Electronic Journal of Statistics
Vol. 11 (2017) 1827–1857
ISSN: 1935-7524
DOI: 10.1214/17-EJS1268

Finite sample bounds for expected number of false rejections under martingale dependence with applications to FDR*

Julia Benditkis and Arnold Janssen

Heinrich-Heine University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany
e-mail: Benditkis@math.uni-duesseldorf; Janssena@uni-duesseldorf.de

Abstract: Much effort has been made to improve the famous step up procedure of Benjamini and Hochberg given by linear critical values $\frac{i\alpha}{n}$. It is pointed out by Gavrilov, Benjamini and Sarkar that step down multiple testing procedures based on the critical values $\beta_i = \frac{i\alpha}{n+1-i(1-\alpha)}$ still control the false discovery rate (FDR) at the upper bound α under basic independence assumptions. Since that result is no longer true for step up procedures and for step down procedures, if the p-values are dependent, a big discussion about the corresponding FDR starts in the literature. The present paper establishes finite sample formulas and bounds for the FDR and the expected number of false rejections for multiple testing procedures using critical values β_i under martingale and reverse martingale dependence models. It is pointed out that martingale methods are natural tools for the treatment of local FDR estimators which are closely connected to the present coefficients β_i. The martingale approach also yields new results and further inside for the special basic independence model.

MSC 2010 subject classifications: Primary 62G10; secondary 62G20.
Keywords and phrases: False Discovery Rate (FDR), Expected Number of False Rejections (ENFR), multiple testing, step up procedure, step down procedure.

Received June 2016.

Contents

1 Introduction 1828
2 Results for step down procedures 1832
  2.1 Control of the expected number of false rejections $\mathbb{E}[V]$ 1832
  2.2 Consequences under Dirac-Martingale configurations 1835
  2.3 Control of the FDR 1835
  2.4 Improvement of the power 1837
3 Simulation example 1838
4 Results for SU Procedures 1839

*Partially suppoted by SAW-Project “Multiplizität, Modellvalidierung und Reproduzierbarkeit in hochdimensionalen Microarray-Daten”.

5 Finite sample results for SU procedures using critical values of Gavrilov et al. 1842
6 The proofs and technical results 1843
Appendix. Examples of martingale models 1854
Acknowledgements 1855
References 1855

1. Introduction

Multiple tests are nowadays well established procedures for judging high dimensional data. The famous Benjamini and Hochberg [2] step up multiple testing procedure given by linear critical values controls the false discovery rate FDR for various dependence models. The FDR is the expectation of the ratio of the number of false rejections and the total number of rejected hypotheses. The linear step up procedure is frequently applied in practice as an FDR-controlling procedure. Gavrilov et al. [11] pointed out that linear critical values can be substituted by

$$\beta_i = \frac{i\alpha}{n+1-i(1-\alpha)},\ i \leqslant n, \tag{1.1}$$

for step down (SD) procedures and the FDR control (i.e. FDR$\leqslant \alpha$) remains true for the basic independence model of the underlying p-values. Note that the present critical values β_i are closely related to critical values given by the asymptotic optimal rejection curve (i.e., the critical values (1.1) differ from the ones that are based on the asymptotic optimal rejection curve by additional summand 1 in denominator) which is obtained by Finner et al. [8]. In the asymptotic set up they derived step up procedures with asymptotic FDR control under various conditions. However, step up multiple testing procedures given by the β_i's do not control the FDR by the desired level α at finite sample size, see for instance Dickhaus [7], Gontscharuk [12].

The intension of the present paper is twofold.

- We like to calculate the FDR of step down and step up procedures more precisely using martingale and reverse martingale arguments. Here we get also new results under the basic independence model.
- On the other hand we can extend the results for dependent p-values which are martingale or reverse martingale dependent. As application finite sample FDR formulas for step down and step up procedures based on (1.1) are derived. We refer to the Appendix for a collection of examples of martingale models.

Sarkar [19] proposed exact formulas for the FDR of step up multiple testing procedures and for some upper bound of the FDR of step down multiple testing procedures.

Martingale arguments were earlier used in Storey et al. [23], Pena et al. [17], Heesen and Janssen [14] for step up and in Benditkis [1] for step down multiple testing procedures.

This paper is organized as follows. Below the basic notations are introduced. Section 2 presents our results for step down procedures. A counterexample, Example 4, motivates to study specific dependence concepts which allow FDR control, namely our martingale dependence model. The FDR formula, see (1.7) below, consists of two terms. In particular, it relies on the expected number of false rejections which is studied in Sections 2.1 and 2.2. Note that the results of Lemma 6 motivate naturally the consideration of martingale methods. Section 2.3 is devoted to the FDR control under dependence which extends the results of Gavrilov et al. [11]. Within the class of step down procedures the first coefficient β_1 is often responsible for the quality of the multiple testing procedure. In Section 2.4 we propose an improvement of the power of SD procedures due to an increase of first critical values without a loss of FDR control.

Step up procedures corresponding to the β's from (1.1) are studied in Sections 4 and 5. We obtain the lower bound for the present FDR which can be greater than α. A couple of examples for martingale models can be found in Appendix. The proofs and additional material are collected in the Section 6.

Basics. Let us consider a multiple testing problem, which consists of n null hypotheses $H_1, ..., H_n$ with associated p-values $p_i,\ i = 1, ..., n$. Assume that all p-values arise from the same experiment given by one data set, where each p_i can be used for testing the traditional null H_i. The p-values vector $p = (p_1, ..., p_n) \in [0, 1]^n$ is a random variable based on an unknown distribution P. Recall that simultaneous inference can be established by so called multiple testing procedure $\phi = \phi(p)$, $\phi = (\phi_1, ..., \phi_n) : [0, 1]^n \to \{0, 1\}^n$, which rejects the null H_i iff, i.e. if and only if, $\phi_i(p) = 1$ holds. The set of hypotheses can be divided in the disjoint union $I_0 \bigcup I_1 = \{1, ..., n\}$ of unknown portions of true null I_0 and false null I_1, respectively. We denote the number of true null by $n_0 = |I_0|$ and the number of false ones by $n_1 = |I_1| = n - n_0$, where $n_0 > 0$ is assumed. Widely used single threshold multiple testing procedures can be represented as

$$\phi_\tau = (\mathbb{I}(p_1 \leqslant \tau), ..., \mathbb{I}(p_n \leqslant \tau))$$

via the indicator function $\mathbb{I}(\cdot)$, where $\tau \in [0, 1]$ is a random critical boundary variable. Thus all null hypotheses with related p-values that are not larger than the threshold τ have to be rejected. Let $p_{1:n} \leqslant p_{2:n} \leqslant \cdots \leqslant p_{n:n}$ denote the ordered values of the p-values p.

Definition 1. *Let $\alpha_1 \leqslant \alpha_2 \leqslant \cdots \leqslant \alpha_n$ be a deterministic sequence of critical values. Set for convenience* $\max\{\emptyset\} = 0$.

(a) The step down (SD) critical boundary variable is given by

$$\tau_{SD} = \max\{\alpha_i : p_{j:n} \leqslant \alpha_j, \quad for \quad all \quad j \leqslant i\}. \tag{1.2}$$

(b) The step up (SU) critical boundary variable is given by

$$\tau_{SU} = \max\{\alpha_i : p_{i:n} \leqslant \alpha_i\}. \tag{1.3}$$

(c) The appertaining multiple testing procedure $\phi_{SD} = \phi_{\tau_{SD}}$ and $\phi_{SU} = \phi_{\tau_{SU}}$ are called step down (SD) procedure, step up (SU) procedure, respectively.

Let $\hat{F}_n$ denote the empirical distribution function of the p-values and let $V = V(\tau) = \sum_{i\in I_0} \mathbb{I}(p_i \leqslant \tau)$, $S = S(\tau) = \sum_{i\in I_1} \mathbb{I}(p_i \leqslant \tau)$ and $R = R(\tau) = \sum_{i=1}^{n} \mathbb{I}(p_i \leqslant \tau) = n\hat{F}_n(\tau)$ be the number V of false rejections w.r.t. τ, the number S of true rejections and the number R of all rejections, respectively. The False Discovery Rate (FDR) of a procedure with critical boundary variable τ is defined as

$$\mathrm{FDR} = \mathbb{E}\left[\frac{V(\tau)}{R(\tau)}\right],$$

with the convention $\frac{0}{0} = 0$. The FDR is often chosen as an error rate control criterion. There is another useful equivalent description of step down procedures.

Remark 2. *Introduce the random variable*

$$\sigma := \min\{\alpha_i : p_{i:n} > \alpha_i\} \wedge \alpha_n, \quad \text{with the convention} \quad \min\{\emptyset\} = 1, \tag{1.4}$$

where $a \wedge b = \min(a, b)$ denotes the minimum of two real numbers a and b.

Then we have $\tau_{SD} \leqslant \sigma$ but the step down procedures $\phi_{SD} = \phi_\sigma$ coincide and $FDR = \mathbb{E}\left[\frac{V(\tau_{SD})}{R(\tau_{SD})}\right] = \mathbb{E}\left[\frac{V(\sigma)}{R(\sigma)}\right]$ holds. The reason for this is that no p-value falls in the interval $(\tau_{SD}, \sigma]$ and $R(\tau_{SD}) = R(\sigma)$ is valid.

There is much interest in multiple testing procedures such that the FDR is controlled by a prespecified acceptable level $\alpha \in (0, 1)$, i.e. to bound the expectation of the portion of false rejections. The well-known so called Benjamini and Hochberg procedures with linear critical values $\alpha_i = \alpha\frac{i}{n}$ lead to the FDR bound

$$\mathrm{FDR} \leqslant \alpha\frac{n_0}{n}$$

for SD and SU procedures under positive dependence, more precisely under positive regression dependence on a subset (PRDS). There are several proposals to exhaust the FDR more accurate by α by an enlarged choice of critical values. A proper choice for SD procedures are α_i

$$\begin{aligned} &0 < \alpha_i \leqslant \beta_i = \frac{i\alpha}{n + 1 - i(1-\alpha)},\ 1 \leqslant i \leqslant n, \\ &\alpha_0 = \alpha_1,\ \beta_0 = \beta_1, \end{aligned} \tag{1.5}$$

which allow the control FDR $\leqslant \alpha$ under the basic independence assumption of the p-values, see Gavrilov et al. [11]. Note that the first critical value of the Benjamini and Hochberg procedure $\frac{\alpha}{n}$ is larger than β_1 for each $\alpha > 0$ and we have $\beta_i > \frac{i\alpha}{n}$ for $\alpha < 0.5$ and $i \geqslant 2$. For $i = 1, ..., n$, $\beta_i = g_\alpha^{-1}\left(\frac{i}{n}\right)$ are inverse values of

$$g_\alpha(t) = \frac{n+1}{n} f_\alpha(t) = \frac{n+1}{n}\frac{t}{t(1-\alpha)+\alpha}, \tag{1.6}$$

where g_α is related to the asymptotic optimal rejection curve f_α, see Finner et al. [8]. It is known that SU procedures given by β_i do not control the FDR for the independence model in general, see Gontscharuk [12], Heesen and Janssen [14]. If the p-values are dependent then the FDR control of the SD procedures based on $\beta_i,\ i \leqslant n$, cannot be guaranteed in general (see Example 4 of Section 2).

Gavrilov et al. [11], Theorem 1A, propose to reduce the critical values β_i in order to get FDR control of SD procedures under positive regression dependence on a subset. Unfortunately, the procedure based on these new reduced critical values may be too conservative. Below we keep the critical values $\alpha_i,\ i \leqslant n$, of (1.5) and introduce dependence assumptions for the p-values which insure the FDR-control for the underlying SD procedures.

The main idea of this paper can be outlined as follows. The FDR of SD and SU procedures based on the critical values β_i equals

$$\mathrm{FDR} = \frac{\alpha}{n+1}\mathbb{E}\left[\frac{V}{\beta_R}\right] + \frac{1-\alpha}{n+1}\mathbb{E}\left[V\right], \tag{1.7}$$

for more details we refer to Benditkis [1]. A monotonicity argument implies the next Lemma.

Lemma 3. *Consider an SD or SU procedure with critical values $(\alpha_i)_i$ given by (1.5). Then*

(a) *FDR* $\leqslant \frac{\alpha}{n+1}\mathbb{E}\left[\frac{V}{\beta_R}\right] + \frac{1-\alpha}{n+1}\mathbb{E}\left[V\right]$.

(b) *The conditions*

$$\mathbb{E}\left[\frac{V}{\beta_R}\right] \leqslant n_0 \quad \textit{and} \tag{1.8}$$

$$\mathbb{E}\left[V\right] \leqslant \frac{\alpha}{1-\alpha}(n_1+1) \tag{1.9}$$

ensure the FDR control, i.e. FDR $\leqslant \alpha$.

