A new adaptive method to control the false discovery rate

By FANG LIU and SANAT K. SARKAR

Temple University

Summary

Benjamini, Krieger and Yekutieli (BKY, 2006) have given an adaptive method of controlling the false discovery rate (FDR) by incorporating an estimate of $n_0$, the number of true null hypotheses, into the FDR controlling method of Benjamini and Hochberg (BH, 1995). The BKY method improves the BH method in terms of the FDR control and power. Benjamini, Krieger and Yekutieli have proved that their method controls the FDR when the $p$-values are independent and provided numerical evidence showing that the control over the FDR continues to hold when the $p$-values have some type of positive dependence. In this paper, we propose an alternative adaptive method via a different estimate of $n_0$. Like the BKY method, this new method controls the FDR under the independence, and can maintain a control over the FDR, as shown numerically, under the same type of positive dependence of the $p$-values. More importantly, as our simulations indicate, the proposed method can often outperform the BKY method in terms of the FDR control and power, particularly when the correlation between the test statistics is moderately low or the proportion of true null hypotheses is very high. When applied to a real microarray data, the new method is seen to pick up a few more significant genes than the BKY method.

Some key words: Adaptive BH methods; False discovery rate; Multiple testing;

1. Introduction

Multiple hypothesis testing plays a pivotal role in analyzing data from modern scientific investigations, such as DNA microarray, functional magnetic resonance imaging (fMRI),
and many other biomedical studies. For instance, identification of differentially expressed genes across various experimental conditions in a microarray study or active voxels in an fMRI study is carried out through multiple testing. Since these investigations typically require tens and thousands of hypotheses to be tested simultaneously, the traditional multiple testing methods, like those designed to control the probability of at least one false rejection, the familywise error rate (FWER), become too conservative to use in these investigations. Benjamini and Hochberg (1995) introduced the false discovery rate (FDR), the expected proportion of false rejections among all rejections, which is less conservative than the FWER and has become the most popular measure of type I error rate in modern multiple testing.

Benjamini and Hochberg (1995) gave a method, referred to as the BH method, for controlling the FDR. The FDR of this method at level $\alpha$ is equal to $n_0 \alpha / n$, where $n_0$ is the number of true null hypotheses, when the underlying test statistics are independent, and less than or equal to $n_0 \alpha / n$ when these statistics are positively dependent in a certain sense [Benjamini and Hochberg (1995), Benjamini and Yekutieli (2001) and Sarkar (2002)]. Since $n_0$ is unknown, by estimating it and modifying the BH method using this estimate can potentially make the BH method less conservative and thus more powerful. A number of such adaptive BH methods have been proposed in the literature, among which the one in Benjamini, Krieger and Yekutieli (2006) has received much attention and will be our main focus in this paper.

We consider estimating $n_0$ using a different estimate than the one considered in Benjamini, Krieger and Yekutieli (2006) before modifying the BH method. Like the BKY method, this new adaptive version of the BH method is proved to control the FDR when the $p$-values are independent and numerically shown to control the FDR under normal distributional setting with equal positive correlation. Moreover, as our simulations indicate, it outperforms the BKY method, in the sense of providing better FDR control and power, when the correlation between the test statistics is moderately low or the proportion of true null hypotheses is quite large.

This paper is organized as follows. We start with a background in the next section for our proposed method providing notations, the definition of the FDR, and some basic formulas. Section 3 revisits some FDR controlling methods, especially adaptive FDR
controlling methods. The new estimate of \( n_0 \) is proposed in Section 4. Our proposed alternative version of adaptive BH method based on this new \( n_0 \) estimate is developed in Section 5. The results of a simulation study conducted to investigate the FDR controlling property and power performance of our proposed method relative to the BKY method are also presented in Section 5. Both BKY and the new adaptive FDR methods are applied to a real microarray data; the comparative results are presented in Section 6. The paper concludes with some final remarks made in Section 7.

2. Notation, definition and formulas

Consider testing \( n \) null hypotheses \( H_1, \ldots, H_n \) simultaneously against certain alternatives using their respective \( p \)-values \( p_1, \ldots, p_n \). A multiple testing of these hypotheses is typically carried out using a stepwise or single-step procedure. Let \( p_{1:n} \leq \cdots \leq p_{n:n} \) be the ordered versions of these \( p \)-values, with \( H_{1:n}, \ldots, H_{n:n} \) being their corresponding null hypotheses. Then, given a non-decreasing set of critical constants \( 0 < \alpha_1 \leq \cdots \leq \alpha_n < 1 \), a step-up procedure rejects the set \( \{ H_{i:n}, \ i \leq i^*_{SU} \} \) and accepts the rest, where \( i^*_{SU} = \max\{1 \leq i \leq n : p_{i:n} \leq \alpha_i\} \), if the maximum exists, otherwise accepts all the null hypotheses. A step-down procedure, on the other hand, rejects the set of null hypotheses \( \{ H_{i:n}, \ i \leq i^*_{SD} \} \) and accepts the rest, where \( i^*_{SD} = \max\{1 \leq i \leq n : p_{j:n} \leq \alpha_j \ \forall \ j \leq i\} \), if the maximum exists, otherwise accepts all the null hypotheses. When the constants are same in a step-up or step-down procedure, it reduces to what is defined as a single-step procedure.

Let \( R \) denote the total number of rejections and \( V \) denote the number of those that are false, the type I errors, while testing \( n \) null hypotheses using a multiple testing method. Then, the FDR of this method is defined by

\[
\text{FDR} = E(\text{FDP}) \text{, where } \text{FDP} = \frac{V}{\max\{R, 1\}}
\]

(1)
is the false discovery proportion. Different formulas for the FDR of a stepwise procedure - step-up, step-down or single-step - have been considered in different papers [see, for example, Benjamin and Yekutielli (2001), Sarkar (2002, 2006)]. However, we will present an alternative expression for the FDR, given recently in Sarkar (2008b), that provides better insight and will be of use in the present paper.
For any multiple testing method,

$$\text{FDP} = \frac{V}{\max\{R, 1\}} = \sum_{i \in I_0} \sum_{r=1}^{n} \frac{1}{r} \mathbb{I}(H_i \text{ is rejected, } R = r),$$

where $I_0$ is the set of indices of true null hypotheses. For a step-up procedure, this expectation can be written more explicitly as follows, with $P_i$ denoting the random variable corresponding to the observed $p$-value $p_i$.

