

# PROCEDURES CONTROLLING GENERALIZED FALSE DISCOVERY RATE

Sanat K. Sarkar<sup>1</sup> and Wenge Guo

*Temple University and National Institute of Environmental Health Sciences*

*Abstract:* Procedures controlling error rates measuring at least  $k$  false rejections, instead of at least one, can potentially increase the ability of a procedure to detect false null hypotheses in situations where one seeks to control  $k$  or more false rejections having tolerated a few of them. The  $k$ -FWER, the probability of at least  $k$  false rejections, is such an error rate that is recently introduced in the literature and procedures controlling it have been proposed. Recently, Sarkar (2007) introduced an alternative, less conservative notion of error rate, the  $k$ -FDR, generalizing the usual notion of false discovery rate (FDR), and proposed a procedure controlling it based on  $k$ -dimensional joint distributions of the null  $p$ -values and assuming the  $MTP_2$  (multivariate totally positive of order two) positive dependence among all the  $p$ -values. In this article, we assume a less restrictive form of positive dependence than the  $MTP_2$  and develop procedure based only on the bivariate, rather than the  $k$ -dimensional, distributions of the null  $p$ -values.

*Key words and phrases:* Arbitrary dependence, average power, clumpy dependence, gene expression, generalized FDR, multiple hypothesis testing, positive regression dependence on subset, stepwise procedure.

## 1 Introduction

We consider in this article the general problem of simultaneously testing of  $n$  null hypotheses  $H_i$ ,  $i = 1, \dots, n$ , against their respective alternatives, using tests that are available for these individual hypotheses. Procedures developed for dealing with this problem have been traditionally based on the idea of controlling the familywise error rate (FWER),

---

<sup>1</sup>The research is supported by NSF Grants DMS-0306366 and DMS-0603868.

the probability of rejecting at least one true null hypothesis [Hochberg and Tamhane (1987)]. However, quite often, especially when large number of hypotheses are simultaneously tested, the notion of FWER is too stringent, allowing little chance to detect many false null hypotheses. Therefore, researchers have focused in the last decade on defining alternative less stringent error rates and developing methods that control them.

The false discovery rate (FDR), the expected proportion of falsely rejected null hypotheses, due to Benjamini and Hochberg (1995) is the first of these alternative error rates that has received considerable attention [Benjamini, Krieger and Yekutieli (2006), Benjamini and Yekutieli (2001, 2005), Finner, Dickhaus and Roters (2007, 2008), Gavrilov, Benjamini and Sarkar (2008), Genovese and Wasserman (2002, 2004), Sarkar (2002, 2004, 2006, 2008a), Storey (2002, 2003) and Storey, Taylor and Siegmund (2004)]. Recently, the ideas of controlling the probabilities of falsely rejecting at least  $k$  null hypotheses, which is the  $k$ -FWER, and the false discovery proportion (FDP) exceeding a certain threshold  $\gamma \in [0, 1)$  have been introduced as alternatives to the FWER and methods controlling these new error rates have been suggested [Dudoit, van der Laan and Pollard (2004), Guo and Rao (2006), Guo and Romano (2007), Hommel and Hoffmann (1987), Lehmann and Romano (2005), Korn, Troendle, McShane and Simon (2004), Romano and Shaikh (2006a, b) and Romano and Wolf (2005, 2007), Sarkar (2007, 2008b) and van der Laan, Dudoit and Pollard (2004)].

Often in practice one is willing to tolerate a few false rejections but wants to control too many of them, say  $k$  or more. A procedure controlling the  $k$ -FWER is more appropriate and better able to make true discoveries than the corresponding FWER ( $k = 1$ ) procedure. However, a less conservative notion of error rate than the  $k$ -FWER exists in this case. It is the  $k$ -FDR, the expected ratio of  $k$  or more false rejections to the total number of rejections, a natural generalization of the FDR, recently proposed by Sarkar (2007). He developed a  $k$ -FDR procedure utilizing  $k$ th order joint null distributions of the  $p$ -values under the independence or the  $MTP_2$  positive dependence, due to Karlin and Rinott (1980), of the

$p$ -values. It is less conservative and uniformly more powerful in detecting  $k$  or more true discoveries than the  $k$ -FWER procedure that Sarkar (2008b) developed as a generalization of Hochberg’s (1988) FWER procedure, and often outperforms the Benjamini-Hochberg (BH) stepup FDR procedure. He has also developed a  $k$ -FDR procedure utilizing  $k$ th order joint null distributions of the  $p$ -values under any form of dependence of the  $p$ -values, which generalizes the Benjamini-Yekutieli (BY) procedure in Benjamini and Yekutieli (2001).

Sarkar and Guo (2008) brought newer insight to the notion of  $k$ -FDR considering a mixture model involving independent  $p$ -values. They provided a simple and intuitive upper bound to the  $k$ -FDR, based on which they introduced conservative point estimates of the  $k$ -FDR before developing through these estimates newer stepup  $k$ -FDR procedures. One of these is a generalized version of the BH FDR procedure and others are adaptive versions of this procedure using a class of estimates of the number of true null hypotheses. The  $k$ -FDR control of these procedures was proved for independent test statistics.

In this article, we go back to the work of Sarkar (2007) and develop  $k$ -FDR procedures relaxing both the  $MTP_2$  condition and the requirement of using the  $k$ -dimensional joint distributions of the null  $p$ -values. More specifically, we assume the positive dependence condition, a weaker version of the  $MTP_2$ , under which a stepwise procedure, stepdown or stepup, with the critical values of the BH procedure (to be simply referred to as the BH stepwise procedure) is known to control the FDR [Benjamini and Yekutieli (2001) and Sarkar (2002)], and generalize this BH stepwise procedure to a  $k$ -FDR stepwise procedure based only the bivariate, instead of the  $k$ -dimensional, distributions of the null  $p$ -values. The positive dependence condition assumed in this paper is, however, slightly weaker than considered originally in the above two papers. We offer two such generalizations in the positive dependence case, one more conservative than the other but easier to implement, both reducing to the same procedure under independence. Based on numerical calculations, we show that, for appropriately chosen values of  $k$ , our  $k$ -FDR stepwise procedure under independence is uniformly more powerful than the corresponding BH stepwise procedure.

Under positive dependence, each of these  $k$ -FDR procedures remains more powerful than the corresponding BH stepwise procedure under weak dependence, but unfortunately loses its edge with increasing dependence. Often in practice, like in microarray analysis and fMRI studies, the  $p$ -values tend to be *clumpy dependent* in that they are more dependent within small groups, strongly or weakly, but are independent between these groups. The  $k$ -FDR procedures in this case are seen to improve their performance. They remain more powerful than the corresponding BH stepwise procedure even for large positive correlations. An alternative stepdown  $k$ -FDR procedure is proposed that performs better than the one in the above stepwise procedure under independence. When applied to the gene expression data in Hedenfalk et al. (2001), it is seen to detect more differentially expressed genes than the BH stepup procedure for different values of  $k$ . Finally, we develop a generalized BY procedure that uniformly outperforms the original BY procedure as a  $k$ -FDR procedure under any form of dependence.