Whereas the FDR is hard to bound under dependence, the inequality (1.8) is known under PRDS and equality holds under reverse martingale structure (including the basic independence model), see Heesen and Janssen [14] for SU procedure. Then it remains to bound the expected number of false rejections $\mathbb{E}\left[V\right]$, which is at least possible for SD procedures under certain martingale dependence assumptions. In the following we always use a general assumption, that the p-values for the true null $(H_i)_{i\in I_0}$ fulfil

$$\mathbb{E}\left[\sum_{i\in I_0}\mathbb{I}(p_i \in [0,t])\right] \leqslant n_0 t \quad \text{for all} \quad t \in [0,1), \tag{1.10}$$

which can be interpreted as "stochastically larger" condition compared with the uniform distribution in the mean for I_0. Thereby, $\mathbb{I}(\cdot)$ is the indicator function.

Now, we define the basic independence assumptions (BIA) that are often used in the FDR-control-framework.

(BIA) We say that p-values fulfil the basic independence model if the vectors of p-values $(p_i)_{i\in I_0}$ and $(p_i)_{i\in I_1}$ are independent, and each dependence structure is allowed for the "false" p-values, i.e. for $(p_i)_{i\in I_1}$. Under true null hypotheses the p-values $(p_i)_{i\in I_0}$ are independent and stochastically larger (or equal) compared to the uniform distribution on $[0,1]$, i.e., $P(p_i \leqslant x) \leqslant x$ for all $x \in [0,1]$ and $i \in I_0$.
If in addition all p-values are i.i.d. uniformly distributed on $[0,1]$ for $i \in I_0$ then we talk about the BIA model with uniform true p-values.

2. Results for step down procedures

In this section we consider a step down procedure with critical values $\beta_i,\ i \leqslant n$, from (1.1). It is well known that this procedure controls the FDR if the p-values fulfil the basic independence assumptions (BIA) (cf. Gavrilov et al. [11]). However, in practice the independence of the single tests corresponding to the present p-values are rare.

For general dependent p-values the FDR of the SD procedure may exceed the level α. The next counter example motivates the consideration of special kinds of dependence in order to establish FDR control.

Example 4. *For $n=3,\ n_0=2,\ n_1=1,\ I_0=\{2,3\}$ and $\alpha=\frac{1}{4}$ consider the SD procedure with critical values $\beta_i=\frac{i\alpha}{n+1-i(1-\alpha)}$. Consider the vector of p-values $(0,U_1,U_2)$ with true p-values defined as follows*

$$
\begin{aligned}
&U_1 \quad \textit{is uniformly distributed on} \quad (0,1).\\
&U_2 = (U_1+\beta_2)\,\mathbb{I}\,(U_1 \leqslant \beta_2) + (U_1-\beta_2)\,\mathbb{I}\,(U_1 \in (\beta_2, 2\beta_2)) + U_1\mathbb{I}\,(U_1 \geqslant 2\beta_2)\,.
\end{aligned}
$$

For such p-values we get

$$
FDR = \frac{2}{3}P(U_1 \leqslant 2\beta_2) = \frac{4\beta_2}{3} = \frac{4}{15} > \frac{1}{4}.
$$

We will start with the expected number of false rejections (ENFR), which was earlier studied by Finner and Roters [10] and Scheer [20].

2.1. *Control of the expected number of false rejections $\mathbb{E}\,[V]$*

The present martingale approach relies on the empirical distribution function $\hat{F}_n$ of the p-values and on the adapted stochastic process

$$
\begin{aligned}
&t \mapsto \hat{\alpha}_n(t) = \frac{t}{1-t}\frac{1-\hat{F}_n(t)}{\hat{F}_n(t)+\frac{1}{n}},\ t \in T, \text{w.r.t. the filtration}\\
&\mathcal{F}_t^T = \sigma\{\mathbb{I}(p_i \leqslant s), s \leqslant t, s,\ i \leqslant n\},\ t \in T, \quad \text{of the p-values.}
\end{aligned}
\tag{2.1}
$$

Thereby, $T \subset [0,1)$ is a parameter space with $0 \in T$. The value $\hat{\alpha}(t)$ is frequently used as a conservative estimator for the FDR on the constant critical boundary value $\tau = t$. Storey et al. [22] use a similar estimator for the FDR(t) of SU procedures if the p-values are independent. A similar estimator is also used by Benjamini, Krieger and Yekutieli [3], Heesen and Janssen [15] and Heesen [13]. It is easy to see that for $\beta_i,\ i \leqslant n$, we get from (1.5)

$$\hat{\alpha}_n(\beta_i) \leqslant \alpha \quad \text{iff} \quad R(\beta_i) \geqslant i-1, \tag{2.2}$$

$$\hat{\alpha}_n(\beta_i) = \alpha \quad \text{iff} \quad R(\beta_i) = i-1, \tag{2.3}$$

since

$$\hat{\alpha}_n(\beta_i) = \alpha \left(\frac{i}{n+1-i}\right)\left(\frac{n-R(\beta_i)}{R(\beta_i)+1}\right).$$

The consequences of these useful relations are summarized.

Lemma 5. *Consider the critical values $(\beta_i)_{i\leqslant n}$ and the critical boundary value σ from (1.4). Then we have*

(a) $\sigma = \min\{\beta_i : \hat{\alpha}_n(\beta_i) \geqslant \alpha, i \leqslant n\} \wedge \beta_n$.
(b) Moreover $\tau_{SD} \leqslant \sigma$ and $\hat{\alpha}_n(\sigma) = \alpha\mathbb{I}(R(\sigma) < n)$ hold.
(c) The random variable σ is a stopping time w.r.t. the filtration $\left(\mathcal{F}_t^T\right)_{t\in T}$ of the p-values with time domain $T = \{0, \beta_1, ..., \beta_n\}$.

It is quite obvious that the maximum coefficients $\beta_i, i \leqslant n$, of the α's in (1.5) and the extreme p-values $p_i = 0,\ i \in I_1$, for all false null are least favourable for bounding $\mathbb{E}\left[V\right]$. First, we focus on the β_i-based SD procedure. An important role plays the process

$$M_t = M_{I_0}(t) = \sum_{i\in I_0} \frac{\mathbb{I}(p_i \leqslant t) - t}{1-t},\ t \in T. \tag{2.4}$$

Lemma 6. *Let $p_i = 0$ for all $i \in I_1$. For the critical values $(\beta_i)_{i\leqslant n}$ from (1.1) we have*

$$\mathbb{E}\left[V(\tau_{SD})\right] \leqslant \frac{\alpha}{1-\alpha}(n_1+1) \quad \textit{iff} \quad \mathbb{E}\left[M_{I_0}(\sigma)\right] \leqslant \alpha(n+1)P(R(\tau_{SD}) = n).$$

The probability $P(R(\tau_{SD}) = n)$ is typically very small. Note that we will show below by martingale arguments that $\mathbb{E}\left[M_{I_0}(\tau_{SD})\right] \leqslant 0$, which implies the crucial condition (1.9).

Next, we introduce a dependence assumption which allows the control of expected number of false rejection of the SD procedure with critical values $\beta_i, i \leqslant n$.

(D1) Let $T \subset [0,1)$ be a set with $0 \in T$. We say that p-values $p_1, ..., p_n$ are $\mathcal{F}_T = (\mathcal{F}_t)_{t\in T}-$ (super-) martingale dependent on a subset J if the stochastic process $M(t) = M_J(t) = \sum_{i\in J}\left(\frac{\mathbb{I}(p_i\leqslant t)-t}{1-t}\right), t \in T$, is a $\mathcal{F}_T-$ (super-)martingale.

Note that the super-martingale model (D1) includes BIA if $J = I_0$. This is well known, see Shorack and Wellner [21] (p. 133), Benditkis [1]. Some examples of martingale dependent random variables can be found in a separate Appendix.

Recall that the general condition (1.10) implies $\mathbb{E}\left[M_{I_0}(0)\right] = 0$, which is always assumed.

Now, we formulate the main result of this subsection under the super-martingale assumption, which will be applied to our equality (1.7).

Theorem 7. *Consider the SD multiple testing procedure with critical values β_i, $i \leqslant n$, given in (1.5). Suppose that the super-martingale assumption (D1) holds with $J = I_0$ and $T = \{0, \beta_1, ..., \beta_n\}$. We get*

$$\mathbb{E}\left[V(\tau_{SD})\right] \leqslant \frac{\alpha}{1-\alpha}(\mathbb{E}\left[S(\tau_{SD})\right] + 1) \leqslant \frac{\alpha}{1-\alpha}(n_1 + 1).$$

The next remark shows that under (D1) we can assume that the p-values which belong to true null are stochastically larger compared with the uniform distribution on $(0,1)$ (cf. Heesen and Janssen [14], Benditkis [1]). Let $U(0,1)$ denote the uniform distribution on the unit interval.

Remark 8. *Let $(p_i)_{i\leqslant n}$ fulfil the martingale assumption (D1) for $J = I_0$ on $T \subset [0,1)$ and let $\pi : I_0 \to I_0$ be some random permutation of the index-set I_0 which is independent of $(p_i)_{i\leqslant n}$.*

(a) If

$$M_{I_0}(t) = \sum_{i\in I_0} \frac{\mathbb{I}(p_i \leqslant t) - t}{1-t} \quad \textit{is a } \mathcal{F}_T - \textit{martingale},$$

then the random variable $Y_i = p_{\pi(i)}, i \in I_0$, is $U(0,1)$-distributed.

(b) If $(p_i)_{i\in I_0}$ fulfil the super-martingale assumption (D1) for $J = I_0$ on $T \subset [0,1]$, then Y_i, $i \in I_0$, are stochastically larger compared with $U(0,1)$.

(c) As long as the boundary critical value τ only depends on the order statistics, the multiple testing procedure ϕ_τ remains unchanged if $(p_i)_{i\leqslant n}$ is substituted by $\left((p_{\pi(i)})_{i\in I_0}, (p_i)_{i\in I_1}\right)$.

(d) It can be shown that under (D1) the (super-)martingale assumption also holds under the filtration given by the exchangeable $\left((Y_i)_{i\in I_0}, (p_i)_{i\in I_1}\right)$. Note that the exchangeability of the $p_{\pi(i)}$, $i \in I_0$, is only needed in the proofs in connection with the PRDS assumption introduced in Section 2.3.

Proof of (a) and (b). Firstly, note that the random variables $Y_i, i \in I_0$ are exchangeable, since σ is an independent uniformly distributed permutation. This implies

$$\mathbb{E}\left[\mathbb{I}(Y_i \leqslant t)\right] = \mathbb{E}\left[\frac{1}{n_0}\sum_{j\in I_0} \mathbb{I}(Y_j \leqslant t)\right] = \mathbb{E}\left[\frac{1}{n_0}\sum_{j\in I_0} \mathbb{I}(p_j \leqslant t)\right]. \tag{2.5}$$

Moreover, we get

$$\mathbb{E}\left[\frac{1}{n_0}\sum_{j\in I_0} \mathbb{I}(Y_j \leqslant t)\right] = \frac{(1-t)}{n_0}\mathbb{E}\left[M_{I_0}(t)\right] + t \leqslant t. \tag{2.6}$$

□

2.2. Consequences under Dirac-Martingale configurations

In this subsection we consider the following assumptions

(i) Martingale dependence assumption (D1) holds with $J = I_0$ and $T = \{0, \beta_1, ..., \beta_n\}$,
(ii) $p_i = 0$ a.s. for all $i \in I_1$.

Structures that fulfil the assumptions (i) and (ii) are called Dirac martingale configurations $\mathrm{DM}(n_1)$. The part (a) of the next lemma proposes exact formulas for ENFR for step down procedures with critical values β_i. Part (b) derives a lower bound for ENFR if the $(p_i)_{i \in I_1}$ are by accident uniformly distributed which is another example for extreme ordering compared with (ii).

It is quite obvious that the Dirac-Martingale configuration is the least favourable for the ENFR of under (D1). Note that in addition the Dirac-uniform configuration is least favourable for the FDR of SU procedure given by β_i (1.5) if n_0 is fixed (for instance Benjamini and Yekutielli [5]).

