# Adaptive Controls of FWER and FDR Under Block <br> Dependence 

By WENGE GUO<br>New Jersey Institute of Technology, Newark, NJ 07102, USA<br>wenge.guo@njit.edu

And SANAT K. SARKAR

Temple University, Philadelphia, PA 19122, USA
sanat@temple.edu

## SUMMARY

Often in multiple testing, the hypotheses appear in non-overlapping, equal sized blocks with the associated $p$-values exhibiting dependence within but not between blocks. We consider adapting the Bonferroni method for controlling the familywise error rate (FWER) and the BenjaminiHochberg method for controlling the false discovery rate (FDR) to such dependence structure without losing their ultimate controls over the FWER and FDR, respectively, in a non-asymptotic setting. We present variants of conventional adaptive Bonferroni and Benjamini-Hochberg methods with proofs of their respective controls over the FWER and FDR. Numerical evidence is presented to show that these new adaptive methods can capture the present dependence structure more effectively than the corresponding conventional adaptive methods. This paper offers a solution to the open problem of constructing adaptive FWER and FDR controlling methods under dependence in a non-asymptotic setting and providing real improvements over the corresponding non-adaptive ones.

## 1. Introduction

In many multiple hypothesis testing problems arising in modern scientific investigations, the hypotheses appear in non-overlapping, equal-sized blocks. Such block formation is often a natural phenomenon due to the underlying experimental process or can be created based on other considerations. For instance, the hypotheses corresponding to (i) the different time-points in a microarray time-course experiment (Guo, Sarkar and Peddada, 2010; Sun and Wei, 2011) for each gene; or (ii) the phenotypes (or the genetic models) with (or using) which each marker is tested in a genome-wide association study (Lei et al., 2006); or (iii) the conditions (or subjects) considered for each voxel in brain imaging (Heller et al. 2007), naturally form a block. While applying multiple testing in astronomical transient source detection from nightly telescopic image consisting of large number of pixels (each corresponding to a hypotheses), Clements, Sarkar and Guo (2011) considered grouping the pixels into blocks of equal size based on telescope 'point spread function.'

A special type of dependence, which we call block dependence, is the relevant dependence structure that one should take into account while constructing multiple testing procedures in presence of such blocks. This dependence can be simply described by saying that the hypotheses or the corresponding $p$-values are mostly dependent within but not between blocks. Also known as the clumpy dependence (Storey, 2003), this has been considered mainly in simulation studies to investigate how multiple testing procedures proposed under independence continue to perform under it (Benjamini, Krieger and Yekutieli, 2006; Finner, Dickhaus, and Roters, 2007; Sarkar, Guo and Finner, 2012, and Storey, Taylor and Siegmund, 2004), not in offering FWER or FDR controlling procedures precisely utilizing it. In this article, we focus on constructing procedures controlling the familywise error rate (FWER) and the false discovery rate (FDR) that incorporate the block dependence in a non-asymptotic setting in an attempt to improve the corresponding procedures that ignore this structure. More specifically, we consider the Bonferroni method for the FWER control and the Benjamini-Hochberg (BH, 1995) method for the FDR control and
adapt them to the data in two ways - incorporating the block dependence and estimating the number of true null hypotheses capturing such dependence.

Adapting to unknown number of true nulls has been a popular way to improve the FWER and FDR controls of the Bonferroni and BH methods, respectively. However, construction of such adaptive methods with proven control of the ultimate FWER or FDR in a non-asymptotic setting and providing real improvements under dependence is an open problem (Benjamini, Krieger and Yekutieli, 2006; Blanchard and Roquaine, 2009). We offer some solutions to this open problem in this paper under a commonly encountered type of dependence, the block dependence.

## 2. PRELIMINARIES

Suppose that $H_{i j}, i=1, \ldots, g ; j=1, \ldots, s$, are the $n=g s$ null hypotheses appearing in $g$ blocks or groups of size $s$ each that are to be simultaneously tested based on their respective $p$-values $P_{i j}, i=1, \ldots, g ; j=1, \ldots, s$. Let $n_{0}$ of these null hypotheses be true, which for notational convenience will often be identified by $\hat{P}_{i j}$ 's. We assume that $\hat{P}_{i j} \sim U(0,1)$ and make the following assumption regarding dependence of $P_{i j}$ 's:

ASSUMPTION 1. (Block Dependence) The $p$-value row vectors $\left(P_{i 1}, \ldots, P_{i s}\right), i=1, \ldots, g$, forming the $g$ blocks of size s each, are independent to each other.

Under this assumption, the null $p$-values are independent between but not within blocks. Regarding dependence within blocks, our assumption will depend on whether we want to control the FWER or the FDR. More specifically, we develop methods controlling the FWER under arbitrary dependence and the FDR under positive dependence of the $p$-values within each block. The positive dependence condition, when assumed for each $i$, will be of the type characterized by the following:

$$
\begin{equation*}
E\left\{\phi\left(P_{i 1}, \ldots, P_{i s}\right) \mid \hat{P}_{i j} \leq u\right\} \uparrow u \in(0,1) \tag{1}
\end{equation*}
$$

for each $\hat{P}_{i j}$ and any (coordinatewise) non-decreasing function $\phi$. This type of positive dependence is commonly encountered and used in multiple testing; see, for instance, Sarkar (2008) for references.

We will be using two types of multiple testing procedure in this paper - stepup and singlestep. Let $\left(P_{i}, H_{i}\right), i=1, \ldots, n$, be the pairs of $p$-value and the corresponding null hypothesis, and $P_{(1)} \leq \cdots \leq P_{(n)}$ be the ordered $p$-values. Given a set of critical constants $0 \leq \alpha_{1} \leq \cdots \leq$ $\alpha_{n} \leq 1$, a stepup test rejects $H_{i}$ for all $i$ such that $P_{i} \leq P_{(R)}$, where $R=\max \{1 \leq i \leq n$ : $\left.P_{(i)} \leq \alpha_{i}\right\}$, provided this maximum exists, otherwise, it accepts all the null hypotheses. A singlestep test rejects $H_{i}$ if $P_{i} \leq c$ for some constant $c \in(0,1)$.

Let $V$ be the number of falsely rejected among all the $R$ rejected null hypotheses in a multiple testing procedure. Then, the FWER and FDR of this procedure are defined respectively by FWER $=\operatorname{pr}(V \geq 1)$ and $\operatorname{FDR}=E(V / \max \{R, 1\})$.

## 3. Estimating $n_{0}$ UNDER BLOCK DEPENDENCE

Let $\mathbf{P}=\left(\left(P_{i j}\right)\right)$ denote the matrix of $p$-values and $\mathbf{H}=\left(\left(H_{i j}\right)\right)$, where $H_{i j}=0$ or 1 according to it is true or false. We consider estimating $n_{0}=\sum_{i=1}^{g} \sum_{j=1}^{s} I\left(H_{i j}=0\right)$ using an estimate $\widehat{n}_{0}(\mathbf{P})$ satisfying the following property while constructing our adaptive methods in the following sections:

PROPERTY 1. $\widehat{n}_{0}(\mathbf{P})$ is non-decreasing in each $P_{i j}$, and

$$
\begin{equation*}
\sum_{i=1}^{g} \sum_{j=1}^{s} I\left(H_{i j}=0\right) E_{D U}\left\{\frac{1}{\widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right)}\right\} \leq 1 \tag{2}
\end{equation*}
$$

where $\mathbf{P}^{(-i)}$ is the $(g-1) \times s$ sub-matrix of $\mathbf{P}$ obtained by deleting its ith row, $\hat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right)$ is obtained from $\hat{n}_{0}(\mathbf{P})$ by replacing the entries in the ith row of $\mathbf{P}$ by zeros, and $E_{D U}$ is the expectation under the Dirac-uniform configuration of $\mathbf{P}^{(-i)}$, that is, when the p-values in $\mathbf{P}^{(-i)}$ that correspond to the false null hypotheses are set to 0 and each of the remaining $p$-values are considered to be uniformly distributed on [0,1].