**Formula 2.1.** For a step-up procedure of testing the $n$ null hypotheses $H_1, \ldots, H_n$ using the critical values $\alpha_1 \leq \cdots \leq \alpha_n$, the FDR is given by

$$\text{FDR} = \sum_{i \in I_0} E \left[ \frac{I(P_i \leq \alpha_{R_{SU,n-1}^{(-i)}(\alpha_2, \ldots, \alpha_n) + 1})}{R_{SU,n-1}^{(-i)}(\alpha_2, \ldots, \alpha_n) + 1} \right],$$

where $R_{SU,n-1}^{(-i)}(\alpha_2, \ldots, \alpha_n)$ is the number of rejections in testing the $n - 1$ null hypotheses other than $H_i$ using the step-up procedure based on their $p$-values and the critical constants $\alpha_2 \leq \cdots \leq \alpha_n$.

By taking $\alpha_i = c$, for all $i = 1, \ldots, n$, in the above formula, one gets the following formula for a single-step procedure that rejects $H_i$ if $p_i \leq c$:

$$\text{FDR} = \sum_{i \in I_0} E \left[ \frac{I(P_i \leq c)}{R_{n-1}^{(-i)}(c) + 1} \right],$$

where $R_{n-1}^{(-i)}(c)$ is the number of rejections in testing the $n - 1$ null hypotheses other than $H_i$ using the single-step procedure based on the $p$-values other than $p_i$.

**Formula 2.2.** For a step-down procedure of testing the $n$ null hypotheses $H_1, \ldots, H_n$ using the critical constants $\alpha_1 \leq \cdots \leq \alpha_n$, the FDR satisfies the following inequality:

$$\text{FDR} \leq \sum_{i \in I_0} E \left[ \frac{I(P_i \leq \alpha_{R_{SD,n-1}^{(-i)}(\alpha_1, \ldots, \alpha_{n-1}) + 1})}{R_{SD,n-1}^{(-i)}(\alpha_1, \ldots, \alpha_{n-1}) + 1} \right],$$

where $R_{SD,n-1}^{(-i)}(\alpha_1, \ldots, \alpha_{n-1})$ is the number of rejections in testing the $n - 1$ null hypotheses other than $H_i$ using the step-down procedure based on their $p$-values and the critical constants $\alpha_1 \leq \cdots \leq \alpha_{n-1}$. 
A number of FDR controlling methods have been proposed in the literature, among which the BH method has received the most attention. In this section, we will briefly review this and some of its adaptive versions.

3.1. The BH Method

The BH method is a step-up procedure with the critical values $\alpha_i = i\alpha/n$, $i = 1, \ldots, n$; that is, it rejects the null hypotheses $H_{1:n}, \ldots, H_{r:n}$ and accepts the rest, where

$$ r = \max \left\{ 1 \leq i \leq n : p_{i:n} \leq \frac{i}{n} \alpha \right\}, $$

provided this maximum exists; otherwise, accepts all the null hypotheses.

These critical values are the same ones as Simes (1986) originally considered while testing the global null hypotheses $H_0 = \bigcap_{i=1}^n H_i$. Simes also proposed to use them in a step-up manner for multiple testing of the $H_i$'s upon rejection of the global null hypothesis. However, as an FWER controlling method at level $\alpha$, it works only in a weak sense, that is, when all the null hypotheses are true, with the $p$-values being either independent (Simes, 1986) or positively dependent in a certain sense [Sarkar and Chang (1997), Sarkar (1998, 2008a)], but it does not work in a strong sense, that is, under any configuration of true and false null hypotheses, even when the $p$-values are independent (Hommel, 1988). Benjamini and Hochberg (1995) showed that this step-up procedure can be used to control the FDR in a strong sense, at least when the $p$-values are independent. In particular, they proved that $\text{FDR} \leq n_0 \alpha/n$ for this method when the $p$-values are independent with each having $U(0,1)$ distribution under the corresponding null hypothesis.

Later, it was proved that the FDR of the BH method is actually equal to $n_0 \alpha/n$ under the independence of the $p$-values [Benjamini and Yekutieli (2001), Finner and Roters (2001), Sarkar (2002, 2008b), Storey, Taylor and Siegmund (2004), of course, assuming that a null $p$-value is distributed as $U(0,1)$], and is less than or equal to $n_0 \alpha/n$ under the following type of positive dependence among the $p$-values:

$$ \mathbb{E} \left\{ \psi (P_1, \ldots, P_n) \mid P_i = u \right\} \text{ is non-decreasing in } u \text{ for each } i \in I_0, $$

for any (coordinatewise) non-decreasing function $\psi$ [Benjamini and Yekutieli (2001), Sarkar (2002, 2008b)]. This is referred to as the positive regression dependence on subset
(PRDS) condition, which is satisfied by a number of multivariate distributions arising in many multiple testing situations, among which the multivariate normal with non-negative correlations is the most common. Other commonly arising multivariate distributions for which the BH method works are multivariate $t$ with the associated multivariate normal with non-negative correlations (when $\alpha \leq 1/2$), absolute valued multivariate $t$ with the associated normals being independent and some type of multivariate $F$ [Benjamini and Yekutieli (2001), Sarkar (2002, 2004)].

Sarkar (2002) proved that the step-down analog of the BH method, that is, the method that rejects the null hypotheses $H_{1:n}, \ldots, H_{r:n}$ and accepts the rest, where

$$r = \max \left\{ 1 \leq i \leq n : p_{j:n} \leq \frac{j}{n} \alpha \text{ for all } j = 1, \ldots, i \right\},$$

provided this maximum exists, otherwise, accepts all the null hypotheses, also controls the FDR under the independence or the same type of positive dependence as above for the $p$-values.

The positive dependence condition required for the FDR control of the BH method or its step-down analog can be slightly relaxed from (4) to the following:

$$E \left\{ \psi(P_1, \ldots, P_n) \mid P_i \leq u \right\} \text{ is non-decreasing in } u \text{ for each } i \in I_0, \quad (5)$$

for any (coordinatewise) non-decreasing function $\psi$ [Finner, Dickhaus and Roters (2009) and Sarkar (2008b)].

If $n_0$ were known, the step-up procedure with the critical values $\alpha_i = i\alpha/n_0$ for $i = 1, \ldots, n$, would control the FDR precisely at the desired level $\alpha$, when the $p$-values are independent. This has been the rationale for considering an adaptive version of the BH method that looks for a way to estimate $n_0$ with $\hat{n}_0$ from the available data and modifies the BH critical values to $\hat{\alpha}_i = i\alpha/\hat{n}_0$ for $i = 1, \ldots, n$. We will briefly review a number of such adaptive BH methods in the following subsections.

