Supplement to “Adaptive False Discovery Rate Control for Heterogeneous Data”

Joshua D. Habiger
Department of Statistics
Oklahoma State University
301G MSCS
Stillwater, OK 74078

June 11, 2015

Abstract

This article contains proofs of theorems, lemmas and corollaries in *Adaptive False Discovery Rate Control for Heterogeneous Data* as well as more detailed discussions of the simulation experiments. Equations, lemmas and figures introduced in this article are prefixed by “S” for supplement.

Keywords: Multiple Testing; P-value; Weighted P-value; Decision Function; Stopping Time; Weak Dependence

1 Proofs of results in Section 3

PROOF OF THEOREM 1. Setting up the Lagrangian

$$L(\boldsymbol{t}, k) = \pi(\boldsymbol{t}, \boldsymbol{p}, \boldsymbol{\gamma}) - k\left[\left(\sum_{m\in\mathcal{M}} t_m\right) - Mt\right]$$

and taking derivative with respect to t_m and setting it equal to 0 yields equation (3). Now, recall we denote the solution to equation (3) with respect to t_m by $t_m(k/p_m, \gamma_m)$ and observe $k \mapsto t_m(k/p_m, \gamma_m)$ is continuous and strictly decreasing in k with $\lim_{k\to\infty} t_m(k/p_m, \gamma_m) = 0$ and $\lim_{k\downarrow 0} t_m(k/p_m, \gamma_m) = 1$ by (A1). Thus, $\bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma}) = M^{-1}\sum_{m\in\mathcal{M}} t_m(k/p_m, \gamma_m)$ is continuous and strictly decreasing in k with $\lim_{k\to\infty} \bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma}) = 0$ and $\lim_{k\downarrow 0} \bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma}) = 1$. Hence, there exists a unique k satisfying $\bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma}) = t$ for any $t \in (0, 1)$ and hence a unique collection $[t_m(k/p_m, \gamma_m), m \in \mathcal{M}]$.

To show that the solution is a maximum, it suffices to show that the sequence of the determinants of the principal minors of the bordered hessian matrix, evaluated at the solution, alternates in sign. The jth principle minor of the bordered Hessian matrix is

$$\boldsymbol{H}_j = \begin{bmatrix} 0 & \mathbf{1}_j^T \\ \mathbf{1}_j & \boldsymbol{D}_j \end{bmatrix}$$

where $\boldsymbol{D}_j$ is a $j \times j$ diagonal matrix with diagonal elements $d_m = \pi''_{\gamma_m}(t_m)$ and $\mathbf{1}_j$ is a vector of 1s of length j. Note that $d_m < 0$ at the solution due to (A1). Now, observe that $|\boldsymbol{H}_1| = -1 < 0$ where $|\cdot|$ denotes the determinant, and for $j \geq 2$, we have the recursive relation

$$|\boldsymbol{H}_j| = d_j|\boldsymbol{H}_{j-1}| + (-1)^j \prod_{m=1}^{j-1}(-d_m). \tag{S1}$$

Because $d_j < 0$, for j an even (odd) integer each expression on the righthand side of equation (S1) is positive (negative). Hence $\{|\boldsymbol{H}_j|, j = 1, 2, ...\}$ alternates in sign. $\|$

Proof of Theorem 2. Observe that $\widetilde{FDP}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma}))$ is continuous in k under (A1). Hence, it suffices to show that $\widetilde{FDP}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma}))$ takes on values 0 and $1 - p_{(M)}$ by the Mean Value Theorem. We first show that

$$\lim_{k \downarrow 0} \widetilde{FDP}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma})) \geq 1 - p_{(M)}.$$

Observe that (A1) implies $t_m \leq \pi_{\gamma_m}(t_m) \leq 1$ for $t_m \in [0, 1]$ and hence

$$\bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma}) \leq \bar{G}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma})) \leq M^{-1}\left[\sum_{m \in \mathcal{M}} (1 - p_m)t_m(k/p_m, \gamma_m) + p_m\right]. \tag{S2}$$

The inequalities in (S2) imply

$$\begin{aligned} \widetilde{FDP}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma})) &= \frac{\sum_{m \in \mathcal{M}}[1 - G_m(t_m(k/p_m, \gamma_m))]}{\sum_{m \in \mathcal{M}}[1 - t_m(k/p_m, \gamma_m)]} \frac{\bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma})}{\bar{G}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma}))} \\ &\geq \frac{\sum_{m \in \mathcal{M}}[1 - p_m][1 - t_m(k/p_m, \gamma_m)]}{\sum_{m \in \mathcal{M}}[1 - t_m(k/p_m, \gamma_m)]} \frac{\bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma})}{\bar{G}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma}))} \\ &\geq \left(1 - p_{(M)}\right) \frac{\bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma})}{\bar{G}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma}))}, \end{aligned}$$

which converges to $1-p_{(M)}$ as $k \downarrow 0$ if

$$\frac{\bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma})}{\bar{G}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma}))} \to 1 \tag{S3}$$

as $k \downarrow 0$. To verify (S3), observe that $\lim_{k\downarrow 0} t_m(k/p_m, \gamma_m) = 1$ by (A1) and hence $\bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma}) \to 1$ as $k \downarrow 0$. This, along with the inequalities in (S2), imply $\bar{G}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma})) \to 1$ as $k \downarrow 0$ and hence (S3) is satisfied.

Now if

$$\lim_{k\to\infty} \frac{\bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma})}{\bar{G}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma}))} = 0, \tag{S4}$$

then by the first inequality in (S2) and the definition of $\widetilde{FDP}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma}))$

$$\widetilde{FDP}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma})) \leq \frac{\bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma})}{\bar{G}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma}))} \to 0$$

as $k \to \infty$ and the proof would be complete. Hence, it suffices to show (S4). But because $t_m(k/p_m, \gamma_m) \downarrow 0$ as $k \to \infty$ and $\pi'_{\gamma_m}(t_m) \to \infty$ as $t_m \downarrow 0$ by (A1), we have

$$\frac{\pi_{\gamma_m}(t_m(k/p_m, \gamma_m))}{t_m(k/p_m, \gamma_m)} \to \infty$$

as $k \to \infty$ by Hôpital's rule. Further for a_m, b_m, $m \in \mathcal{M}$ any positive constants,

$$\frac{\sum_{m\in\mathcal{M}} a_m}{\sum_{m\in\mathcal{M}} b_m} = \sum_{m\in\mathcal{M}} \frac{a_m}{b_m}\left(\frac{b_m}{\sum_{m\in\mathcal{M}} b_m}\right) \geq \min\left\{\frac{a_m}{b_m}, m \in \mathcal{M}\right\}.$$