The following two results provide examples of estimate of $n_{0}$ satisfying Property 1 under Assumption 1.

Result 1. Consider the estimate

$$
\begin{equation*}
\widehat{n}_{0}^{(1)}=\frac{n-R(\lambda)+s}{1-\lambda} \tag{3}
\end{equation*}
$$

for any $(2 g+3)^{-\frac{2}{g+2}} \leq \lambda<1$, where $R(\lambda)=\sum_{i=1}^{g} \sum_{j=1}^{s} I\left(P_{i j} \leq \lambda\right)$ is the number of $p$ values in $\mathbf{P}$ not exceeding $\lambda$. It satisfies Property 1 under Assumption 1.

Result 2. Consider the estimate

$$
\widehat{n}_{0}^{(2)}=\frac{s[(g-k) s+1]}{1-P_{k s: g s}},
$$

for any $1 \leq k \leq g$, where $P_{k s: g s}$ is the $(k s)$ th ordered among all the $n=g s p$-values in $\mathbf{P}$. It satisfies Property 1 under Assumption 1.

Remark 1. When $s=1$, the estimates $\widehat{n}_{0}^{(1)}$ and $\widehat{n}_{0}^{(2)}$ reduce to the ones considered in the contexts of adaptive FDR control (Benjamini, Krieger and Yekutieli, 2006; Blanchard and Roquain, 2009; Sarkar, 2008 and Storey, Taylor and Siegmund, 2004) and adaptive FWER control (Finner and Gontscharuk; 2009, Guo, 2009 and Sarkar, Guo and Finner, 2012). Of course, Result 1 holds for any $\lambda \in(0,1)$ when $s=1$. For notational convenience, we use $\widehat{n}_{0}^{(0)}$ to denote Storey et al.'s estimate, i.e.,

$$
\widehat{n}_{0}^{(0)}=\frac{n-R(\lambda)+1}{1-\lambda} .
$$

REmARK 2. Although the estimate $\widehat{n}_{0}^{(2)}$ has been theoretically proved to satisfy Property 1 under Assumption 1, our numerical results (not shown here) illustrate that $\widehat{n}_{0}^{(2)}$ is very conservative as an estimate of $n_{0}$. Therefore, in this paper, we concentrate on the estimate $\widehat{n}_{0}^{(1)}$ and propose new adaptive multiple testing methods mainly based on $\widehat{n}_{0}^{(1)}$.

To prove Result 1, we use the following lemmas, with proofs given or outlined in Appendix:

Lemma 1. Given a $p \times q$ matrix $\mathbf{A}=\left(\left(a_{i j}\right)\right)$, where $a_{i j}=0$ or 1 and $\sum_{i=1}^{p} \sum_{j=1}^{q} a_{i j}=m$, the entries of $\mathbf{A}$ can be always rearranged to form a new matrix $\mathbf{B}=\left(\left(b_{i j}\right)\right)$ in such a way that, for each $j=1, \ldots, q$, the entries in the $j$ th column of $\mathbf{B}$ are the entries of $\mathbf{A}$ in different rows, $\sum_{i=1}^{p} b_{i j}=\left\lfloor\frac{m}{q}\right\rfloor$ or $\left\lfloor\frac{m}{q}\right\rfloor+1$, and $\sum_{i=1}^{p} \sum_{j=1}^{q} b_{i j}=m$.

Lemma 2. For any set of positive real numbers $a_{1}, \ldots, a_{m}$, the following inequality holds:

$$
\frac{1}{\sum_{i=1}^{m} a_{i}} \leq \frac{1}{m^{2}} \sum_{i=1}^{m} \frac{1}{a_{i}} .
$$

Lemma 3. If $X \sim \operatorname{Bin}(n, \pi)$, then

$$
E\left(\frac{1}{X+1}\right)=\frac{1-(1-\pi)^{n+1}}{(n+1) \pi} .
$$

Lemma 4. The function $f(x)=(2 x+3)^{-\frac{2}{x+2}}$ is increasing in $x \geq 1$ and $f(x) \leq f(1)$ for all $0 \leq x \leq 1$.

Proof of Result 1. First, consider the expectation

$$
E_{D U}\left\{\frac{1}{\widehat{n}_{0}^{(1)}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right)}\right\}
$$

in terms of $\mathbf{P}^{(-i)}$. Let $\mathbf{H}^{(-i)}$ be the sub-matrix of $\mathbf{H}$ corresponding to $\mathbf{P}^{(-i)}$. Since this expectation remains unchanged under the type of rearrangements considered in Lemma 1 for $\mathbf{H}^{(-i)}$, we can assume without any loss of generality that the number of true null $p$-values in the $j$ th column of $\mathbf{P}^{(-i)}$ is $n_{0 j}^{(-i)}=\left\lfloor\frac{n_{0}-m_{i}}{s}\right\rfloor$ or $\left\lfloor\frac{n_{0}-m_{i}}{s}\right\rfloor+1$ for each $j=1, \ldots, s$, where $m_{i}=\sum_{j=1}^{s} I\left(H_{i j}=0\right)$.

Let $\widehat{W}_{j}^{(-i)}(\lambda)=\sum_{i^{\prime}(\neq i)=1}^{g} I\left(H_{i^{\prime} j}=0, P_{i^{\prime} j}>\lambda\right)$, for $j=1, \ldots, s$. Under Assumption 1 and the Dirac-uniform configuration of $\mathbf{P}^{(-i)}, \widehat{W}_{j}^{(-i)}(\lambda) \sim \operatorname{Bin}\left(n_{0 j}^{(-i)}, 1-\lambda\right)$. So, we have

$$
\begin{align*}
& E_{D U}\left\{\frac{1}{\widehat{n}_{0}^{(1)}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right)}\right\}=E\left\{\frac{1-\lambda}{\sum_{j=1}^{s}\left[\widehat{W}_{j}^{(-i)}(\lambda)+1\right]}\right\} \\
\leq & \frac{1}{s^{2}} \sum_{j=1}^{s} E\left\{\frac{1-\lambda}{\widehat{W}_{j}^{(-i)}(\lambda)+1}\right\}=\frac{1}{s^{2}} \sum_{j=1}^{s} \frac{1-\lambda^{n_{0 j}^{(-i)}+1}}{n_{0 j}^{(-i)}+1}, \tag{4}
\end{align*}
$$

with the first inequality following from Lemma 2 and the second equality following from Lemma 3.