3.2. The Adaptive BH Method of Benjamini & Hochberg

Benjamini and Hochberg (2000) introduced this adaptive BH method for independent $p$-values based on an estimate of $n_0$ developed using the so called the lowest slope (LSL) method.

When all the null hypotheses are true and the test statistics are independent, the $p$-values should be iid as $U(0,1)$ with the expectations of the ordered $p$-values as
\( E(P_{i:n}) = i/(n+1), \ i = 1, \ldots, n. \) Therefore, the plot of \( p_{i:n} \) versus \( i \) should exhibit a linear relationship, along the line with the slope \( S = 1/(n+1) \) and passing through the origin and the point \((n+1, 1)\) (assuming \( p_{n+1:n} = 1 \)).

When \( n_0 \leq n \), the \( p \)-values corresponding to the false null hypotheses tend to be small, so they concentrate on the left side of the above plot. The relationship over the right side of the plot remains approximately linear with the slope \( \beta = 1/(n_0+1) \). Therefore, using a suitable set of the largest \( p \)-values, a straight line through the point \((n+1, 1)\) can be fitted with slope \( \hat{\beta} \) and \( n_0 \) can be estimated as \( \hat{n}_0 = 1/\hat{\beta} \). Benjamini and Hochberg (2000) suggested estimating \( n_0 \) using the LSL method and the corresponding adaptive BH method as follows:

1. Apply the original BH method. If none is rejected, accept all hypotheses and stop; otherwise, continue.
2. Calculate the slopes \( S_i = 1 - p_{i:n}/(n+1-i) \).
3. Start with \( i = 1 \), proceed as long as \( S_i \geq S_{i-1} \) and stop when the first time \( S_j < S_{j-1} \).
   Let \( \hat{n}_0^{BH} = \min[n, 1/S_j + 1] \).
4. Apply the BH method with \( \alpha_i = i\alpha/\hat{n}_0^{BH} \).

Though there is no theoretical proof that this version of adaptive BH method guarantees an FDR control, simulation studies indicate that it does.

### 3.3. The Adaptive BH Method of Storey, Taylor and Siegmund

Storey, Taylor and Siegmund (2004) used the following estimate of \( n_0 \):

\[ \hat{n}_0^{STS}(\lambda) = \frac{n - R(\lambda) + 1}{1 - \lambda}, \]  
where \( R(\lambda) = \#\{P_i \leq \lambda\}, \) for some \( \lambda \in [0, 1) \), and considered the adaptive method with the critical values \( \alpha_i = \min\{i\alpha/\hat{n}_0^{STS}, \lambda\}, \ i = 1, \ldots, n. \) It controls the FDR under the independence of the \( p \)-values [Benjamini, Krieger and Yekutieli (2006), Storey, Taylor and Siegmund (2004), Sarkar (2004, 2008b)], as well as under certain form of weak dependence asymptotically as \( n \to \infty \) [Storey, Taylor and Siegmund (2004)].

This adaptive BH method is closely connected to Storey’s (2002) estimation based approach to controlling the FDR. Storey (2002) derived a class of point estimates of the
FDR for a single-step test that rejects $H_i$ if $p_i \leq t$, for some fixed threshold $t$, under the following model:

**Mixture Model.** Let $P_i$ denote the random $p$-value corresponding to $p_i$ and $H_i = 0$ or 1 according as the associated null hypothesis is true or false. Let $(P_i, H_i), i = 1, \ldots, n,$ be independently and identically distributed with $\Pr(P_i \leq u \mid H_i) = (1 - H_i)u + H_i F_1(u)$, $u \in (0, 1)$, for some continuous cdf $F_1(u)$, and $\Pr(H_i = 0) = \pi_0 = 1 - \Pr(H_i = 1)$.

Having proved that the FDR of the above single-step test for this mixture model is given by

$$\text{FDR}(t) = \frac{\pi_0 t}{F(t)} \Pr\{R(t) > 0\},$$

where

$$F(t) = \Pr(P_i \leq t) = \pi_0 t + (1 - \pi_0) F_1(t),$$

[see also Liu and Sarkar (2009)], Storey (2002) proposed the following class of point estimates of the FDR$(t)$:

$$\hat{\text{FDR}}(t) = \frac{\hat{\pi}_0(\lambda) t}{\hat{F}(t)}, \quad \lambda \in [0, 1)$$

where

$$\hat{F}(t) = \frac{1}{n} \max\{R(t), 1\} \quad \text{and} \quad \hat{\pi}_0(\lambda) = \frac{\hat{n}_0}{n} = \frac{n - R(\lambda)}{n(1 - \lambda)}.$$  

This estimate of $n_0$ was originally suggested by Schweder and Spjotvoll (1982) in a different context. Storey (2002) showed that $E(\hat{\text{FDR}}(t)) \geq \text{FDR}(t)$, that is, $\hat{\text{FDR}}(t)$ is conservatively biased as an estimate of FDR$(t)$, which he argued is desirable, because by controlling it one can control the true FDR$(t)$. He suggested using

$$t_\alpha = \sup\{0 \leq t \leq 1 : \hat{\text{FDR}}(t) \leq \alpha\}$$

(11)

to threshold the $p$-values, that is, to use it as the cut-off point below which a $p$-value should be declared significant at a level $\alpha$. He pointed out that if one approximates $t_\alpha$ by $p_{i_\alpha(\lambda); n}$, that is, rejects the null hypotheses $H_{1:n}, \ldots, H_{i_\alpha(\lambda); n}$, where

$$\hat{i}_\alpha(\lambda) = \max\{1 \leq i \leq n : \hat{\text{FDR}}(p_i; n) \leq \alpha\},$$

then one gets the BH method when $\lambda = 0$. For $\lambda \neq 0$, thresholding the $p$-values at $p_{i_\alpha(\lambda); n}$ is same as using an adaptive BH method. Unfortunately, however, the FDR of such an
The adaptive BH method is not less than or equal to $\alpha$, even under independence, unless the $\hat{n}_0$ in (10) is modified, which Storey, Taylor and Siegmund (2004) did.