Lemma 9 (Some exact formulas for the ENFR)**.** *Suppose that the martingale assumption (i) hold. Let τ_{SD} be the critical boundary value, which corresponds to critical values β_i.*
(a) Assume additionally (ii) then

$$E_1 := \mathbb{E}_{DM(n_1)}\left[V(\tau_{SD})\right] = \frac{\alpha}{1-\alpha}(n_1+1) - \frac{\alpha}{1-\alpha}(n+1)P_{DM(n_1)}(V(\tau_{SD}) = n_0). \tag{2.7}$$

(b) Let $(p_1, ..., p_n)$ be exchangeable and martingale dependent on I_1, i.e., $M_{I_1}(t) = \left(\frac{S(t)-n_1 t}{1-t}\right)_{t \in T}$ is an $\mathcal{F}_T-$martingale. Then, each p_i, $i \leqslant n$, is uniformly distributed and

$$E_2 := \mathbb{E}_{U(0,1)}\left[V(\tau_{SD})\right] = \frac{\alpha}{1-\alpha}\frac{n_0}{n} - \frac{\alpha}{1-\alpha}\frac{(n+1)n_0}{n}P_{U(0,1)}(R(\tau_{SD}) = n). \tag{2.8}$$

(c) If $P(p_i \leqslant t) \geqslant t$ for all $i \in I_1$ and all $t \in [0,1]$ then

$$E_2 \leqslant \mathbb{E}\left[V(\tau_{SD})\right] \leqslant E_1.$$

2.3. Control of the FDR

As mentioned in Lemma 3 the control (1.9) of the ENFR is not enough for the FDR control. We have to bound $\mathbb{E}\left[\frac{V(\tau_{SD})}{\tau_{SD}}\right]$ by n_0. To do this we need further assumptions.

(D2) The p-values are said to be positive regression dependent on a subset J (PRDS) if
$$x \mapsto \mathbb{E}\left[f(p_1, ..., p_n) \mid p_i = x\right]$$

is increasing in x for each $i \in J$ and any coordinate-wise increasing, integrable function $f : [0,1]^n \to \mathbb{R}$ (cf. Finner et al. [9], Benjamini and Yekutielli [5].)

Remark 10. *The assumption (D2) implies that*

$$x \mapsto \mathbb{E}\left[g(p_1, ..., p_n) \mid p_i \leqslant x\right]$$

is increasing in x for each $i \in J$ and any coordinate-wise increasing, integrable function $g : [0,1]^n \to \mathbb{R}$ (see Dickhaus [7], Benditkis [1]).

The dependence assumption (D2) is well-known in the FDR-framework. Benjamini and Yekutielli [5] proved that the Benjamini and Hochberg linear step up procedure controls the FDR under such kind of positive dependence. Gavrilov et al. [11] have shown that in this case the FDR of the step down procedure using critical values $\beta_i,\ i \leqslant n$, may exceed the pre-chosen level α. Theorem 11 proves the FDR control of that SD procedure under the additional super-martingale assumption. We collected some examples of dependence structures satisfying the both conditions in Appendix. For an example of dependence structure which is PRDS and not martingale dependent we refer to Gavrilov et al. [11] (Subsection 4.1).

Theorem 11. *Let $(p_i)_{i\in I}$ fulfil the super-martingale assumption (D1) with $T = \{0, \beta_1, ..., \beta_m\}$ and the PRDS assumption (D2) on I_0. Then we have for β_i-based SD procedure*

$$FDR_{\tau_{SD}} = \mathbb{E}\left[\frac{V(\tau_{SD})}{R(\tau_{SD})}\right] \leqslant \alpha.$$

The next lemma is a technical tool for the proof of Theorem 11.

Lemma 12. *Let $(p_i)_{i\leqslant n}$ fulfil (D2) on I_0. For the SD procedure based on the critical values $\beta_i, i \leqslant n$, we have*

$$\mathbb{E}\left[\frac{V(\tau_{SD})}{\tau_{SD}}\right] \leqslant n_0.$$

Remark 13. *(a) Lemma 12 remains true for any random variable $\tau = \tau(p)$, which is a non-increasing function of $p_i,\ i \in I_0$, and has a finite range of values $\{a_1, ..., a_m\},\ 0 < a_1 \leqslant a_2 \leqslant ... \leqslant a_m$ for some $m \in \mathbb{N}$. That means that*

$$\mathbb{E}\left[\frac{V(\tau)}{\tau}\right] \leqslant n_0$$

can be always bounded under PRDS. The exact structure of the random variable τ is not important. The inequality remains true for SD as well for SU procedures.

(b) Theorem 11 remains true for any additional $\mathcal{F}_{\tilde{T}}-$ stopping time $\tilde{\tau}$ with $\tilde{T} = \{0, \tilde{\beta}_1, ..., \tilde{\beta}_m\},\ m \in \mathbb{N},\ 0 \leqslant \tilde{\beta}_1 \leqslant ... \leqslant \tilde{\beta}_m < 1$, which is a non-increasing function of $p_i,\ i \in I_0$ if $\tilde{\tau} \leqslant \sigma$ holds.

The next theorem shows that we can relinquish the PRDS assumption (D2) under some modification of the assumption (D1).

Theorem 14. *Let M_{I_0} be a martingale w.r.t. to the new filtration $\mathcal{F}_T^f = \sigma\left(\mathbb{I}(p_i \leqslant s), p_j, s \leqslant t, s \in T, i \in I_0, j \in I_1\right)$, $t \in T$, with $T = \{0, \beta_1, ..., \beta_n\}$. Then, we get*

$$\mathbb{E}\left[\frac{V(\tau_{SD})}{R(\tau_{SD})}\right] \leqslant \frac{\alpha}{1-\alpha} - \left(\mathbb{E}\left[\frac{V(\tau_{SD})}{S(\tau_{SD})+1}\right] - \mathbb{E}\left[\frac{V(\tau_{SD})}{R(\tau_{SD})}\right]\right) \leqslant \frac{\alpha}{1-\alpha}$$

for the SD procedure based on $\beta_i, i \leqslant n$.

Observe that the filtrations $\mathcal{F}_T$ and $\mathcal{F}_T^f$ are different. The martingale condition w.r.t. $\mathcal{F}_T^f$ holds if M_{I_0} is a martingale conditioned under the outcomes $(p_i)_{i\in I_1}$, which is weaker than BIA.

Although the presented bound $\frac{\alpha}{1-\alpha}$ is slightly larger than α, the inequality can get a gain if the ratio $\frac{V(\tau_{SD})}{S(\tau_{SD})+1}$ is compared with the false discovery proportion $\frac{V(\tau_{SD})}{R(\tau_{SD})}$.

2.4. Improvement of the power

In this subsection we concentrate on the power of FDR-controlling procedures, which can be characterized by the value $\frac{\mathbb{E}[S(\tau)]}{n_1}$ for $n_1 > 0$. Let us consider a SD procedure with arbitrary critical values $\alpha_i, i \leqslant n$, which controls the FDR. Then we can increase the corresponding critical boundary value τ_{SD} and improve the power of this procedure without loss of the FDR control under the PRDS assumption. Note that the result seems to be new also for the BIA model.

Lemma 15. *Assume the following:*

1. *the random variables $(p_i)_{i\in\{1,\dots,n\}}$ satisfy (D2) on I_0,*
2. *$(p_i)_{i\in I_0}$ and $(p_i)_{i\in I_1}$ are stochastically independent,*
3. *let each $p_i, i \in I_0$ be stochastically larger than $U(0,1)$,*
4. *the SD procedure using critical values $\alpha_i, i \leqslant n$, controls the FDR at level α under 1.-3.*

Then a SD procedure using critical values $c_i = \max(\alpha_i, 1-(1-\alpha)^{\frac{1}{n}})$, $i \leqslant n$, controls the FDR at level α.

Remark 16. *(a) The critical value $1-(1-\alpha)^{\frac{1}{n}}$ is the smallest critical value of the SD procedure which was proposed by Benjamini and Liu [4]. The procedure of Benjamini and Liu controls the FDR under BIA, see Benjamini and Liu [4], and under PRDS assumption, see Sarkar [18].*

(b) Due to Lemma 15 we can increase the first critical value of the SD procedure based on the critical values $\alpha_i, i \leqslant n$, from (1.5) in order to improve the power without loss of the FDR control.

(c) The critical values $c_i = \max(\beta_i, 1-(1-\alpha)^{\frac{1}{n}})$ with β_i, $i \leqslant n$ from (1.5) are already larger than the critical values proposed by Benjamini and Liu [4].

(d) *More general results about increased critical values can be found in Benditkis [1], Chap.4.*

To prove Lemma 15 we need the following technical result, that shows some similarities of PRDS structure and independence.

Lemma 17. *Let $(b_i)_{i\in I_0}$ be some set of real numbers with $b_i \in [0,1]$, $i \in I_0$. If p-values $(p_i)_{i\leqslant n}$ are PRDS on I_0 then*

$$P\left(\bigcap_{i\in I_0}\{p_i > b_i\}\right) \geqslant \prod_{i\in I_0} P\left(p_i > b_i\right). \tag{2.9}$$

If "false" p-values $f = (f_1, .., f_{n_1}) = (p_i)_{i\in I_1}$ are specified then we denote the conditional expectation by $\mathbb{E}_f\left[\cdot\right] = \mathbb{E}\left[\cdot|(p_i)_{i\in I_1}\right]$.

3. Simulation example

We conducted a simulation study to investigate the FDR level under martingale dependence of adaptive procedures numerically. Therefore we compare the SD procedure based on (1.5) with the enlarged critical values by Lemma 15, the linear Benjamini Hochberg SD and SU procedures, that use the critical values $\frac{i\alpha}{n}$, $i = 1, ..., n$, and the adaptive $\lambda-$based Storey's SU procedure, based on the critical values $\lambda \wedge \frac{i}{\hat{n}_0(\lambda)}\alpha$ with $\hat{n}_0(\lambda) = n\frac{1-\hat{F}_n(\lambda)+\frac{1}{n}}{1-\lambda}$. The number of tests is set to $n = 100, 500$ and 800. The fraction of the true null hypotheses $\pi_0 = n_0/n$ is assumed to be $\pi_0 = 1/4, 1/2, 3/4$ and $(n-1)/n$, respectively, the tuning parameter $\lambda = 0.8$ and the FDR-controlling level $\alpha = 0.15$. Our computations are based on $L = 10000$ replications. We investigate the following configurations, which is MD and PRDS at the same time.

The random variables $X_1, ..., X_n$ are independent, 2-parameter exponentially distributed with scale parameter $\tilde{\lambda} = 1$ and location parameter $\vartheta_i \leqslant 5, i = 1, ..., n$. X_0 is independent from each $X_i, i = 1, ..., n$ and X_0 is exponential distributed with scale parameter $\tilde{\lambda} = 1$ and location parameter $\vartheta_0 = 5$. We consider the test problem

$$H_i : \vartheta_i = 5 \qquad \text{vs.} \qquad H_i^c : \vartheta_i < 5. \tag{3.1}$$

The location parameter $\vartheta_i, i \in I$ is generated as

$$\vartheta_i = 5\mathbb{I}(i \in I_0) + \zeta_i\mathbb{I}(i \in I_1), i \in I, \quad \text{with} \quad \zeta_i \quad \text{i.i.d.} \quad U(0, 0.5).$$

The p-values are generated in the following way. We define the test statistics

$$D_i = X_0 \wedge X_i,$$

and the corresponding p-values by $p_i = G_{H_0}(D_i)$. Thereby, G_{H_0} is the distribution function of the test statistics D_i, $i = 1, ..., n$, under the null hypotheses. We have

$$G_{H_0}(t) = P_{H_0}(X_0 \wedge X_i \leqslant t)$$

$$= 1 - P_{H_0}(X_i > t)P(X_0 > t)$$
$$= 1 - (\mathbb{I}(t < 5) + \mathbb{I}(5 \leqslant t)\exp(-(t-5)))^2$$
$$= \mathbb{I}(5 \leqslant t)(1 - \exp(-2(t-5))).$$

We summarize the results of the aforementioned simulation in the following table for $\alpha = 0.15$.

TABLE 1

Comparison of the FDR for the β_i-based-SD procedure with (1.5) (β_i-SD), the adaptive Storeys SU procedure (StSU),the linear SD procedure (BHsd) and the linear SU procedure (BHsu) under the martingale dependence.

n	π_0	β_i-SD	BHsd	BHsu	StSu
$n = 100$	99/100	0.0734	0, 0728	0.1417	0.39
	3/4	0.093	0.071	0.1148	0.306
	1/2	0.103	0.057	0.074	0.212
	1/4	0.102	0.033	0.037	0.124
$n = 500$	499/500	0.0782	0.0781	0.152	0.404
	3/4	0.091	0.069	0.108	0.302
	1/2	0.105	0.058	0.079	0.216
	1/4	0.104	0.033	0.037	0.125
$n = 800$	799/800	0.074	0.073	0.145	0.395
	3/4	0.093	0.072	0.11	0.3
	1/2	0.104	0.057	0.074	0.207
	1/4	0.103	0.033	0.037	0.125

As we can see from Table 1 the linear SU exhausts the level $\alpha = 0.15$ well if almost all hypotheses are true, as in the BIA case. If the portion π_0 of the true nulls is smaller than $\frac{3}{4}$, the β_i-SD seems to exhaust the level α better than the BH SU although a SD procedures are compared with SU ones. The FDR of the linear SD is smaller than the FDR of the β_i-SD for all values of $n_0(n)$. The FDR of the adaptive $\lambda-$based procedure of Storey lies above the level α in most cases, hence, this procedure can not be used for such kind of dependence. Of course it is known that Storey's procedure is not robust under dependence. Further simulations can be found in Benditkis [1].