Let $n_{0}-m_{i}=\left(a_{i}+\beta_{i}\right) s$, for some non-negative integer $a_{i}$ and $0 \leq \beta_{i}<1$. Note that

$$
\begin{equation*}
a_{i} s \leq n_{0} \leq\left(a_{i}+\beta_{i}+1\right) s \tag{5}
\end{equation*}
$$

Also, $\left(1-\beta_{i}\right)$ proportion of the $s$ values $n_{0 j}^{(-i)}, j=1, \ldots, s$, are all equal to $a_{i}$ and the remaining $\beta_{i}$ proportion are all equal to $a_{i}+1$. So, the right-hand side of (4) is equal to

$$
\begin{aligned}
& \frac{1}{s}\left[\frac{1-\beta_{i}}{a_{i}+1}\left(1-\lambda^{a_{i}+1}\right)+\frac{\beta_{i}}{a_{i}+2}\left(1-\lambda^{a_{i}+2}\right)\right] \leq \frac{1}{s}\left[\frac{1-\beta_{i}}{a_{i}+1}+\frac{\beta_{i}}{a_{i}+2}\right]\left(1-\lambda^{a_{i}+2}\right) \\
= & \frac{\left(a_{i}+2-\beta_{i}\right)\left(1-\lambda^{a_{i}+2}\right)}{s\left(a_{i}+1\right)\left(a_{i}+2\right)} \leq \frac{\left(a_{i}+1+\beta_{i}\right)\left(a_{i}+2-\beta_{i}\right)\left(1-\lambda^{a_{i}+2}\right)}{n_{0}\left(a_{i}+1\right)\left(a_{i}+2\right)} \\
= & \frac{1}{n_{0}}\left[1+\frac{\beta_{i}\left(1-\beta_{i}\right)}{\left(a_{i}+1\right)\left(a_{i}+2\right)}\right]\left(1-\lambda^{a_{i}+2}\right) \leq \frac{1}{n_{0}}\left[1+\frac{1}{4\left(a_{i}+1\right)\left(a_{i}+2\right)}\right]\left(1-\lambda^{a_{i}+2}\right) .
\end{aligned}
$$

The desired inequality (2) then holds for this estimate if

$$
\left[1+\frac{1}{4\left(a_{i}+1\right)\left(a_{i}+2\right)}\right]\left(1-\lambda^{a_{i}+2}\right) \leq 1
$$

which is true if and only if

$$
\begin{equation*}
\lambda \geq\left[1+4\left(a_{i}+1\right)\left(a_{i}+2\right)\right]^{-\frac{1}{a_{i}+2}}=\left(2 a_{i}+3\right)^{-\frac{2}{a_{i}+2}} \tag{6}
\end{equation*}
$$

Let $f\left(a_{i}\right)=\left(2 a_{i}+3\right)^{-\frac{2}{a_{i}+2}}$. As seen from (5), $a_{i} \leq n_{0} / s \leq g$, thus, the inequality $f(g) \geq f\left(a_{i}\right)$ holds for all $a_{i} \geq 0$, since $f(g) \geq f\left(a_{i}\right)$ if $a_{i} \geq 1$ and $f(g) \geq f(1) \geq f\left(a_{i}\right)$ if $0 \leq a_{i} \leq 1$, due to Lemma 4. So, the inequality (6) holds if $\lambda \geq(2 g+3)^{-2 /(g+2)}$. This proves the result.

Result 2 can be proved with the help of the following lemma, a proof of which again will be outlined in Appendix:

LEMMA 5. Given a set of numbers $a_{1}, \ldots, a_{m}$, consider their ordered values $-\infty \leq a_{1: m} \leq$ $\cdots \leq a_{m: m}$. The following inequality holds: $a_{r: m} \geq a_{r-l: m-l}$ for all $r, l \leq m$, where $a_{r-l: m-l}$ is
the $(r-l)$ th ordered component in any subset of size $m-l$ of $a_{1}, \ldots, a_{m}$, and $a_{r-l: m-l}=-\infty$ if $r-l \leq 0$.

Proof of Result 2. As in proving Result 1, we can assume without any loss of generality that the number of true null $p$-values in the $j$ th column of $\mathbf{P}^{(-i)}$ is $n_{0 j}^{(-i)}=\left\lfloor\frac{n_{0}-m_{i}}{s}\right\rfloor$ or $\left\lfloor\frac{n_{0}-m_{i}}{s}\right\rfloor+1$, for each $j=1, \ldots, s$.

Let $P_{(k-1) s:(g-1) s}^{(-i)}$ be the $[(k-1) s]$ th ordered $p$-value among those in $\mathbf{P}^{(-i)}$. Then,

$$
\begin{equation*}
E_{D U}\left\{\frac{1}{\widehat{n}_{0}^{(2)}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right)}\right\}=\frac{1}{s} E_{D U}\left\{\frac{1-P_{(k-1) s:(g-1) s}^{(-i)}}{(g-k) s+1}\right\} \tag{7}
\end{equation*}
$$

Now, let $P_{k-1: g-1}^{(-i j)}$ denote the $(k-1)$ th ordered $p$-value among those in $j$ th column of $\mathbf{P}^{(-i)}$. Apply Lemma 5 with $m=(g-1) s, r=(k-1) s$, and $l=(g-1)(s-1)$ to see that $P_{(k-1) s:(g-1) s}^{(-i)} \geq P_{\tilde{k}-1: g-1}^{(-i j)}$, with $\tilde{k}-1=(k-1) s-(g-1)(s-1)$, i.e., $\tilde{k}=g-(g-k) s$, for each $j=1, \ldots, s$. Thus, noting that $(g-k) s+1=g-\tilde{k}+1$, we have

$$
\begin{align*}
E_{D U}\left\{1-P_{(k-1) s:(g-1) s}^{(-i)}\right\} & \leq \min _{1 \leq j \leq s} E_{D U}\left\{1-P_{\tilde{k}-1: g-1}^{(-i j)}\right\}=\min _{1 \leq j \leq s} \min \left\{1, \frac{g-\tilde{k}+1}{n_{0 j}^{(-i)}+1}\right\} \\
& \leq \min _{1 \leq j \leq s}\left\{\frac{g-\tilde{k}+1}{n_{0 j}^{(-i)}+1}\right\} \tag{8}
\end{align*}
$$

The equality follows from the fact that $E_{D U}\left\{1-P_{\tilde{k}-1: g-1}^{(-i j)}\right\}=\frac{g-\tilde{k}+1}{n_{0 j}^{(-i)}+1}$ if $n_{0 j}^{(-i)}>g-\tilde{k}$ and $=1$, otherwise.

As in proving Result 1 , let $n_{0}-m_{i}=\left(a_{i}+\beta_{i}\right) s$, for some non-negative integer $a_{i}$ and $0 \leq$ $\beta_{i}<1$. Then, since $n_{0 j}^{(-i)}=a_{i}$ or $=a_{i}+1$ for each $j=1, \ldots, s, n_{0 j}^{(-i)}=a_{i}+1$ as $\beta_{i}>0$ for at least one $j=1, \ldots, s$, and $n_{0} \leq\left(a_{i}+\beta_{i}+1\right) s \leq\left(a_{i}+1\right) s+s I\left(\beta_{i}>0\right)$, we note that

$$
\begin{equation*}
\frac{1}{s} \min _{1 \leq j \leq s}\left\{\frac{g-\tilde{k}+1}{n_{0 j}^{(-i)}+1}\right\}=\frac{g-\tilde{k}+1}{\left(a_{i}+1\right) s+s I\left(\beta_{i}>0\right)} \leq \frac{g-\tilde{k}+1}{n_{0}} \tag{9}
\end{equation*}
$$

From (7)-(9), we finally get

$$
E_{D U}\left\{\frac{1}{\widehat{n}_{0}^{(2)}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right)}\right\} \leq 1
$$

the desired result.