3-4. The Adaptive BH Method of Benjamini, Krieger and Yekutieli

Unlike Storey (2002) or Storey, Taylor and Siegmund (2004) where $n_0$ is estimated based on the number of significant $p$-values observed in a single-step test with an arbitrary critical value $\lambda$, Benjamini, Krieger and Yekutieli (2006) considered estimating $n_0$ from the BH method at level $\alpha/(1 + \alpha)$. Their adaptive version of the BH method, the BKY method, runs as follows:

1. Apply the BH method at level $q = \frac{\alpha}{1 + \alpha}$. Let $r_1$ be the number of rejections. If $r_1 = 0$, accept all the null hypotheses and stop; if $r_1 = n$, reject all the null hypotheses and stop; otherwise continue to the next step.

2. Estimate $n_0$ as

$$\hat{n}_0^{BKY} = \frac{n - r_1}{1 - q} = (n - r_1)(1 + \alpha).$$

3. Apply the BH method with the critical values $\alpha_i = i\alpha/\hat{n}_0^{BKY}$, $i = 1, \ldots, n$.

As Benjamini, Krieger and Yekutieli (2006) have proved, the BKY method controls the FDR at $\alpha$ under independence of the $p$-values. While it is less powerful than the adaptive method proposed in Storey, Taylor and Siegmund (2004) when the $p$-values are independent, simulation studies have shown that with the test statistics generated from multivariate normals with common positive correlations it can also control the FDR [Benjamini, Krieger and Yekutieli (2006) and Romano, Shaikh and Wolf (2008)].

Benjamini, Krieger and Yekutieli (2006) also extended the BKY method to a multiple-stage procedure (MST) by repeating the two-stage procedure as long as more hypotheses are rejected, which is stated as follows:

1. Let $r = \max\{i : \text{for all } j \leq i, \text{ there exists } l \geq j \text{ so that } p_{vn} \leq \alpha l/[n + 1 - j(1 - \alpha)]\}$. 

2. If such an $r$ exists, reject $p_{1:n}, \ldots, p_{r:n}$; otherwise reject no hypotheses.

This multiple-stage procedure is a combination of step-up and step-down procedures. They offered no analytical proof of its FDR control. Benjamini, Krieger and Yekutieli (2006) also mentioned that a multiple-stage step-down procedure (MST) can be developed by choosing $l = j$ in MST. They provided numerical results showing that the MST
method can also control the FDR, the theoretical justification of which is given later in Gavrilov, Benjamini and Sarkar (2009) to be reviewed in the following section.

3.5. The Adaptive Method of Gavrilov, Benjamini and Sarkar

As mentioned above, Gavrilov, Benjamini and Sarkar (2009) re-examined the multiple-stage step-down procedure, the MSD procedure, mentioned in Benjamini, Krieger and Yekutieli (2006) and proved that this multiple-stage step-down procedure can control the FDR under the independence of the $p$-values. The following is the MSD method:

Find $k = \max\{1 \leq i \leq n : p_{j:n} \leq j\alpha/(n + 1 - j(1 - \alpha)) \text{ for all } j = 1, \ldots, i\}$ and reject $H_{1:n}, \ldots, H_{k:n}$ if $k$ exists; otherwise reject no hypotheses.

Although it has been referred to as a multiple-stage step-down procedure by Benjamini, Krieger and Yekutieli (2006), it is actually, as Sarkar (2008b) argued, an adaptive version of the step-down analog of the BH method considered in Sarkar (2002). To see this, first note that, under the same setup involving the mixture model and a constant rejection threshold $t$ for each $p$-value as in Storey (2002) or Storey, Taylor and Siegmund (2004), one can consider estimating $n_0$ based on the number of significant $p$-values compared to the $t$ itself, rather than a different arbitrary constant $\lambda$. In other words, by considering the Storey, Taylor and Siegmund (2004) type estimate of $n_0 = n\pi_0$ with $\lambda = t$ and using this estimate in $\hat{\text{FDR}}_{\lambda}(t)$, Storey’s original estimate of the FDR$(t)$, one can develop the following alternative estimate of FDR$(t)$:

$$
\hat{\text{FDR}}^*(t) = \frac{[n - R(t) + 1]t}{(1 - t) \max\{R(t), 1\}}.
$$

A step-down procedure developed through this estimate, that is, the one that rejects $H_{1:n}, \ldots, H_{r:n}$ where

$$
r = \max\{1 \leq i \leq n : \hat{\text{FDR}}^*(p_{j:n}) \leq \alpha \text{ for all } j = 1, \ldots, i\}
$$

$$
= \max\{1 \leq i \leq n : \frac{p_{j:n}}{1 - p_{j:n}} \leq \frac{j\alpha}{n - j + 1} \text{ for all } j = 1, \ldots, i\},
$$

(14)

which is same as the MSD, is an adaptive version of the step-down analog of the BH method.

There are some other methods to estimate $n_0$ in the literature, such as parametric beta-uniform mixture model by Pounds and Morris (2003), the Spacing LOESS Histogram
(SPLOSH) method by Pounds and Cheng (2004), the nonparametric MLE method by Langaas and Lindqvist (2005), the moment generating function approach by Broberg (2005), and the resampling strategy by Lu and Perkins (2007). These other \( n_0 \) estimates could also be used while developing adaptive versions of the BH method or its step-down analog. However, whether or not any of these can control the FDR theoretically, at least when the \( p \)-values are independent, is an important open problem.

4. A NEW ESTIMATE OF \( n_0 \)

We present in this section the new estimate of \( n_0 \) and the results of a simulation study comparing this estimate to \( \hat{n}_0^{STS} \) and \( \hat{n}_0^{BKY} \) before we use it to propose our version of adaptive BH method in the next section.

4.1. The Estimate

Our estimate of \( n_0 \) is developed somewhat along the line of that in the BKY method. However, instead of deriving it from the number of significant \( p \)-values in the original BH method at level \( q = \alpha/(1 + \alpha) \), which is being done in the BKY method, we consider deriving it from the number of significant \( p \)-values in the step-down analog of the BH method at the same level \( q \) but using a formula that is similar to that in Storey, Taylor and Siegmund (2004). More specifically, our proposed estimate of \( n_0 \) is given by:

\[
\hat{n}_0^{NEW}(\gamma) = \frac{n - k + 1}{1 - \gamma_{k+1}}, \tag{15}
\]

where \( k \) is the number of rejections in the step-down version of the BH method with the critical values \( \gamma_i = i\gamma/n \), for \( i = 1, \ldots, n \), where \( \gamma = \alpha/(1 + \alpha) \) and \( \gamma_{n+1} \in [\gamma, (1 + \gamma)/2) \). The choice of \( \gamma_{n+1} \) in this particular interval is dictated by our main result proved in the section that for such \( \gamma_{n+1} \) the FDR of the corresponding adaptive BH method can be controlled at \( \alpha \), at least when the \( p \)-values are independent.