Hence,

$$A(k) \equiv \frac{\sum_{m\in\mathcal{M}} \pi_{\gamma_m}(t_m(k/p_m, \gamma_m))}{\sum_{m\in\mathcal{M}} t_m(k/p_m, \gamma_m)} \geq \min\left\{\frac{\pi_{\gamma_m}(t_m(k/p_m, \gamma_m))}{t_m(k/p_m, \gamma_m)}, m \in \mathcal{M}\right\} \to \infty$$

as $k \to \infty$ which implies

$$\begin{aligned}\frac{\bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma})}{\bar{G}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma}))} &= \left[\sum_{m\in\mathcal{M}} \frac{(1-p_m)t_m(k/p_m, \gamma_m)}{\bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma})} + \frac{p_m \pi_{\gamma_m}(t_m(k/p_m, \gamma_m))}{\bar{t}_M(k, \boldsymbol{p}, \boldsymbol{\gamma})}\right]^{-1} \\ &\leq \left[M(1-p_{(M)}) + Mp_{(1)}A(k)\right]^{-1} \to 0\end{aligned}$$

as $k \to \infty$, where $p_{(1)} \equiv \min\{\boldsymbol{p}\}$. $\|$

2 Proofs of results in Section 5

Proof of Lemma 1. The proof combines techniques from the proofs of Theorem 3 in Storey et al. (2004) and Theorem 9 in Peña et al. (2011). First, observe that because $u = \lambda$, $0 \leq \hat{t}_\alpha^\lambda \leq \lambda$ by definition and that if $\hat{t}_\alpha^\lambda = 0$ then $FDR(\hat{t}_\alpha^\lambda \boldsymbol{w}) = 0$ trivially. Let us focus on the setting where $0 < \hat{t}_\alpha^\lambda \leq \lambda$. By the definition of $\hat{t}_\alpha^\lambda$, $\widehat{FDP}^\lambda(\hat{t}_\alpha^\lambda \boldsymbol{w}) \leq \alpha$ which gives $R(\hat{t}_\alpha^\lambda) \geq \hat{M}_0(\lambda \boldsymbol{w})\hat{t}_\alpha^\lambda/\alpha$ by the definition of $\widehat{FDP}^\lambda(\cdot)$. Hence,

$$
\begin{aligned}
FDR(\hat{t}_\alpha^\lambda \boldsymbol{w}) &= E\left[\frac{V(\hat{t}_\alpha^\lambda \boldsymbol{w})}{R(\hat{t}_\alpha^\lambda \boldsymbol{w})}\right] \leq E\left[\alpha \frac{1}{\hat{M}_0(\lambda \boldsymbol{w})} \frac{V(\hat{t}_\alpha^\lambda \boldsymbol{w})}{\hat{t}_\alpha^\lambda}\right] \quad \text{(S5)} \\
&\leq E\left[\frac{\alpha}{\hat{M}_0(\lambda \boldsymbol{w})} \frac{V(\lambda \boldsymbol{w})}{\lambda}\right], \quad \text{(S6)}
\end{aligned}
$$

where (S6) is established as follows. First, if $\hat{t}_\alpha^\lambda = \lambda$, it is true trivially. Now suppose that $0 < \hat{t}_\alpha^\lambda < \lambda$. Define filtration $\mathcal{F}_t = \sigma\{\boldsymbol{\delta}(s\boldsymbol{w}), 0 < t \leq s \leq \lambda\}$ and observe that $\hat{t}_\alpha^\lambda$ is a stopping time with respect to $\mathcal{F}_t$ (with time running backwards). Further, for $0 < t \leq \lambda$, $V(t\boldsymbol{w})/t$ is a reverse martingale with respect to $\mathcal{F}_t$. This can be verified by noting that for $0 < s \leq t \leq \lambda$

$$
\begin{aligned}
E\left[\frac{V(s\boldsymbol{w})}{s} | \mathcal{F}_t\right] &= \frac{1}{s} \sum_{m \in \mathcal{M}_0} E\left[\delta_m(sw_m) | \mathcal{F}_t\right] \\
&= \frac{1}{s} \sum_{m \in \mathcal{M}_0} \delta_m(tw_m) E[\delta_m(sw_m) | \delta_m(tw_m) = 1, \mathcal{F}_t] \\
&= \frac{1}{s} \sum_{m \in \mathcal{M}_0} \delta_m(tw_m) E[\delta_m(sw_m) | \delta_m(tw_m) = 1] \\
&= \frac{1}{s} \sum_{m \in \mathcal{M}_0} \delta_m(tw_m) \frac{sw_m}{tw_m} \\
&= \sum_{m \in \mathcal{M}_0} \frac{\delta_m(tw_m)}{t} \\
&= \frac{V(t\boldsymbol{w})}{t},
\end{aligned}
$$

where first equality follows by the definition of $V(\cdot)$ and the second is due to the fact that $\delta_m(sw_m) = 0$ if $\delta_m(tw_m) = 0$ by the NS assumptions. The third equality is satisfied due to (A3). The forth equality follows by the fact that $\Pr([\delta_m(sw_m) = 1] \cap [\delta_m(tw_m) = 1]) =$

$E[\delta_m(sw_m)] = sw_m$ for $m \in \mathcal{M}_0$ and $s \leq \lambda$ under the NS assumptions and under (A2). The forth and fifth equalities follow from some algebra and the definition of $V(\cdot)$, respectively. Hence, by the law of iterated expectation and the Optional Stopping Theorem (Doob, 1953)

$$\begin{aligned} E\left[\frac{\alpha}{\hat{M}_0(\lambda \boldsymbol{w})}\frac{V(\hat{t}_\alpha^\lambda \boldsymbol{w})}{\hat{t}_\alpha^\lambda}\right] &= E\left\{\frac{\alpha}{\hat{M}_0(\lambda \boldsymbol{w})}E\left[\frac{V(\hat{t}_\alpha^\lambda \boldsymbol{w})}{\hat{t}_\alpha^\lambda}|\mathcal{F}_\lambda\right]\right\} \\ &= E\left[\frac{\alpha}{\hat{M}_0(\lambda \boldsymbol{w})}\frac{V(\lambda \boldsymbol{w})}{\lambda}\right]. \end{aligned}$$

Hence, we have established (S6).

Now, note that $M - R(\lambda \boldsymbol{w}) = M_0 - V(\lambda \boldsymbol{w}) + [M_1 - \sum_{\mathcal{M}_1} \delta_m(\lambda w_m)] \geq M_0 - V(\lambda \boldsymbol{w})$. Further observe that $V \mapsto V(\lambda \boldsymbol{w})/[M_0 - V(\lambda \boldsymbol{w}) + 1]$ is convex. Hence, by Theorem 3 in Hoeffding (1956) and with $p = \lambda \bar{w}_0$

$$\begin{aligned} E\left[\frac{V(\lambda \boldsymbol{w})}{M_0 - V(\lambda \boldsymbol{w}) + 1}\right] &\leq \sum_{k=0}^{M_0} \frac{k}{M_0 - k + 1}\binom{M_0}{k} p^k (1-p)^{M_0 - k} \\ &= \frac{p}{1-p}(1 - p^{M_0}). \end{aligned}$$

The last equality follows from basic calculations. Thus,

$$\begin{aligned} E\left[\alpha \frac{1}{\hat{M}_0(\lambda \boldsymbol{w})}\frac{V(\lambda \boldsymbol{w})}{\lambda}\right] &= E\left[\alpha \frac{(1-\lambda)}{M - R(\lambda \boldsymbol{w}) + 1}\frac{V(\lambda \boldsymbol{w})}{\lambda}\right] \\ &\leq \alpha \frac{(1-\lambda)}{\lambda} E\left[\frac{V(\lambda \boldsymbol{w})}{M_0 - V(\lambda \boldsymbol{w}) + 1}\right] \\ &= \alpha \frac{(1-\lambda)}{\lambda}\frac{p}{1-p}(1 - p^{M_0}). \end{aligned}$$

