$. If $R_t$ is continuous at $\tau$, then for large $j$, $R_{t_j} = R_\tau$ and letting $j \to \infty$ yields $q_n(\tau, \Gamma(\tau)) \leq \alpha R_\tau$. If $R_t$ has a jump at $\tau$, then for large $j$, $R_{t_j} = R_\tau - 1$ and letting $j \to \infty$ yields

$$\text{qbin} \left( \Gamma(\tau); R_\tau - 1, \frac{n\tau}{R_\tau - 1} \right) \leq \alpha R_\tau + 1$$

$$\implies \text{qbin} \left( \Gamma(\tau); R_\tau, \frac{n\tau}{R_\tau - 1} \right) \leq \alpha R_\tau + 2.$$
Then \( q_n(\tau, \Gamma(\tau)) - \alpha R_\tau \leq \text{qbin}(\Gamma(\tau); R_\tau, n\tau/(R_\tau - 1)) - \alpha R_\tau \leq 2. \)

The proof that for the BH procedure (4.4) is standard so is omitted. □

**Proof of Lemma S2.4** Let \( \mathcal{F}_t = \sigma(1 \{\xi_i \leq s\}, s \in [t, 1], i = 1, \ldots, n). \) Then for \( t \) running backward from 1 to 0, \( \mathcal{F}_t \) consist a filtration and for both procedure (5.1) and the BH procedure (4.4), \( \tau \) is a stopping time with respect to the filtration. In particular, \( \{\tau \geq t\} \cap \{R_\tau - k\} \in \mathcal{F}_t. \) Let \( i_1, \ldots, i_{R_\tau} \) be the random indices of those \( \xi_i \) that are strictly less than \( t. \) By the independence of \((\xi_1, H_1), \ldots, (\xi_n, H_n)\) and \( V_\tau = H_{i_1} + \cdots + H_{i_{R_\tau}}, \) it is not difficult to see that for any \( t, A \in \mathcal{F}_t, k \geq 0, \) and \( n_1 < \cdots < n_k, \) conditioning on \( E = \{R_\tau = k, i_1 = n_1, \ldots, i_k = n_k\}, V_\tau \) and \( A \) are independent, i.e. \( P(\{V_\tau = v\} \cap A \mid E) = P(\{V_\tau = v\} \mid E) \times P(A \mid E). \) Consequently, \( P(V_\tau = v \mid \tau = t, E) = P(V_\tau = v \mid E). \) By Proposition 2.1, the right end is \( P(S = v), \) with \( S \sim \text{Bin}(k, n/\Gamma(t)). \) Since the conditional probability does not involve \( n_1, \ldots, n_k, \) then \( P(V_\tau = v \mid \tau = t, R_\tau = k) = P(S = v). \) □

### S3 Supercritical case with increasing sparsity of false nulls

Let \( \zeta_k^{(n)} \) be defined as in (S2.4). We need two lemmas in order to prove Theorem 5.2.

**Lemma S3.1** Given \( p_0 \in (0, 1), \) for any \( \epsilon > 0, \)

\[
\lim_{n \to \infty} \sup_{p \in [p_0, 1]} P(\{|X_{1,p} + \cdots + X_{n,p} - np| \geq \epsilon n\}) = 0.
\]

where for each \( p, X_{1,p}, X_{2,p}, \ldots \) are iid \( \sim \) Bernoulli(\( p \)).
Lemma S3.2 If $k_n \leq n$ satisfies $k_n \to \infty$, then

$$
\sup_{k_n \leq k \leq n} \left| \frac{\zeta_{n,k}}{k/n} - 1 \right| \to 0.
$$

Proof of Theorem 5.2 Assume the lemmas are true for now. Fix $\epsilon > 0$ such that $(1 - \epsilon)^2 \lim_{n \to \infty} \alpha^{(n)}(x) > \alpha$. Then by condition (5.5), for $n$ large enough, $(1 - \epsilon)^2 / F^{(n)}(0) = (1 - \epsilon)^2 \alpha^{(n)} / (1 - \pi_n) > \alpha$.