## 4. Adaptive FWER CONTROL UNDER BLOCK DEPENDENCE

Our proposed method is based on the idea of adapting the Bonferroni method to the block dependence structure with ultimate control of the FWER in a non-asymptotic setting. Given an estimate $\hat{n}_{0}$ of $n_{0}$ obtained from the available $p$-values, the Bonferroni method can be adapted to the data through $\widehat{n}_{0}$ by rejecting $H_{i j}$ if $P_{i j} \leq \alpha / \widehat{n}_{0}$; see, for instance, Finner and Gontscharuk (2009). Our method is such an adaptive version of the Bonferroni method, but based on the estimate of $n_{0}$ introduced above that captures the block dependence.

## Definition 1. (Adaptive Bonferroni under block dependence)

1. Define an estimate $\widehat{n}_{0}(\mathbf{P})$ satisfying Property 1.
2. Reject $H_{i j}$ if $P_{i j} \leq \alpha / \widehat{n}_{0}(\mathbf{P})$.

THEOREM 1. Consider the block dependence structure in which the p-values within each block are arbitrarily dependent. The FWER of the above adaptive Bonferroni method is strongly controlled at $\alpha$ under such block dependence.

Proof. The FWER of this method is given by

$$
\begin{align*}
\text { FWER } & =\operatorname{pr}\left\{\bigcup_{i=1}^{g} \bigcup_{j=1}^{s}\left(P_{i j} \leq \frac{\alpha I\left(H_{i j}=0\right)}{\widehat{n}_{0}(\mathbf{P})}\right)\right\} \leq \sum_{i=1}^{g} \sum_{j=1}^{s} \operatorname{pr}\left\{P_{i j} \leq \frac{\alpha I\left(H_{i j}=0\right)}{\widehat{n}_{0}(\mathbf{P})}\right\} \\
& \leq \sum_{i=1}^{g} \sum_{j=1}^{s} \operatorname{pr}\left\{P_{i j} \leq \frac{\alpha I\left(H_{i j}=0\right)}{\widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right)}\right\} \leq \alpha \sum_{i=1}^{g} \sum_{j=1}^{s} E_{D U}\left\{\frac{I\left(H_{i j}=0\right)}{\widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right)}\right\} \\
& \leq \alpha . \tag{10}
\end{align*}
$$

In (10), the first inequality follows from the Bonferroni inequality, the second and third follow from the non-decreasing property of $\widehat{n}_{0}$ and that $\widehat{P}_{i j} \sim U(0,1)$ and the assumption of block dependence, and the fourth follows from condition (2) satisfied by $\widehat{n}_{0}$. Thus, the desired result is proved.

Corollary 1. The adaptive Bonferroni methods of the above type based on the estimates $\widehat{n}_{0}^{(1)}$ and $\widehat{n}_{0}^{(2)}$ strongly control the FWER at $\alpha$ under the block dependence considered in Theorem 1.

## 5. AdAptive FDR control under block dependence

Our proposed method in this section is based on the idea of adapting the BH method to the block dependence structure without losing the ultimate control over the FDR in a non-asymptotic setting. This adaptation is done in two steps. First, we adjust the BH method to the block dependence structure and then develop its oracle version given the number of true nulls. Second, we consider the data-adaptive version of this oracle method by estimating $n_{0}$ using our estimate that also captures the block dependence.

Towards adjusting the BH method to the block structure, we note that it is natural to first identify blocks that are significant by applying the BH method to simultaneously test the intersection null hypotheses $\tilde{H}_{i}=\bigcap_{j=1}^{s} H_{i j}, i=1, \ldots, g$, based on some block specific $p$-values, and then go back to each significant block to see which hypotheses in that block are significant. Let $\tilde{P}_{i}$, $i=1, \ldots, g$, be the block $p$-values obtained by combining the $p$-values in each block through a combination function. Regarding the choice of this combination function, we note that the combination test for $\tilde{H}_{i}$ based on $\tilde{P}_{i}$ must allow simultaneous testing of the individual hypotheses $H_{i j}, j=1, \ldots, s$, with a strong control of the FWER. This limits our choice to the Bonferroni adjusted minimum $p$-value; see also Guo, Sarkar and Peddada (2010). With these in mind, we consider adjusting the BH method as follows:

- Choose $\tilde{P}_{i}=s \min _{1 \leq j \leq s} P_{i j}$ as the $i$ th block $p$-value, for $i=1, \ldots, g$.
- Order the block $p$-values as $\tilde{P}_{(1)} \leq \cdots \leq \tilde{P}_{(g)}$, and find $B=\max \left\{1 \leq i \leq g: \tilde{P}_{(i)} \leq\right.$ $i \alpha / g\}$.
- Reject $H_{i j}$ for all $(i, j)$ such that $\tilde{P}_{i} \leq \tilde{P}_{(B)}$ and $s P_{i j} \leq B \alpha / g$, provided the above maximum exists, otherwise, accept all the null hypotheses.

The number of false rejections in this adjusted BH method is given by

$$
V=\sum_{i=1}^{g} \sum_{j=1}^{s} I\left(H_{i j}=0, P_{i j} \leq B \alpha / n\right)
$$

So, with $R$ as the total number of rejections, the FDR of this method under block dependence is

$$
\begin{align*}
\mathrm{FDR} & =\sum_{i=1}^{g} \sum_{j=1}^{s} I\left(H_{i j}=0\right) E\left(\frac{I\left(P_{i j} \leq B \alpha / n\right)}{\max \{R, 1\}}\right) \\
& \leq \sum_{i=1}^{g} \sum_{j=1}^{s} I\left(H_{i j}=0\right) E\left(\frac{I\left(P_{i j} \leq B \alpha / n\right)}{\max \{B, 1\}}\right) \tag{11}
\end{align*}
$$

since $R \geq B$. For each $(i, j)$,

$$
\begin{equation*}
\frac{I\left(P_{i j} \leq B \alpha / n\right)}{\max \{B, 1\}}=\sum_{b=1}^{g} \frac{I\left(P_{i j} \leq b \alpha / n, B^{(-i)}=b-1\right)}{b} \tag{12}
\end{equation*}
$$

where $B^{(-i)}$ is the number of significant blocks detected by the adjusted BH method based on $\left\{\tilde{P}_{1}, \ldots, \tilde{P}_{g}\right\} \backslash\left\{\tilde{P}_{i}\right\}$, the $g-1$ block $p$-values other than the $\tilde{P}_{i}$, and the critical values $i \alpha / g$, $i=2, \ldots, g$. Taking expectation in (12) under the block dependence and applying it to (11), we see that

$$
\begin{aligned}
\mathrm{FDR} & \leq \frac{\alpha}{n} \sum_{i=1}^{g} \sum_{j=1}^{s} I\left(H_{i j}=0\right) \sum_{b=1}^{g} \operatorname{pr}\left(B^{(-i)}=b-1\right)=\alpha / n \sum_{i=1}^{g} \sum_{j=1}^{s} I\left(H_{i j}=0\right) \\
& =n_{0} \alpha / n
\end{aligned}
$$