## 2 Preliminaries

Suppose that  $H_1, \dots, H_n$  are the null hypotheses that we want to simultaneously test using the corresponding  $p$ -values  $P_1, \dots, P_n$ , respectively. Let  $P_{(1)} \leq \dots \leq P_{(n)}$  be the ordered  $p$ -values and  $H_{(1)}, \dots, H_{(n)}$  the associated null hypotheses. Then, given a non-decreasing set of critical constants  $0 < \alpha_1 \leq \dots \leq \alpha_n < 1$ , a stepdown multiple testing procedure rejects the set of null hypotheses  $\{H_{(i)}, i \leq i_{SD}^*\}$  and **rejects** the rest, where  $i_{SD}^* = \max\{i : P_{(j)} \leq \alpha_j \forall j \leq i\}$ , if the maximum exists, otherwise accepts all the null hypotheses. A stepup procedure, on the other hand, rejects the set  $\{H_i, i \leq i_{SU}^*\}$  and accepts the rest, where  $i_{SU}^* = \max\{i : P_{(i)} \leq \alpha_i\}$ , if the maximum exists, otherwise accepts all the null hypotheses. A stepwise (stepdown or stepup) procedure with the same constant is referred to as a single-step procedure.

The constants in a stepwise procedure are determined subject to the control at a pre-specified level  $\alpha$  of a suitable error rate. Let  $V$  denote the number of falsely rejected null

hypotheses. Then, the  $k$ -FWER is defined as  $k\text{-FWER} = P\{V \geq k\}$ , with 1-FWER being the original FWER. The following lemma provides a general theory towards determining the constants in a stepwise procedure providing a control of the  $k$ -FWER. It is assumed that there are  $n_0$  true null hypotheses and  $\hat{P}_1, \dots, \hat{P}_{n_0}$  are the  $p$ -values that correspond to these null hypotheses.

**Lemma 2.1** *Given a stepwise procedure involving  $P_1, \dots, P_n$  and the critical values  $0 < \alpha_1 \leq \dots \leq \alpha_n < 1$ , consider the corresponding stepwise procedure in terms of the null  $p$ -values  $\hat{P}_1, \dots, \hat{P}_{n_0}$  and the critical values  $\alpha_{n-n_0+1} \leq \dots \leq \alpha_n$ . Let  $V_n$  denote the number of false rejections in the original stepwise procedure and  $\hat{R}_{n_0}$  denote the number of rejections in the stepwise procedure involving the null  $p$ -values. Then, we have for any fixed  $k \leq n_0 \leq n$ ,*

$$\{V_n \geq k\} \subseteq \{\hat{R}_{n_0} \geq k\} .$$

PROOF. Let  $\hat{P}_{(1)} \leq \dots \leq \hat{P}_{(n_0)}$  be the ordered values of  $\{\hat{P}_1, \dots, \hat{P}_{n_0}\}$ . Note that  $\hat{P}_{(j)} \leq P_{(n-n_0+j)}$  for all  $1 \leq j \leq n_0$ . Therefore, the original stepwise procedure rejects less number of null hypotheses, and hence has less number of falsely rejected null hypotheses, than the corresponding stepwise procedure where the  $p$ -values corresponding to the false null hypotheses are all very close to zero. Thus, the lemma follows. ■

**Remark 2.1** According to Lemma 2.1, construction of a stepwise procedure providing a control of the  $k$ -FWER at  $\alpha$  basically reduces to that of finding the constants in that procedure guaranteeing the inequality  $P\{\hat{R}_{n_0} \geq k\} \leq \alpha$  for all  $k \leq n_0 \leq n$ . It unifies the arguments used separately towards constructing stepdown  $k$ -FWER and stepup  $k$ -FWER procedures in Lehmann and Romano (2005) and Romano and Shaikh (2006a). It essentially implies that the least favorable configuration for the  $k$ -FWER corresponds to the situation where  $p$ -values corresponding to false null hypotheses are all zero.

Sarkar (2007) has recently introduced the concept of  $k$ -FDR. Let  $R$  denote the total

number of rejections. Then, it is defined as  $k\text{-FDR} = E(k\text{-FDP})$ , where

$$k\text{-FDP} = \begin{cases} \frac{V}{R} & \text{if } V \geq k \\ 0 & \text{otherwise.} \end{cases}$$

That is, the  $k\text{-FDR}$  is the expected ratio of  $k$  or more false rejections to the total number of rejections. When  $k = 1$ , it reduces to the original FDR. The  $k\text{-FDR}$  is a less conservative notion of error rate than the FDR, as  $k\text{-FDR} \leq \text{FDR}$ . Using it, instead of the FDR, when one is willing to control at least  $k$  false rejections rather than at least one, is a natural generalization of the idea of using the  $k\text{-FWER}$  instead of the FWER.

We will be assuming in this article that each  $P_i$  under its respective true null hypothesis  $H_i$  is  $U(0, 1)$  and jointly the  $p$ -values are positively dependent in the following sense:

$$E \left\{ \phi(P_1, \dots, P_n) \mid \hat{P}_i \leq u \right\} \uparrow u \in (0, 1), \quad (1)$$

for every  $\hat{P}_i$  and any increasing (coordinatewise) function  $\phi$ . The condition (1) is more relaxed than the positive regression dependence on subset (PRDS) condition, that is,  $E \left\{ \phi(P_1, \dots, P_n) \mid \hat{P}_i = u \right\} \uparrow u \in (0, 1)$ , used in Benjamini and Yekutieli (2001) and Sarkar (2002), and is satisfied by the  $p$ -values arising in a number of multiple testing situations. In particular, it is satisfied by the  $p$ -values corresponding to multivariate normal test statistics with a common non-negative correlation that we consider in our numerical calculations later. Since the  $k\text{-FDR}$  is 0, and hence trivially controlled, by any procedure if  $n_0 < k$ , we will assume throughout this paper that  $k \leq n_0 \leq n$ .

### 3 $k$ -FDR procedures under independence or positive dependence

We will develop in this section stepwise procedures that control the  $k$ -FDR, for  $k \geq 2$ , under the condition (1). To this end, we first have

$$\begin{aligned}
k\text{-FDR} &= E \left\{ \frac{V}{R} \cdot I(V \geq k) \right\} \\
&= E \left\{ \sum_{r=k}^n \frac{1}{r} \sum_{i=1}^{n_0} I(\hat{P}_i \leq \alpha_r, V \geq k, R = r) \right\} \\
&= \sum_{i=1}^{n_0} \sum_{r=k}^n \frac{1}{r} \Pr \left\{ \hat{P}_i \leq \alpha_r, V \geq k, R = r \right\} \\
&= \sum_{i=1}^{n_0} \sum_{r=k}^n \frac{\alpha_r}{r} \Pr \left\{ V \geq k, R = r \mid \hat{P}_i \leq \alpha_r \right\} . \tag{2}
\end{aligned}$$

With  $r \vee k = \max\{k, r\}$ , let  $\alpha_r = (r \vee k)\alpha_k/k$ , for some fixed  $0 < \alpha_k < 1$ . Then, the  $k$ -FDR in (2) simplifies to

$$k\text{-FDR} = \frac{\alpha_k}{k} \sum_{i=1}^{n_0} \sum_{r=k}^n P \left\{ V \geq k, R = r \mid \hat{P}_i \leq \alpha_r \right\} .$$

Now, for any  $r > k$  and  $i \leq n_0$ ,

$$\begin{aligned}
&\Pr \left\{ V \geq k, R = r \mid \hat{P}_i \leq \alpha_r \right\} \\
&= \Pr \left\{ V \geq k, R \geq r \mid \hat{P}_i \leq \alpha_r \right\} - \Pr \left\{ V \geq k, R \geq r+1 \mid \hat{P}_i \leq \alpha_r \right\} \\
&\leq \Pr \left\{ V \geq k, R \geq r \mid \hat{P}_i \leq \alpha_{r-1} \right\} - \Pr \left\{ V \geq k, R \geq r+1 \mid \hat{P}_i \leq \alpha_r \right\} .
\end{aligned}$$