The result follows by plugging $\lambda \bar{w}_0$ in for p in the last expression. $\|$

Proof of Theorem 3. From Lemma 1 and because $\bar{w}_0 \leq w_{(M)}$,

$$FDR(\hat{t}_{\alpha^*}^\lambda \boldsymbol{w}) \leq \alpha^* \bar{w}_0 \frac{1-\lambda}{1-\lambda \bar{w}_0} = \alpha \frac{\bar{w}_0}{w_{(M)}}\frac{1-\lambda w_{(M)}}{1-\lambda \bar{w}_0} \leq \alpha. \|$$

3 Proofs of results in Section 6

Before proving Theorem 4 the following Glivenko-Cantelli-type Lemma regarding the uniform convergence of the FDP estimators and the FDP is presented. For similar results in the unweighted adaptive setting see Theorem 6 in Storey et al. (2004) or see the proof of Theorem 2 in Genovese et al. (2006) for the weighted, but unadaptive, setting. See also Finner et al. (2009); Fan et al. (2012) and references therein for additional results on almost sure convergence of the FDP.

Lemma S1. *Fix $\delta \in (0, u)$. Under (A2) and (A4) - (A6),*

$$\sup_{\delta \le t \le u} |\widehat{FDP}^{0}_{M}(t\boldsymbol{w}_M) - FDP^0_\infty(t)| \to 0,$$

$$\sup_{\delta \le t \le u} |\widehat{FDP}^{\lambda}_{M}(t\boldsymbol{w}_M) - FDP^\lambda_\infty(t)| \to 0,$$

and

$$\sup_{\delta \le t \le u} |FDP_M(t\boldsymbol{w}_M) - FDP_\infty(t)| \to 0$$

almost surely.

Proof. In what follows we denote $\max\{R(t\boldsymbol{w}_M), 1\}$ by $R(t\boldsymbol{w}_M)$ for short. Observe $R(t\boldsymbol{w}_M)$ is nondecreasing in t almost surely by the NS assumptions and $G(t)$ is strictly increasing in t for $0 \le t \le u$ by (A6). Hence, for any $\delta \in (0, u)$,

$$\sup_{\delta \le t \le u} \left|\widehat{FDP}^{0}_{M}(t\boldsymbol{w}_M) - FDP^0_\infty(t)\right| = \sup_{\delta \le t \le u} \left|\frac{t}{R(t\boldsymbol{w}_M)/M} - \frac{t}{G(t)}\right|$$

$$= \sup_{\delta \le t \le u} \left|\frac{t\,[G(t) - R(t\boldsymbol{w}_M)/M]}{G(t)R(t\boldsymbol{w}_M)/M}\right| \le \frac{\sup_{\delta \le t \le u} |G(t) - R(t\boldsymbol{w}_M)/M|}{G(\delta)R(\delta\boldsymbol{w}_M)/M}$$
$$\to \frac{0}{G(\delta)^2} = 0$$

almost surely, where the numerator converges to 0 by the Glivenko-Cantelli Theorem and the denominator converges to $G(\delta)^2$ by (A4) and the Continuous Mapping Theorem.

As for the second claim, denote $\hat{a}^\lambda_{0,M} = \hat{M}_0(\lambda_M \boldsymbol{w}_M)/M$ and $a^\lambda_{0,\infty} = [1 - G(\lambda)]/[1 - \lambda]$.

Additionally observe

$$\widehat{FDP}^{\lambda}_{M}(t\boldsymbol{w}_M) = \hat{a}^{\lambda}_{0,M}\widehat{FDP}^{0}_{M}(t\boldsymbol{w}) \quad \text{and} \quad FDP^{\lambda}_{\infty}(t) = a^{\lambda}_{0,\infty}FDP^{0}_{\infty}(t),$$

Then using the triangle inequality

$$\begin{aligned}
&\sup_{\delta\le t\le u}\left|\widehat{FDP}^{\lambda}_{M}(t\boldsymbol{w}_M) - FDP^{\lambda}_{\infty}(t)\right| = \sup_{\delta\le t\le u}\left|\hat{a}^{\lambda}_{0,M}\widehat{FDP}^{0}_{M}(t\boldsymbol{w}_M) - a^{\lambda}_{0,\infty}FDP^{0}_{\infty}(t)\right| \\
&\le \left|\hat{a}^{\lambda}_{0,M} - a^{\lambda}_{0,\infty}\right| \times \sup_{\delta\le t\le u}\left|\widehat{FDP}^{0}_{M}(t\boldsymbol{w}_M)\right| + \ a^{\lambda}_{0,\infty} \times \sup_{\delta\le t\le u}\left|\widehat{FDP}^{0}_{M}(t\boldsymbol{w}_M) - \widehat{FDP}^{0}_{\infty}(t)\right| \\
&< 2\epsilon + \epsilon,
\end{aligned}$$

where the last inequality is satisfied for all large enough M for any $\epsilon > 0$. To verify the last inequality note that $\hat{a}^{\lambda}_{0,M} \to a^{\lambda}_{0,\infty}$ almost surely by (A2), (A4) and the Continuous Mapping Theorem, and hence $|\hat{a}^{\lambda}_{0,M} - a^{\lambda}_{0,\infty}| < \epsilon$ for all large enough M. Further, for all large enough M,

$$\sup_{\delta\le t\le u}\widehat{FDP}^{0}_{M}(t\boldsymbol{w}_M) < \sup_{\delta\le t\le u} FDP^{0}_{\infty}(t) \ + \epsilon \le 2$$

by the first claim of the Lemma and (A6). Additionally, $G(\lambda) \ge \lambda$ by (A6) and consequently $a^{\lambda}_{0,\infty} \le 1$. Lastly, $\sup_{\delta\le t\le u}|\widehat{FDP}^{0}_{M}(t\boldsymbol{w}_M) - FDP^{0}_{\infty}(t)| < \epsilon$ for all large enough M by the first claim of the Lemma.

To prove the third claim, we first show that

$$\begin{aligned}
&\sup_{\delta\le t\le u}|FDP_M(t\boldsymbol{w}_M) - FDP_\infty(t)| \\
&\le \sup_{\delta\le t\le u}\left|\frac{V(t\boldsymbol{w}_M)}{R(t\boldsymbol{w}_M)} - \frac{a_0\mu_0 t}{R(t\boldsymbol{w}_M)/M}\right| + \sup_{\delta\le t\le u}\left|\frac{a_0\mu_0 t}{R(t\boldsymbol{w}_M)/M} - \frac{a_0\mu_0 t}{G(t)}\right| \\
&= \sup_{\delta\le t\le u}\frac{M}{R(t\boldsymbol{w})}\left|\frac{V(t\boldsymbol{w}_M)}{M} - a_0\mu_0 t\right| \qquad \text{(S7)} \\
&\quad + a_0\mu_0\sup_{\delta\le t\le u}\left|\widehat{FDP}^{0}_{M}(t\boldsymbol{w}_M) - FDP^{0}_{\infty}(t)\right|. \qquad \text{(S8)}
\end{aligned}$$