The results presented in the following section favoring \( \hat{n}_0^{NEW} \) as an estimate of \( n_0 \) over \( \hat{n}_0^{BKY} \) provide some rationale for our choice of this new estimate.

4.2. Simulation Study

We ran a simulation study to investigate numerically how \( \hat{n}_0^{NEW} \) performs compared to \( \hat{n}_0^{STS} \) (with \( \lambda = 0.5 \)) and \( \hat{n}_0^{BKY} \) as an estimate of \( n_0 \). We generated \( n \) dependent random variables \( X_i \sim N(\mu_i, 1) \), \( i = 1, \ldots, n \), with a common non-negative correlation
Table 1. The estimated mean and variance of $\hat{n}_0^{NEW}$, $\hat{n}_0^{STS}$ and $\hat{n}_0^{BKY}$ for the cases of $n = 5000$, $n_0 = 2500$ and $\rho = 0, 0.25, 0.5$ and 0.75.

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NEW</td>
<td>BKY</td>
<td>STS</td>
<td>NEW</td>
<td>BKY</td>
</tr>
<tr>
<td>$\rho = 0$</td>
<td>4996</td>
<td>5242</td>
<td>3296</td>
<td>33.87</td>
<td>55.82</td>
</tr>
<tr>
<td>$\rho = 0.25$</td>
<td>4927</td>
<td>5166</td>
<td>3284</td>
<td>69133</td>
<td>81661</td>
</tr>
<tr>
<td>$\rho = 0.5$</td>
<td>4862</td>
<td>5088</td>
<td>3280</td>
<td>279038</td>
<td>325556</td>
</tr>
<tr>
<td>$\rho = 0.75$</td>
<td>4881</td>
<td>5008</td>
<td>3276</td>
<td>448641</td>
<td>712831</td>
</tr>
</tbody>
</table>

$\rho$, and determined their $p$-values for testing $\mu_i = 0$ against $\mu_i > 0$. We repeated this 10,000 times by setting $n$ at 5000, $\rho$ at 0, 0.25, 0.5, 0.75, the proportion of the true null hypotheses $\pi_0$ at 0, 0.25, 0.5, 0.75 and 1, the value of $\mu_i$ for each false null hypothesis at 1, and the value of $\alpha$ at 0.05. Each time, we calculated the values of the three estimates. From these 10,000 values, we constructed the boxplot and calculated the estimated mean and variance for each estimate. We present these boxplots in Figure 1 and the estimated means and variances in Table 1 only for $\pi_0 = 0.5$, as they provide very similar comparative pictures for other values of $\pi_0$.

As seen from Figure 1 and Table 1, $\hat{n}_0^{NEW}$ is a better estimate of $n_0$ than $\hat{n}_0^{BKY}$. Looking at $\hat{n}_0^{STS}$ and comparing it to the other two, one notices that although it is more centrally located at the true $n_0$, it is more variable, and the variability increases quite dramatically with increasing $\rho$. The variabilities of both $\hat{n}_0^{NEW}$ and $\hat{n}_0^{BKY}$, on the other hand, remain relatively more stable with increasing $\rho$.

The above findings seem to suggest that the adaptive BH method based on our estimate $\hat{n}_0^{NEW}$ may perform well compared to that based on $\hat{n}_0^{BKY}$ in some situations. Moreover, both these adaptive BH methods seem to behave similarly in terms of the FDR control and power compared to that based on $\hat{n}_0^{STS}$. For instance, like the BKY method, the adaptive BH method based on $\hat{n}_0^{NEW}$ which controls the FDR under independence, which we will prove in the next section, can also control the FDR under positive dependence, which we will verify also in the next section.
5. New Adaptive Method to Control the FDR

In this section, we present our adaptive version of the BH method based on the estimate $\hat{n}_0^{NEW}$ of $n_0$. We will prove that the FDR of this adaptive BH method is controlled under independence of the $p$-values and numerically show that this control continues to hold even when the $p$-values are positively dependent under normal distributional setting with equal positive correlation. The performance of this adaptive procedure is examined by comparing it to the BKY procedure.

5.1. The New Adaptive BH Method

The following is our proposed adaptive BH method:

**Procedure 5.1.**

1. Observe $R_{SD}(\gamma_1, \ldots, \gamma_n)$, the number of rejections in a step-down method with the critical values $\gamma_i = i\gamma/n$, $i = 1, \ldots, n$, with $\gamma = \alpha/(1 + \alpha)$, and calculate

$$
\hat{n}_0^{NEW} = \frac{n - R_{SD}(\gamma_1, \ldots, \gamma_n) + 1}{1 - \gamma R_{SD}(\gamma_1, \ldots, \gamma_n) + 1},
$$

with an arbitrary $\gamma_{n+1} \in [\gamma, (1 + \gamma)/2)$.

2. Apply the step up procedure with the critical values $\alpha_i = i\alpha/\hat{n}_0^{NEW}$, $i = 1, \ldots, n$, for testing the null hypotheses.

**Theorem 5.1** Procedure 5.1 controls the FDR at $\alpha$ when the $p$-values are independent.

The following two lemmas will facilitate our proof of this theorem. These lemmas will be proved later in this section.

**Lemma 5.1.** Let $U \sim U(0,1)$. Then, for any non-increasing function $\phi(U) > 0$ and a constant $c > 0$, we have

$$
E \left\{ \frac{I(U \leq c\phi(U))}{\phi(U)} \right\} \leq c.
$$

(17)

**Lemma 5.2.** Let $R_{n-1, SD}(-i, c_1, \ldots, c_{n-1})$ be the number of rejections in a step-down method based on the $n - 1$ $p$-values other than $p_i$, where $i \in I_0$, and a set of critical
values $0 < c_1 \leq \cdots \leq c_{n-1} < 1$. Then, under independence of the p-values, we have

$$
\sum_{i \in I_0} E \left\{ \frac{1 - c_{R_{SD,n-1}(c_1,\ldots,c_{n-1})+1}}{n - R_{SD,n-1}(c_1,\ldots,c_{n-1})} \right\} \leq 1 - Pr \{ P_{1:n} \leq c_1, \ldots, P_{n:n} \leq c_n \},
$$

(18)

for an arbitrary, fixed $c_n \in [c_{n-1}, 1)$.