The inequality follows from the assumption (1), since  $\{V \geq k, R \geq r\}$  is a decreasing set

for a stepwise procedure. Thus, we get

$$\begin{aligned}
k\text{-FDR} &\leq \frac{\alpha_k}{k} \sum_{i=1}^{n_0} \Pr \left\{ V \geq k, R = k \mid \hat{P}_i \leq \alpha_k \right\} + \\
&\quad \frac{\alpha_k}{k} \sum_{i=1}^{n_0} \sum_{r=k+1}^n \Pr \left\{ V \geq k, R = r \mid \hat{P}_i \leq \alpha_r \right\} \\
&\leq \frac{\alpha_k}{k} \sum_{i=1}^{n_0} \Pr \left\{ V \geq k, R = k \mid \hat{P}_i \leq \alpha_k \right\} + \\
&\quad \frac{\alpha_k}{k} \sum_{i=1}^{n_0} \Pr \left\{ V \geq k, R \geq k+1 \mid \hat{P}_i \leq \alpha_k \right\} \\
&= \frac{\alpha_k}{k} \sum_{i=1}^{n_0} \Pr \left\{ V \geq k \mid \hat{P}_i \leq \alpha_k \right\} = \frac{1}{k} \sum_{i=1}^{n_0} \Pr \left\{ V \geq k, \hat{P}_i \leq \alpha_k \right\}. \quad (3)
\end{aligned}$$

Applying Lemma 2.1 to (3), we get

$$k\text{-FDR} \leq \frac{1}{k} \sum_{i=1}^{n_0} P \left\{ \hat{R}_{n_0} \geq k, \hat{P}_i \leq \alpha_k \right\}.$$

Let  $\hat{R}_{n_0-1}^{(-i)}$  denote the number of rejections in the corresponding stepwise procedure based on the ordered values  $\hat{P}_{(1)}^{(-i)} \leq \dots \leq \hat{P}_{(n_0-1)}^{(-i)}$  of  $\{\hat{P}_1, \dots, \hat{P}_{n_0}\} \setminus \{\hat{P}_i\}$  and the  $n_0 - 1$  critical values  $\alpha_{n-n_0+2} \leq \dots \leq \alpha_n$ . Then, for any  $k \leq r \leq n_0$ , we notice that, when our procedure is a stepup procedure

$$\begin{aligned}
\left\{ \hat{R}_{n_0} = r, \hat{P}_i \leq \alpha_k \right\} &= \left\{ \hat{P}_{(r)} \leq \alpha_{n-n_0+r}, \hat{P}_{(r+1)} > \alpha_{n-n_0+r+1}, \dots, \right. \\
&\quad \left. \hat{P}_{(n_0)} > \alpha_n, \hat{P}_i \leq \alpha_k \right\} \\
&= \left\{ \hat{P}_{(r-1)}^{(-i)} \leq \alpha_{n-n_0+r}, \hat{P}_{(r)}^{(-i)} > \alpha_{n-n_0+r+1}, \dots, \right. \\
&\quad \left. \hat{P}_{(n_0-1)}^{(-i)} > \alpha_n, \hat{P}_i \leq \alpha_k \right\} \\
&= \left\{ \hat{R}_{n_0-1}^{(-i)} = r-1, \hat{P}_i \leq \alpha_k \right\}; \quad (4)
\end{aligned}$$

whereas, for a stepdown procedure, we have

$$\begin{aligned}
\left\{ \hat{R}_{n_0} = r, \hat{P}_i \leq \alpha_k \right\} &= \left\{ \hat{P}_{(1)} \leq \alpha_{n-n_0+1}, \dots, \hat{P}_{(r)} \leq \alpha_{n-n_0+r}, \right. \\
&\quad \left. \hat{P}_{(r+1)} > \alpha_{n-n_0+r+1}, \hat{P}_i \leq \alpha_k \right\} \\
&\subseteq \left\{ \hat{P}_{(1)}^{(-i)} \leq \alpha_{n-n_0+2}, \dots, \hat{P}_{(r-1)}^{(-i)} \leq \alpha_{n-n_0+r}, \right. \\
&\quad \left. \hat{P}_{(r)}^{(-i)} > \alpha_{n-n_0+r+1}, \hat{P}_i \leq \alpha_k \right\} \\
&= \left\{ \hat{R}_{n_0-1}^{(-i)} = r-1, \hat{P}_i \leq \alpha_k \right\}. \tag{5}
\end{aligned}$$

Thus, for a stepwise (stepup or stepdown) procedure, we get

$$\begin{aligned}
k\text{-FDR} &\leq \frac{1}{k} \sum_{i=1}^{n_0} \sum_{r=k}^{n_0} \Pr \left\{ \hat{R}_{n_0-1}^{(-i)} = r-1, \hat{P}_i \leq \alpha_k \right\} \\
&= \frac{1}{k} \sum_{i=1}^{n_0} \sum_{j(\neq i)=1}^{n_0} \sum_{r=k}^{n_0} \frac{1}{r-1} \Pr \left\{ \hat{R}_{n_0-1}^{(-i)} = r-1, \right. \\
&\quad \left. \hat{P}_i \leq \alpha_k, \hat{P}_j \leq \alpha_{n-n_0+r} \right\}. \tag{6}
\end{aligned}$$

As explained in (4) and (5),

$$\begin{aligned}
&\left\{ \hat{R}_{n_0-1}^{(-i)} = r-1, \hat{P}_i \leq \alpha_k, \hat{P}_j \leq \alpha_{n-n_0+r} \right\} \\
&= \left\{ \hat{R}_{n_0-2}^{(-i,-j)} = r-2, \hat{P}_i \leq \alpha_k, \hat{P}_j \leq \alpha_{n-n_0+r} \right\},
\end{aligned}$$

for a stepup procedure, and

$$\begin{aligned}
&\left\{ \hat{R}_{n_0-1}^{(-i)} = r-1, \hat{P}_i \leq \alpha_k, \hat{P}_j \leq \alpha_{n-n_0+r} \right\} \\
&\subseteq \left\{ \hat{R}_{n_0-2}^{(-i,-j)} = r-2, \hat{P}_i \leq \alpha_k, \hat{P}_j \leq \alpha_{n-n_0+r} \right\},
\end{aligned}$$

for a stepdown procedure, where  $\hat{R}_{n_0-2}^{(-i,-j)}$  is the number of rejections in the corresponding stepwise procedure involving  $\left\{ \hat{P}_1, \dots, \hat{P}_{n_0} \right\} \setminus \left\{ \hat{P}_i, \hat{P}_j \right\}$  and critical values  $\alpha_{n-n_0+3} \leq \dots \leq$

$\alpha_n$ . Thus,

$$k\text{-}FDR \leq \frac{(n - n_0 + k)\alpha_k}{k^2(k - 1)} \sum_{i=1}^{n_0} \sum_{j(\neq i)=1}^{n_0} \sum_{r=k}^{n_0} \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} = r - 2, \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+r} \right\}.$$