First, we show that for some $K_0 > 0$, $P(R_{\tau} < K_0) \to 1$. Let $m_n(t) = R_t \lor k_n$. Then

$$
n(t \lor \xi_{n,k_n}) \geq \frac{nF^*_n(\zeta_{m_n(t)})}{m_n(t)}.
$$

(S3.1)

By the selection of $k_n$, $m_n(t) \to \infty$. Then by the convexity of $F^*_n(x)$ (because $F_n$ is concave) and Lemma S3.2,

$$
\frac{nF^*_n(\zeta_{m_n(t)})}{m_n(t)} \geq \frac{F^*_n((1 - \epsilon)m_n(t)/n)}{m_n(t)/n} \geq (1 - \epsilon)(F^*_n)'(0) = \frac{1 - \epsilon}{F'_n(0)} \geq \frac{\alpha}{1 - \epsilon}
$$

(S3.2)

and hence by Lemma S3.1, there is $K_0 > 0$, such that for all $K \geq K_0$,

$$
\text{qbin} \left( 1 - \gamma; K, \frac{1 - \epsilon}{F'_n(0)} \right) > (1 - \epsilon)^2 K \frac{F^*_n(0)}{F'_n(0)} \geq \alpha K.
$$

Combined with (S3.1) and (S3.2), this implies

$$
P \left( \bigcap_{t:R_t \geq K_0} \left\{ \text{qbin} \left( 1 - \gamma; R_t, \frac{n(t \lor \xi_{n,k_n})}{R_t \lor k_n} \right) > \alpha R_t \right\} \right) \to 1.
$$

As a result, $P(R_{\tau} < K_0) \to 1$.

Now suppose condition (5.6) is satisfied. We show parts (a)--(d) in sequel.
(a) Note that

\[ R_\tau = \max \left\{ k \geq 1 : \text{qbin} \left( 1 - \gamma; k, \frac{n \xi_{n;k_n}^{(k\vee k_n)}}{k \vee k_n} \wedge 1 \right) \leq \alpha k \right\} \]

Because \( P(R_\tau < K_0) \to 1 \) and \( k_n \to \infty \), \( P(R_\tau = R'_n) \to 1 \), where

\[ R'_n = \max \{ k \geq 1 : \text{qbin} \left( 1 - \gamma; k, p_n \wedge 1 \right) \leq \alpha k \} \]

with \( p_n = n \xi_{n;k_n} / k_n \). Since \( p_n \xrightarrow{P} \alpha_* \), by the properties of qbin as listed before the proof of Lemma S2.3, part (a) then follows from

\[ P(\text{qbin}(1 - \gamma; k, p_n) \in \{ \text{qbin}(1 - \gamma; k, \alpha_*), \text{qbin}(1 - \gamma; k, \alpha_*) + 1 \}) \to 1. \]

(b) Since \( F_n \) is concave, given \( R_\tau = \ell > 0 \) and \( \tau = t \), \( V_\tau \) is stochastically dominated by \( \text{Bin}(\ell, (1 - \pi_n)t/F_n(t)) \), but stochastically dominates \( \text{Bin}(\ell, (1 - \pi_n)/F'_n(0)) \). Because \( \tau \xrightarrow{P} 0 \), \( (1 - \pi_n)\tau /F_n(\tau) \xrightarrow{P} \alpha_* \). Part (b) therefore follows.

(c) Let \( Z_1, Z_2, \ldots \) be iid \( \sim \text{Bernoulli}(\alpha_*) \). If \( \gamma < \gamma_* \), then \( \ell_0 = 0 \), or else there were \( k > 0 \) such that \( P(Z_1 + \ldots + Z_k \leq \alpha k) \geq 1 - \gamma \). Then \( \gamma \geq 1 - P(Z_1 + \ldots + Z_k \leq \alpha k) \geq \gamma_* \), which is a contradiction. It is clear that \( \ell_1 = 0 \). Therefore, by part (a), \( R_\tau \xrightarrow{P} 0 \). On the other hand, if \( \gamma > \gamma_* \), then \( \ell_0 > 0 \). By part (b), \( P(V_\tau \leq \alpha R_\tau | R_\tau = \ell_0) \to P(Z_1 + \ldots + Z_\ell_0 \leq \alpha \ell_0) \geq 1 - \gamma \), and hence \( \lim_n P(V_\tau/R_\tau > \alpha | R_\tau = \ell_0) \leq \gamma \). The case \( R_\tau = \ell_1 \) can be similarly shown as long as \( \ell_1 > 0 \). This completes the proof of (c).