The inequality is a consequence of the triangle inequality and the definitions of $FDP_\infty(t)$ and $FDP_M(t\boldsymbol{w}_M)$. The expression in (S7) is verified by factoring out $R(t\boldsymbol{w}_M)/M$ in the first expression on the previous line while the expression in (S8) follows from factoring out $a_0\mu_0$ in the second expression and by the definitions of $\widehat{FDP}^{0}_{M}(t\boldsymbol{w}_M)$ and $FDP^{0}_{\infty}(t)$. Now, the quantity in (S8) converges to 0 almost surely because $a_0\mu_0$ is bounded under

(A5) and by the first claim of the Lemma. To show that the first expression converges to 0 almost surely, first note for any $t \in (\delta, u]$, because $R(t\boldsymbol{w}_M)$ is nondecreasing in t, $R(t\boldsymbol{w}_M)/M > G(\delta/2)$ and hence that

$$\frac{M}{R(t\boldsymbol{w}_M)} < \frac{1}{G(\delta/2)}$$

for all large enough M. Hence, if

$$\sup_{\delta \le t \le u} \left| \frac{V(t\boldsymbol{w}_M)}{M} - a_0\mu_0 t \right| \to 0 \tag{S9}$$

almost surely, then

$$\sup_{\delta \le t \le u} \frac{M}{R(t\boldsymbol{w})} \left| \frac{V(t\boldsymbol{w}_M)}{M} - a_0\mu_0 t \right| \le \frac{\epsilon}{G(\delta/2)}$$

for all large enough M and the proof would be completed since ϵ is arbitrary and δ is fixed. To show (S9), first observe that $E[V(t\boldsymbol{w}_M)]/M_0 = \bar{w}_{0,M}t$ under the NS conditions. Also note that by the triangle inequality

$$\begin{aligned}
\sup_{\delta \le t \le u} \left| \frac{V(t\boldsymbol{w}_M)}{M_0} - \mu_0 t \right| &\le \sup_{\delta \le t \le u} \left| \frac{V(t\boldsymbol{w}_M)}{M_0} - \bar{w}_{0,M}t \right| + \sup_{\delta \le t \le u} |\bar{w}_{0,M}t - \mu_0 t| \\
&\le \sup_{\delta \le t \le u} \left| \frac{V(t\boldsymbol{w}_M)}{M_0} - \bar{w}_{0,M}t \right| + u\,|\bar{w}_{0,M} - \mu_0| \to 0
\end{aligned}$$

almost surely, where the first quantity converges to 0 by the Glivenko-Cantelli Theorem and the second quantity converges to 0 because $\bar{w}_{0,M} \to \mu_0$ almost surely under (A5) and because $u \le 1$. Thus, again using the triangle inequality

$$\begin{aligned}
&\sup_{\delta \le t \le u} \left| \frac{V(t\boldsymbol{w}_M)}{M} - a_0\mu_0 t \right| = \sup_{\delta \le t \le u} \left| \frac{V(t\boldsymbol{w}_M)}{M_0} \left[\frac{M_0}{M} + a_0 - a_0 \right] - a_0\mu_0 t \right| \\
&\le \left| \frac{M_0}{M} - a_0 \right| \sup_{\delta \le t \le u} \left| \frac{V(t\boldsymbol{w}_M)}{M_0} \right| + a_0 \sup_{\delta \le t \le u} \left| \frac{V(t\boldsymbol{w}_M)}{M_0} - \mu_0 t \right| \to 0
\end{aligned}$$

almost surely, where the first quantity converges to 0 because $M_0/M \to a_0$ almost surely under (A5) and because $V(t\boldsymbol{w}_M)/M_0 \le 1$, while the second quantity converges to 0 because $a_0 \le 1$ and $V(t\boldsymbol{w}_M)/M_0 \to \mu_0 t$. Hence we have established (S9). □

Proof of Theorem 4. Let us first focus on the equalities. Suppose that $t^0_{\alpha,\infty} < u$. Then $FDP^0_\infty(t^0_{\alpha,\infty}) = \alpha$ by the definition of $t^0_{\alpha,\infty}$ and by (A6). Additionally due to (A6), for any

$\epsilon > 0$ there exists a $0 < \delta < \epsilon$ such that

$$FDP^0_\infty(t^0_{\alpha,\infty} + \delta) < \alpha + \epsilon.$$

Now, Lemma S1 gives $\widehat{FDP}^0_M(t\boldsymbol{w}_M) < FDP^0_\infty(t^0_{\alpha,\infty} + \delta)$ for $0 \leq t < t^0_{\alpha,\infty} + \delta$ and all large enough M. Hence, this and (A6) imply

$$\hat{t}^0_{\alpha,M} = \sup\left[0 \leq t \leq u : \widehat{FDP}^0_M(t\boldsymbol{w}_M) \leq \alpha\right] \leq t^0_{\alpha,\infty} + \delta < t^0_{\alpha,\infty} + \epsilon$$

for all large enough M. Similar arguments give $\hat{t}^0_{\alpha,M} > t^0_{\alpha,\infty} - \epsilon$ for all large enough M. Now if $t^0_{\alpha,\infty} = u$ then

$$t^0_{\alpha,\infty} - \epsilon \leq \hat{t}^0_{\alpha,M} \leq t^0_{\alpha,\infty} = u$$

for all large enough M. Hence, $|\hat{t}^0_{\alpha,M} - t^0_{\alpha,\infty}| < \epsilon$ for all large enough M and we conclude $\hat{t}^0_{\alpha,M} \to t^0_{\alpha,\infty}$ almost surely. As for the second equality, $FDP^\lambda_\infty(t) = a^\lambda_{0,\infty} FDP^0_\infty(t)$ is also continuous and strictly increasing by (A6) and consequently identical argument apply. Thus $\hat{t}^\lambda_{\alpha,M} \to t^\lambda_{\alpha,\infty}$ almost surely.

As for the inequality, note that (A6) implies $\lambda \leq G(\lambda)$ which implies

$$a^\lambda_{0,\infty} = \frac{1 - G(\lambda)}{1 - \lambda} \leq 1. \tag{S10}$$

Hence,

$$FDP^\lambda_\infty(t) = a^\lambda_{0,\infty} FDP^0_\infty(t) \leq FDP^0_\infty(t) \tag{S11}$$

for every $t \in (0, u]$. This, (A6) and the definitions of $FDP^0_\infty(\cdot)$, $t^0_{\alpha,\infty}$ and $t^\lambda_{\alpha,\infty}$ imply $t^0_{\alpha,\infty} \leq t^\lambda_{\alpha,\infty}$. $\|$