The inequality follows from the fact that

$$\frac{n - n_0 + r}{r - 1} \leq \frac{n - n_0 + k}{k - 1},$$

for all  $k \leq r \leq n_0$ . Since for  $r > k$ ,

$$\begin{aligned} & \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} = r - 2, \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+r} \right\} \\ = & \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} \geq r - 2, \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+r} \right\} - \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} \geq r - 1, \right. \\ & \left. \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+r} \right\} \\ \leq & \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} \geq r - 2, \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+r-1} \right\} - \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} \geq r - 1, \right. \\ & \left. \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+r} \right\}, \end{aligned}$$

with the inequality following due to the property (1) and the fact that  $\{\hat{R}_{n_0-2}^{(-i,-j)} \geq r -$

$2, \hat{P}_i \leq \alpha_k\}$  is decreasing, we have

$$\begin{aligned}
& \sum_{r=k}^{n_0} \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} = r-2, \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+r} \right\} \\
\leq & \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} = k-2, \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+k} \right\} + \\
& \sum_{r=k+1}^{n_0} \left[ \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} \geq r-2, \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+r-1} \right\} - \right. \\
& \left. \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} \geq r-1, \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+r} \right\} \right] \\
= & \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} = k-2, \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+k} \right\} + \\
& \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} \geq k-1, \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+k} \right\} \\
= & \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} \geq k-2, \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+k} \right\}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
& k\text{-FDR} \\
\leq & \frac{(n-n_0+k)\alpha_k}{k^2(k-1)} \sum_{i=1}^{n_0} \sum_{j(\neq i)=1}^{n_0} \Pr \left\{ \hat{R}_{n_0-2}^{(-i,-j)} \geq k-2, \hat{P}_i \leq \alpha_k \mid \right. \\
& \left. \hat{P}_j \leq \alpha_{n-n_0+k} \right\} \\
\leq & \frac{(n-n_0+k)\alpha_k}{k^2(k-1)} \sum_{i=1}^{n_0} \sum_{j(\neq i)=1}^{n_0} \Pr \left\{ \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+k} \right\} \\
= & \frac{1}{k(k-1)} \sum_{i=1}^{n_0} \sum_{j(\neq i)=1}^{n_0} \Pr \left\{ \hat{P}_i \leq \alpha_k, \hat{P}_j \leq \alpha_{n-n_0+k} \right\}. \tag{7}
\end{aligned}$$

Thus, we have the following theorem as one of our main results.

**Theorem 3.1** *For a stepwise procedure with the  $p$ -values satisfying (1) and critical values given by  $\alpha_i = \{i \vee k\}\alpha_k/k$ ,  $i = 1, \dots, n$ , for some fixed  $2 \leq k \leq n$  and  $0 < \alpha_k < 1$ , we have*

$$k\text{-FDR} \leq \max_{k \leq n_0 \leq n} \left\{ \frac{1}{k(k-1)} \sum_{i=1}^{n_0} \sum_{j(\neq i)=1}^{n_0} H_{ij} \left( \alpha_k, \frac{(n-n_0+k)\alpha_k}{k} \right) \right\}, \tag{8}$$

where  $H_{ij}(u, v) = P\{\hat{P}_i \leq u, \hat{P}_j \leq v\}$ , for  $i \neq j$ .

**Remark 3.1** When  $k = 1$ , the right-hand side of (3) reduces to  $n_0\alpha_1$ , providing the inequality  $\text{FDR} \leq n_0\alpha_1$ . It is being sharpened in Theorem 3.1 by using joint distributions of the  $p$ -values when  $k \geq 2$ . This result for  $k = 1$  is known in the literature [Benjamini and Hochberg (1995), Benjamini and Yekutieli (2001) and Sarkar (2002)], but under conditions slightly more restrictive than (1). Theorem 3.1 highlights the point, which was emphasized in Sarkar (2007), that the concept of  $k$ -FDR, like that of the  $k$ -FWER, allows one to explicitly utilize the joint distribution of the  $p$ -values. However, while Sarkar (2007) used the  $k$ th order joint null distributions of the  $p$ -values, here we use only the bivariate null distributions. Moreover, we are relying on a less restrictive assumption on the dependence of the  $p$ -values.

We can invoke the same positive dependence property shared by every pair  $(\hat{P}_i, \hat{P}_j)$  and apply the inequality

$$P\{\hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_{n-n_0+k}\} \leq \{ \hat{P}_i \leq \alpha_k \mid \hat{P}_j \leq \alpha_k \}$$

to the second to last line in (7) to obtain the following corollary to Theorem 3.1, providing an upper bound to the  $k$ -FDR (with  $k \geq 2$ ) that is more relaxed and easier to implement than the one in (8).

**Corollary 3.1** *For the stepwise procedure in Theorem 3.1 and under the conditions stated therein, we have, for  $k \geq 2$ ,*

$$k\text{-FDR} \leq \max_{k \leq n_0 \leq n} \left\{ \frac{n - n_0 + k}{k^2(k-1)} \sum_{i=1}^{n_0} \sum_{j(\neq i)=1}^{n_0} H_{ij}(\alpha_k, \alpha_k) \right\}. \quad (9)$$

The  $k$ -FDR of the stepwise procedure in Theorem 3.1 can be controlled at  $\alpha$  under the conditions stated in the theorem by equating the upper bound in (8) or (9) to  $\alpha$  and solving the resulting equation for  $\alpha_k$ . Of course, one needs to know the bivariate distributions of

all pairs of null  $p$ -values. For instance, when the null  $p$ -values are exchangeable with  $H$  as the common and known bivariate cdf, the  $\alpha_k$  can be obtained satisfying the equation

$$\max_{k \leq n_0 \leq n} \left\{ \frac{n_0(n_0 - 1)}{k(k - 1)} H \left( \alpha_k, \frac{(n - n_0 + k)\alpha_k}{k} \right) \right\} = \alpha , \quad (10)$$

or, can be obtained little more conservatively by solving

$$\frac{D(k, n)H(\alpha_k, \alpha_k)}{k^2(k - 1)} = \alpha , \quad (11)$$

where

$$D(k, n) = \max_{k \leq n_0 \leq n} \{n_0(n_0 - 1)(n - n_0 + k)\} .$$

In particular, when the  $p$ -values are independent, we have the following:

**Proposition 3.1** *Consider a stepwise procedure with the critical values  $\alpha_i = (i \vee k)\beta/n$ ,  $i = 1, \dots, n$ , with  $\beta = n\sqrt{(k - 1)\alpha/D(k, n)}$ . It controls the  $k$ -FDR at  $\alpha$  when the  $p$ -values are independent.*

**Remark 3.2** It is important to note that in a stepwise procedure with a control of the  $k$ -FDR, the first  $k - 1$  critical values can be chosen arbitrarily without affecting the  $k$ -FDR, as in the case of a  $k$ -FWER stepwise procedure. Nevertheless, as argued in Lehmann and Romano (2005) and Sarkar (2007, 2008b), keeping these critical values constant at the  $k$ th critical value would be the best option. Thus, even though one can use the stepwise BH procedure with the critical values  $\alpha_i = i\alpha/n$ ,  $i = 1, \dots, n$ , as a  $k$ -FDR procedure, since it controls the FDR under the conditions assumed in this paper (see Remark 3.1), one may consider improving it by modifying the critical values to  $\alpha_i = (i \vee k)\alpha/n$ ,  $i = 1, \dots, n$ . Nevertheless, as seen in Theorem 3.1, it does not take full advantage of the notion of  $k$ -FDR in the sense that its critical values are not determined subject to a direct control of the  $k$ -FDR, and hence can potentially be improved. Proposition 3.1 helps us develop

Table 1: *Minimum values  $k$  for different  $n$  under independence with  $\alpha = 0.05$ .*

$n$	100	200	500	1000	2000	5000	10000
$k$	2	3	5	9	17	39	77

such an improvement for some choices of  $(n, k)$  at least under the independence.