Remark 18. *We can see from Table 1 that the FDR of the β_i-SD increases when π_0 decreases down to $\frac{1}{2}$. This phenomenon does not seem to be unusual for this SD procedure under positive dependence, see for example Gavrilov et al. [11], p.627, Benditkis [1], pp. 95-97. Note that for $\pi_0 = \frac{99}{100}$ the influence of $p_i,\ i \in H_1$, is small which causes relatively small value of R and β_R is close to the stopping rule of the BH SD procedure. However, as π_0 decreases down to $\frac{1}{2}$ the positive dependence between $p_i,\ i \in H_0$, and $p_i,\ i \in H_1$, becomes more influential. We think that stronger positive dependence is here responsible for the present effect.*

4. Results for SU Procedures

It is well known that the FDR of the SU procedures with critical values $\beta_i,\ i \leqslant n$, see (1.7), may exceed the prespecified level α. In particular, by Lemma 3.25 of

Gontscharuk [12] the worst case FDR is greater than α in the limit $n \to \infty$. The reason for this is that $\mathbb{E}\left[V(\tau_{SU})\right]$ may exceed the bound $\frac{\alpha}{1-\alpha}(n_1+1)$ under some Dirac uniform configurations, i.e. there exist values of n_0, such that $\mathbb{E}\left[V(\tau_{SU})\right] \geqslant \frac{\alpha}{1-\alpha}(n_1+1)$, if $p_i = 0$, for all $i \in I_1$ and $p_j,\ j \in I_0$ are i.i.d. uniformly $U(0,1)$ distributed. Below the critical values β_i are slightly modified in order to get finite sample FDR control. Main tools for the proof are reverse martingale arguments which were already applied by Heesen and Janssen [14] for step up procedures, which extend results for BIA models. Introduce the reverse filtration

$$\mathcal{G}_t^T = \sigma\{(\mathbb{I}(p_i \leqslant s), 1 \leqslant i \leqslant n, s \geqslant t), s, t \in T\}$$

given by the p-values.

(R) Let $T \subset (0,1]$ be a set with $1 \in T$. We say that p-values $p_1, ..., p_n$ are $\mathcal{G}_t^T$–reverse super-martingale dependent if $\frac{V(t)}{t} = \frac{\sum_{i\in I_0} \mathbb{I}(p_i \leqslant t)}{t}$ is a $\left(\mathcal{G}_t^T\right)_{t\in T}$-reverse super-martingale.

Lemma 19. *Consider R-super-martingale dependent p-values for an index set $1 \in T \subset [\delta, 1]$ for some $\delta > 0$. Let τ be any $(\mathcal{G}_t^T)_{t\in T}$ reverse stopping time with values τ in T. Then we have*

$$\mathbb{E}\left[\frac{V(\tau)}{\tau}\right] \leqslant n_0 \tag{4.1}$$

with equality "=" if the reverse super-martingale is a reverse-martingale.

Remark 20. *The inequality (4.1) is also fulfilled under the so called "dependency control condition", which was proposed by Blanchard and Roquain [6]. Note that the assumption (R) and the dependency control condition do not imply each other.*

Lemma 19 applies to various SU procedures.

Example 21. *Consider critical values $0 < a_1 \leqslant a_2 \leqslant ... \leqslant a_n < 1$ and an index set T, $\{a_1, a_2, ..., a_n, 1\} \subset T \subset [\delta, 1]$ for some constant $0 < \delta \leqslant a_1$.*

(a) (SU procedures given by $(a_i)_i$.) The variable

$$\tau = \max(a_i : p_{i:n} \leqslant a_i) \vee a_1 \tag{4.2}$$

is a reverse stopping time with $\tau_{SU} \vee a_1 = \tau$, $R(\tau_{SU}) = R(\tau)$ and also $V(\tau_{SU}) = 0$ if $\tau_{SU} \neq \tau$. Thus

$$\mathbb{E}\left[\frac{V(\tau_{SU})}{\tau_{SU}}\right] = \mathbb{E}\left[\frac{V(\tau)}{\tau}\right] \leqslant n_0. \tag{4.3}$$

(b) (Truncation of the SU procedure given by (a).) Assume the R-super-martingale condition for $T = [\eta, 1]$ and $0 < \eta \leqslant a_1$. Imagine that the statistician likes to reject

- *at most k hypotheses, $1 \leqslant k \leqslant n$, but all H_i with p-values $p_i \leqslant \eta$.*

– *Introduce* $\tau_0 = \max\{t \in [0,1] : \hat{F}_n(t) \leqslant \frac{k}{n}\}$ *and the reverse stopping time*

$$\tilde{\tau} = (\tau_0 \wedge \tau_{SU}) \vee \eta. \tag{4.4}$$

Then, the inequality (4.3) holds when τ *is replaced by* $\tilde{\tau}$*. Obviously, also* $\mathbb{E}\left[V(\tilde{\tau})\right] \leqslant \mathbb{E}\left[V(\tau)\right]$ *follows. In case* $a_i = \alpha_i$*, see (1.5), the condition* $\mathbb{E}\left[V(\tau)\right] \leqslant \frac{\alpha}{1-\alpha}(n_1+1)$ *thus, would imply control for* $FDR_{SU}(\tau)$ *as well as for* $FDR_{SU}(\tilde{\tau})$*.*

Below a finite sample exact SU procedure under the BIA model is presented, which can be established by numerical methods or Monte Carlo tools. Consider new coefficients

$$a_i = \frac{i\alpha}{n+1-i\delta},\ i \leqslant n,\ 0 \leqslant \delta < 1-\alpha. \tag{4.5}$$

Let $\mathcal{P}_{BI(n)}$ denote all distributions of p-values at sample size n under the basic independence BIA regime. Then, the worst case FDR of the SU procedure given by (4.5) under the parameters (n,δ) is

$$\sup_{\mathcal{P}_{BI(n)}} FDR(n,\delta) = \max_{0\leqslant n_1<n} FDR_{DU(n_1)}(n,\delta) \tag{4.6}$$

given by a Dirac uniform configuration, where $FDR_{DU(n_1)}(n,\delta)$ denotes the step up FDR under $\mathrm{DU}(n_1)$ with uniformly distributed p-values p_i for $i \in I_0$. Recall from Heesen and Janssen [14] that there exists a unique parameter $\kappa_n = \delta \in (0, 1-\alpha)$ for the coefficients (4.5) with

$$\sup_{\mathcal{P}_{BI(n)}} FDR(n,\kappa_n) = \alpha \tag{4.7}$$

with larger (smaller) worst case FDR for $\delta > \kappa_n$ ($\delta < \kappa_n$, respectively). The solution κ_n can be found by checking the maximum (4.6) of a finite number of constellations.

The next theorem introduces the asymptotics of the present SU procedures under the basic independence model.

Theorem 22. *Consider a sequence of SU procedures with critical values* $a_i = a_i(\delta_n)$, $1 \leqslant i \leqslant n$, *given by (4.5) with* $0 \leqslant \delta_n < 1-\alpha$.

(a) Under the condition $\limsup_{n\to\infty} \delta_n < 1-\alpha$ *we have*

$$\limsup_{n\to\infty} \sup_{\mathcal{P}_{BI(n)}} FDR(n,\delta_n) = \alpha. \tag{4.8}$$

(b) Assume that $\delta_n \to \delta \in (0,1-\alpha)$ *and let the portion* $\frac{n_0}{n} \leqslant c$ *be limited by some constant* $\alpha < c < 1$*. Then*

$$\limsup_{n\to\infty} \sup_{\mathcal{P}_{BI(n)},\ n_0\leqslant cn} FDR(n,\delta_n) = \frac{cx(\delta)}{1-c+cx(\delta)} < \alpha, \tag{4.9}$$

where $x(\delta) = \frac{\left((c\alpha+\delta(1-c)-1)^2-4(1-c)c\alpha\delta\right)^{1/2}-c\alpha-\delta(1-c)+1}{2c\delta}$.

Remark 23. *Theorem 22 together with the finite sample adjusted SU procedures at parameter κ_n, see (4.7), can be viewed as a finite sample contribution to the program of Finner et al. [8], who got the asymptotically optimal rejection curve for SU procedures.*

5. Finite sample results for SU procedures using critical values of Gavrilov et al.

Consider below a SU procedure using critical values $\beta_i = \frac{i\alpha}{n+1-i(1-\alpha)},\ i \leqslant n$. As mentioned above, this SU procedure may exceed the FDR level α under some Dirac uniform configurations (cf. Gontscharuk [12], Heesen and Janssen [14]).

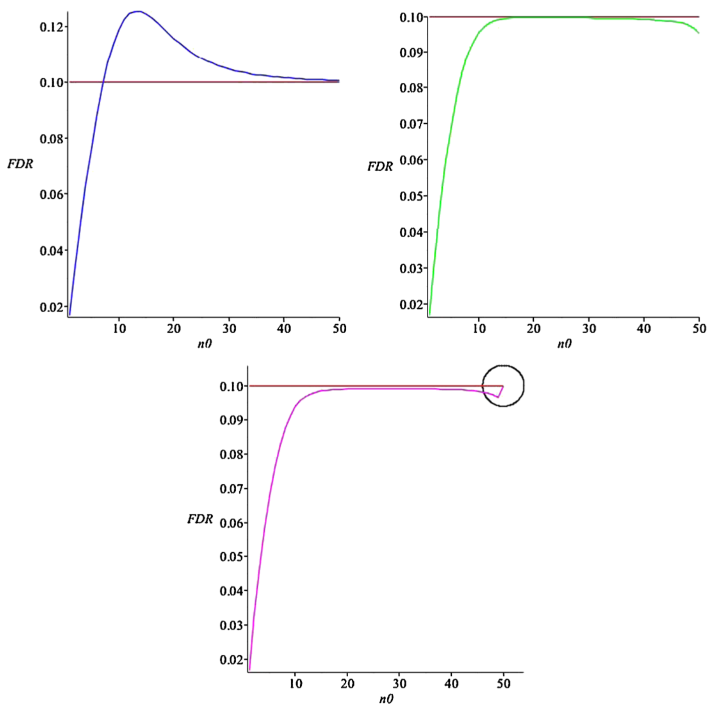

Fig 1. *The FDR under Dirac uniform configuration, as function of n_0, $1 \leqslant n_0 \leqslant n$, of SU (blue line), SD (green line) and SD with improved by Lemma 15 first critical value (magenta line) for Dirac uniform configuration of the procedures based on the set of critical values β_i, $i \leqslant n$ with $n = 50$, $\alpha = 0.1$.*

Remark 24. *Note that Figure 1 relies on the exact formula that is based on Bolshev's recursion (Shorack and Wellner [21]) and Monte Carlo simulations are not required. We refer to Gontscharuk ([12], p.44) for more details.*

As we can see from the Figure 1, the FDR of the SU procedure may be larger than prechosen level α in contrast to the SD procedure based on the same critical values $\beta_i, i \leqslant n$. The next theorem gives an explanation in terms of the ENFR. The range $n_0 \in [10, 50]$ covers the most important region where the inequality of the Theorem 25 works.

Theorem 25. *Let $(p_i)_{i\in I_0}$ fulfil the reverse martingale dependence assumption (R) and $p_j = 0$ whenever $j \in I_1$. For $f(n) = \frac{2\alpha(n+1)^2}{n+3}$ we get for SU procedure based on the critical values (1.1)*

(a) $\mathbb{E}\left[V(\tau_{SU})\right] \geqslant \frac{\alpha}{1-\alpha}(n_1+1)$ for all $n_0 \geqslant f(n)$,
(b) $FDR_{\tau_{SU}} \geqslant \alpha$ for all $n_0 \geqslant f(n)$.

Moreover, we have ">" in (a) and (b) if $n_0 > f(n)$.

In the concrete situation of Figure 1 we observe $f(50) = 9.8$, which is visible in the first graphic.