Proof of Theorem 5. By Lemma S1 and (A6), for $0 < s < t \leq u$

$$\begin{aligned} FDP_M(t\boldsymbol{w}_M) - FDP_M(s\boldsymbol{w}_M) \quad &> \\ a_0\mu_0 t/G(t) - a_0\mu_0 s/G(s) - 2 \sup_{0 \leq t \leq u} &|FDP_M(t\boldsymbol{w}_M) - a_0\mu_0 t/G(t)| \\ &\to a_0\mu_0[t/G(t) - s/G(s)] > 0 \end{aligned}$$

almost surely. Claim (C1) is then a consequence of Theorem 4 and the Continuous Mapping Theorem. To verify Claims (C2) and (C3), first observe that by the triangle inequality

$$\begin{aligned} & |FDP_M(\hat{t}^{\lambda}_{\alpha,M}\boldsymbol{w}_M) - FDP_\infty(t^{\lambda}_{\alpha,\infty})| \\ \leq \ & |FDP_M(\hat{t}^{\lambda}_{\alpha,M}\boldsymbol{w}_M) - FDP_\infty(\hat{t}^{\lambda}_{\alpha,M})| + |FDP_\infty(\hat{t}^{\lambda}_{\alpha,M}) - FDP_\infty(t^{\lambda}_{\alpha,\infty})|. \end{aligned}$$

The first quantity converges to 0 almost surely by Lemma S1 and the second quantity converges to 0 almost surely by Theorem 4 and the Continuous Mapping Theorem. Hence, $FDP_M(\hat{t}^{\lambda}_{\alpha,M}\boldsymbol{w}_M) \to FDP_\infty(t^{\lambda}_{\alpha,\infty})$ almost surely. Thus to prove Claims (C2) and (C3) it suffices to show that $FDP_\infty(t^{\lambda}_{\alpha,\infty}) \leq \alpha$ if $\mu_0 \leq 1$, with equality when $G(t)$ is a DU distribution with $\mu_0 = 1$ and $FDP^{\lambda}_\infty(u) \geq \alpha$. To show this, consider the following:

$$\begin{aligned} FDP_\infty(t^{\lambda}_{\alpha,\infty}) &= a_0\mu_0 \frac{t^{\lambda}_{\alpha,\infty}}{G(t^{\lambda}_{\alpha,\infty})} \\ &\leq a_0 \frac{t^{\lambda}_{\alpha,\infty}}{G(t^{\lambda}_{\alpha,\infty})} \\ &\leq \frac{1-G(\lambda)}{1-\lambda} \frac{t^{\lambda}_{\alpha,\infty}}{G(t^{\lambda}_{\alpha,\infty})} \\ &= FDP^{\lambda}_\infty(t^{\lambda}_{\alpha,\infty}) \\ &\leq \alpha. \end{aligned}$$

The first equality is due to the definition of $FDP_\infty(\cdot)$. The first inequality is satisfied when $\mu_0 \leq 1$ and is an equality when $\mu_0 = 1$. As for the second inequality, note that $G(\lambda) \leq a_0\lambda + 1 - a_0$ when $\mu_0 \leq 1$ and $G(\lambda) = a_0\lambda + 1 - a_0$ under a DU distribution with $\mu_0 = 1$. Consequently

$$a_0 = \frac{1 - [a_0\lambda + 1 - a_0]}{1-\lambda} \leq \frac{1-G(\lambda)}{1-\lambda}$$

when $\mu_0 \leq 1$ and the inequality is an equality when G is a DU distribution with $\mu_0 = 1$. The last equality is satisfied by the definition of $FDP^{\lambda}_\infty(\cdot)$. The last inequality is satisfied by the definition of $t^{\lambda}_{\alpha,\infty}$ and is an equality when G is a DU distribution with $\mu_0 = 1$ and $FDP_\infty(u) \geq \alpha$ because these conditions imply $FDP_\infty(u) = FDP^{\lambda}_\infty(u) \geq \alpha$. That is, $FDP_\infty(u)$ is continuous and monotone and takes on value α. Hence, $FDP_\infty(t^{\lambda}_{\alpha,\infty}) \leq \alpha$ if $\mu_0 \leq 1$ with equality if G is a DU distribution with $\mu_0 = 1$ and $FDP_\infty(u) \geq \alpha$. $\|$

Proof of Theorem 6. Under the conditions of the theorem

$$\begin{aligned} Cov(W_{m,M}, \theta_{m,M}) &= E[W_{m,M}|\theta_{m,M} = 1]E[\theta_{m,M}] - E[W_{m,M}]E[\theta_{m,M}] \\ &= E[\theta_{m,M}](E[W_{m,M}|\theta_{m,M} = 1] - 1). \end{aligned}$$

Hence, $Cov(W_{m,M}, \theta_{m,M}) \geq 0$ implies $E[W_{m,M}|\theta_{m,M} = 1] \geq 1$ and consequently $E[W_{m,M}|\theta_{m,M} = 0] \leq 1$, with equality if $Cov(W_{m,M}, \theta_{m,M}) = 0$. Hence $E[\bar{W}_{0,M}|\boldsymbol{\theta}_M \neq \mathbf{1}_M] = \mu_0 \leq 1$ with equality if $Cov(W_{m,M}, \theta_{m,M}) = 0$. The result follows because $\bar{W}_{0,M} \to \mu_0$ almost surely. ‖

Proof of Corollary 1. Observe that $u = 1$ and $\lambda < 1$ is fixed. Hence (A2) is satisfied and (A4) - (A6) are satisfied by the conditions of the Theorem. Therefore Claim (C1) holds by Theorem 5. Now, additionally note that $\mu_0 = 1$ if $\boldsymbol{w}_M = \mathbf{1}_M$ and that $FDP_\infty(1) = a_0 \geq \alpha$ under the conditions of the Theorem. Thus Claims (C2) and (C3) hold by Theorem 5. ‖

Before proving Theorem 7 we provide Lemma S2. It will be used to verify that optimal weights are weakly dependent so that decision functions satisfy the weak dependence structure defined in (A4) - (A5). Below, denote $t_0(k) = E[t_m(k/p_m, \gamma_m)]$ and denote $G(t_0(k)) = E[\delta_m(t_m(k/p_m, \gamma_m))]$, where the expectations are taken over all random quantities, i.e. over $(Z_m, \theta_m, p_m, \gamma_m)$ for some fixed $k > 0$. Further, define

$$\widetilde{FDP}_\infty(t_0(k)) = \frac{1 - G(t_0(k))}{1 - t_0(k)} \frac{t_0(k)}{G(t_0(k))}$$

and

$$k^* = \inf\{k : \widetilde{FDP}_\infty(t_0(k)) = \alpha\},$$

and denote

$$\tilde{w}_{m,\infty}(k^*, p_m, \gamma_m) = U_m t_m(k^*/p_m, \gamma_m)/t_0(k^*).$$

Lemma S2. *Suppose that* $\Pr(p_m \leq 1 - \alpha)$. *Under Model 1 and (A1),* $k^*_M \to k^*$ *almost surely and*

$$\tilde{w}_{m,M}(k^*_M, \boldsymbol{p}, \boldsymbol{\gamma}) \to \tilde{w}_{m,\infty}(k^*, p_m, \gamma_m)$$

almost surely.