(d) For both procedure (5.1) and the BH procedure (4.4), in order to show that their respective powers tend to 0, by \( P(R \leq K_0) \to 1 \), it is enough to show \( n - N_0 \xrightarrow{P} \infty \). Denote \( s_n = F_n^*(k_n/n) \). Since \( M_n := \# \{ i \leq n : \xi_i^{(n)} \leq s_n, H_i^{(n)} = 1 \} \sim \text{Bernoulli}(n, \pi_n G_n(s_n)) \) and \( M_n \leq n - N_0 \), it is enough to
show $n\pi_n G_n(s_n) \to \infty$. Since $k_n/n = F_n(s_n) = (1 - \pi_n)s_n + \pi_n G_n(s_n)$ and $s_n/F_n(s_n) \to \alpha_* < 1$, 
\[
\frac{\pi_n G_n(s_n)}{k_n/n} = 1 - \frac{(1 - \pi_n)s_n}{F_n(s_n)} \to 1 - \alpha_* > 0,
\]
yielding $n\pi_n G_n(s_n) \sim (1 - \alpha_*)k_n \to \infty$. □

**Proof of Lemma S3.1** It is enough to show
\[
\lim_{n \to \infty} \sup_{p \in [p_0,1]} P(X_1,p + \cdots + X_n,p > (p + \epsilon)n) = 0, \quad \text{and} \quad \lim_{n \to \infty} \sup_{p \in [p_0,1]} P(X_1,p + \cdots + X_n,p < (p - \epsilon)n) = 0.
\]
We will only show the first limit. The second one can be shown similarly.

Clearly, when $p \geq 1 - \epsilon$, $P(X_1,p + \cdots + X_n,p > (p + \epsilon)) = 0$. If $p < 1 - \epsilon$, then by Chernoff’s inequality, $P(X_1,p + \cdots + X_n,p > (p + \epsilon)n) \leq e^{-nI(p)}$, where $I(p) = \sup_{t>0} ((p + \epsilon)t - \Lambda_p(t))$, with $\Lambda_p(t) = \log(1 - p + pe^t)$. Since $\Lambda_p(t)$ is convex and $\Lambda_p'(0) = p$, $I(p) > 0$. It can be verified that $I(p)$ is continuous on $[0,1-\epsilon)$. Letting $I(p) = \infty$ for $p \geq 1 - \epsilon$, it follows that $\inf_{p \geq p_0} I(p) > 0$, which implies the limit. □

**Proof of Lemma S3.2** $\xi_{n:1}, \ldots, \xi_{n:n}$ have the same joint distribution as 
\[
\left( \frac{S_1}{S_{n+1}}, \ldots, \frac{S_n}{S_{n+1}} \right)
\]
where $S_k = U_1 + \cdots + U_k$ and $U_1, U_2, \ldots$ are iid $\sim \text{Exp}(1)$. By the WLLN, $S_{n+1}/n \xrightarrow{P} 1$. Therefore, it is enough to show $\sup_{k \geq k_n} |S_k/k - 1| \xrightarrow{P} 0$, which follows from the strong law of large numbers (SLLN). □
S4  Nonsparse case

**Proof of Theorem 4.1**  The proof of part (a) is omitted because it follows closely Genovese & Wasserman (2002). For part (b), let $R'$ be the number of projections in (4.3). Then by Proposition 4.1, $P(R' > k_n) \to 0$. Since $P(R > 0) \leq P(R' > k_n)$, part (b) follows.  

To prove Theorem 4.2, we need the following standard result for empirical processes.