The procedure in Proposition 3.1 will be uniformly less conservative and more powerful than the corresponding stepwise procedure with the BH critical values if  $\beta > \alpha$ , that is, if

$$\frac{n^2(k-1)}{D(k, n)} > \alpha. \quad (12)$$

Since

$$D(k, n) = \tilde{n}_0(\tilde{n}_0 - 1)(n + k - \tilde{n}_0),$$

with

$$\tilde{n}_0 = \frac{n+k+1}{3} \left\{ 1 + \left[ 1 - \frac{3(n+k)}{(n+k+1)^2} \right]^{\frac{1}{2}} \right\},$$

one can determine the minimum value of  $k$  for which the inequality in (12) holds for given  $n$  and  $\alpha$ . Table 1 presents such values of  $k$  for some values of  $n$  and  $\alpha = 0.05$ . For instance, with 1000 null hypotheses, if one is willing to allow at most  $k-1$  false rejections and seeks to control the  $k$ -FDR at  $\alpha = 0.05$ , our procedure in Proposition 3.1 does provide a better control of the  $k$ -FDR than the corresponding stepwise BH procedure under the independence of the  $p$ -values as long as  $k \geq 9$ .

The stepdown procedure in Proposition 3.1 can be improved under independence as follows. Consider a stepdown procedure with critical values  $\alpha_i = \{i \vee k\}\beta/n$ ,  $i = 1, \dots, n$ , for a fixed  $0 < \beta < 1$ . From the first line in (6), we see that for this procedure

$$k\text{-FDR} \leq \frac{n_0\beta}{n} P\{R_{n_0-1} \geq k-1\},$$

where  $R_{n_0-1}$  is the number of rejections in the stepdown procedure based on  $\hat{P}_{(1):n_0-1} \leq \dots \leq \hat{P}_{(n_0-1):n_0-1}$ , the ordered versions of any  $n_0 - 1$  of the  $n_0$  null  $p$ -values, and the corresponding critical values  $\alpha_{n-n_0+2} \leq \dots \leq \alpha_n$ . Let

$$G_{k,n}(u) = P \{U_{(k)} \leq u\} = \sum_{j=k}^n \binom{n}{j} u^j (1-u)^{n-j} ,$$

the cdf of the  $k$ th order statistic based on  $n$  iid  $U(0, 1)$ . Then, since

$$\begin{aligned} P \{R_{n_0-1} \geq k - 1\} &= P \left\{ \hat{P}_{1:n_0-1} \leq \alpha_{n-n_0+2}, \dots, \hat{P}_{(k-1):n_0-1} \leq \alpha_{n-n_0+k} \right\} \\ &\leq G_{k-1, n_0-1}(\alpha_{n-n_0+k}) , \end{aligned}$$

we have the following:

**Proposition 3.2** *Consider a stepdown procedure with the critical values  $\alpha_i = (i \vee k)\beta/n$ ,  $i = 1, \dots, n$ , where*

$$\frac{\beta}{n} \max_{k \leq n_0 \leq n} \left\{ n_0 G_{k-1, n_0-1} \left( \frac{(n - n_0 + k)\beta}{n} \right) \right\} = \alpha . \quad (13)$$

*It controls the  $k$ -FDR at  $\alpha$  when the  $p$ -values are independent.*

Table 2 presents values of  $\beta$  in Proposition 3.2 for some values of  $(n, k)$  and  $\alpha = 0.05$ . It shows that the stepdown procedure in Proposition 3.2 indeed performs better than the stepdown procedure in Proposition 3.1 under independence. It continues to be a more powerful  $k$ -FDR procedure than the stepdown BH procedure even for values of  $k$  smaller than those for the stepdown procedure in Proposition 3.1.

With positively dependent  $p$ -values, since the upper bound in Theorem 3.1 or its corollary is greater than or equal to the corresponding upper bound under independence, the same critical values in Proposition 3.1 may not continue to control the  $k$ -FDR. In this case, as noted before, one needs to compute  $\alpha_k$ , or equivalently  $\beta$ , where  $\alpha_k = k\beta/n$ , from

Table 2: Values  $\beta$  for the stepdown procedure in Proposition 3.2 for different  $(n, k)$  with  $\alpha = 0.05$ .

	$n$						
	100	200	500	1000	2000	5000	10000
$k = 2$	0.105	0.081	0.064	0.057	0.054	0.052	0.051
$k = 5$	0.169	0.115	0.077	0.063	0.057	0.053	0.051
$k = 10$	0.274	0.175	0.10	0.073	0.061	0.054	0.052

these upper bounds using the bivariate distributions  $H_{ij}$ . For  $p$ -values generated from multivariate normal test statistics with a common non-negative correlation  $\rho$ , values of  $\beta$  are numerically computed using these bounds for some values of  $(n, k)$  and  $\alpha = 0.05$  and are presented in Table 3, where  $\beta \equiv \beta_1 = n\alpha_k/k$  with  $\alpha_k$  in (8) and  $\beta \equiv \beta_2 = n\alpha_k/k$  with  $\alpha_k$  in (9).

With significantly larger critical values than the corresponding BH stepwise procedure, our proposed  $k$ -FDR stepwise procedures, corresponding to  $\beta_1$  and  $\beta_2$ , are seen to be quite powerful compared to the corresponding stepwise BH procedure when the  $p$ -values are independent or weakly but positively dependent. However, as the  $p$ -values become more and more positively dependent our procedures unfortunately lose their edge over the corresponding stepwise BH procedure.

Often in practice, like in microarray analysis and fMRI studies, the  $p$ -values tend to be *clumpy dependent* in that they are more dependent within groups than between groups. Suppose that there are  $g$  independent groups and, for each  $i = 1, \dots, n_0$ ,  $J_i$  is the set of indices of null hypotheses in the group containing the  $\hat{P}_i$ . Clearly,  $J_i \equiv J_j$  if  $i$  and  $j$  belong to the same group. The upper bounds in Theorem 3.1 and Corollary 3.1 can be expressed,

Table 3: Values of  $\beta_1$  and  $\beta_2$  for different  $(n, k)$  and non-negative  $\rho$  with  $\alpha = 0.05$ .