Remark 26. *As supplement the following lower $FDR_{\tau_{SU}}$ bound can obtained for small n_0 which is not covered by Theorem 25. Under the Dirac uniform assumptions we have*

$$\frac{n_0}{n}\left(\frac{1}{1+\frac{1}{\alpha}\left(\frac{n_0}{n}(1-\alpha)+\frac{1}{n}\right)}\right) \leqslant FDR_{\tau_{SU}} \leqslant \frac{n_0}{n},$$

where $\frac{n_0}{n}$ is a universal upper FDR bound. Note that for large n and $n_0 << n$ the value $\left(\frac{n_0}{n}\right)^{-1}$ FDR is close to 1 which is here much higher than for the standard BH procedures, where the ration is just α. We obtain that the slope at 0 becomes maximal as $n \to \infty$.

Proof of Remark 26. The upper bound holds since $V \mapsto \frac{V}{n_1+V} \leqslant \frac{n_0}{n}$ is increasing. The lower bound is given by the inequality $\mathbb{E}\left[V\right] \geqslant n_0\beta_{n_1}$ and (1.7). □

6. The proofs and technical results

The proof of Lemma 3 is obvious.

Proof of Lemma 5. (a) Firstly, consider the case $\{\beta_i : p_{i:n} > \beta_i,\ i \leqslant n\} \neq \emptyset$ and define $j^* = \min\{i : p_{i:n} > \beta_i,\ i \leqslant n\}$. Then, we get $\sigma = \beta_{j^*}$ due to the definition of σ. This implies

$$\beta_{j^*} < p_{j^*:n} \text{ and } \ \beta_i \geqslant p_{i:n} \ \text{ for all } \ i \leqslant j^*-1.$$

Consequently we get

$$p_{j^*-1} \leqslant \beta_{j^*-1} < \beta_{j^*} < p_{j^*:n},$$

which implies $R(\beta_{j^*-1}) = R(\beta_{j^*}) = j^* - 1$. Due to (2.3) we get

$$j^* = \min\{i : \hat{\alpha}_n(\beta_i) = \alpha,\ i \leqslant n\},$$

which completes the proof for this case. The case $\{\beta_i : p_{i:n} > \beta_i,\ i \leqslant n\} = \emptyset$ is obvious since $\hat{\alpha}(t) = 0$ if $\hat{F}_n(t) = 1$.

The first statement of (b) is obvious and coincides with Remark 2. If there is any index $i \leqslant n$ with $R(\beta_i) = i - 1$, then $\hat{\alpha}_n(\sigma) = \alpha$ due to (2.2) and (a). Otherwise we have $\hat{\alpha}_n(\beta_i) < \alpha$ for all $i \leqslant n$ and $R(\sigma) = R(\beta_n) = n$ holds. This implies $\hat{\alpha}_n(\sigma) = 0$. Consequently, we get $\hat{\alpha}_n(\sigma) = \alpha \mathbb{I}(R(\sigma) < n)$.

The part (c) is obvious. □

Since τ_{SD} is not a stopping time w.r.t. $\mathcal{F}_T$ we will turn to the critical boundary σ in order to apply Lemma 5.

Proof of Lemma 6 and Theorem 7. Firstly, note that we have $V(\sigma) = V(\tau_{SD})$, as well as $S(\sigma) = S(\tau_{SD})$. Due to Lemma 5 (b) we have $\hat{\alpha}_n(\sigma) = \alpha \mathbb{I}(R(\sigma) < n)$. Further, we obtain

$$(1 - \hat{\alpha}_n(\sigma))(R(\sigma) + 1) = M_{I_0}(\sigma) + \frac{S(\sigma) - n_1\sigma}{1 - \sigma} + 1 \tag{6.1}$$

by (2.1) and (2.4) which is a fundamental equation connection between $\hat{\alpha}(\cdot)$ and $M_{I_0}(\cdot)$.

In case $S(\sigma) = n_1$ of Lemma 6 we have

$$(1 - \alpha)V(\sigma) = M_{I_0}(\sigma) + \alpha(n_1 + 1) - \alpha(n + 1)\mathbb{I}(V(\sigma) = n_0), \tag{6.2}$$

which implies the equivalence in Lemma 6.

Under the conditions of Theorem 7 we have $\mathbb{E}\left[M_{I_0}(\sigma)\right] \leqslant 0$ by the optional sampling Theorem of stopped super-martingales. Thus, the fact that we have $\frac{S(\sigma) - n_1\sigma}{1 - \sigma} \leqslant S(\sigma)$ implies

$$(1 - \alpha)\mathbb{E}\left[V(\tau_{SD})\right] = (1 - \alpha)\mathbb{E}\left[V(\sigma)\right] \leqslant \mathbb{E}\left[M_{I_0}(\sigma)\right] + \alpha\mathbb{E}\left[S(\sigma) + 1\right] \leqslant \alpha(n_1 + 1) \tag{6.3}$$

due to (6.1) and Remark 2. □

Proof of Lemma 9. (a) Consider again the equality (6.2). If expectations are taken, the optional sampling theorem applies to $M_{I_0}(\sigma)$, which proves the result by Lemma 5 (a).

(b) Analogous to the case (a) we have

$$(1 - \alpha)(V(\sigma) + S(\sigma) + 1) + \alpha(n + 1)\mathbb{I}(R(\sigma) = n) = M_{I_0}(\sigma) + M_{I_1}(\sigma) + 1, \tag{6.4}$$

thereby $M_{I_1}(t) = \frac{\sum_{i \in I_1} \mathbb{I}(p_i \leqslant t) - n_1 t}{1 - t}$ is a $\mathcal{F}_t^T$−martingale. Equality (6.4) implies

$$(1 - \alpha)\mathbb{E}\left[V(\sigma) + S(\sigma)\right] = \alpha - \alpha(n + 1)P(R(\sigma) = n) \tag{6.5}$$

by taking the expectation $\mathbb{E}$ and applying the Optional Sampling Theorem. The equality $\mathbb{E}\left[S(\sigma)\right] = \frac{n_1}{n_0}\mathbb{E}\left[V(\sigma)\right]$ (which follows from the assumption that all p-values are identically distributed) completes the proof of part (b) of this lemma.

(c) The proof follows immediately from the observation that under martingale dependence the critical boundary value τ_{SD}, and, consequently, $V(\tau_{SD})$, becomes maximal under assumptions of part (a) and minimal under assumptions of part (b). □

Remark 27. *1. The proof of the next Lemma 12 uses the technique which was proposed by Finner and Roters [9] for the proof of FDR-control of Benjamini and Hochberg procedure under PRDS.*

2. As long as we are concerned with the super-martingale assumption (D1) we may assume w.l.o.g. that $(p_i)_{i\in I_0}$ are identically distributed and stochastically larger than $U(0,1)$, c.f. Remark 8. These technical tools are only used for the subsequent proofs of Sections 2.2 - 2.4 in connection with PRDS. The reference to Remark 8 is not cited again in each step of the proofs.

Proof of Lemma 12. Let us define $\beta_0 = 0$ for technical reasons and denote $(U_j)_{j\leqslant n_0} := (p_i)_{i\in I_0}$. Firstly, note that $\tau_{SD} = \beta_R$ holds obviously. Thereby, $R = R(\tau_{SD})$ is the number of rejections of the SD procedure with deterministic critical values $\beta_i,\ i \leqslant n$. We obtain the following sequence of (in)equalities:

$$\mathbb{E}\left[\frac{V(\tau_{SD})}{\tau_{SD}}\right] = \sum_{j=1}^{n_0} \mathbb{E}\left[\frac{\mathbb{I}(U_j \leqslant \beta_R)}{\beta_R}\right] \tag{6.6}$$

$$= \sum_{j=1}^{n_0}\sum_{i=1}^{n} \mathbb{E}\left[\frac{\mathbb{I}(U_j \leqslant \beta_i)}{\beta_i}\mathbb{I}(\beta_R = \beta_i)\right] \tag{6.7}$$

$$= \sum_{j=1}^{n_0}\sum_{i=1}^{n} \mathbb{E}\left[\frac{\mathbb{I}(U_j \leqslant \beta_i)}{\beta_i}\left(\mathbb{I}(\beta_R \leqslant \beta_i) - \mathbb{I}(\beta_R \leqslant \beta_{i-1})\right)\right] \tag{6.8}$$

$$= \sum_{j=1}^{n_0}\sum_{i=1}^{n} \frac{P(U_j \leqslant \beta_i)}{\beta_i}\mathbb{E}\left[\frac{\mathbb{I}(U_j \leqslant \beta_i)\left(\mathbb{I}(\beta_R \leqslant \beta_i) - \mathbb{I}(\beta_R \leqslant \beta_{i-1})\right)}{P(U_j \leqslant \beta_i)}\right] \tag{6.9}$$

$$\leqslant \sum_{j=1}^{n_0}\sum_{i=1}^{n} \left(\mathbb{E}\left[\mathbb{I}(\beta_R(p) \leqslant \beta_i)|U_j \leqslant \beta_i\right] - \mathbb{E}\left[\mathbb{I}(\beta_R(p) \leqslant \beta_{i-1})|U_j \leqslant \beta_i\right]\right) \tag{6.10}$$

$$\leqslant \sum_{j=1}^{n_0}\sum_{i=1}^{n} \left(\mathbb{E}\left[\mathbb{I}(\beta_R \leqslant \beta_i)|U_j \leqslant \beta_i\right] - E\left[\mathbb{I}(\beta_R \leqslant \beta_{i-1})|U_j \leqslant \beta_{i-1}\right]\right) \tag{6.11}$$

$$= \sum_{j=1}^{n_0} \mathbb{E}\left[\mathbb{I}(\beta_R \leqslant \beta_n)|U_j \leqslant \beta_n\right] = n_0. \tag{6.12}$$

The inequality in (6.10) is valid since $U_1, ..., U_{n_0}$ are stochastically greater than $U(0,1)$. Obviously, the inequality in (6.11) holds because the present function $x \mapsto \mathbb{I}(\beta_R(p) \leqslant \beta_{i-1} \mid U_i \leqslant x)$ is increasing in x for all $i \in \{1, ..., n_0\}$ and since $U_1, ..., U_{n_0}$ are assumed to be PRDS. Consequently, using the telescoping sum we obtain the first equality in (6.12). The proof is completed because $\beta_R(p) \leqslant \beta_n$ by definition of $\beta_R = \tau_{SD}$. □

Proof of Theorem 11. Combining Lemma 3, Theorem 7 and Lemma 12 yields the statement. ☐

To prove Theorem 14 we need the following technical result. For the sake of simplicity we use the designation $M = M_{I_0}$.

Lemma 28. *Under the assumptions of Theorem 14*

$$E\left[\frac{M(\sigma)}{S(\sigma)+1}\right] \leqslant 0.$$

Proof of Lemma 28. First, note that the process $\beta_i \mapsto S(\beta_i)$ is always $\mathcal{F}^f_T$-measurable. If we put $M = M_{I_0}$ then

$$\begin{aligned}
\mathbb{E}\left[\frac{M(\sigma)}{S(\sigma)+1}\right] &= \mathbb{E}\left[\sum_{i=1}^n \frac{M(\beta_i)}{S(\beta_i)+1}\mathbb{I}(\sigma=\beta_i)\right] \\
&= \mathbb{E}\left[\sum_{i=1}^n \frac{M(\beta_i)}{S(\beta_i)+1}\left(\mathbb{I}(\sigma\leqslant\beta_i)-\mathbb{I}(\sigma\leqslant\beta_{i-1})\right)\right] \\
&= \sum_{i=1}^n \left(\mathbb{E}\left[\frac{M(\beta_i)}{S(\beta_i)+1}\mathbb{I}(\sigma\leqslant\beta_i)\right] - \mathbb{E}\left[\mathbb{E}\left[\frac{M(\beta_i)}{S(\beta_i)+1}\mathbb{I}(\sigma\leqslant\beta_{i-1})|\mathcal{F}^f_{\beta_{i-1}}\right]\right]\right).
\end{aligned} \tag{6.13}$$

Since σ is a $\mathcal{F}^f_T$-stopping time, $S(\beta_i)$ is a measurable w.r.t. $\mathcal{F}^f_{\beta_{i-1}}$ and using martingale property $\mathbb{E}\left[M(\beta_i)|\mathcal{F}^f_{\beta_{i-1}}\right] = M(\beta_{i-1})$ we get

$$\mathbb{E}\left[\mathbb{E}\left[\frac{M(\beta_i)}{S(\beta_i)+1}\mathbb{I}(\sigma\leqslant\beta_{i-1})|\mathcal{F}^f_{\beta_{i-1}}\right]\right] = \mathbb{E}\left[\frac{M(\beta_{i-1})}{S(\beta_i)+1}\mathbb{I}(\sigma\leqslant\beta_{i-1})\right].$$