Proof. Note that $0 < \alpha \leq 1 - p_{(M)}$ with probability 1 so that k^*_M is well defined for $M = 1, 2, ...$ by Theorem 2. Further, observe that $t_m(k^*/p_m, \gamma_m)$, $m = 1, 2, ...$ are i.i.d. continuous random variables taking values in $[0, 1]$ under Model 1. Hence, by the Strong Law of Large numbers $\bar{t}_M(k^*, \boldsymbol{p}, \boldsymbol{\gamma}) \to t_0(k^*)$ almost surely. Likewise, $\bar{G}_M(\boldsymbol{t}(k^*, \boldsymbol{p}, \boldsymbol{\gamma}) \to G(t_0(k^*))$ almost surely and by the Continuous Mapping Theorem

$$\widetilde{FDP}_M(\boldsymbol{t}(k^*, \boldsymbol{p}, \boldsymbol{\gamma})) \to \widetilde{FDP}_\infty(t_0(k^*))$$

almost surely. Because further $\widetilde{FDP}_M(\boldsymbol{t}(k, \boldsymbol{p}, \boldsymbol{\gamma}))$ and $\widetilde{FDP}_\infty(t_0(k))$ are both continuous in k by (A1), we have from the Continuous Mapping Theorem and the definitions of k^*_M and k^* that $k^*_M \to k^*$ almost surely. Thus,

$$\begin{aligned} \tilde{\boldsymbol{w}}_{m,M}(k^*_M, \boldsymbol{p}, \boldsymbol{\gamma}) &= U_m \boldsymbol{w}_{m,M}(k^*_M, \boldsymbol{p}, \boldsymbol{\gamma}) \\ &= U_m \frac{t_m(k^*_M/p_m, \gamma_m)}{\bar{t}_M(k^*_M, \boldsymbol{p}, \boldsymbol{\gamma})} \\ &\to U_m \frac{t_m(k^*/p_m, \gamma_m)}{t_0(k^*)} = \tilde{w}_{m,\infty}(k^*, p_m, \gamma_m) \end{aligned}$$

almost surely by the Continuous Mapping Theorem.

□

Proof of Theorem 7. First we verify (A2). Observe $\lambda_M = \bar{t}_M(k^*_M, \boldsymbol{p}, \boldsymbol{\gamma}) \to t_0(k^*)$ almost surely by the Strong Law of Large Numbers and the Continuous Mapping Theorem, where recall $t_0(k^*) = E[t_m(k^*/p_m, \gamma_m)]$. Thus, by the definition of $\tilde{w}_{m,M}$

$$\lim_{M\to\infty} \tilde{w}_{m,M} = \lim_{M\to\infty} \frac{U_m t_{m,M}(k^*_M/p_m, \gamma_m)}{\bar{t}_M(k^*_M, \boldsymbol{p}, \boldsymbol{\gamma})} \leq \frac{1}{t_0(k^*)}$$

almost surely, where the last inequality is due to the Continuous Mapping Theorem, Lemma (S2) and because $U_m t_m(k^*_M, \boldsymbol{p}, \boldsymbol{\gamma}) \leq 1$ almost surely by construction. That is, (A2) is satisfied with $\lambda = u = 1/t_0(k^*)$.

Before verifying (A4) - (A6) we introduce some notation. Denote

$$G^{k^*}(t) = E[\delta_m(t\tilde{w}_{m,\infty}(k^*, p_m, \gamma_m))]$$

where the expectation is taken over all random quantities, i.e. taken over $(Z_m, \theta_m, p_m, \gamma_m, U_m)$. Further we sometimes suppress $\boldsymbol{p}$ and $\boldsymbol{\gamma}$ and write $\tilde{w}_{m,\infty}(k^*) = \tilde{w}_{m,\infty}(k^*, p_m, \gamma_m)$, $\tilde{w}_{m,M}(k^*) = \tilde{w}_{m,M}(k^*, \boldsymbol{p}, \boldsymbol{\gamma})$ and $\tilde{\boldsymbol{w}}_M(k^*) = [\tilde{w}_{m,M}(k^*), m \in \mathcal{M}]$. Further, denote $\tilde{\boldsymbol{w}}_\infty(k^*) = [\tilde{w}_{m,\infty}(k^*), m \in \mathcal{M}]$.

Now consider (A4). Observe that $\delta_m(t\tilde{w}_{m,\infty}(k^*)), m = 1, 2, ...$ are i.i.d. Bernoulli random variables with success probability $G^{k^*}(t)$ under Model 1 so that

$$\frac{R(t\tilde{\boldsymbol{w}}_\infty(k^*))}{M} = \frac{\sum_{m\in\mathcal{M}} \delta_m(t\tilde{w}_{m,\infty}(k^*))}{M} \to G^{k^*}(t)$$

almost surely by the Strong Law of Large Numbers. Further, by the NS assumptions, Lemma S2, and because $G^{k^*}(t)$ is continuous, we have that for any $\epsilon > 0$ there exists an $\epsilon' > 0$ such that

$$\begin{aligned} \frac{R(t\tilde{\boldsymbol{w}}_M(k^*_M))}{M} &= \frac{\sum_{m\in\mathcal{M}} \delta_m(t\tilde{w}_{m,M}(k^*_M))}{M} < \frac{\sum_{m\in\mathcal{M}} \delta_m(t[\tilde{w}_{m,\infty}(k^*) + \epsilon'])}{M} \\ &< G^{k^*}(t + t\epsilon') < G^{k^*}(t) + \epsilon \end{aligned}$$

for all large enough M. Similar arguments give

$$\frac{R(t\tilde{\boldsymbol{w}}_M(k^*_M))}{M} > G^{k^*}(t) - \epsilon$$

for all large enough M. Hence, $R(t\tilde{\boldsymbol{w}}_M(k^*))/M \to G^{k^*}(t)$ almost surely. Then because $k^*_M \to k^*$ almost surely by Lemma S2, $R(t\tilde{\boldsymbol{w}}_M(k^*_M))/M \to G^{k^*}(t)$ almost surely by the Continuous Mapping Theorem.

As for (A5), recall the NS conditions give $E[\delta_m(t_m)|\theta_m = 0] = t_m$. Hence, taking the expectation over all random quantities, we have by the law of iterated expectation

$$E[(1 - \theta_m)\delta_m(t\tilde{w}_{m,\infty}(k^*, p_m, \gamma_m))] = a_0\mu_0 t,$$

where $a_0 = E[1 - \theta_m]$ and $\mu_0 = E[\tilde{w}_{m,\infty}(k^*, p_m, \gamma_m)|\theta_m = 0]$. Then, arguments akin to those in the proof of (A4) give

$$\frac{V(t\tilde{\boldsymbol{w}}_M(k^*_M))}{M} = \frac{M_0}{M}\frac{V(t\tilde{\boldsymbol{w}}_M(k^*_M))}{M_0} \to a_0\mu_0 t$$

almost surely.