If $n_{0}$ were known, the FDR control of the adjusted BH method could be made tighter, from $n_{0} \alpha / n$ to $\alpha$, by incorporating $n_{0}$ into it as follows: Let $g_{0}=n_{0} / s$, replace $g$ by $g_{0}$ to redefine $B$ as $B_{0}=\max \left\{1 \leq i \leq g: \tilde{P}_{(i)} \leq i \alpha / g_{0}\right\}$, reject $H_{i j}$ for all $(i, j)$ such that $\tilde{P}_{i} \leq \tilde{P}_{\left(B_{0}\right)}$ and $s P_{i j} \leq B_{0} \alpha / g_{0}$, provided this maximum exists, otherwise, accept all the null hypotheses. This is the oracle form of the adjusted BH method, which motivates us to present our proposed adaptive BH method in the following:

Definition 2. (Adaptive BH under block dependence)

1. Consider an estimate $\widehat{n}_{0}(\mathbf{P})$ satisfying Property 1, and define $\widehat{g}_{0}=\widehat{n}_{0} / s$.
2. Find $B^{*}=\max \left\{1 \leq i \leq g: \tilde{P}_{(i)} \leq i \alpha / \widehat{g}_{0}\right\}$.
3. Reject $H_{i j}$ for all $(i, j)$ such that $\tilde{P}_{i} \leq \tilde{P}_{\left(B^{*}\right)}$ and $s P_{i j} \leq B^{*} \alpha / \widehat{g}_{0}$, provided the maximum in Step 2 exists, otherwise, accept all the null hypotheses.

Theorem 2. Consider the block dependence structure in which the $p$-values are positively dependent as in (1) within each block. The FDR of the above adaptive BH is strongly controlled at $\alpha$ under such block dependence.

Proof. Proceeding as in finding the FDR of the adjusted BH method, we first have

$$
\begin{equation*}
\mathrm{FDR} \leq \sum_{i=1}^{g} \sum_{j=1}^{s} I\left(H_{i j}=0\right) \sum_{b=1}^{g} \frac{\operatorname{pr}\left(P_{i j} \leq b \alpha / \widehat{n}_{0}(\mathbf{P}), B^{*(-i)}=b-1\right)}{b}, \tag{13}
\end{equation*}
$$

where $B^{*(-i)}$ is the number of significant blocks detected by the BH method based on the $g-1$ block $p$-values $\left\{\tilde{P}_{1}, \ldots, \tilde{P}_{g}\right\} \backslash\left\{\tilde{P}_{i}\right\}$ and the critical values $i \alpha / \widehat{g}_{0}, i=2, \ldots, g$. For each $(i, j)$,

$$
\begin{align*}
& \frac{1}{b} I\left(H_{i j}=0\right) \sum_{b=1}^{g} \operatorname{pr}\left(P_{i j} \leq b \alpha / \widehat{n}_{0}(\mathbf{P}), B^{*(-i)}=b-1\right) \\
\leq & \frac{1}{b} I\left(H_{i j}=0\right) \sum_{b=1}^{g} \operatorname{pr}\left(P_{i j} \leq b \alpha / \widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right), B^{*(-i)}=b-1\right) \\
\leq & \alpha E\left\{\frac{I\left(H_{i j}=0\right)}{\widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right)} \sum_{b=1}^{g} \operatorname{pr}\left(B^{*(-i)}=b-1 \mid P_{i j} \leq b \alpha / \widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right), \mathbf{P}^{(-i)}\right)\right\} . \tag{14}
\end{align*}
$$

Now,

$$
\begin{aligned}
& \sum_{b=1}^{g} \operatorname{pr}\left(B^{*(-i)}=b-1 \mid P_{i j} \leq b \alpha / \widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right), \mathbf{P}^{(-i)}\right) \\
= & \sum_{b=1}^{g} \operatorname{pr}\left(B^{*(-i)} \geq b-1 \mid P_{i j} \leq b \alpha / \widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right), \mathbf{P}^{(-i)}\right)- \\
& \sum_{b=1}^{g-1} \operatorname{pr}\left(B^{*(-i)} \geq b \mid P_{i j} \leq b \alpha / \widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right), \mathbf{P}^{(-i)}\right)
\end{aligned}
$$

$$
\begin{align*}
& \leq \sum_{b=1}^{g} \operatorname{pr}\left(B^{*(-i)} \geq b-1 \mid P_{i j} \leq b \alpha / \widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right), \mathbf{P}^{(-i)}\right)- \\
& \quad \sum_{b=1}^{g-1} \operatorname{pr}\left(B^{*(-i)} \geq b \mid P_{i j} \leq(b+1) \alpha / \widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right), \mathbf{P}^{(-i)}\right) \\
& =\operatorname{pr}\left(B^{*(-i)} \geq 0 \mid P_{i j} \leq \alpha / \widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right), \mathbf{P}^{(-i)}\right)=1 \tag{15}
\end{align*}
$$

The inequality in (15) holds since $\left(P_{i 1}, \ldots, P_{i s}\right)$ is independent of $\mathbf{P}^{(-i)}$ and $I\left(B^{*(-i)} \geq b\right)$ is decreasing in $P_{i j}$ 's, the conditional probability

$$
\operatorname{pr}\left(B^{*(-i)} \geq b \mid P_{i j} \leq b \alpha / \widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right), \mathbf{P}^{(-i)}\right)
$$

is of the form $E\left\{\phi\left(P_{i 1}, \ldots, P_{i s}\right) \mid P_{i j} \leq b u\right\}$, for some decreasing function $\phi$ and constant $u>0$, and hence is decreasing in $b$ due to the positive dependence condition assumed in the theorem.

From (13)-(15), we finally get

$$
\begin{equation*}
\mathrm{FDR} \leq \alpha \sum_{i=1}^{g} \sum_{j=1}^{s} E\left\{\frac{I\left(H_{i j}=0\right)}{\widehat{n}_{0}\left(\mathbf{P}^{(-i)}, \mathbf{0}\right)}\right\} \leq \alpha \tag{16}
\end{equation*}
$$

which proves the desired result.

COROLLARY 2. The adaptive BH method of the above type based on the estimates $\widehat{n}_{0}^{(1)}$ and $\widehat{n}_{0}^{(2)}$ strongly control the FDR at $\alpha$ under the block dependence considered in Theorem 2.

REMARK 3. Blanchard and Roquain (2009) first presented an adaptive BH method that continues to control the FDR under the same dependence assumption of the $p$-values as made for the original BH method. Their idea is to estimate $n_{0}$ independently through an FWER controlling method before incorporating that into the original BH method. While this adaptive BH method would be applicable to our present context, it does not capture the group structure of the data. Moreover, their simulation studies only show an improvement of their adaptive BH method over the original BH method in very limited situations. Hu, Zhao and Zhou (2010) considered adjusting the BH method in presence of group structure by weighting the $p$-values according to
the relative importance of each group before proposing its adaptive version by estimating these weights. We should emphasize that this version of the adaptive BH method is known to control the FDR only in an asymptotic setting and under weak dependence.