Define $\beta_0 = 0$. Now, we can continue the chain of equalities (6.13) as follows.

$$\begin{aligned}
&\sum_{i=1}^n \left(\mathbb{E}\left[\frac{M(\beta_i)}{S(\beta_i)+1}\mathbb{I}(\sigma\leqslant\beta_i)\right] - \mathbb{E}\left[\frac{M(\beta_{i-1})}{S(\beta_{i-1})+1}\mathbb{I}(\sigma\leqslant\beta_{i-1})\right]\right) \\
&+\sum_{i=1}^n \left(\mathbb{E}\left[\frac{M(\beta_{i-1})}{S(\beta_{i-1})+1}\mathbb{I}(\sigma\leqslant\beta_{i-1})\right] - \mathbb{E}\left[\frac{M(\beta_{i-1})}{S(\beta_i)+1}\mathbb{I}(\sigma\leqslant\beta_{i-1})\right]\right)
\end{aligned} \tag{6.14}$$

$$= \sum_{i=1}^n \left(\mathbb{E}\left[\frac{S(\beta_i)-S(\beta_{i-1})}{(S(\beta_{i-1})+1)(S(\beta_i+1))}\mathbb{E}_f\left[M(\beta_{i-1})\mathbb{I}(\sigma\leqslant\beta_{i-1})\right]\right]\right) \tag{6.15}$$

because the first term in (6.14) is equal to zero due to the telescoping sum since $\mathbb{E}[M(\beta_n)] = 0$. Now, we will show that

$$\mathbb{E}\left[M(\beta_{i-1})\mathbb{I}(\sigma\leqslant\beta_{i-1})\right] \leqslant 0 \tag{6.16}$$

for all $i \leqslant n$. Indeed, we have

$$\mathbb{E}\left[M(\beta_{i-1})\mathbb{I}(\sigma\leqslant\beta_{i-1})\right] = -\mathbb{E}\left[M(\beta_{i-1})\mathbb{I}(\sigma>\beta_{i-1})\right]. \tag{6.17}$$

Further, by the definition of σ we know that $\hat{\alpha}_n(t) \leqslant \alpha$ for all $t \leqslant \sigma,\ t \in T$. Hence, due to (2.2) we get $\mathbb{I}(\sigma > \beta_{i-1})) = \prod_{j=0}^{i-1} \mathbb{I}\left(R(\beta_j) \geqslant j-1\right)$. On the other hand, we can conclude from the definition of the process M_{I_0} that $\mathbb{I}\left(R(\beta_j) \geqslant j-1\right) = \mathbb{I}\left(M(\beta_j) \geqslant c(j)\right)$. Thereby, constants $c_j,\ j \leqslant i-1$, are defined as $c_j = \frac{j-1-S(\beta_j)-n_0\beta_j}{1-\beta_j}$. Consequently, (6.16) is equivalent to

$$\mathbb{E}\left[M(\beta_{i-1}) \prod_{j=0}^{i-1} \mathbb{I}\left(M(\beta_j) \geqslant c(j)\right))\right] \geqslant 0, \tag{6.18}$$

which follows immediately from the following Lemma 29 (by setting $X = M(\beta_{i-1})$ and $A = \prod_{j=0}^{i-2} \mathbb{I}(M(\beta_j) \geqslant c(j)))$ and Lemma 30. □

Lemma 29. *Let X be a random variable and A be a measurable set and $c \in \mathbb{R}$ be some constant. The inequality $\mathbb{E}\left[\mathbb{I}_A X\right] \geqslant 0$ implies $\mathbb{E}\left[\mathbb{I}_A X \mathbb{I}(X > c)\right] \geqslant 0$.*

Proof of Lemma 29. The case $c \geqslant 0$ is obvious. If $c < 0$ we get

$$0 \leqslant \mathbb{E}\left[\mathbb{I}_A X\right] = \mathbb{E}\left[\mathbb{I}_A X \mathbb{I}(X > c)\right] + \mathbb{E}\left[\mathbb{I}_A X \mathbb{I}(X \leqslant c)\right] \leqslant \mathbb{E}\left[\mathbb{I}_A X \mathbb{I}(X > c)\right],$$

which implies the proof of this lemma. □

Lemma 30. *Under the assumptions of Theorem 14 we have*

$$\mathbb{E}\left[M(\beta_{i-1}) \prod_{j=0}^{i-2} \mathbb{I}(M(\beta_j) \geqslant c(j))\right] \geqslant 0 \textit{ for } i \geqslant 2. \tag{6.19}$$

Proof of Lemma 30. The proof is based on induction. Firstly, we define now $k := i-1$.

Let $k = 2$ then (6.19) is equivalent to

$$\mathbb{E}\left[M(\beta_2)\mathbb{I}(M(\beta_1) \geqslant c(1))\right] \geqslant 0,$$

which is true due to the following chain of equalities.

$$\begin{aligned} &\mathbb{E}\left[M(\beta_2)\mathbb{I}(M(\beta_1) \geqslant c(1))\right] = \mathbb{E}\left[\mathbb{E}\left[M(\beta_2)\mathbb{I}(M(\beta_1) \geqslant c(1))|\mathcal{F}^f_{\beta_1}\right]\right] \\ &= \mathbb{E}\left[\mathbb{I}(M(\beta_1) \geqslant c(1))\mathbb{E}\left[M(\beta_2)|\mathcal{F}^f_{\beta_1}\right]\right] = \mathbb{E}\left[M(\beta_1)\mathbb{I}(M(\beta_1) \geqslant c(1))\right]. \end{aligned} \tag{6.20}$$

The two last equalities in (6.20) are valid due to the measurability of $M(\beta_1)$ w.r.t. $\mathcal{F}^f_{\beta_1}$ and martingale property $\mathbb{E}\left[M(\beta_2)|\mathcal{F}^f_{\beta_1}\right] = M(\beta_1)$.

Assume that

$$\mathbb{E}\left[M(\beta_k) \prod_{j=0}^{k-1} \mathbb{I}(M(\beta_j) \geqslant c(j))\right] \geqslant 0 \tag{6.21}$$

holds for $k \geqslant 2$. Then we prove that

$$\mathbb{E}\left[M(\beta_{k+1})\prod_{j=0}^{k}\mathbb{I}(M(\beta_j) \geqslant c(j))\right] \geqslant 0$$

is also true. We have

$$\begin{aligned}
&\mathbb{E}\left[M(\beta_{k+1})\prod_{j=0}^{k}\mathbb{I}(M(\beta_j) \geqslant c(j))\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[M(\beta_{k+1})\prod_{j=0}^{k}\mathbb{I}(M(\beta_j) \geqslant c(j))|\mathcal{F}^f_{\beta_k}\right]\right] \\
&= \mathbb{E}\left[M(\beta_k)\prod_{j=0}^{k}\mathbb{I}(M(\beta_j) \geqslant c(j))\right] \\
&= \mathbb{E}\left[\left(M(\beta_k)\prod_{j=0}^{k-1}\mathbb{I}(M(\beta_j) \geqslant c(j))\right)\mathbb{I}(M(\beta_k) \geqslant c(k))\right] \geqslant 0.
\end{aligned}$$

The last inequality follows from (6.21) and Lemma 29. □

Now, we are able to prove Theorem 14.

Proof of Theorem 14. Let us remind (6.1), which implies

$$(1-\alpha)V(\sigma) \leqslant M_{I_0}(\sigma) + \alpha(S(\sigma)+1). \tag{6.22}$$

Dividing by $S(\sigma)+1$ yields

$$(1-\alpha)\frac{V(\sigma)}{S(\sigma)+1} \leqslant \frac{M_{I_0}(\sigma)}{S(\sigma)+1} + \alpha.$$

Taking the expectation $\mathbb{E}[\cdot]$ and using Lemma 28 deliver the result. □

Proof of Lemma 17. W.l.o.g. let us assume that $I_0 = \{1, ..., n\}$. For any other subset I_0 the proof works in the same way. First, we show the inequality

$$\mathbb{E}\left[\prod_{i=2}^{n}\mathbb{I}(p_i > b_i) \mid p_1 > b_1\right] \geqslant \mathbb{E}\left[\prod_{i=2}^{n}\mathbb{I}(p_i > b_i) \mid p_1 \leqslant b_1\right]. \tag{6.23}$$

To do this let F denote the marginal distribution function of p_1. Here, define $f(u) = \mathbb{E}\left[\prod_{i=2}^{n}\mathbb{I}(p_i > b_i)|p_1 = u\right]$, then we have

$$\mathbb{E}\left[\prod_{i=2}^{n}\mathbb{I}(p_i > b_i)\mathbb{I}(p_1 > b_1)\right] = \int_{b_1}^{1} f(u)dF(u)$$

and

$$\mathbb{E}\left[\prod_{i=2}^{n}\mathbb{I}(p_i > b_i)\mathbb{I}(p_1 \leqslant b_1)\right] = \int\limits_0^{b_1} f(u)dF(u).$$

Then (6.23) is equivalent to

$$\frac{\int\limits_{b_1}^{1} f(u)dF(u)}{1-F(b_1)} \geqslant \frac{\int\limits_{0}^{b_1} f(u)dF(u)}{F(b_1)}. \tag{6.24}$$

From the mean value theorem for Riemann-Stieltjes integrals we can deduce that there exist some values ξ_1 and ξ_2 with

$$\begin{aligned} \xi_1 &= \frac{\int\limits_{b_1}^{1} f(u)dF(u)}{1-F(b_1)}, \qquad \inf_{t\in(b_1,1)} f(t) \leqslant \xi_1 \leqslant \sup_{t\in(b_1,1)} f(t), \\ \xi_2 &= \frac{\int\limits_{0}^{b_1} f(u)dF(u)}{F(b_1)}, \qquad \inf_{t\in(0,b_1)} f(t) \leqslant \xi_2 \leqslant \sup_{t\in(0,b_1)} f(t). \end{aligned} \tag{6.25}$$

Since f is an increasing function of u, (6.25) yields $\xi_1 \geqslant \xi_2$, hence we get (6.24).

Further, we obtain

$$P(p_1 > b_1, p_2 > b_2, ..., p_n > b_n) = \mathbb{E}\left[\prod_{i=1}^{n}\mathbb{I}(p_i > b_i)\right] \tag{6.26}$$

$$= P(p_1 > b_1)\mathbb{E}\left[\prod_{i=2}^{n}\mathbb{I}(p_i > b_i) \mid p_1 > b_1\right] \geqslant P(p_1 > b_1)\mathbb{E}\left[\prod_{i=2}^{n}\mathbb{I}(p_i > b_i)\right] \tag{6.27}$$

$$= P(p_1 > b_1)P\left(\bigcap_{i=2}^{n}(p_i > b_i)\right) \geqslant ... \geqslant \prod_{i=1}^{n} P(p_i > b_i). \tag{6.28}$$

The inequality in (6.27) holds due to the PRDS-assumption since according to the law of total probability we have

$$\begin{aligned} &\mathbb{E}\left[\prod_{i=2}^{n}\mathbb{I}(p_i > b_i)\right] = \mathbb{E}\left[\prod_{i=2}^{n}\mathbb{I}(p_i > b_i) \mid p_1 > b_1\right] - \\ &P(p_1 \leqslant b_1)\underbrace{\left(\mathbb{E}\left[\prod_{i=2}^{n}\mathbb{I}(p_i > b_i) \mid p_1 > b_1\right] - \mathbb{E}\left[\prod_{i=2}^{n}\mathbb{I}(p_i > b_i) \mid p_1 \leqslant b_1\right]\right)}_{\geqslant 0 \text{ due to } (6.23)}. \end{aligned} \tag{6.29}$$

□

For the two following proofs we define the p-values, which belong to the true null by $(U_i)_{i\in\{1,\dots,n_0\}} = (p_j)_{j\in I_0}$. Now, we are able to prove Lemma 15 by conditioning under the portion f belonging to the false null.

Proof of Lemma 15. Consider an arbitrary FDR-controlling SD procedure that uses critical values $\alpha_i, i \leqslant n$. Let $j^* := \max\{i : f_j \leqslant 1-(1-\alpha)^{\frac{1}{n}} \text{ for all } j \leqslant i\}$ and consider two possible cases.