$n$	$k$	$\rho = 0.0$		$\rho = 0.05$		$\rho = 0.10$		$\rho = 0.15$		$\rho = 0.20$	
		$\beta_1$	$\beta_2$	$\beta_1$	$\beta_2$	$\beta_1$	$\beta_2$	$\beta_1$	$\beta_2$	$\beta_1$	$\beta_2$
50	2	0.079	0.079	0.066	0.062	0.055	0.050	0.046	0.039	0.038	0.031
100	4	0.096	0.096	0.081	0.077	0.068	0.061	0.058	0.049	0.048	0.039
200	8	0.103	0.103	0.087	0.083	0.074	0.066	0.063	0.053	0.053	0.043
500	20	0.107	0.107	0.091	0.086	0.077	0.069	0.065	0.056	0.055	0.045
1000	40	0.108	0.108	0.092	0.087	0.078	0.070	0.066	0.056	0.056	0.045
2000	80	0.109	0.109	0.093	0.088	0.079	0.071	0.067	0.057	0.056	0.046
5000	200	0.109	0.109	0.093	0.088	0.079	0.071	0.067	0.057	0.057	0.046
10000	400	0.109	0.109	0.093	0.088	0.079	0.071	0.067	0.057	0.057	0.046

respectively, in this case as

$$k\text{-FDR} \leq \max_{k \leq n_0 \leq n} \left\{ \frac{n - n_0 + k}{k^2(k-1)} \sum_{i=1}^{n_0} \sum_{j(\neq i) \in J_i} \left[ \frac{k}{n - n_0 + k} H_{ij} \left( \alpha_k, \frac{(n - n_0 + k)\alpha_k}{k} \right) + (n_0 - |J_i|)\alpha_k^2 \right] \right\}, \quad (14)$$

and

$$\leq \max_{k \leq n_0 \leq n} \left\{ \frac{n - n_0 + k}{k^2(k-1)} \sum_{i=1}^{n_0} \sum_{j(\neq i) \in J_i} [H_{ij}(\alpha_k, \alpha_k) + (n_0 - |J_i|)\alpha_k^2] \right\}, \quad (15)$$

where  $|J_i|$  is the cardinality of  $J_i$ . Generating  $n$   $p$ -values from multivariate normal test statistics with a common non-negative correlation  $\rho$  within independent groups each containing  $n/g$   $p$ -values, we repeat our numerical calculations to examine how our stepwise

Table 4: Values of  $\beta_1$  and  $\beta_2$  for different  $(n, k, g)$  and non-negative  $\rho$  under clumpy dependence with  $\alpha = 0.05$ .

$n$	$k$	$g$	$\rho = 0.0$		$\rho = 0.2$		$\rho = 0.5$		$\rho = 0.8$	
			$\beta_1$	$\beta_2$	$\beta_1$	$\beta_2$	$\beta_1$	$\beta_2$	$\beta_1$	$\beta_2$
100	4	5	0.096	0.096	0.090	0.087	0.074	0.055	0.051	0.021
200	8	10	0.103	0.103	0.096	0.079	0.079	0.058	0.054	0.022
500	20	20	0.107	0.107	0.101	0.098	0.086	0.065	0.063	0.026
1000	40	25	0.108	0.108	0.105	0.102	0.094	0.078	0.076	0.038
2000	80	50	0.109	0.109	0.105	0.103	0.094	0.078	0.077	0.038
5000	200	50	0.109	0.109	0.108	0.107	0.103	0.095	0.095	0.066
10000	400	100	0.109	0.109	0.108	0.107	0.103	0.095	0.095	0.066

procedures perform compared to the corresponding stepwise BH procedure in this case. Table 4 presents the values of  $\beta_1$  and  $\beta_2$  for some values of  $(n, k, g, \rho)$  and  $\alpha = 0.05$ . This time, our stepwise procedures are seen to uniformly dominate the corresponding stepwise BH procedure even for positive correlations as large as 0.5. For correlation larger than 0.5, while the procedure corresponding to  $\beta_1$  continues to uniformly dominate the corresponding stepwise BH procedure, the procedure corresponding to  $\beta_2$  works well only when the number of null hypotheses is large.

We investigate the extent of power improvement our  $k$ -FDR stepup procedures offer over the stepup BH procedure. We simulated the average power, the expected proportion of false null hypotheses that are rejected, for each of these stepup procedures. Figure 1 presents this power comparison, with our procedures corresponding to  $\beta_1$  and  $\beta_2$  labelled  $k$ -FDR SU 1 and  $k$ -FDR SU 2, respectively, and the BH procedure labelled FDR BH. Each simulated power was obtained by (i) generating  $n = 200$  dependent normal random variables  $N(\mu_i, 1)$ ,  $i = 1, \dots, n$ , with a common correlation  $\rho = 0.1$  and with  $n_1$  of the 200

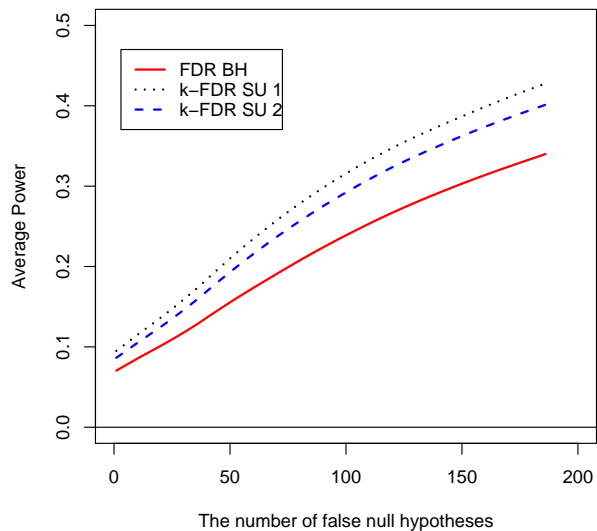


Figure 1: Power of two  $k$ -FDR stepup procedures in the case of positive dependence with parameters  $n = 200, k = 8, d = 2, \rho = 0.1$  and  $\alpha = 0.05$ .

$\mu_i$ 's being equal to  $d = 2$  and the rest 0, (ii) applying the corresponding stepup procedure with  $k = 8$  to the generated data to test  $H_i : \mu_i = 0$  against  $K_i : \mu_i > 0$  simultaneously for  $i = 1, \dots, 200$  at  $\alpha = 0.05$ , and (iii) repeating steps (i) and (ii) 1,000 times before observing the proportion of the  $n_1$  false  $H_i$ 's that are correctly declared significant. As we can see from this figure, our proposed  $k$ -FDR stepup procedures are uniformly more powerful than the BH stepup procedure under weak dependence, with the power difference getting significantly higher with increasing number of false null hypotheses.

## 4 An application to gene expression data

Hereditary breast cancer is known to be associated with mutations in BRCA1 and BRCA2 proteins. Hedenfalk et al. (2001) report that a group of genes are differentially expressed between tumors with BRCA1 mutations and tumors with BRCA2 mutations. The data, which are publicly available from the web site [http://research.nhgri.nih.gov/microarray/NEJM\\_Supplement/](http://research.nhgri.nih.gov/microarray/NEJM_Supplement/), consist of 22 breast cancer samples, among which  $n_1 = 7$  are BRCA1

Table 5: *Numbers of differentially expressed genes for the data in Hedenfalk et al. (2001) using the  $k$ -FDR stepdown procedure in Proposition 3.2.*

	$k = 2$	$k = 5$	$k = 10$	$k = 20$	$k = 30$	$k = 40$	$k = 50$
$\alpha = 0.05$	74	75	76	82	94	110	130
$\alpha = 0.07$	103	110	124	131	139	159	162

mutants,  $n_2 = 8$  are BRCA2 mutants, and  $n_3 = 7$  are sporadic (not used in this illustration). Expression levels in terms of florescent intensity ratios of a tumor sample to a common reference sample, are measured for 3,226 genes using cDNA microarrays. If any gene has one ratio exceeding 20, then this gene is eliminated, since a value of 20 is so high that it does not seem trustworthy for this data set. Such preprocessing leaves  $m = 3,170$  genes.