## 6. Simulation studies

We performed simulation studies to investigate the following questions:

Q1. How does the newly suggested adaptive Bonferroni method under block dependence based on the estimate $\widehat{n}_{0}^{(1)}$ perform in terms of the FWER control and power with respect to the block size $s$, the parameter $\lambda$, and the strength of dependence among the $p$-values compared to the original Bonferroni method and the existing adaptive Bonferroni method based on the estimate $\widehat{n}_{0}^{(0)}$ ?

Q2. How does the newly suggested adaptive BH method under block dependence based on the estimate $\widehat{n}_{0}^{(1)}$ perform in terms of the FDR control and power with respect to the block size $s$, the parameter $\lambda$, and the strength of dependence among the $p$-values compared to the original BH method and the existing adaptive BH method based on the estimate $\widehat{n}_{0}^{(0)}$ ?

Two types of dependence, block and total, were considered for the $p$-values and simulated using multivariate normal test statistics. For block dependence, a covariance matrix providing $g$ independent groups of size $s$ each and having a common non-negative correlation $\rho$ within each group was used; whereas, for total dependence, a compound symmetric covariance matrix with a common non-negative correlation $\rho$ was used.

Figures 1 and 2 answer Q1, while Figures 3 and 4 answer Q2. More specifically, Figures 1 and 3 answer Q1 and Q2, respectively, in terms of the group size and the strength of block or total dependence among the $p$-values, and Figure 2 (or 4 ) presents the performance of the new adaptive Bonferroni (or BH ) method in terms of the FWER (or FDR) control under block (or total) dependence relative to the existing adaptive Bonferroni (or BH ) method for different values of $\lambda$ and strengths of block (or total) dependence among the $p$-values. The reason we


Fig. 1. Simulated FWER and (average) power, the expected proportion of false nulls that are rejected, for each of the three multiple testing methods - the original Bonferroni method (Bonf.) and the two adaptive Bonferroni methods (adBon1, based on $\widehat{n}_{0}^{(0)}$; adBon2, based on $\widehat{n}_{0}^{(1)}$ ) with $\lambda=0.5$. These were obtained by (i) generating $n=100$ dependent normal random variables $N\left(\mu_{i}, 1\right), i=1, \ldots, n$, grouped into $g=50,25$ or 10 groups each of size $s=2,4$ or 10 , respectively, with a block dependent (the first two rows) or compound symmetric (the last two rows) covariance matrix, and half of the $\mu_{i}$ 's in each group being equal to 0 while the rest being equal to $d=\sqrt{10}$; (ii) applying each method to the generated data to test $H_{i}: \mu_{i}=0$ against $K_{i}: \mu_{i} \neq 0$ simultaneously for $i=1, \ldots, 100$, at level $\alpha=0.05$., and (iii) repeating steps (i) and (ii) 2, 000 times. [Bonf - solid; adBon1 - dot-dashes; adBon2 - long dashes]
do not present in Figure 2 (or 4) the results for the total (or block) dependence case is that our simulation studies did not show remarkable difference of the FWERs (or FDRs) between the two adaptive Bonferroni (or BH) methods in that case.

The following are the observations:
From Figure 1: Under block dependence, when the group size is small, both adaptive Bonferroni methods control the FWER when the $\rho$ within each block is close to either zero or one but become liberal, although slightly, when this $\rho$ gets away from zero and one. However, when the group size is relatively large, the new adaptive Bonferroni method maintains a control over the FWER whatever be the $\rho$, whereas the existing adaptive Bonferroni method can still lose control over the FWER for some values of $\rho$. When the $p$-values are totally dependent with a unform pairwise dependence, the FWER of the new adaptive method becomes smaller with increasing


Fig. 2. Simulated FWERs of the original Bonferroni (Bonf) and the two adaptive Bonferroni methods (adBon1, based on $\widehat{n}_{0}^{(0)}$; adBon2, based on $\widehat{n}_{0}^{(1)}$ ) with $\lambda=0.3,0.5,0.7$ or 0.9 . These were obtained by (i) generating $n=120$ block dependent $p$-values from standardized normal test statistics with $n_{0}=40$ (row 1) or 80 (row 2), grouped into $g=40$ groups, with one $\mu_{i}$ in each group being equal to 0 while the rest being equal to $d=\sqrt{10}$ when $n_{0}=40$, and two $\mu_{i}$ 's in each group being equal to 0 while the rest being equal to $d=\sqrt{10}$ when $n_{0}=$ 80; (ii) applying each method to the generated data to test $H_{i}: \mu_{i}=0$ against $K_{i}: \mu_{i} \neq 0$ simultaneously for $i=1, \ldots, 120$, at level $\alpha=0.05$, and (iii) repeating steps (i) and (ii) 2, 000 times.. [Bonf - solid; adBon1 -dot-dashes; adBon2 - long dashes.]
group size, whereas the FWER of the existing adaptive Bonferroni method almost remains unchanged. When the group size is small or moderate, these two adaptive methods both lose control of the FWER except when $\rho$ is close to zero or one; however, when $s$ is large, the new adaptive method regains control of the FWER, but the existing method still loses control of the FWER. So, considering the power performances of the two adaptive Bonferroni methods along with their FWER control, it is clear that the new method is a better choice as an adaptive version of the Bonferroni method under block dependence than the existing one, particularly when the group size is relatively large. It controls the FWER with reasonable power irrespective of the strength of dependence, not only under block dependence but also when there is a total dependence among all the $p$-values.

From Figure 2: When the value of $\lambda$ is small, the FWERs of both adaptive Bonferroni methods slightly exceed $\alpha$ except when $\rho$ is close to one. However, when $\lambda$ is chosen to be relatively large, the FWER of the new adaptive method is controlled at $\alpha$ with increasing $\rho$, whereas the existing adaptive Bonferroni method still loses control of the FWER. With larger proportion of true nulls,


Fig. 3. Simulated FDR and (average) power, the expected proportion of false nulls that are rejected, for each of the three multiple testing methods - the original Benjamini-Hochberg $(\mathrm{BH})$ and the two adaptive BH methods (adBH1, based on $\widehat{n}_{0}^{(0)}$; adBH2, based on $\widehat{n}_{0}^{(1)}$ ) with $\lambda=0.5$. These were obtained by (i) generating $n=240$ dependent normal random variables $N\left(\mu_{i}, 1\right), i=1, \ldots, n$, grouped into $g=120,80$ or 60 groups with a block dependent (the first two rows) or compound symmetric (the last two rows) covariance matrix, half of the $\mu_{i}$ 's in each group being equal to 0 while the rest being equal to $d=\sqrt{10}$ when $s=2$ or 4 , one $\mu_{i}$ being equal to 0 while the rest being equal to $d=\sqrt{10}$ in each of the first 40 groups, and two $\mu_{i}$ 's being equal to 0 while the rest being equal to $d=\sqrt{10}$ in each of the remaining 40 groups, when $s=3$; (ii) applying each method to the generated data to test $H_{i}: \mu_{i}=0$ against $K_{i}: \mu_{i} \neq 0$ simultaneously for $i=1, \ldots, 100$, at level $\alpha=0.05$, and (iii) repeating steps (i) and (ii) 2,000 times. [BH - solid; adBH1 - dot-dashes; adBH2 long dashes.]
the FWERs of both adaptive methods become larger, but the new adaptive method keeps its
FWER controlled at $\alpha$ with a large $\lambda$.
From Figure 3: Under block dependence, the simulated FDRs and average powers for the two adaptive BH methods remain unchanged with increasing $\rho$. For different $s$ and $\rho$, both these adaptive BH methods seem to be more powerful FDR controlling methods than the conventional BH method. However, while the two adaptive BH methods are equally powerful when $s$ is small, the new adaptive method seems to become less powerful with increasing $s$. When the $p$-values are totally dependent with a uniform pairwise dependence, the FDR of the new adaptive BH method becomes smaller and gets controlled over a wider range of values of $\rho$ with increasing $s$; however, the existing adaptive BH method seems to always lose control over the FDR for