1. Let $j^* = 0$. In this case we have under PRDS assumption

$$\mathbb{E}_f\left[\frac{V}{R}\mathbb{I}(V>0)\right] \leqslant \mathbb{E}_f\left[\mathbb{I}(V>0)\right] = P(U_{1:n_0} \leqslant 1-(1-\alpha)^{\frac{1}{n}})$$
$$= 1 - P(\bigcap_{i=1}^{n_0}\{U_i > 1-(1-\alpha)^{\frac{1}{n}}\}) \leqslant \alpha,$$

where the last inequality is valid due to Lemma 17.

2. Let $j^* > 0$. Define the vector $f_0^* = (0,\dots,0,f_{j^*+1},\dots,f_{n_1})$, where the j^* first coordinates are replaced by 0. We get

$$\mathbb{E}_f\left[\frac{V}{R}\mathbb{I}(V>0)\right] \leqslant \mathbb{E}_f\left[\frac{V}{j^*+V}\right] = \mathbb{E}_{f_0^*}\left[\frac{V(\tau_{SD})}{R(\tau_{SD})}\right] \leqslant \alpha. \tag{6.30}$$

Thereby, τ_{SD} is the critical boundary value corresponding to the SD procedure with critical values $\alpha_i,\ i \leqslant n$. □

Proof of Lemma 19. The optional stopping theorem for reverse martingales implies

$$\mathbb{E}\left[\frac{V(\tau)}{\tau}\right] \leqslant \frac{V(1)}{1} = n_0. \tag{6.31}$$

In the case of reverse martingales we have an equality. □

It is quite obvious that the variables (4.2) and (4.4) are stopping times and Lemma 19 can be applied.

The proof of Theorem 22 requires some preparations. Consider a wider class of rejection curves given by positive parameters b and $\alpha,\ \delta \geqslant 0$ and

$$g(t) = \frac{tb}{\delta t + \alpha},\ 0 \leqslant t \leqslant 1,\ b > \delta + \alpha. \tag{6.32}$$

Note that the condition $g(1) > 1$ is necessary for proper SU procedures with critical values

$$a_i := g^{-1}\left(\frac{i}{n}\right) = \frac{\alpha i}{nb - i\delta}. \tag{6.33}$$

By the choice $b = \frac{n+1}{n}$ and $\delta = 1-\alpha$ the coefficients β_i are included. Thus we arrive at the following equation for the FDR of (6.33)

$$\text{FDR} = \mathbb{E}\left[\frac{V}{R}\right] = \frac{\alpha}{nb}\mathbb{E}\left[\frac{V}{a_R}\right] + \frac{\delta}{nb}\mathbb{E}\left[V\right] \tag{6.34}$$

for each multiple testing procedure. In contrast to SD procedures the term $\mathbb{E}\left[\frac{V}{a_R}\right]$ can be bounded under the R-super-martingale condition, cf. Heesen and Janssen [14].

Remark 31. *Consider SU procedures for parameters (δ, b, α) under R-super-martingale models with fixed portion $n_1 < n$ of false p-values.*

(a) The R-martingale models are least favourable for bounding $\mathbb{E}\left[\frac{V}{a_R}\right]$.
(b) Consider two settings $(p_i)_{i \leqslant n}$ and $(q_i)_{i \leqslant n}$ of R-martingale models. If $p_i \leqslant q_i$ holds for all $i \in I_1$ then $\mathbb{E}_p[V] \leqslant \mathbb{E}_q[V]$ and $FDR_p \leqslant FDR_q$ holds.
(c) Under BIA with uniformly distributed p-values under the null, as well as under the reverse martingale dependence (R), we have:
  (i) The Dirac uniform configuration $DU(n_1)$ is least favourable for $\mathbb{E}[V]$ and FDR.
  (ii) Suppose that $\frac{\alpha}{b}$ or $\frac{\delta}{b}$ increases. Then the coefficients a_i increase and the FDR and $\mathbb{E}[V]$ do not decrease.

Proof of Theorem 22. (a) Proposition 4.1 of Heesen and Janssen [14] establishes the asymptotic lower bound:

$$\sup_{\mathcal{P}_{BI(n)}} \mathrm{FDR}(n, \delta_n) \geqslant \min_{i \leqslant n} \frac{n a_i}{i} = \frac{n\alpha}{n+1-\delta_n} \to \alpha$$

as $n \to \infty$. To obtain the upper bound we can first exclude all coefficients $\delta_n = 0$, which correspond to a Benjamini and Hochberg procedure with level $\frac{n\alpha}{n+1}$. Fix some value γ with $\limsup_{n\to\infty} \delta_n < \gamma < 1-\alpha$ and introduce the rejection curve $g_\gamma(t) = \frac{t}{\gamma t + \alpha}$. For all $\delta_n < \gamma$ the $\mathrm{FDR}(n, \delta_n)$ of the a_i's can now be compared with the FDR of the SU procedure with critical values $g_\gamma^{-1}\left(\frac{i}{n}\right)$. By (6.34) and Remark (31) we have for each regime

$$\mathrm{FDR}(n, \delta_n) \leqslant \mathrm{FDR}_n(g_\gamma^{-1})$$

using obvious notations. The worst case asymptotics is given by Theorem 5.1 of Heesen and Janssen [14]

$$\limsup_{n\to\infty} \sup_{\mathrm{BI}(n)} \mathrm{FDR}_n(g_\gamma^{-1}) = K,$$

where

$$K = \sup\left\{\frac{x}{1-x}\frac{1-g_\gamma(x)}{g_\gamma(x)} : 0 < x \leqslant g_\gamma^{-1}(1)\right\} = \alpha.$$

(b) Similarly as above the $\mathrm{FDR}(n, \delta_n)$ is bounded below and above by the FDR of SU given by rejection curves. Choose constants $0 < \gamma_1 < \delta < \gamma_2 < 1-\alpha$ and $b > 1$ and consider $\delta_n \in (\gamma_1, \gamma_2)$ and large n with $\frac{n+1}{n} \leqslant b$. Introduce $g_{\gamma_2}(t) = \frac{t}{\gamma_2 t + \alpha}$ and $g_{\gamma_1,b}(t) = \frac{tb}{\gamma_1 t + \alpha}$. Again we have

$$\mathrm{FDR}_n(g_{\gamma_1,b}) \leqslant \mathrm{FDR}(n, \delta_n) \leqslant \mathrm{FDR}_n(g_{\gamma_2}).$$

Let $x(\gamma_2) \in (0,1)$ denote the unique solution of the equation

$$g_{\gamma_2}(x) = (1-c) + xc. \tag{6.35}$$

If we repeat the proof of Proposition 5.1 of Heesen and Janssen [14] we have

$$\limsup_{n\to\infty} \mathrm{FDR}_n(g_{\gamma_2}) = \sup\{\frac{x}{1-x}\frac{1-g_{\gamma_2}(x)}{g_{\gamma_2}(x)} : x(\gamma_2) \leqslant x \leqslant g_{\gamma_2}^{-1}(1)\}$$

which is equal to $\frac{cx(\gamma_2)}{(1-c)+cx(\gamma_2)}$. Similarly, a lower bound of (4.9) is $\frac{cx(\gamma_1,b)}{(1-c)+cx(\gamma_1,b)}$ with solution $x(\gamma_1, b) \in (0,1)$ of (6.35). If now $\gamma_1 \uparrow \delta,\ \gamma_2 \downarrow \delta$ and $b \downarrow \delta$, the bounds turn to the value $\frac{cx(\delta)}{1-c+cx(\delta)}$ given by the solution $x(\delta) \in (0,1)$ of (6.35). ☐

Proof of Theorem 25. Define the process $\left(\tilde{M}(t)\right)_{t\in T} = \left(\frac{V(t)}{t} - n_0\right)_{t\in T}$, which is, obviously, a centered reverse martingale w.r.t $\mathcal{G}_t^T$ due to the reverse martingale assumption. Now, we remind that for the step-wise procedure using critical values $(\beta_i)_{i\leqslant n}$ the following equality is valid by (6.2):

$$(1-\alpha)V(\tau_{SU}) = \tilde{M}(\tau_{SU})\frac{\tau}{1-\tau} - \alpha(n+1)\mathbb{I}\left(V(\tau_{SU}) = n_0\right) + \alpha(n_1+1)$$

under the Dirac distribution of "false" p-values $(p_i)_{i\in I_1}$. Thus, we have to show

$$\mathbb{E}\left[\tilde{M}(\tau_{SU})\frac{\tau_{SU}}{1-\tau_{SU}}\right] \geqslant \alpha(n+1)P(V(\tau_{SU})) \quad \text{for} \quad n_0 \geqslant f(n)$$

to prove the part (a). First note, that because of $V(\tau_{SU}) = V(\tau_{SU} \vee \beta_1)$ it is enough to prove the statement of this theorem for the reverse stopping time $\tilde{\tau}_{SU} = \tau_{SU} \vee \beta_1$, where $\tilde{\tau}_{SU} \in \{\beta_1, ..., \beta_n\}$. Fix an $\varepsilon = \varepsilon(n) > 0$ with $\beta_n + \varepsilon < 1$ and define a fictive additional coefficient $\beta_{n+1} = \beta_n + \varepsilon$. Then, we have

$$\mathbb{E}\left[\tilde{M}(\tilde{\tau}_{SU})\frac{\tilde{\tau}_{SU}}{1-\tilde{\tau}_{SU}}\right] = \sum_{i=1}^n \mathbb{E}\left[\tilde{M}(\beta_i)\frac{\beta_i}{1-\beta_i}\mathbb{I}(\tilde{\tau}_{SU} = \beta_i)\right] \tag{6.36}$$

$$= \sum_{i=1}^n \mathbb{E}\left[\tilde{M}(\beta_i)\frac{\beta_i}{1-\beta_i}\mathbb{I}(\tilde{\tau}_{SU} \geqslant \beta_i) - \tilde{M}(\beta_i)\frac{\beta_i}{1-\beta_i}\mathbb{I}(\tilde{\tau}_{SU} \geqslant \beta_{i+1})\right]. \tag{6.37}$$

Consider the term $\mathbb{E}\left[\tilde{M}(\beta_i)\frac{\beta_i}{1-\beta_i}\mathbb{I}(\tilde{\tau}_{SU} \geqslant \beta_{i+1})\right]$. Since $\tilde{\tau}_{SU}$ is a reverse stopping time w.r.t. filtration $\mathcal{G}_t^T$ the value $\mathbb{I}(\tilde{\tau}_{SU} \geqslant \beta_{i+1})$ is $\mathcal{G}_{\beta_{i+1}}$−measurable. Further, using the reverse martingale property of $\tilde{M}(t)$ we get

$$\begin{aligned}\mathbb{E}\left[\tilde{M}(\beta_i)\frac{\beta_i}{1-\beta_i}\mathbb{I}(\tilde{\tau}_{SU} \geqslant \beta_{i+1})\right] &= \frac{\beta_i}{1-\beta_i}\mathbb{E}\left[\mathbb{I}(\tilde{\tau}_{SU} \geqslant \beta_{i+1})\mathbb{E}\left[\tilde{M}(\beta_i)|\mathcal{G}_{\beta_{i+1}}\right]\right] \\ &= \frac{\beta_i}{1-\beta_i}\mathbb{E}\left[\mathbb{I}(\tilde{\tau}_{SU} \geqslant \beta_{i+1})\tilde{M}(\beta_{i+1})\right].\end{aligned}$$

Consequently, continuing the chain of equalities (6.36)-(6.37) we get

$$
\begin{aligned}
&\mathbb{E}\left[\tilde{M}(\tilde{\tau}_{SU})\frac{\tilde{\tau}_{SU}}{1-\tilde{\tau}_{SU}}\right] \\
&= \sum_{i=1}^{n}\left(\frac{\beta_i}{1-\beta_i}\mathbb{E}\left[\tilde{M}(\beta_i)\mathbb{I}(\tilde{\tau}_{SU}\geqslant\beta_i)\right]-\frac{\beta_i}{1-\beta_i}\mathbb{E}\left[\tilde{M}(\beta_{i+1})\mathbb{I}(\tilde{\tau}_{SU}\geqslant\beta_{i+1})\right]\right)
\end{aligned}
\tag{6.38}
$$

$$
\begin{aligned}
&= \sum_{i=1}^{n}\left(\frac{\beta_i}{1-\beta_i}\mathbb{E}\left[\tilde{M}(\beta_i)\mathbb{I}(\tilde{\tau}_{SU}\geqslant\beta_i)\right]-\frac{\beta_{i+1}}{1-\beta_{i+1}}\mathbb{E}\left[\tilde{M}(\beta_{i+1})\mathbb{I}(\tilde{\tau}_{SU}\geqslant\beta_{i+1})\right]\right) \\
&+ \sum_{i=1}^{n}\left(\left(\frac{\beta_{i+1}}{1-\beta_{i+1}}-\frac{\beta_i}{1-\beta_i}\right)\mathbb{E}\left[\tilde{M}(\beta_{i+1})\mathbb{I}(\tilde{\tau}_{SU}\geqslant\beta_{i+1})\right]\right) \\
&= \sum_{i=1}^{n}\left(\left(\frac{\beta_{i+1}}{1-\beta_{i+1}}-\frac{\beta_i}{1-\beta_i}\right)\mathbb{E}\left[\tilde{M}(\beta_{i+1})\mathbb{I}(\tilde{\tau}_{SU}\geqslant\beta_{i+1})\right]\right).
\end{aligned}
\tag{6.39}
$$