We tested each gene for differential expression between these two tumor types by using a two-sample  $t$ -test statistic. For each gene, the base 2 logarithmic transformation of the ratio is obtained before we compute the two-sample  $t$ -test statistic. We then calculate the raw  $p$ -value by using a permutation method from Storey and Tibshirani (2003) with the times of permutation  $B = 2,000$ . Finally, we adjust these raw  $p$ -values by using the BH stepup procedure and our proposed  $k$ -FDR stepdown procedure in Proposition 3.2. At  $\alpha = 0.05$  and  $0.07$ , the BH procedure results in 73 and 95 significant genes, respectively, while those numbers for our  $k$ -FDR method are presented in Table 5 for  $k = 2, 5, 10, 20, 30, 40$  and  $50$ . The  $k$ -FDR stepdown procedure is seen to always detect more differentially expressed genes than the BH procedure for these values of  $k$  and  $\alpha$ .

## 5 $k$ -FDR procedure under arbitrary dependence

We now consider developing a stepup procedure with a control of the  $k$ -FDR under any form of dependence of the  $p$ -values, which will uniformly dominate the BY proce-

cedure with the critical values  $\alpha_i = i\alpha/n \sum_{j=1}^n \frac{1}{j}$ ,  $i = 1, \dots, n$ , or its modification obtained by keeping the first  $k - 1$  critical values same as the  $k$ th one. Define  $p_{ijr} = P \left\{ \hat{P}_i \in [\alpha_{j-1}, \alpha_j], V \geq k, R = r \right\}$ , given a set critical values  $0 = \alpha_0 < \alpha_1 < \dots < \alpha_n$ . Then, from (2) we have,

$$\begin{aligned}
k\text{-FDR} &= \sum_{r=k}^n \sum_{i=1}^{n_0} \frac{1}{r} \Pr \left\{ \hat{P}_i \leq \alpha_r, V \geq k, R = r \right\} = \sum_{r=k}^n \sum_{i=1}^{n_0} \sum_{j=1}^r \frac{1}{r} p_{ijr} \\
&= \sum_{i=1}^{n_0} \sum_{j=1}^n \sum_{r=k \vee j}^n \frac{1}{r} p_{ijr} \leq \sum_{i=1}^{n_0} \sum_{j=1}^n \frac{1}{k \vee j} \sum_{r=k \vee j}^n p_{ijr} \\
&\leq \sum_{i=1}^{n_0} \sum_{j=1}^n \frac{1}{k \vee j} \Pr \left\{ \hat{P}_i \in [\alpha_{j-1}, \alpha_j], V \geq k \right\} \\
&\leq \sum_{i=1}^{n_0} \sum_{j=1}^n \frac{\alpha_j - \alpha_{j-1}}{k \vee j}. \tag{16}
\end{aligned}$$

Let  $\alpha_i = (i \vee k)\alpha_k/k$ , for some fixed  $0 < \alpha_k < 1$ . Then,

$$k\text{-FDR} \leq n_0 \frac{\alpha_k}{k} \left\{ 1 + \sum_{j=k+1}^n \frac{1}{j} \right\}. \tag{17}$$

Thus, we have the following theorem.

**Theorem 5.1** *A stepup procedure with the critical values*

$$\alpha_i = \frac{(i \vee k)\alpha_k}{k \left\{ 1 + \sum_{j=k+1}^n \frac{1}{j} \right\}}, \quad i = 1, \dots, n,$$

for a fixed  $0 < \alpha_k < 1$ , controls the  $k$ -FDR at  $\alpha$ .

**Remark 5.1** It would be interesting to see if there exists a joint distribution of  $p$ -values under which the  $k$ -FDR of a stepup procedure with critical values of the form  $\alpha_i = (i \vee k)\alpha_k/k$ ,  $i = 1, \dots, n$ , for some  $0 < \alpha_k < 1$ , is equal to the upper bound in (17). Indeed, there is such a joint distribution, which is presented in the Appendix.

## APPENDIX A: THE CONSTRUCTION OF A JOINT DISTRIBUTION

This appendix presents the construction a joint distribution of  $p$ -values under which the  $k$ -FDR of the stepup procedure with critical values of the form  $\alpha_i = (i \vee k)\alpha_k/k$ ,  $i = 1, \dots, n$ , for some  $0 < \alpha_k < 1$ , attains the upper bound in (17).

Let  $I = \{1, 2, \dots, n\}$  denote the set of indices of all null hypotheses, with  $I_0$  and  $I_1$  being those of true and false null hypotheses, respectively. The main idea is to construct a joint distribution under which for each  $i \in I_0$  and  $j \geq k$ , each  $p$ -value  $\hat{P}_i \sim U[0, 1]$  and the event  $\{\hat{P}_i \in [\alpha_{j-1}, \alpha_j]\}$  is same as the event  $\{V \geq k, R = j\}$ . That is, if  $j = r$ ,

$$\begin{aligned} p_{ijr} &= P\{\hat{P}_i \in [\alpha_{j-1}, \alpha_j], V \geq k, R = r\} \\ &= P\{\hat{P}_i \in [\alpha_{j-1}, \alpha_j]\} = \alpha_j - \alpha_{j-1} . \end{aligned}$$

Otherwise  $p_{ijr} = 0$ . Then, by the second line in (16), we have

$$k\text{-FDR} = \sum_{i=1}^{n_0} \sum_{j=1}^n \sum_{r=k \vee j}^n \frac{1}{r} p_{ijr} = n_0 \frac{\alpha_k}{k} \left\{ 1 + \sum_{j=k+1}^n \frac{1}{j} \right\} .$$

The construction of the joint distribution proceeds as follows: let  $U_k, \dots, U_{n+1}$  be  $n-k+2$  uniformly distributed random variables such that  $U_k \sim U[0, \alpha_k]$ ,  $U_{n+1} \sim U[\alpha_n, 1]$ , and  $U_i \sim U[\alpha_{i-1}, \alpha_i]$ ,  $i = k+1, \dots, n$ . Let  $M$  be a random variable taking values  $k, \dots, n+1$  with the following probability distribution.

$$P\{M = m\} = \begin{cases} n_0 \frac{\alpha_k}{k} & \text{if } m = k , \\ n_0 \frac{\alpha_k}{k} \cdot \frac{1}{m} & \text{if } k+1 \leq m \leq n_0 , \\ \frac{\alpha_k}{k} & \text{if } n_0 + 1 \leq m \leq n , \\ 1 - \left\{ n + n_0 \sum_{j=k+1}^{n_0} \frac{1}{j} \right\} \frac{\alpha_k}{k} & \text{if } m = n+1 . \end{cases}$$

Here, we assume  $\{n + n_0 \sum_{j=k+1}^{n_0} \frac{1}{j}\} \frac{\alpha_k}{k} \leq 1$ , which is easily satisfied. We want to associate a  $p$ -value to each of the indices  $1, 2, \dots, n$ . The association proceeds in 3 distinct phases.

**Phase 1.** For each given  $M = m \in \{k, \dots, n_0\}$ , choose  $m$  indices  $i_1, i_2, \dots, i_m$  randomly without replacement from  $I_0$ . Each of these chosen indices has the same  $p$ -value  $P_{i_j} = U_m$ . Each of the indices  $s$  in  $I - \{i_1, i_2, \dots, i_m\}$  has the same  $p$ -value  $P_s = U_{n+1}$ .

**Phase 2.** For each given  $M = m \in \{n_0 + 1, \dots, n\}$ , choose all  $n_0$  indices from  $I_0$  and  $(m - n_0)$  indices randomly without replacement from  $I_1$ . Each of these indices is associated with the same  $p$ -value  $U_m$ . Each of the remaining unaccounted indices from  $I_1$  is associated with the same  $p$ -value  $U_{n+1}$ .