Fig. 4. Simulated FDRs of the original Benjamini-Hochberg (BH) and the two adaptive BH methods (adBH1, based on $\widehat{n}_{0}^{(0)}$; adBH2, based on $\widehat{n}_{0}^{(1)}$ ) with $\lambda=0.3,0.5,0.7$ or 0.9 . These were obtained by (i) generating $n=250$ dependent $p$-values from dependent standardized normal test statistics with a compound symmetric covariance matrix grouped into $g=50$ groups, with $n_{0}=125$ (row 1 ) or 200 (row 2 ), two $\mu_{i}$ 's being equal to 0 and the rest being equal to $d=\sqrt{10}$ in each of the first 25 groups, while three $\mu_{i}$ 's being equal to 0 and the rest being equal to $d=\sqrt{10}$ in each of the remaining 25 groups, when $n_{0}=125$; four $\mu_{i}$ 's in each group being equal to 0 and the rest being equal to $d=\sqrt{10}$ when $n_{0}=200$; (ii) applying each method to the generated data to test $H_{i}: \mu_{i}=0$ against $K_{i}: \mu_{i} \neq 0$ simultaneously for $i=1, \ldots, 250$, at level $\alpha=0.05$, and (iii) repeating steps (i) and (ii) 2,000 times.[Bonf - solid; adBon1 - dot-dashes; adBon2 - long dashes.]
any positive values of $\rho$. Considering the power performances of the two adaptive BH methods along with their FDR control, it seems that the new method is a better choice as an adaptive version of the BH method under block dependence than the existing one. It controls the FDR with reasonable power irrespective of the strength of dependence under block dependence and for a wider range of values for the pairwise dependence under total dependence.

From Figure 4: With small proportions of true nulls, the FDR of the new adaptive BH method seems to be controlled at $\alpha$ either when the value of $\lambda$ is very large or when $\rho$ is small or moderate and $\lambda$ is not so large. However, the existing adaptive BH method always loses control of the FDR for any chosen value of $\lambda$ when $\rho$ is not close to zero. With larger proportion of true nulls, although the FDRs of these two adaptive methods become larger, when $\lambda$ is chosen to be very large, the FDR of the new adaptive BH method is controlled at $\alpha$ for small and moderate $\rho$ and is only slightly larger than $\alpha$ for large $\rho$, and when $\lambda$ is chosen to be not so large, the FDR of the
new adaptive method is pretty close to that of the existing adaptive BH method and is less than $\alpha$ only for small $\rho$.

## 7. CONCLUDING REMARKS

Construction of adaptive multiple testing methods with proven control of the ultimate FWER or FDR under dependence in a non-asymptotic setting is an open problem. In this paper, we offered a solution to this open problem under a commonly encountered type of dependence, the block dependence. We have developed new adaptive Bonferroni and BH methods with proven FWER and FDR control, respectively, under the assumption of block dependence, with numerical evidence that they often provide real improvements over the conventional Bonferroni and BH methods.

There is, however, a scope of doing further investigations. For instance, in our simulation studies, we evaluated the performances of our suggested adaptive Bonferroni and BH methods under certain types of positive dependence. It might be interesting to provide some insight into the performances of these adaptive methods under other dependence settings. Figure 5 provides a few numerical results on the FWER control for the adaptive Bonferroni method in the setting of block dependence where the test statistics in each group is negatively dependent. We see from this figure that in this setting of negative block dependence, the simulated FWER of our newly suggested adaptive Bonferroni method is controlled at $\alpha$ for different values of $\lambda$, such as $0.5,0.7$ and 0.9 . Figure 5 also reveals that there is no need to impose any restriction on the choice of $\lambda$ for the adaptive Bonferroni method under such setting and even the existing adaptive Bonferroni method based on Storey et al's estimate can control the FWER at $\alpha$ under negative block dependence. It would be interesting to see if this numerical finding can be justified theoretically. In this paper, we assume that all block sizes are the same in the setting of block dependence. However, in some real applications, the block sizes might be different. To exploit the general block dependence capturing such dependency, it seems that our proposed estimate $\widehat{n}_{0}^{(1)}$ of $n_{0}$ needs to be generalized. Based on such generalized estimate $\widehat{n}_{0}^{(1)}$, we should be able to develop


Fig. 5. Simulated FWERs of original Bonferroni and two adaptive Bonferroni methods based on the estimates $\widehat{n}_{0}^{(0)}$ and $\widehat{n}_{0}^{(1)}$ (Bonf - solid; adBon1 - dot-dashes; adBon2 - long dashes) applied to block dependent p-values generated from 100 standardized normal test statistics with negative equicorrelation $\rho$ within each block and $n_{0}=50, d=\sqrt{10}, s=2, \lambda=0.5,0.7$, or 0.9.
adaptive Bonferroni and BH methods for controlling the FWER and FDR, respectively, in the setting of general block dependence. It would be interesting to see if the techniques and approaches we have developed in this paper can be applied to this general setting.

It is interesting to note that if each block of null hypotheses is interpreted as a family of null hypotheses, then the problem of multiple testing under block dependence is equivalent to a problem of simultaneously testing multiple families of hypotheses for controlling the FWER or FDR over all hypotheses together. Such problem is often seen in large scale data analysis in modern scientific investigations, such as DNA microarray and fMRI studies. When the paper is close to be finished, Professor Benjamini brought to our attention that their recent Arxiv preprint, Benjamini and Bogomolov (2011), just discussed the problem of testing multiple families of hypotheses. However, their objective is to control an average error rate over the selected families including average FWER and FDR rather than the overall FWER and FDR. It would be interesting to investigate the connection between our theory and methods and those in the above paper. We are going to do that in a different communication.

## Acknowledgements

The research of the first author is supported by the NSF Grant DMS-1006021 and the research of the second author is supported by the NSF Grant DMS-1006344.

## APPENDIX

Proof of Lemma 1. Let $s=\left(s_{1}, \ldots, s_{q}\right)$ be the column sum vector of A, that is, $s_{j}=$ $\sum_{i=1}^{p} a_{i j}, j=1, \ldots, q$, and $\sum_{j=1}^{q} s_{j}=m$. Without any loss of generality, we can assume that $s_{1} \geq \ldots \geq s_{q}$. Consider a given column sum vector $s^{*}=\left(s_{1}^{*}, \ldots, s_{q}^{*}\right)$ satisfying $s_{1}^{*} \geq \ldots \geq s_{q}^{*}$, where $s_{j}^{*}=\left\lfloor\frac{m}{q}\right\rfloor$ or $\left\lfloor\frac{m}{q}\right\rfloor+1$ for $j=1, \ldots, q$, and $\sum_{j=1}^{q} s_{j}^{*}=m$.