Proof of Theorem 5.1. Using Formula 2.1, we first note that

$$
FDR = \sum_{i \in I_0} E \left[ \frac{I \left( P_i \leq \frac{\alpha R_{SU,n-1}(\alpha_2,\ldots,\alpha_n)+1}{n - R_{SU,n-1}(\alpha_2,\ldots,\alpha_n)} \right)}{\alpha R_{SU,n-1}(\alpha_2,\ldots,\alpha_n)+1} \right] - \sum_{i \in I_0} E \left[ \frac{I \left( P_i \leq \frac{\alpha R_{SU,n-1}(\alpha_2,\ldots,\alpha_n)+1}{n - R_{SU,n-1}(\alpha_2,\ldots,\alpha_n)} \right) \alpha \right],
$$

(19)

with

$$
\alpha_i = \frac{i \alpha \left( 1 - \gamma R_{SD}(\gamma_1,\ldots,\gamma_n)+1 \right)}{n - R_{SD}(\gamma_1,\ldots,\gamma_n)+1} = \begin{cases} 
\frac{i \alpha [n - \gamma (R_{SD}(\gamma_1,\ldots,\gamma_n)+1)]}{n-1} & \text{if } R_{SD}(\gamma_1,\ldots,\gamma_n) = 0, \ldots, n-1, \\
\frac{i \alpha (1 - \gamma_{n+1})}{n} & \text{if } R_{SD}(\gamma_1,\ldots,\gamma_n) = n,
\end{cases}
$$

(20)

$i = 1, \ldots, n$.

Now, notice that $R_{SD}(\gamma_1,\ldots,\gamma_n)$, with fixed $(\gamma_1,\ldots,\gamma_n)$, is a decreasing function of each of the p-values, and as a function of $R_{SD}(\gamma_1,\ldots,\gamma_n)$, $\alpha_i$ is an increasing function if $\gamma \leq n/(n+2)$ and $\gamma_{n+1} \leq (1+\gamma)/2$. But, $\gamma \leq n/(n+2)$ means that $\alpha \leq n/2$, which is obviously true, since $n \geq 2$. Thus, as long as $\gamma_{n+1} \leq (1+\gamma)/2$, each $\alpha_i$ is a (componentwise) decreasing function of $P = (P_1,\ldots,P_n)$. So, by letting $P_i \to 0$ in $\alpha_1$ we see that

$$
\alpha_1 \leq \frac{\alpha \left( 1 - \gamma R_{SD,n-1}(\gamma_2,\ldots,\gamma_n)+1 \right)}{n - R_{SD,n-1}(\gamma_2,\ldots,\gamma_n)},
$$

(21)

since $R_{SD}(\gamma_1,\ldots,\gamma_n \to R_{SD,n-1}(\gamma_2,\ldots,\gamma_n)+1$ as $P_i \to 0$. Let us define $g(P) = R_{SU,n-1}(\alpha_2,\ldots,\alpha_n)+1$ and $h(P^{(-i)})$ equal the right-hand side of (21), with $P^{(-i)} = (P_1,\ldots,P_n) \setminus \{P_i\}$. Then, we have

$$
FDR \leq \sum_{i \in I_0} E \left[ \frac{I \left( P_i \leq g(P) h(P^{(-i)}) \right)}{g(P)} \right].
$$
\[
\begin{align*}
&= \sum_{i \in I_0} E \left[ E \left\{ \frac{I \left( P_i \leq g(P) h(P^{(-i)}) \right)}{g(P)} \bigg| P^{(-i)} \right\} \right] \\
&\leq \alpha [1 - Pr \{ P_{1:n} \leq \gamma_2, \ldots, P_{n:n} \leq \gamma_{n+1} \}] \\
&\leq \alpha,
\end{align*}
\]

with the second and third inequalities following from Lemmas 5.1 and 5.2 respectively.

Thus, the theorem is proved.

We will now give proofs of Lemmas 5.1 and 5.2.

**Proof of Lemma 5.1.** Consider the function \( \psi(u) = u - c\phi(u) \). Since this is non-decreasing, there exists a constant \( c^* \) such that \( \{ \psi(u) \leq 0 \} \subseteq \{ u \leq c^* \} \) and \( \psi(c^*) \leq 0 \), that is, \( c^* \leq c\phi(c^*) \). Since \( \phi(u) \geq \phi(c^*) \) when \( u \leq c^* \), we have

\[
E \left\{ \frac{I(U \leq c\phi(U))}{\phi(U)} \right\} \leq E \left\{ \frac{I(U \leq c^*)}{\phi(c^*)} \right\} = \frac{c^*}{\phi(c^*)} \leq \frac{c\phi(c^*)}{\phi(c^*)} = c.
\]

Thus, the lemma is proved.

**Proof of Lemma 5.2.**

\[
\begin{align*}
&\sum_{i \in I_0} E \left\{ \frac{1 - c\gamma_{R_{SD,n-1}(c_1, \ldots, c_{n-1})+1}}{n - R_{SD,n-1}(c_1, \ldots, c_{n-1})} \right\} \\
&= \sum_{i \in I_0} \sum_{r=0}^{n-1} \frac{1}{n - r} Pr \left\{ R_{SD,n-1}(c_1, \ldots, c_{n-1}) = r \right\} \\
&= \sum_{i \in I_0} \sum_{r=0}^{n-1} \frac{1}{n - r} Pr \left\{ P_{1:n-1}^{(-i)} \leq c_1, \ldots, P_{r:n-1}^{(-i)} \leq c_r, P_{r+1:n-1}^{(-i)} > c_{r+1}, P_i > c_{r+1} \right\} \\
&\leq \sum_{i=1}^{n} \sum_{r=0}^{n-1} \frac{1}{n - r} Pr \left\{ P_{1:n-1}^{(-i)} \leq c_1, \ldots, P_{r:n-1}^{(-i)} \leq c_r, P_{r+1:n-1}^{(-i)} > c_{r+1}, P_i > c_{r+1} \right\} \\
&= \sum_{r=0}^{n-1} Pr \left\{ P_{1:n} \leq c_1, \ldots, P_{r:n} \leq c_r, P_{r+1:n} > c_{r+1} \right\} \\
&= 1 - Pr \left\{ P_{1:n} \leq c_1, \ldots, P_{n:n} \leq c_n \right\},
\end{align*}
\]

where \( P_{1:n-1}^{(-i)} \leq \ldots \leq P_{n-1:n-1}^{(-i)} \) are the ordered components of \( P^{(-i)} \). The third equality in (23) follows from results on ordered random variables given in Sarkar (2002). Thus, the lemma is proved.
5.2. Simulation Study

A simulation study was performed to compare the FDR control and power of our proposed method with those of the BKY method. The study consisted of two parts, the first part was designed for small number of hypotheses, while the second part was designed for relatively large number of hypotheses as seen in most applications of the FDR.