For (A6), first observe that $G^{k^*}(t) = a_0\mu_0 t + (1-a_0)G_1(t)$ for $t \le u$, where

$$G_1(t) = E\left[\pi_{\gamma_m}(t\tilde{w}_{m,\infty}(k^*, p_m, \gamma_m))\right]$$

and the expectation is taken over all random quantities. Clearly $t \mapsto G_1(t)$ is concave and twice differentiable because $t \mapsto \pi_{\gamma_m}(t)$ is twice differentiable almost surely by (A1). To see that $t/G(t) \to 0$ as $t \downarrow 0$ note that $G_1'(t) \to \infty$ as $t \downarrow 0$ because $\pi'_{\gamma_m}(t) \to \infty$ as $t \downarrow 0$ almost surely by (A1). Hence,

$$\frac{t}{G^{k^*}(t)} = \frac{t}{a_0\mu_0 t + (1-a_0)G_1(t)} \to 0$$

as $t \downarrow 0$ by an application of Hôptial's rule. Clearly, $\lim_{t\uparrow u} t/G^{k^*}(t) \to u/G^{k^*}(u)$ because $G^{k^*}(t)$ is continuous. To see that $u/G^{k^*}(u) \le 1$ we establish the following:

$$\begin{aligned}
G^{k^*}(u) &= E[\delta_m(u\tilde{w}_{m,\infty}(k^*))] \\
&= a_0 E[\delta_m(u\tilde{w}_{m,\infty}(k^*))|\theta_m = 0] \\
&\quad +(1-a_0)E[\delta_m(u\tilde{w}_{m,\infty}(k^*))|\theta_m = 1] \\
&= a_0 E[u\tilde{w}_{m,\infty}(k^*)] + (1-a_0)E[\pi_{\gamma_m}(u\tilde{w}_{m,\infty}(k^*))] \\
&\ge a_0 E[u\tilde{w}_{m,\infty}(k^*)] + (1-a_0)E[u\tilde{w}_{m,\infty}(k^*)] \\
&= E[u\tilde{w}_{m,\infty}(k^*)] \\
&= uE[\tilde{w}_{m,\infty}(k^*)] = u.
\end{aligned}$$

The first equality is by the definition of $G^{k^*}(u)$ while the second equality is due to the law of iterated expectation. The third is a consequence of the definition of $\pi_{\gamma_m}(t)$ and the NS assumptions. The inequality is satisfied because $\pi_{\gamma_m}(t) \ge t$ almost surely for every $t \in [0,1]$ under (A1). The forth equality is obvious. As for the fifth, recall $E[U_m|p_m, \gamma_m] = 1$, $\tilde{w}_{m,\infty}(k^*) = U_m w_{m,\infty}(k^*)$ and that $E[w_{m,\infty}(k^*)] = 1$. Hence, by the law of iterated expectation $E[\tilde{w}_{m,\infty}(k^*)] = E[w_{m,\infty}(k^*)] = 1$.

To verify that $\mu_0 \le 1$ we make use of Theorem 6 and write $W_m = w_{m,M}(k^*_M, \boldsymbol{p}, \boldsymbol{\gamma})$ and $\tilde{W}_m = U_m W_m$ for short. First let us focus on $Cov(W_m, \theta_m)$. From the law of iterated expectation,

$$Cov(W_m, \theta_m) = E[Cov(W_m, \theta_m | p_m)] + Cov(E[W_m | p_m], p_m). \tag{S12}$$

Observe that

$$\begin{aligned} Cov(W_m, \theta_m | p_m) &= E[W_m \theta_m | p_m] - E[W_m | p_m] E[\theta_m | p_m] \\ &= p_m E[W_m | p_m] - p_m E[W_m | p_m] = 0 \end{aligned}$$

which implies that the first expectation in (S12) is 0. To compute the second expectation, first observe $\pi'_{\gamma_m}(t_m)$ is continuous and strictly decreasing and hence the solution to $\pi'_{\gamma_m}(t_m) = a$, denoted $t_m(a, \gamma_m)$, is continuous and strictly decreasing in a almost surely by (A1). Hence $t_m(k^*_M / p_m, \gamma_m)$ is strictly increasing and continuous in p_m almost surely. Thus,

$$E[W_m | \boldsymbol{p}, \boldsymbol{\gamma}] = E\left[M \frac{t_m(k^*_M / p_m, \gamma_m)}{t_m(k^*_M / p_m, \gamma_m) + \sum_{j \neq m} t_j(k^*_M / p_j, \gamma_j)} \middle| \boldsymbol{p}, \boldsymbol{\gamma} \right]$$

is also increasing in p_m almost surely because the function $x/(x+a)$ for a any positive constant is increasing in x for $x > 0$. This implies $E[W_m | p_m]$ is also increasing in p_m almost surley which implies $Cov(E[W_m | p_m], p_m) \geq 0$. As for $\tilde{W}_m = U_m W_m$,

$$Cov(\tilde{W}_m, \theta_m) = E[Cov(U_m W_m, \theta_m | W_m)] + Cov(E[U_m W_m | W_m], E[\theta_m | W_m])$$

by the law of iterated expectation. But

$$E[Cov(U_m W_m, \theta_m | W_m)] = E[W_m Cov(U_m, \theta_m | W_m)] = 0$$

because $Cov(U_m, \theta_m | W_m)$ is 0 by construction. Additionally,

$$Cov(E[U_m W_m | W_m], E[\theta_m | W_m]) = Cov(W_m, E[\theta_m | W_m]) \geq 0$$

because $Cov(W_m, \theta_m) \geq 0$. Hence, $Cov(\tilde{W}_m, \theta_m) \geq 0$ and thus, by Theorem 6, $\mu_0 \leq 1$. $\|$

Proof of Theorem 8. First recall from the proof of Theorem 7 (where here we take

$U_m = 1$ almost surely for every m) that $\lambda_M = \bar{t}_M(k_M^*) \to t_0(k^*)$ Hence, we have

$$FDP_\infty^\lambda(t) = \frac{1 - G^{k^*}(t_0(k^*))}{1 - t_0(k^*)} \frac{t}{G^{k^*}(t)}.$$

Further observe that because $t/G^{k^*}(t)$ is strictly increasing by (A6), then $t_0(k^*) = t_{\alpha,\infty}^\lambda$ by the definition of $t_{\alpha,\infty}^\lambda$. Hence $\bar{t}_M(k_M^*) \to t_0(k^*) = t_{\alpha,\infty}^\lambda$ almost surely. $\|$

Proof of Corollary 2 First observe that $Cov(w_{m,M}, \theta_{m,M}) = 0$ and hence $\mu_0 = 1$ by Theorem 6. It therefore suffices to show that (A4) - (A6) are satisfied. But $\delta_m(tw_{m,M})$, $m = 1, 2, ...$ are i.i.d. Bernoulli random variables under Model 1 and the conditions of the Theorem. Hence, $R(t\boldsymbol{w}_M)/M \to G(t)$ for $G(t) = E[\delta_m(t\boldsymbol{w}_{m,M})]$ almost surely by the Strong Law of Large Numbers and (A4) is satisfied. Likewise $(1 - \theta_{m,M})\delta_m(tw_{m,M})$, $m = 1, 2, ...$ are i.i.d. random variable with mean $a_0 t$ under the NS assumptions and the conditions of the Theorem. Hence,

$$\frac{V(t\boldsymbol{w}_M)}{M} = \frac{1}{M} \sum_{m \in \mathcal{M}} (1 - \theta_{m,M})\delta_m(tw_{m,M}) \to a_0 t$$

almost surely by the Strong Law of Large Numbers and (A5) is satisfied. Condition (A6) is verified using arguments identical to those used in verifying (A6) in the proof of Theorem 7 with $G^{k^*}(t) = G(t)$ and $w_{m,M} = \tilde{w}_{m,\infty}(k^*)$. $\|$

Proof of Corollary 3. Observe that $\pi(\boldsymbol{t}, \boldsymbol{p}, \boldsymbol{\gamma})$ is proportional to $\pi(\boldsymbol{t}, \boldsymbol{1}, \boldsymbol{\gamma})$ and hence the maximization of $\pi(\boldsymbol{t}, \boldsymbol{p}, \boldsymbol{\gamma})$ with respect to $\boldsymbol{t}$ is free of p_m. Thus $\tilde{w}_{m,M}(k, \boldsymbol{p}, \boldsymbol{\gamma})$ is independent of p_m and hence independent of θ_m. The result then follows from Theorems 6 and 7. $\|$

4 Simulation details

4.1 Simulations 1 - 4

In Simulation 1, observe that the FDP is increasing in a for both adaptive procedures. For example the FDP of each procedure is 0.021 when $a = 1$ but is 0.039 when $a = 5$. This

Table 1: The average CDP (FDP) for the UU, UA, WU, and WA procedures in Simulations 1 - 4.