We prove that $s^{*}$ is majorized by $s$; that is, for each $k=1, \ldots, q$,

$$
\begin{equation*}
\sum_{j=k}^{q} s_{j}^{*} \geq \sum_{j=k}^{q} s_{j} . \tag{17}
\end{equation*}
$$

Suppose the inequality (17) does not hold for some $k=1, \ldots, q$, and $k_{1}=\max \left\{k: \sum_{j=k}^{q} s_{j}^{*}<\right.$ $\left.\sum_{j=k}^{q} s_{j}\right\}$. Since $s_{k_{1}}>s_{k_{1}}^{*}$, for each $j=1, \ldots, k_{1}-1, s_{j} \geq s_{k_{1}} \geq s_{k_{1}}^{*}+1 \geq\left\lfloor\frac{m}{q}\right\rfloor+1 \geq s_{j}^{*}$, implying that

$$
\sum_{j=1}^{q} s_{j}=\sum_{j=1}^{k_{1}-1} s_{j}+\sum_{j=k_{1}}^{q} s_{j}>\sum_{j=1}^{k_{1}-1} s_{j}^{*}+\sum_{j=k_{1}}^{q} s_{j}^{*}=m,
$$

which is a contradiction. So, $s^{*}$ is majorized by $s$.
By Theorem 2.1 of Ryser (1957), one can rearrange the 1's in the rows of A to construct a $p \times q$ matrix which has the column sum vector $s^{*}$. Thus, the desired result follows.

Outlines of proofs of Lemmas 2-5. Lemma 2 follows from the well-known inequality between the arithmetic and harmonic means, or using the Jensen inequality. For Lemma 3, one can see Liu and Sarkar (2010). Lemma 5 can be proved by successively using the fact $a_{r: m} \geq a_{r-1: m-1}$ for all $r \leq m$.

To prove Lemma 4, let $g(x)=\ln f(x)=-\frac{2}{x+2} \ln (2 x+3)$. Note that $g^{\prime}(x)=2 \ln (2 x+$ $3)-\frac{4 x+8}{2 x+3}$ and $g^{\prime \prime}(x)=\frac{4}{2 x+3}+\frac{4}{(2 x+3)^{2}}>0$ for $x \geq 0$. Thus, $g(x)$ is a convex function for $x \geq 0$. Observe that $g^{\prime}(0)<0, g^{\prime}(1)>0$, and $g(0)<g(1)$. So, $g^{\prime}(x)>0$ for $x \geq 1$ and $g(x) \leq$ $\max \{g(0), g(1)\}=g(1)$ for $0 \leq x \leq 1$. Thus, the desired result follows.

## REFERENCES

[1] Benjamini, Y. \& Bogomolov, M. (2011). Adjusting for selection bias in testing multiple families of hypotheses. Arxiv preprint, math.ST/1106.3670v1.
[2] Benjamini, Y. \& 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.
[3] Benjamini, Y, Krieger, K. \& Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. Biometrika 93, 491-507.
[4] Benjamini, Y. \& Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. Ann. Statist. 29, 1165-1188.
[5] Blanchard, G. \& Roquain, E. (2009). Adaptive FDR control under independence and dependence. $J$. Mach. Learn. 10, 2837-2871.
[6] Clements, N., Sarkar, S. \& Guo, W. (2011). Astronomical transient detection using grouped p-values and controlling the false discovery rate. Unpublished manuscript.
[7] Finner, H., Dickhaus, T. \& Roters, M. (2007). Dependency and false discovery rate: Asymptotics. Ann. Statist. 35, 1432-1455.
[8] Finner, H. \& Gontscharuk, V. (2009). Controlling the familywise error rate with plug-in estimator for the proportion of true null hypotheses. J. Roy. Statist. Soc., Ser. B 71, 1031-1048.
[9] Gavrilov, Y., Benjamini, Y. \& Sarkar, S. K. (2009). An adaptive step-down procedure with proven FDR control. Ann. Statist. 37, 619-629.
[10] Guo, W. (2009). A note on adaptive Bonferroni and Holm procedures under dependence. Biometrika, 96, 1012-1018.
[11] Guo, W., Sarkar, S. \& Peddada, S. (2010). Controlling false discoveries in multidimensional directional decisions, with applications to gene expression data on ordered categories. Biometrics 66, 485-492.
[12] Heller R., Golland Y., Malach R. \& Benjamini Y. (2007). Conjunction Group Analysis: An alternative to mixed/random effect analysis. NeuroImage 37, 1178-1185.
[13] Hochberg, Y. \& Benjamini, Y. (1990). More powerful procedures for multiple significance testing. Statist. Med. 9, 811-818.
[14] Hochberg, Y., \& Tamhane, A. C. (1987). Multiple Comparison Procedures. Wiley: New York.
[15] Hu, J., Zhao, H. \& Zhou, H. (2010). False Discovery Rate Control with Groups. J. American Statistical Association 105, 1215-1227.
[16] Lei S., Chen Y., Xiong D., Li L. \& Deng H. (2006) Ethnic difference in osteoporosis-related phenotypes
and its potential underlying genetic determination. Journal of Musculoskelet Neuronal Interact. 6, 36-46.
[17] Liu, F. \& Sarkar, S. K. (2010). A note on estimating the false discovery rate under mixture model. Journal of Statistical Planning and Inference 140, 1601-1609.
[18] Liu, F. \& Sarkar, S. K. (2011). A new adaptive method to control the false discovery rate. Series in Biostatistics 4, World Scientific, 3-26.
[19] Romano, J. P., Shaikh, A. M. \& Wolf, M. (2008). Control of the false discovery rate under dependence using the bootstrap and subsampling. TEST 17, 417-442.
[20] RySER, H. J. (1957). Combinatorial properties of matrices of zeros and ones. Canad. J. Math. 9, 371-377.
[21] SARKAR, S. K. (1998). Some probability inequalities for ordered MTP ${ }_{2}$ random variables: a proof of the Simes conjecture. Ann. Statist. 26, 494-504.
[22] Sarkar, S. K. (2002). Some results on false discovery rate in stepwise multiple testing procedures. Ann. Statist. 30, 239-257.
[23] SARKAR, S. K. (2006). False discovery and false nondiscovery rates in single-step multiple testing procedures. Ann. Statist. 34, $394-415$.
[24] Sarkar, S. K. (2008). On methods controlling the false discovery rate (with discussions). Sankhya 70, 135168.
[25] Sarkar, S. K. \& Chang, C-K. (1997). The Simes method for multiple hypothesis testing with positively dependent test statistics. J. Amer. Statist. Assoc. 92, 1601-1608.
[26] Sarkar, S., Guo, W. \& Finner, H. (2012). On adaptive procedures controlling the familywise error rate. Journal of Statistical Planning and Inference, 142, 65-78.
[27] Schweder, T. \& SpJ $\phi$ Tvoll, E. (1982). Plots of p-values to evaluate many tests simulataneously. Biometrika 69, 493-502.
[28] Storey J. (2003). Comment on 'Resampling-based multiple testing for DNA microarray data analysis' by Ge, Dudoit, and Speed. Test 12, 52-60.
[29] Storey, J. D., Taylor, J. E. \& 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.
[30] Sun W \& WEI Z. (2011). Multiple testing for pattern identification, with applications to microarray time course experiments. Journal of the American Statistical Association 106, 73-88.