In the first part of the study, we generated $n$ dependent random variables $X_i \sim N(\mu_i, 1)$, $i = 1, \ldots, n$, with a common non-negative correlation $\rho$, and applied both the BKY and our proposed methods to test $\mu_i = 0$ against $\mu_i > 0$, simultaneously for $i = 1, \ldots, n$ at a level $\alpha$. We repeated this 10,000 times by setting $n$ at 4, 8, 16, 64, 128, 256 and 512, the value of $\rho$ at 0, 0.1, 0.25 and 0.5, the proportion of the true null hypotheses $\pi_0$ at 0, 0.25, 0.5, 0.75 and 1, $\alpha$ at 0.05, and $\mu_i$ at 1 for each false null hypothesis, to simulate the FDR and average power (the expected proportion of alternative $\mu_i$'s that are correctly identified) for both methods.

Figure 2 compares the FDR control and Table 2 lists the ratios of power of both methods to the ‘Oracle’ method when $n = 32, 128$ and 512. The ‘Oracle’ method is the BH method based on the critical values $\alpha_i = i\alpha/n_0$, which controls the FDR at the exact level $\alpha$ under the independence of the test statistics. Obviously, it is not implementable in practice as $n_0$ is unknown, but it serves as a benchmark against which other methods can be compared. As seen in Figure 2, our proposed method, which is known to control the FDR at the desired level $\alpha = 0.05$ under independence, can continue to maintain a control over the FDR even under positive dependence, like the BKY method, although ours is often less conservative. Also in terms of power, as seen from Table 2, our method appears to be more powerful than the BKY method in most of the cases considered, especially when the correlation is not very high.

The second part of the study was conducted by setting $n = 5000$. The simulated FDR and power were also based on 10,000 iterations. The comparison between simulated FDR of the two methods is presented in Figure 3. Again, there is evidence that our method can continue to control the FDR under positive dependence, at least when the $p$-values are equally correlated. The power comparisons in this case are displayed in Figures 4 and 5. Figure 4 indicates that the proposed method is more powerful than the BKY when the
Table 2. Estimated power for $n = 32, 128, 512$ and $\rho = 0, 0.1, 0.25, 0.5$

<table>
<thead>
<tr>
<th>$\pi_0$ = 0.25</th>
<th>$\pi_0$ = 0.5</th>
<th>$\pi_0$ = 0.75</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho = 0$</td>
<td>$\rho = 0.1$</td>
<td>$\rho = 0.25$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>new/oracle</td>
<td>0.1881</td>
<td>0.3201</td>
</tr>
<tr>
<td></td>
<td>0.1160</td>
<td>0.1769</td>
</tr>
<tr>
<td></td>
<td>0.0737</td>
<td>0.1479</td>
</tr>
<tr>
<td></td>
<td>0.4661</td>
<td>0.4956</td>
</tr>
<tr>
<td></td>
<td>0.3935</td>
<td>0.4142</td>
</tr>
<tr>
<td></td>
<td>0.3111</td>
<td>0.3550</td>
</tr>
<tr>
<td></td>
<td>0.7553</td>
<td>0.7868</td>
</tr>
<tr>
<td></td>
<td>0.7188</td>
<td>0.7060</td>
</tr>
<tr>
<td></td>
<td>0.7088</td>
<td>0.6408</td>
</tr>
<tr>
<td>BKY/oracle</td>
<td>0.1867</td>
<td>0.2293</td>
</tr>
<tr>
<td></td>
<td>0.1105</td>
<td>0.1714</td>
</tr>
<tr>
<td></td>
<td>0.0682</td>
<td>0.1403</td>
</tr>
<tr>
<td></td>
<td>0.4618</td>
<td>0.4937</td>
</tr>
<tr>
<td></td>
<td>0.3746</td>
<td>0.3979</td>
</tr>
<tr>
<td></td>
<td>0.2925</td>
<td>0.3354</td>
</tr>
<tr>
<td></td>
<td>0.7460</td>
<td>0.7796</td>
</tr>
<tr>
<td></td>
<td>0.6914</td>
<td>0.6794</td>
</tr>
<tr>
<td></td>
<td>0.6692</td>
<td>0.6624</td>
</tr>
<tr>
<td>$\rho = 0.1$</td>
<td>new/oracle</td>
<td>0.2281</td>
</tr>
<tr>
<td></td>
<td>0.1769</td>
<td>0.1479</td>
</tr>
<tr>
<td></td>
<td>0.0682</td>
<td>0.1403</td>
</tr>
<tr>
<td></td>
<td>0.4661</td>
<td>0.4956</td>
</tr>
<tr>
<td></td>
<td>0.3935</td>
<td>0.4142</td>
</tr>
<tr>
<td></td>
<td>0.3111</td>
<td>0.3550</td>
</tr>
<tr>
<td></td>
<td>0.7553</td>
<td>0.7868</td>
</tr>
<tr>
<td></td>
<td>0.7188</td>
<td>0.7060</td>
</tr>
<tr>
<td></td>
<td>0.7088</td>
<td>0.6408</td>
</tr>
<tr>
<td>BKY/oracle</td>
<td>0.2293</td>
<td>0.2293</td>
</tr>
<tr>
<td></td>
<td>0.1714</td>
<td>0.1403</td>
</tr>
<tr>
<td></td>
<td>0.0682</td>
<td>0.1403</td>
</tr>
<tr>
<td></td>
<td>0.4618</td>
<td>0.4937</td>
</tr>
<tr>
<td></td>
<td>0.3746</td>
<td>0.3979</td>
</tr>
<tr>
<td></td>
<td>0.2925</td>
<td>0.3354</td>
</tr>
<tr>
<td></td>
<td>0.7460</td>
<td>0.7796</td>
</tr>
<tr>
<td></td>
<td>0.6914</td>
<td>0.6794</td>
</tr>
<tr>
<td></td>
<td>0.6692</td>
<td>0.6624</td>
</tr>
<tr>
<td>$\rho = 0.25$</td>
<td>new/oracle</td>
<td>0.2846</td>
</tr>
<tr>
<td></td>
<td>0.2530</td>
<td>0.2373</td>
</tr>
<tr>
<td></td>
<td>0.5291</td>
<td>0.4949</td>
</tr>
<tr>
<td></td>
<td>0.4582</td>
<td>0.7542</td>
</tr>
<tr>
<td></td>
<td>0.7502</td>
<td>0.6979</td>
</tr>
<tr>
<td>BKY/oracle</td>
<td>0.2908</td>
<td>0.2495</td>
</tr>
<tr>
<td></td>
<td>0.2305</td>
<td>0.5333</td>
</tr>
<tr>
<td></td>
<td>0.4829</td>
<td>0.4423</td>
</tr>
<tr>
<td></td>
<td>0.7467</td>
<td>0.7294</td>
</tr>
<tr>
<td></td>
<td>0.6650</td>
<td></td>
</tr>
<tr>
<td>$\rho = 0.5$</td>
<td>new/oracle</td>
<td>0.3618</td>
</tr>
<tr>
<td></td>
<td>0.3391</td>
<td>0.3288</td>
</tr>
<tr>
<td></td>
<td>0.5333</td>
<td>0.4829</td>
</tr>
<tr>
<td></td>
<td>0.4423</td>
<td>0.7467</td>
</tr>
<tr>
<td></td>
<td>0.7294</td>
<td>0.6650</td>
</tr>
<tr>
<td>BKY/oracle</td>
<td>0.3755</td>
<td>0.3438</td>
</tr>
<tr>
<td></td>
<td>0.3305</td>
<td>0.6096</td>
</tr>
<tr>
<td></td>
<td>0.5725</td>
<td>0.5478</td>
</tr>
<tr>
<td></td>
<td>0.7876</td>
<td>0.7781</td>
</tr>
<tr>
<td></td>
<td>0.7432</td>
<td></td>
</tr>
</tbody>
</table>