**Phase 3.** Given  $M = n + 1$ , each of the indices in  $I_0$  is associated with the same  $p$ -value  $U_{n+1}$ , and each of the indices in  $I_1$  is associated with the same  $p$ -value zero.

It is easy to verify that, for each  $i \in I_0$  and  $j \geq k$ , unconditionally, each  $p$ -value  $\hat{P}_i \sim U[0, 1]$  and when  $\hat{P}_i \in [\alpha_{j-1}, \alpha_j]$ , there are totally  $j$   $p$ -values and at least  $k$   $p$ -values corresponding to true null hypotheses less than  $\alpha_j$ , i.e., the event  $\{\hat{P}_i \in [\alpha_{j-1}, \alpha_j]\}$  is same as the event  $\{V \geq k, R = j\}$ .

## References

- [1] BENJAMINI, Y. and HOCHBERG, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B* **57** 289-300.
- [2] BENJAMINI, Y., KRIEGER, A. M. and YEKUTIELI, D. (2006). Adaptive linear step-up false discovery rate controlling procedures. *Biometrika* **93** 491-507.
- [3] BENJAMINI, Y. and YEKUTIELI, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.* **29** 1165-1188.
- [4] BENJAMINI, Y. and YEKUTIELI, D. (2005). False discovery rate-adjusted multiple confidence intervals for selected parameters. *J Amer. Statist. Assoc.* **100** 71-93.

- [5] DUDOIT, S., VAN DER LAAN, M. and POLLARD, K. (2004). Multiple testing: Part I. Single-step procedures for control of general type I error rates. *Statist. App. Gen. Mol. Bio.* **3**(1): Article 13.
- [6] FINNER, H., DICKHAUS, T. and ROTERS, M. (2007). Dependency and false discovery rate: Asymptotics. *Ann. Statist.* **35** 1432-1455.
- [7] FINNER, H., DICKHAUS, T. and ROTERS, M. (2008). On the false discovery rate and an asymptotically optimal rejection curve. *Ann. Statist.* To appear.
- [8] GAVRILOV, Y., BENJMAINI, Y. and SARKAR, S. K. (2008). An adaptive step-down procedure with proven FDR control under independence. *Ann. Statist.* To appear.
- [9] GENOVESE, C. and WASSERMAN, L. (2002). Operating characteristics and extensions of the false discovery rate procedure. *J. Roy. Statist. Soc. Ser. B* **64** 499-517.
- [10] GENOVESE, C. and WASSERMAN, L. (2004). A stochastic process approach to false discovery control. *Ann. Statist.* **32** 1035-1061.
- [11] GUO, W. and RAO, M. B. (2006). On generalized closure principle for generalized familywise error rates. Unpublished Report.
- [12] GUO, W. and ROMANO, J. P. (2007). A generalized Sidak-Holm procedure and control of generalized error rates under independence. *Statist. App. Gen. Mol. Bio.* **6** (1): Article 3.
- [13] HEDENFALK, I., DUGGAN, D., CHEN, Y., RADMACHER, M., BITTNER, M., SIMON, R., MELTZER, P., GUSTERSON, B., ESTELLER, M., KALLIONIEMI, OP, WILFOND, B., BORG, A. and TRENT, J. (2001). Gene-expression profiles in hereditary breast cancer. *New Eng. J. Med.* **344** 539-548.
- [14] HOCHBERG, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* **75** 800-802.

- [15] HOCHBERG, Y. and TAMHANE, A. C. (1987). *Multiple Comparison Procedures*. Wiley, NY.
- [16] HOLM, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Statist.* **6** 65-70.
- [17] HOMMEL, G. and HOFFMANN, T. (1987). Controlled uncertainty. In *Multiple Hypothesis Testing* P. Bauer, G. Hommel and E. Sonnemann, eds. 154-161, Springer, Heidelberg.
- [18] KARLIN, S. and RINOTT, Y. (1980). Classes of orderings of measures and related correlation inequalities I: Multivariate totally positive distributions. *J. Mult. Anal.* **10** 467-498.
- [19] KORN, E., TROENDLE, T., MCSHANE, L. and SIMON, R. (2004). Controlling the number of false discoveries: Application to high-dimensional genomic data. *J. Statist. Plann. Inf.* **124** 279-398.
- [20] LEHMANN, E. L. and ROMANO, J. P. (2005). Generalizations of the familywise error rate. *Ann. Statist.* **33** 1138-1154.
- [21] ROMANO, J. P. and SHAIKH, A. M. (2006a). Stepup procedures for control of generalizations of the familywise error rate. *Ann. Statist.* **34** 1850-1873.
- [22] ROMANO, J. P. and SHAIKH, A. M. (2006b). On stepdown control of the false discovery proportion. In *The Second E.L. Lehmann Symposium - Optimality* J. Rojo, eds. 40-61, IMS Lecture Notes - Monograph Series **49**.
- [23] ROMANO, J. P. and WOLF, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica* **73** 1237-1282.
- [24] ROMANO, J. P. and WOLF, M. (2007). Control of generalized error rates in multiple testing. *Ann. Statist.* **35** 1378-1408.

- [25] SARKAR, S. K. (2002). Some results on false discovery rate in stepwise multiple testing procedures. *Ann. Statist.* **30** 239-257.
- [26] SARKAR, S. K. (2004). FDR-controlling stepwise procedures and their false negatives rates. *J. Statist. Plann. Inf.* **125** 119-137.
- [27] SARKAR, S. K. (2006). False discovery and false non-discovery rates in single-step multiple testing procedures. *Ann. Statist.* **34** 394-415.
- [28] SARKAR, S. K. (2007). Stepup procedures controlling generalized FWER and generalized FDR. *Ann. Statist.* **35** 2405-2420.
- [29] SARKAR, S. K. (2008a). Two-stage stepup procedures controlling FDR. *J. Statist. Plann, Inf.* **138**, 1072-1084.
- [30] SARKAR, S. K. (2008b). Generalizing Simes' test and Hochberg's stepup procedure. *Ann. Statist.* **36** 337-363.
- [31] SARKAR, S. K. and GUO, W. (2008). On a generalized false discovery rate. *Ann. Statist.* To appear.
- [32] STOREY, J. D. (2002). A direct approach to false discovery rates. *J. Roy. Statist. Soc. Ser. B* **64** 479-498.
- [33] STOREY, J. D. (2003). The positive false discovery rate: A Bayesian interpretation and the q-value. *Ann. Statist.* **31** 2013-2035.
- [34] STOREY, J. D., TAYLOR, J. E. and SIEGMUND, D. (2004). Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: A unified approach. *J. Roy. Statist. Soc. Ser. B* **66** 187-205.
- [35] STOREY, J. D. and TIBSHIRANI, R. (2003). Statistical significance for genomewide studies. *Proc. Nat. Acad. Sci.* **100** 9440-9445.

[36] VAN DER LAAN, M., DODOIT, S. and POLLARD, K. (2004). Augmentation procedures for control of the generalized family-wise error rate and tail probabilities for the proportion of false positives. *Stat. App. Gen. Mol. Bio.* **3**:(1), Article 15.

Department of Statistics

Temple University

E-mail: (sanat@temple.edu)

Biostatistics Branch

National Institute of Environmental Health Sciences

E-mail: (wenge.guo@gmail.com)