**Lemma S4.1** Suppose $\tau_n$ is a sequence of random variables taking values in $[0,1]$, such that for some $u \in (0,1)$, $\tau_n \xrightarrow{d} u$ as $n \to \infty$. Then, letting $\pi_0 = 1 - \pi$,

$$\frac{V_{\tau_n} - n\pi_0\tau_n}{\sqrt{n\pi_0u(1-\pi_0u)}} \xrightarrow{d} N(0,1).$$

Following the proof for the sparse case, for procedure (4.7), define

$$\tau_n = \sup \left\{ t \in [0,1] : \text{qbin}\left( \Gamma(t); R_t, \frac{\pi_0n(t \lor \xi_n/k_n)}{R_t \lor k_n} \land 1 \right) \leq \alpha(R_t \lor 1) \right\}$$

and for the BH procedure (4.3), define $\tau_n = \sup \{ t \in [0,1] : \pi_0 nt \leq R_t \}$. Then following the same notations, $R = R_\tau$ and $V = V_\tau$.

The proof of Theorem 4.2 follows closely that of Theorem 5.1, so we only give its sketch.

**Proof of Theorem 4.2**  (a) Following Genovese & Wasserman (2002), $\tau \xrightarrow{P} u^*$, with $u^* \in (0,1)$ the only positive solution to $\pi_0 u = \alpha F(u)$. By the definition of $R$ and $P(\xi_i \leq u^*, H_i = 1) = F(u^*)$, from WLLN, it follows that $R/n \xrightarrow{P} F(u^*) > 0$ and hence pFDR $\sim$ FDR $= \alpha$. Furthermore, $\Gamma_i(t)$ can be replaced with $\Gamma(t)$ and by Lemma S2.3,

$$P(V_\tau \leq \alpha R_\tau) = P\left( V_\tau \leq \text{qbin}\left( \Gamma(\tau); R_\tau, \frac{\pi_0n\tau}{R_\tau} \land 1 \right) + \mathcal{O}_p(1) \right)$$

S-14
Denote the binomial quantile on the right hand side by \( K \). Applying the CLT to the binomial distributions, from \( \tau \overset{d}{\to} u^* \), it follows that

\[
\frac{K - \pi_0 n \tau}{\sqrt{n \pi_0 u^*(1 - \alpha)}} \overset{p}{\to} \Phi^*(\Gamma(u^*)).
\]

Combining this with Lemma S4.1 yields

\[
P(V_{\tau} \leq \alpha R_{\tau}) \to P \left( \sqrt{\frac{1 - \pi_0 u^*}{1 - \alpha}} Z \leq \Phi^*(\Gamma(u^*)) \right) = 1 - \gamma.
\]

(b) Since \( k_n/n \to 0 \) whereas \( R/n \overset{d}{\to} F(u^*) > 0 \), part (b) easily follows. \( \square \)

**Proof of Proposition 4.1**  
(a) Following the proof of Theorem 4.2 (a),

\[
P(V_{\tau} \leq \alpha R_{\tau}) \to P \left( \sqrt{\frac{1 - \pi_0 u^*}{1 - \alpha}} Z \leq 0 \right) = \frac{1}{2},
\]

where, for the BH procedure (4.3),

\[
\tau = \tau_n = \sup \{ t \in [0, 1] : \pi(1 - \pi) nt \leq R_t \}.
\]

(b) See Chi (2006). \( \square \)

The proof of Theorem 4.3 is almost identical to that of Theorem 5.2 and so is omitted.

**Proof of Proposition 4.2**  
When \( \alpha \in (\alpha_s, 1 - \pi) \), then \( \tau \overset{p}{\to} u^* \). Because \( R - V = \# \left\{ k \leq n : H_k = 1, \xi_k^{(n)} \leq \tau \right\} \) and \( n - N_0 = \# \left\{ k \leq n : H_k = 1 \right\} \), by the WLLN, \( \psi_n \overset{p}{\to} P(\xi \leq u^* | H = 1) = G(u^*) \). On the other hand, when \( \alpha < \alpha_s \), then for procedures (4.6) and (4.7), by Theorem 4.1 and 4.2, it is apparent that \( \psi_n = O_p(1/n) \), and for the BH procedure (4.3), from Chi (2006), \( \psi_n \overset{p}{\to} 0 \) as well. \( \square \)