correlation between the test statistics is moderately low. Figure 5 compares the power of the two methods under the condition of high proportion of true null, $\pi_0 \geq 0.9$, which is often the case in modern multiple testing situations. The proposed method seems to be more powerful than the BKY method in such situations.

In conclusion, the simulation study seems to indicate that the new proposed method can control the FDR under positive dependence of the $p$-values. It is more powerful than the BKY method under positive but not very high correlations between the test statistics. When there is a large proportion of true null hypotheses, the new method appears to perform better than the BKY method even in the case of high correlations.

6. An Application to Breast Cancer Data

We applied both the new adaptive BH and the BKY methods to the breast cancer data of Hedenfalk et al. (2001) available at http://www.nejm.org/general/content/supplemental/hedenfalk/index.html; see also Storey and Tibshirani (2003) from http://genomine.org/qvalue/results.html. The results are presented in this section.

The data consists of 3,226 genes on 7 BRCA1 arrays, 8 BRCA2 arrays and 7 sporadic tumors. The goal of the study is to establish differences in gene expression patterns between these tumor groups. Here we analyzed this data with permutation $t$-test to compare BRCA1 and BRCA2. The data were entered into R and all analyses were done using R. As Storey and Tibshirani (2003) did, if any gene had one or more measurement
(log2 expression value) exceeding 20, then this gene was eliminated. This left \( n = 3170 \) genes for permutation \( t \)-test.

We tested each gene for differential expression between BRCA1 and BRCA2 by using a two-sample \( t \)-test. The \( p \)-values were calculated using a permutation method as in Storey and Tibshirani (2003). We did \( B = 100 \) permutations for each gene and got a set of null statistics \( t_{10b}^0, \ldots, t_{n0b}^0, b = 1, \ldots, B \). The \( p \)-value of the permutation \( t \)-test for gene \( i \) was calculated by

\[
p_i = \frac{\sum_{b=1}^{B} \# \{ j : |t_{ij}^0| \geq |t_{ij}|, j = 1, \ldots, n \}}{nB}
\]

The new adaptive method identifies 94 significant genes at the 0.05 level of false discovery rate, whereas, the BKY method gets 93 significant genes. This additional significant gene picked up by our method is intercellular adhesion molecule 2 (clone 471918).

7. Concluding Remarks

Adaptive BH methods other than those reviewed here have been proposed in the literature; see, for instance, Sarkar (2008b). Among these, the BKY method has received much attention since there is numerical evidence that it can continue to control the FDR under some form of positive dependence among the test statistics. The new adaptive BH method we propose in this article competes well with the BKY method. Like the BKY method, it controls the FDR with independent \( p \)-values and, as can be seen numerically, continue to maintain the control with the same type of positively dependent \( p \)-values as in the BKY method. More importantly, it can perform better than the BKY method in some instances, especially when the proportion of true null hypotheses is very large, which happens in many applications.

We have considered using \( \lambda = 0.5 \) in the STS procedure, since this is what Storey, Taylor and Siegmund (2004) have suggested, even though it may not control the FDR under positive dependence, and \( \alpha/(1 + \alpha) \) for \( \gamma \) in our procedure, since this is what Benjamini, Krieger and Yekutieli (2006) have also considered for the \( q \) in their procedure. All these procedures can be proven to control the FDR under independence if other different values are chosen for \( \lambda, q \) and \( \gamma \). But, the BKY as well as our procedures may
not continue to control the FDR under positive dependence with these other values of \( \gamma \) and \( q \).

ACKNOWLEDGEMENT

The research is supported by NSF Grant DMS-0603868.

REFERENCES


using the bootstrap and subsampling. *TEST* 17, 417–442.


Fig. 1. The simulated distribution of $\hat{n}_{0}^{NEW}$, $\hat{n}_{0}^{STS}$ and $\hat{n}_{0}^{BKY}$ for the cases of $n = 5000$, $n_0 = 2500$ and $\rho = 0, 0.25, 0.5$ and 0.75. Each box displays the median and quartiles as usual. The whiskers extend to the 5th and 95th percentiles. The circles are the extreme values, i.e. the 0.01th and 99.99th percentiles.
Fig. 2. Estimated FDR values for $n = 16, 32, \ldots, 512$ and $\rho = 0, 0.1, 0.25, 0.5$. Legend: NEW — solid line; BKY — dashed line.
Fig. 3. Estimated FDR values for $n = 5000$ and $\rho = 0, 0.1, 0.25, 0.5$. Legend: NEW — solid line; BKY — dashed line.
Fig. 4. Estimated power for $n = 5000$ and $\rho = 0, 0.1, 0.25, 0.5$. Legend: NEW — solid line; BKY — dashed line.
Fig. 5. Estimated power for $n = 5000$, $\pi_0 = 0.9, 0.92, 0.94, 0.96, 0.98$ and $\rho = 0, 0.1, 0.25, 0.5, 0.75$ and 0.9. Legend: NEW — solid line; BKY — dashed line.