	Simulation 1			Simulation 2		
	a			a		
	1	3	5	1	3	5
UU	0.007(0.021)	0.390(0.025)	0.709(0.025)	0.007(0.024)	0.390(0.025)	0.709(0.025)
WU	0.007(0.021)	0.397(0.025)	0.731(0.025)	0.011(0.008)	0.434(0.014)	0.756(0.016)
UA	0.007(0.021)	0.437(0.031)	0.761(0.039)	0.007(0.024)	0.430(0.031)	0.761(0.041)
WA	0.007(0.021)	0.442(0.032)	0.793(0.039)	0.011(0.008)	0.473(0.018)	0.814(0.026)

	Simulation 3			Simulation 4		
	a			a		
	1	3	5	1	3	5
UU	0.007(0.023)	0.391(0.025)	0.709(0.025)	0.007(0.025)	0.391(0.025)	0.710(0.025)
WU	0.013(0.007)	0.404(0.015)	0.719(0.016)	0.006(0.023)	0.354(0.025)	0.682(0.025)
UA	0.007(0.023)	0.430(0.031)	0.757(0.039)	0.007(0.025)	0.425(0.030)	0.756(0.039)
WA	0.013(0.007)	0.439(0.019)	0.774(0.027)	0.006(0.023)	0.387(0.030)	0.727(0.039)

is to be expected as both adaptive procedures are α-exhaustive (see Corollaries 1 and 3) and hence we should expect the FDP to be near 0.05 in high power settings, i.e. for large a. Additionally, the largest gain in power (in terms of the average CDP) of the weighted adaptive procedure over the unweighted adaptive procedure occurs when effect sizes are most heterogeneous. When $a = 5$ the average CDP of the WA procedure is 0.793 while the average CDP of the UA procedure is 0.761. When data are homogeneous (a = 1), the CDPs of the procedures are identical.

In Simulation 2, data generating mechanisms are even more heterogeneous as now the p_ms also vary. General conclusions regarding the CDP are the same, with the advantages of the weighted procedures over their unweighted counterparts being more pronounced. For example, the average CDP of the WAMDF for $\gamma_m \overset{i.i.d.}{\sim} Un(1, 5)$ increased from 0.793 to 0.814 when allowing p_ms to vary, while for the UA procedure the CDP is still 0.761. We also observe that for $a = 5$ the average FDP of the WA procedure is only 0.026 while the average FDP of the UA procedure is closer to 0.05; it is 0.039. This is to be expected because, even though the WAMDF will dominate the UA procedure in terms of the average CDP, the UA procedure is α-exhaustive while the WAMDF need not be in this setting.

Now consider non-optimal weights in Simulations 3 and 4. Roeder and Wasserman (2009) concluded that, in the unadaptive setting (the UU and WU procedures), weighted MDFs are robust with respect to weight misspecification in that they generally yield about as many or more rejected null hypotheses as unweighted procedures as long as weights are "reasonably guessed" and yield slightly less rejected null hypotheses when weights are poorly guessed. Simulations 3 and 4 confirm their results and further illustrates that the robustness property applies to adaptive procedures. For example, comparing the unadaptive procedures in Simulation 3, we see that the average CDP of the WU(UU) procedures are 0.013(0.007), 0.404(0.391), and 0.719(0.709) for $a = 1, 3, 5$, respectively. The average CDP of the WA(UA) procedure is 0.013(0.007), 0.439(0.430), and 0.774(0.757), for $a = 1, 3, 5$, respectively. That is, when weights are positively correlated with optimal weights, weighted procedures still perform slightly better than their unweighted counterparts. In the worst case scenario setting in Simulation 4, where weights are independently generated, the FDP is still controlled by the WA procedure, but some loss in power over its unweighted counterpart is observed. For example, the CDP of the WA(UA) procedure is 0.006(0.007), 0.386(0.425), and 0.727(0.756) when $\gamma = 1, 3$, and 5, respectively, while the average FDP of the WA(UA) procedure is 0.025(0.023), 0.030(0.030), and 0.039(0.039) when $\gamma = 1, 3$, and 5, respectively.

4.2 Simulation 5

The average CDP ratio (weighted/unweighted) vs. the average CDP of the weighted procedure is depicted in Figure S1 for all settings. Observe that the CDP ratio is greater than or equal to 1 for each value of p and α as long as the CDP is at least 0.2.

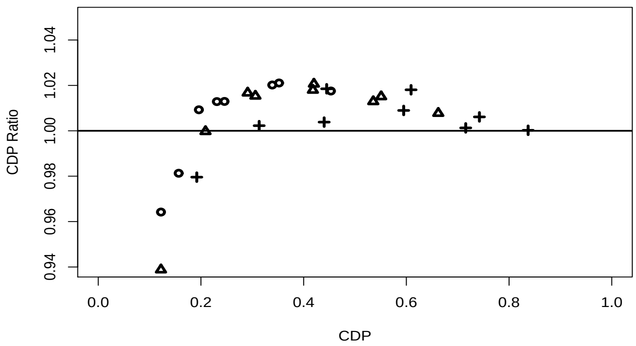

Figure S1: The ratio of the average CDP (weighted/unweighted) vs. the average CDP of the weighted procedure for $p = 0.2(o), p = 0.5(\triangle)$, and $p = 0.8(+)$ for $\bar{\gamma} = 1.75, 2, 2.25$.

References

Doob, J. (1953). Stochastic Processes. New York: John Wiley and Sons.

Fan, J., X. Han, and W. Gu (2012). Estimating false discovery proportion under arbitrary covariance dependence. J. Amer. Statist. Assoc. 107(499), 1019–1035.

Finner, H., T. Dickhaus, and M. Roters (2009). On the false discovery rate and an asymptotically optimal rejection curve. Ann. Statist. 37(2), 596–618.

Genovese, C., K. Roeder, and L. Wasserman (2006). False discovery control with p-value weighting. Biometrika 93(3), 509–524.

Hoeffding, W. (1956). On the distribution of the number of successes in independent trials. Ann. Math. Statist. 27, 713–721.

Peña, E. A., J. D. Habiger, and W. Wu (2011). Power-enhanced multiple decision functions controlling family-wise error and false discovery rates. Ann. Statist. 39(1), 556–583.

Roeder, K. and L. Wasserman (2009). Genome-wide significance levels and weighted hypothesis testing. Statist. Sci. 24(4), 398–413.

Storey, J. D., J. E. Taylor, and D. Siegmund (2004). Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: A unified approach. J. R. Stat. Soc. Ser. B. Stat. Methodol. 66(1), 187–205.