The first sum in (6.39) vanishes due to telescoping summation and the fact that $\mathbb{E}\left[\tilde{M}(\beta_1)\right]=0$. Note also that $\mathbb{E}\left[\tilde{M}(\beta_{n+1})\mathbb{I}(\tilde{\tau}_{SU}\geqslant\beta_{n+1})\right]=0$ because $\mathbb{I}(\tilde{\tau}_{SU}\geqslant\beta_{n+1})=0$. Now, we have to show

$$
\sum_{i=1}^{n-1}B_i\geqslant\alpha(n+1)P(V(\tilde{\tau}_{SU})=n_0)\quad\text{for}\quad n_0\leqslant f(n),
\tag{6.40}
$$

where

$$
B_i=\left(\frac{\beta_{i+1}}{1-\beta_{i+1}}-\frac{\beta_i}{1-\beta_i}\right)\mathbb{E}\left[\tilde{M}(\beta_{i+1})\mathbb{I}(\tilde{\tau}_{SU}\geqslant\beta_{i+1})\right],\ i\leqslant n-1.
$$

First, note that all B_i are non-negative for all $i\leqslant n-1$, due to Lemma 29. Moreover, note that

$$
\tilde{\tau}_{SU}=\beta_n\iff V(\beta_n)=n_0
$$

under the Dirac distribution of $(p_i)_{i\in I_1}$ and consider the last summand

$$
\begin{aligned}
B_{n-1}&=\left(\frac{\beta_n}{1-\beta_n}-\frac{\beta_{n-1}}{1-\beta_{n-1}}\right)\mathbb{E}\left[\tilde{M}(\beta_n)\mathbb{I}(\tilde{\tau}_{SU}\geqslant\beta_n)\right] \\
&=\left(\frac{\beta_n}{1-\beta_n}-\frac{\beta_{n-1}}{1-\beta_{n-1}}\right)\mathbb{E}\left[\left(\frac{n_0}{\beta_n}-n_0\right)\mathbb{I}(\tilde{\tau}_{SU}=\beta_n)\right] \\
&=\left(1-\frac{\beta_{n-1}}{1-\beta_{n-1}}\frac{1-\beta_n}{\beta_n}\right)n_0P\left(V(\tilde{\tau}_{SU})=n_0\right) \\
&=\frac{n+3}{2(n+1)}n_0P\left(V(\tilde{\tau}_{SU})=n_0\right)\geqslant\alpha(n+1)P(V(\tilde{\tau}_{SU})=n_0)
\end{aligned}
$$

if $n_0\geqslant f(n)=\frac{2\alpha(n+1)^2}{n+3}$. This completes the proof of part (a).

(b) The second part follows immediately from (a) and from the formula for FDR of SU procedure based on the set of critical values $(\beta_i)_{i\leqslant n}$ under the reverse martingale model:

$$\mathrm{FDR} = \frac{\alpha n_0}{n+1} + \frac{1-\alpha}{n+1}\mathbb{E}\left[V\right]. \qquad \square$$

Appendix. Examples of martingale models

The family of (super-)martingales is a rich class of models which is briefly reviewed below. In this section we present a couple of examples. Further examples can be found in Heesen and Janssen [14] and Benditkis [1]. For convenience let us describe the model in this section by distributions P on $[0,1]^n$, where the coordinates $(p_1,...,p_n) \in [0,1]^n$ represent p-values. We restrict ourselves to martingale models

$$\left([0,1]^n, P, (\mathcal{F}_t^T)_{t\in T}\right). \tag{6.41}$$

Let $\mathcal{M}_{I_0,I_1}^T$ denote the set of martingale models P on $[0,1]^n$ for fixed portion $I_0 \neq \emptyset,\ I_1$ of $\{1,...,n\}$ and $\{0\} \subset T \subset [0,\eta]$ for some $0 < \eta < 1$.

Obviously, there is a one to one correspondence between martingales and reverse martingales via the transformation

$$\tilde{p}_i = 1 - p_i,\ \tilde{T} = \{1-t : t \in T\},\ \mathcal{G}_t^{\tilde{T}} := \sigma(\mathbb{I}(p_i \geqslant s),\ s \geqslant t,\ s,t \in \tilde{T}) \tag{6.42}$$

of (6.41). Note also that $(p_i)_{i\in I_0}$ follow special copula models if each $p_i,\ i \in I_0$, is uniformly distributed on $(0,1)$.

To warm up consider first some useful elementary examples which will be combined below.

Example 32. *(a) (Marshall and Olkin type dependence (see Marshall and Olkin [16])) Let $X_1,...,X_n$ be i.i.d. continuous distributed real random variables and Y be a continuously distributed real random variable independent of $X_1,...,X_n$. Consider $Z_i := \min(X_i,Y)$ and $\tilde{Z}_i := \max(X_i,Y)$ for $1 \leqslant i \leqslant n$. The transformed true p-values $p_i := H(Z_i)$ and $\tilde{p}_i := \tilde{H}(Z_i),\ i = 1,...,n$ fulfil the martingale property, reverse martingale property, respectively. Thereby, H and $\tilde{H}$ are distribution functions of Z_i and $\tilde{Z}_i,\ i \leqslant n$.*

(b) (Block models) Suppose that the index set

$$\{1,...,n\} = \sum_{j=1}^{k}(I_{0,j} + I_{1,j})$$

splits in k disjoint portions of $I_{0,j}$ the true and $I_{1,j}$ false null. Let $U_1,...,U_k$ denote i.i.d uniformly distributed random variables on $(0,1)$. Suppose that $\left((U_1,(p_i)_{i\in I_{1,1}}),(U_2,(p_i)_{i\in I_{1,2}}),...,(U_k,(p_i)_{i\in I_{1,k}})\right)$ are independent martingale models of dimension $|I_{1,j}|+1$ for $j \leqslant k$. The U's can be duplicated by the definition

$$p_i = U_j \text{ if } i \in I_{0,j},\ I_0 = \sum_{j=1}^{k} I_{0,j},$$

and we arrive at a martingale model where $(p_i)_{i \in \sum_{j=1}^{k} I_{1,j}}$ *are already defined.*

Let us summarize further results.

Example 33. *(a)* $M^T_{I_0,I_1}$ *is closed under convex combinations.*

(b) New martingale models can be produced by stopped martingales via stopping times and the optional switching device, see Heesen and Janssen [14], p.685.

(c) (Martingales and financial models) Let $T \subset [0,\eta]$, $\eta < 1$ *be a set with* $0 \in T$. *Introduce the price process*

$$X_t : [0,1)^n \to [0,\infty),\ X_t(p_1,...,p_n) = \sum_{i \in I_0} \left(\frac{\mathbb{I}(p_i \leqslant t) - t}{1-t} + K \right),\ t \in T$$

on T *for some constant* $K \geqslant \max(\frac{s}{1-s})$, $s \in T$. *Then the process* $t \mapsto X_t$ *can be viewed as a discounted price process for time points* $t \in T$. *The existence of martingale measures for* $(X_t, \mathcal{F}_t)_{t\in T}$ *on the domain* $[0,1]^n$ *is well studied in mathematical finances, see Shiryaev [24]. When the parameter set* T *is finite it turns out that the space of probability measures on* $[0,1]^n$, *making that process to be a martingale, is of infinite dimension.*

(d) (Super-martingales) It is well known that the process

$$\sum_{i \in I_0} \left(\frac{\mathbb{I}(p_i \leqslant t) - t}{1-t} \right) = M_t + A_t,\ t \in T \tag{6.43}$$

admits a Doob-Meyer decomposition given by a $(\mathcal{F}_t)_{t\in T}$ *martingale* $t \mapsto M_t$ *and a compensator* $t \mapsto A_t$ *which is predictable with* $A_0 = 0$. *Note that (6.43) is a supermartingale if* $t \mapsto A_t$ *is non-increasing.*

Acknowledgements

The authors are grateful to Helmut Finner, Veronika Gontscharuk, Marsel Scheer, Philipp Heesen and Marc Ditzhaus for many stimulating discussions.

References

[1] Benditkis, J. (2015) Martingale Methods for Control of False Discovery Rate and Expected Number of False Rejections. *Dissertation. Heinrich Heine University Duesseldorf* http://docserv.uni-duesseldorf.de/servlets/DocumentServlet?id=35438

[2] Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc.* **B**, **57**, 289–300. MR1325392

[3] Benjamini, Y., Krieger, A. and Yekutieli, D. (2005) Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, **93**, 491–507. MR2261438
[4] Benjamini, Y. and Liu, W. (1999) Step-down multiple hypothesis procedure that controls the false discovery rate under independence. *J. Statist. Planng Inf.*, **82**, 163–170. MR1736441
[5] Benjamini, Y. and Yekutieli, D. (2001) The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.*, **29**, 1165–1188. MR1869245
[6] Blanchard, G. and Roquain, E. (2008) Two simple sufficient conditions for FDR control. *Elect. J. Statis.*, **2**, 963–992. MR2448601
[7] Dickhaus, T. (2008) False discovery rate and asymptotics. *Dissertation, Heinrich-Heine-Universität Düsseldorf.*
[8] Finner, H., Dickhaus, T. and Roters, M. (2009) On the false discovery rate and an asymptotically optimal rejection curve. *Ann. Statist.*, **37**, 596–618. MR2502644
[9] Finner, H. and Roters, M. (2001) On the false discovery rate and expected type I errors. *Biom. J.*, **43**, 985–1005. MR1878272
[10] Finner, H. and Roters, M. (2002) Multiple hypotheses testing and expected number of type I errors. *Ann. Statist.*, **30**, 220–238. MR1892662
[11] Gavrilov, Y., Benjamini, Y. and Sarkar, S. K. (2009) An adaptive step-down procedure with proven FDR control. *Ann. Statist.*, **37**, 619–629. MR2502645
[12] Gontscharuk, V. (2010) Asymptotic and exact results on FWER and FDR in multiple hypotheses testing. *Dissertation, Heinrich Heine Universität Düsseldorf.*
[13] Heesen, P. (2014) Adaptive step up tests for the false discovery rate (FDR) under independence and dependence *Dissertation. Heinrich Heine University Duesseldorf* `http://docserv.uni-duesseldorf.de/servlets/DocumentServlet?id=33047`
[14] Heesen, P. and Janssen, A. (2015) Inequalities for the false discovery rate (FDR) under dependence. *Electron. J. Statist.*, **9**, **1**, 679–716. MR3331854
[15] Heesen, P. and Janssen, A. (2016) Dynamic adaptive multiple tests with finite sample FDR control. *J. Statist. Plann. Inference*, **168**, 38–51 MR3412220
[16] Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution. *J. Amer. Stat. Assoc.* **62(317)** 30–44. MR0215400
[17] Pena, E. A., J. D. Habiger, and W. Wu (2011). Power-enhanced multiple decision functions controlling family-wise error and false discovery rates. *Ann. Statist.*, **39(1)**, 556–583. MR2797856
[18] Sarkar, S. K. (2002) Some results on false discovery rate in stepwise multiple testing procedures. *Ann. Statist.*, **30**, 239–257. MR1892663
[19] Sarkar, S. K. (2008) On methods controlling the false discovery rate. *Sankhya: The Indian Journal of Statistics*, **Series A**, 135–168. MR2551809
[20] Scheer, M. (2012) Controlling the number of false rejections in multiple hypotheses testing. *Dissertation, Heinrich Heine Universität Düsseldorf.*

http://docserv.uni-duesseldorf.de/servlets/DocumentServlet?id=23691

[21] Shorack, G. P. and Wellner, J. A. (1986) Empirical Processes with Applications to Statistics. *John Wiley and Sons. Inc., New York* . MR0838963

[22] Storey, J. D. (2002) A direct approach to false discovery rates. *J. R. Statist. Soc.* **B**, **64**, 479–498. MR1924302

[23] 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. R. Statist. Soc.* **B**, **66**, 187–205. MR2035766

[24] Shiryaev, Albert N (1999) Essentials of Stochastic Finance: Facts, Models, Theory. *World Scientific, 1999* MR1695318